MME 2016

34th International Conference

Mathematical Methods in Economics

September 6th – 9th, 2016, Liberec, Czech Republic

CONFERENCE PROCEEDINGS

34th International Conference

Mathematical Methods in Economics

MME 2016

Conference Proceedings

Liberec, Czech Republic
September 6th – 9th, 2016

ISBN 978-80-7494-296-9

Faculty of Economics, Faculty of Mechanical Engineering
Technical University of Liberec

Programme Committee

doc. RNDr. Helena Brožová, CSc.
doc. Ing. Jan Fábry, Ph.D.
prof. RNDr. Ing. Petr Fiala, CSc., MBA
doc. Mgr. Ing. Martin Dlouhý, Dr. MSc.
doc. Ing. Jana Hančlová, CSc.
prof. Ing. Josef Jablonský, CSc.
Ing. František Koblasa, Ph.D.
doc. RNDr. Ing. Ladislav Lukáš, CSc.
doc. Dr. Ing. František Manlig
prof. RNDr. Jan Pelikán, CSc.
doc. Dr. Ing. Miroslav Plevný
prof. RNDr. Jaroslav Ramík, CSc.
Ing. Karel Sladký, CSc.
Ing. Eva Šlaichová, Ph.D.
doc. Ing. Tomáš Šubrt, Ph.D.
doc. RNDr. Jana Talašová, CSc.
prof. RNDr. Milan Vlach, DrSc.
prof. RNDr. Karel Zimmermann, DrSc.
prof. Ing. Mikuláš Luptáčik, Ph.D.
prof. Dr. Sigitas Vaitkevičius
prof. Ing. Miroslav Žižka, Ph.D.

Editors

Aleš Kocourek
Miroslav Vavroušek

Technical editor

Miroslav Vavroušek

Cover layout

Aleš Kocourek

Contents

The Role of Productivity in the Czech Republic: Does Balassa-Samuelson Effect Still Exist?

Emil Adámek[1]

Abstract. The Purchasing Power Parity (PPP) theory claims that price level (or its changes in the case of relative version of PPP) is the only determinant of the development of exchange rate. Nevertheless, it does not provide good empirical results, especially in the case of transitive economies. One of possible explanation is the existence of Balassa-Samuelson (BS) effect. It is one of the theoretical approaches which try to explain the development of exchange rate and price level. It proposes that faster growth of productivity leads to appreciation of domestic currency and/or to growth of domestic price level. The aim of this paper is to assess the role of BS effect in the Czech Republic. Nevertheless, this approach includes some strong assumptions such as validity of PPP theory for tradeable goods or, in the case of the "domestic" BS effect, the assumption that wages tend to be equalized across tradable and non-tradable sectors. In this paper, the significance of BS effect is tested even if some of these conditions are not valid. We use quarterly data (2000 – 2015) in regression analysis. The Ordinary Least Squares (OLS) method is used to assess the influence of productivity on inflation rate.

Keywords: balassa-samuelson effect, price level, inflation rate, exchange rate, productivity, regression analysis, czech republic

JEL Classification: C22, E52, E58
AMS Classification: 62-07

1 Introduction

Balassa-Samuelson (BS) effect was described in [1] and [14]. According to this approach the prices of tradable goods and services are determined by arbitrage opportunity, so the relative version of PPP is valid for them. On the other hand PPP is not valid in case of non-tradable goods and services (because arbitrage is not possible). The theory states that faster growth of productivity in tradable sector leads to faster growth of prices in non-tradable sector.

Nevertheless, this effect seems to be almost insignificant in the case of developed countries, because the productivity growth difference between two developed countries is negligible. On the other hand, transitioned economies have experienced rapid productivity growth. Moreover, the growth in the tradable sector exceeded the growth in the non-tradable one. That is why the discussions about BS effect related to transition countries has arisen in nineties. A lot of studies try to estimate its influence in the case of Central European (CE) or Central and East European (CEE) countries. The aim of this paper is to assess the impact of BS effect in the Czech Republic.

2 Theoretical Backgrounds

2.1 Computing BS effect

While computing BS effect, at first, the price levels (p) are decomposed into traded (p^T) and non-traded (p^N) components in domestic and foreign (*) countries:

$$p_t = \alpha p_t^T + (1 - \alpha) p_t^N, \tag{1}$$

$$p_t{}^* = \alpha^* p_t^{T^*} + (1 - \alpha^*) p_t^{N^*}, \tag{2}$$

where t denotes time and parameter α is share of traded components in consumption basket. The real exchange rate (q) can be expressed as the relative price of goods produced abroad to those produced in domestic country:

[1] VŠB-Technical University of Ostrava, Faculty of Economics, Department of National Economy, Sokolská tř. 33, 701 21 Ostrava 1, Czech Republic, e-mail: emil.adamek@vsb.cz

$$q_t = (e_t + p_t^*) - p_t, \tag{3}$$

where e stands for nominal exchange rate (CZK/EUR in this paper). By substituting equations (1) and (2) into (3), the changes of real exchange can be written as:

$$\Delta q_t = \left(\Delta e_t + \Delta p_t^{T^*} - \Delta p_t^T\right) + \left[(1-\alpha^*)\left(\Delta p_t^{N^*} - \Delta p_t^{T^*}\right) - (1-\alpha)(\Delta p_t^N - \Delta p_t^T)\right]. \tag{4}$$

PPP claims that prices of tradable goods can be expressed as:

$$\Delta p_t^T = \Delta e_t + \Delta p_t^{T^*}. \tag{5}$$

If it was true, the first term on the right side of equation (4) would disappear:

$$\Delta q_t = (1-\alpha^*)\left(\Delta p_t^{N^*} - \Delta p_t^{T^*}\right) - (1-\alpha)(\Delta p_t^N - \Delta p_t^T). \tag{6}$$

Assuming that output of economy (Y) is determined by a Cobb-Douglas production function, it can be expressed as:

$$Y_t^T = A_t^T (L_t^T)^\gamma (K_t^T)^{1-\gamma}, \tag{7}$$

$$Y_t^N = A_t^N (L_t^N)^\delta (K_t^N)^{1-\delta}, \tag{8}$$

where A represents productivity, L stands for labor, K is capital and parameters γ, δ denote labor intensities in tradable and non-tradable sector respectively. Under perfect competition the profit of firm (Π) can be written as:

$$\Pi_t = Y_t^T(L_t^T, K_t^T) + P_{N/T} Y_t^N(L_t^N, K_t^N) - (W_t^T L_t^T + W_t^N L_t^N) - R_t(K_t^T, K_t^N), \tag{9}$$

where $P_{N/T}$ is the ratio of prices in non-tradable to tradable sector, W denotes wages (costs of labor) and R represents interest rate (costs of capital). Function described by equation (9) will be maximized while conditions (10), (11) and (12) are valid:

$$\frac{\partial Y_t^T}{\partial K_t^T} = P_{N/T} \frac{\partial Y_t^N}{\partial K_t^N} = R_t, \tag{10}$$

$$\frac{\partial Y_t^T}{\partial L_t^T} = W_t^T, \tag{11}$$

$$P_{N/T} \frac{\partial Y_t^N}{\partial L_t^N} = W_t^N, \tag{12}$$

By substituting equations (7) and (8) into these conditions, they can be written (in the logarithm form) as:

$$r_t = \log(1-\gamma) + a_t^T - \gamma(k_t^T - l_t^T) = p_{N/T} + \log(1-\delta) + a_t^N - \delta(k_t^N - l_t^N), \tag{13}$$

$$w_t^T = \log(\gamma) + a_t^T + (1-\gamma)(k_t^T - l_t^T), \tag{14}$$

$$w_t^N = p_{N/T} + \log(\delta) + a_t^N + (1-\delta)(k_t^N - l_t^N). \tag{15}$$

By solving equation (13) and substituting results into (14) and (15) (assuming that (14) = (15)), the relative prices in non-traded sector can be expressed as:

$$p_{N/T} = p_t^N - p_t^T = \left\{\delta\left[\log\gamma + \frac{1-\gamma}{\gamma}\log(1-\gamma) - \log\delta - \frac{1-\delta}{\delta}\log(1-\delta) + r_t\left(\frac{1-\delta}{\delta} - -\frac{1-\gamma}{\gamma}\right)\right]\right\} + \frac{\delta}{\gamma}a_t^T - a_t^N. \tag{16}$$

Considering that c, which is term in { } and includes interest rates and factor productivities, is constant; equation (16) can be written as:

$$p_{N/T} = p_t^N - p_t^T = c + \frac{\delta}{\gamma}a_t^T - a_t^N. \tag{17}$$

Equation (17) is the core of all estimated models. Particular forms of models are described in the Section 3.1.

2.2 Review of Empirical Literature

The paper [6] analyses the impact of economic catching-up on annual inflation rates in the European Union with a special focus on the new member countries of CEE. Using an array of estimation methods, he shows that the Balassa–Samuelson effect is not an important driver of inflation rates. Authors [2] investigate the relative price and relative wage effects of a higher productivity in the traded sector compared with the non-traded sector in a two-sector open economy model. They highlight the role of wages and imperfect mobility of working force. The paper [7] studies the Balassa–Samuelson effect in nine CEE countries. Using panel cointegration techniques, they find that the productivity growth differential in the open sector leads to inflation in non-tradable goods. The paper [12] uses time series and panel regression analyses to estimate role of BS effect in four CE countries (the Czech Republic, Slovakia, Poland and Hungary). She finds empirical evidence of BS effect in these countries. Author [5] finds long-term cointegration relationship between BS effect and relative prices in the Czech Republic, Hungary, Poland, Slovakia and Slovenia during transition process.

3 Methodology and Data

3.1 Estimated Models

Models used in this paper are based on equation (17). The assumption that interest rates are constant seems to be strong. That is why all models will be tested with term representing interest rate differential as well (models with interest rate differential are labeled as X1, where X denotes model specification and X = A, B, C). As it was mentioned above, equation (17) presumes also that the wages in tradable and non-tradable sectors ((14) and (15)) are equal. This presumption responds to so called *Baumol-Bowen* effect or *domestic Balassa-Samuelson effect*. This specification corresponding to Model A used in this paper. By substituting (17) into (4), the inflation rate differential in domestic and foreign country can be expressed as:

$$\Delta p_t - \Delta p_t{}^* = \Delta e_t + (1-\alpha)\left(\frac{\delta}{\gamma}\Delta a_t^T - \Delta a_t^N\right) - (1-\alpha^*)\left(\frac{\delta^*}{\gamma^*}\Delta a_t^{T^*} - \Delta a_t^{N^*}\right). \tag{18}$$

Assuming that labor intensities in tradable and non-tradable are the same the Model A can be defined as:

$$(\Delta p^{CZ} - \Delta p^{EA})_t = \beta_1 \Delta e_t{}^{CZ} + \beta_2[(1-\alpha^{CZ})(\Delta a_T^{CZ} - \Delta a_N^{CZ})_t - (1-\alpha^{EA})(\Delta a_T^{EA} - -\Delta a_N^{EA})_t] + \varepsilon_t \tag{19}$$

where *CZ* and *EA* represents the Czech Republic and euro area, β are estimated coefficients and ε is error term. Throughout the paper parameter β_2 stands for BS effect and term in [] is labeled as *bs*. The assumptions of Model A are:

- labor intensity in tradable sector equals labor intensity in non-tradable sector,
- capital is mobile,
- labor is homogenous (in the tradable and non-tradable sectors) and mobile within country, but not internationally,
- PPP holds for tradable goods.

The second form of model (Model B) relaxes the assumption that labor is homogenous and completely mobile within country, therefore wages not necessarily tend to be equal in tradable and non-tradable sector. By computing capital-labor ratio in equations (14) and (15) by substituting into (13), the relative prices can be expressed as:

$$p_{N/T} = p_t^N - p_t^T = c + \frac{1-\delta}{1-\gamma}a_t^T - a_t^N - (1-\delta)(w_t^T - w_t^N). \tag{20}$$

The last term on the right side of equation (20) represents wage differential. By substituting (20) into (4) the relative prices differential can be written as:

$$\Delta p_t - \Delta p_t{}^* = \Delta e_t + (1-\alpha)\left(\frac{1-\delta}{1-\gamma}\Delta a_t^T - \Delta a_t^N\right) - (1-\alpha^*)\left(\frac{1-\delta^*}{1-\gamma^*}\Delta a_t^{T^*} - \Delta a_t^{N^*}\right) + (1-\delta)(1-\alpha)(\Delta w_t^{T*} - w_t^{N*}) - (1-\delta)(1-\alpha)(\Delta w_t^T - \Delta w_t^N), \tag{21}$$

and Model B is given by equation:

$$(\Delta p^{CZ} - \Delta p^{EA})_t = \beta_1 \Delta e_t{}^{CZ} + \beta_2[(1-\alpha^{CZ})(\Delta a_T^{CZ} - \Delta a_N^{CZ})_t - (1-\alpha^{EA})(\Delta a_T^{EA} - \Delta a_N^{EA})_t] +$$

$$\beta_3[(1-\alpha^{EA})(\Delta w_T^{EA}-\Delta w_N^{CZ})_t-(1-\alpha^{CZ})(\Delta w_T^{CZ}-\Delta w_N^{CZ})_t]+\varepsilon_t. \tag{22}$$

where β_3 measures the influence of wages differential and term connected with wages is label as w_{dif}. The assumptions of Model B are:

- labor intensity in tradable sector equals labor intensity in non-tradable sector,
- capital is mobile,
- labor is heterogeneous (in the tradable and non-tradable sectors) and/or not completely mobile within country,
- PPP holds for tradable goods.

The third specification deals with problem of validity of PPP. There are reasons (such as that there is reversible causality between exchange rate and price level or that the theory neglects expectations (see [13]) why PPP does not have to be valid. Authors such as [11] or [15] doubted its validity in the case of the Czech Republic using econometric techniques. If equation (5) does not hold, equation (18) will be as follows:

$$\Delta p_t-\Delta p_t{}^{*}=\Delta p_t{}^{T}-\Delta p_t{}^{T*}+(1-\alpha)\left(\frac{\delta}{\gamma}\Delta a_t^T-\Delta a_t^N\right)-(1-\alpha^*)\left(\frac{\delta^*}{\gamma^*}\Delta a_t^{T^*}-\Delta a_t^{N^*}\right). \tag{23}$$

Possible problem with this specification is that the price differential is explained by the price differential in tradable sector; hence the danger of possible endogeneity arises. Model C is therefore defined as:

$$(\Delta p^{CZ}-\Delta p^{EA})_t=\beta_1(\Delta p_T^{CZ}-\Delta p_T^{EA})_t+\beta_2[(1-\alpha^{CZ})(\Delta a_T^{CZ}-\Delta a_N^{CZ})_t-(1-\alpha^{EA})(\Delta a_T^{EA}--\Delta a_N^{EA})_t]+\varepsilon_t. \tag{24}$$

The assumptions of Model C are:

- labor intensity in tradable sector equals labor intensity in non-tradable sector,
- capital is mobile,
- labor is homogenous (in the tradable and non-tradable sectors) and mobile within country, but not internationally,
- PPP does not hold for tradable goods.

Mentioned above the interest rate differential is added to all 3 specifications of model. It was computed as follows:

$$i_{dif}=\Delta i^{CZ}-\Delta i^{EA}. \tag{25}$$

where i is interest rate and represents additional term in each model. All models were estimated using regression analysis.

3.2 Data

Prices

There are two variables which measure prices in this paper. At first it is aggregate price level (P), which is dependent variable in all forms of model. It is measured as GDP deflator. The other variable measures prices in traded sector (P_T) in Model C and it is defined as Producer Price Index (PPI) in manufactured products. The usage of two different price indexes should lower the danger of endogeneity of model. The higher inflation rate in tradable sector should cause rising of aggregate inflation rate. Data for both variables were gathered form Eurostat Database [8].

Exchange rate

Exchange rate (E) is defined as bilateral nominal exchange rate (CZK/EUR) at the end of period in this paper. Higher values (depreciation) should cause rising of inflation rate. Data were obtained from Eurostat Database [8].

Productivity and Wages

It is important to define tradable and non-tradable sector for variables dealing with productivity and wages. In the literature, there is serious discussion about which sectors are tradable. The lack of available data implies that most authors consider either all industry [5] or [9] or just manufacturing [10] as tradable sector in case of transition countries. Manufacturing is considered as tradable sector for the Czech Republic in this paper for both productivity and wages. The remaining sectors are considered as non-tradable.

Productivity (A) is defined as value added and divided by number of employees in relevant sector. Wages (W) are computed as ratio of wages and salaries to number of employees in relevant sector. Increase in both productivity and wages should lead to rising of inflation rate differential. Both series were sessional adjusted using Census X12 technique. Data were gathered from Eurostat Database [8] based on "Nomenclature générale des Activités économiques dans les Communautés Européennes" (NACE) classification.

Interest rates

One of assumption of traditional approach to BS effect is that term c in equation (17) is constant. Some authors, such as [12] try to improve their models including interest rate differential into that equation. Because both Czech National Bank (CNB) and European Central Bank (ECB) targeting inflation rate; the policy rates (I) are used instead of money market rates in this paper. This variable is the only one which should have negative influence on inflation rate differential. Data (at the end of period) were obtained from ARAD – Data Series System [3] and from ECB – Database [4].

4 Results

4.1 Results of Unit Root test

Since a regression analysis is used in this paper, it is important that all variables should be stationary. That is why results of unit root test are presented in this section. Augmented Dickey-Fuller (ADF) test is used to assess stationarity of all variables. The lag length is based on Schwarz Information Criterion. The results are depicted in Table 1. It is obvious that none of these variables has a unit root and can be used in regression analysis.

variable	**t-statistic**	**probability**
Δe	-7,0098	0,0000
bs	-10,6439	0,0000
w_{dif}	-5,1376	0,0001
i_{dif}	-6,7683	0,0000
$\Delta p_T^{CZ} - \Delta p_T^{EA}$	-5,0596	0,0001
$\Delta p^{CZ} - \Delta p^{EA}$	-5,8363	0,0000

Table 1 Results of the ADF test

4.2 Estimated Models

In this section the outputs of models are presented. Since variables can influence each other with a lag, all models were tested with delays as well. Since the horizon of monetary policy in the Czech Republic come up to six quarters, each variable was lagged up to 6 periods (t = 0, 1, …, 6). Only the best results are depicted in the Table 2. All variables except for interest rate should have positive impact on independent variable. Models labeled with number include monetary policy rates as well. Nevertheless, this variable turned out to be insignificant in all specifications.

variable	**A**	**A1**	**B**	**B1**	**C**	**C1**
Δe (4)	0,2560***	0,2607***	0,1689*	0,1730*	x	x
***Bs* (4)**	**0,1928**	**0,1650**	**0,2369****	**0,2067***	**0,1508**	**0,1210**
i_{dif}	x	-0,0110	x	-0,0122	x	-0,0115
w_{dif} (1)	x	x	0,4023***	0,4074***	x	x
$\Delta p_T^{CZ} - \Delta p_T^{EA}$(2)	x	x	x	x	0,6125**	0,6324***
R^2	0,123	0,145	0,331	0,358	0,112	0,136
DW	1,410	1,422	1,957	1,976	1,431	1,458
F-statistic	3,4943**	2,7717*	8,0815***	6,7050***	3,1493*	2,5766*

Note: Number of lags is in the brackets; values of coefficients are depicted for each variable; OLS method was used in all models; each model includes intercept; *, **, *** denote 10%, 5% and 1% probability respectively.

Table 2 Results of estimated models

As concerns Model A, the impact of BS effect was insignificant. Also, model is significant only at 5% (1% while interest rates are including) significance level. Moreover, the low value of DW-statistic indicates positive autocorrelation.

Model B includes also wages. In this specification BS effect is significant. Moreover, all variables in this specification have a right sign and are significant. The variables explain 33.1 % of variation of inflation rate differential. DW-statistic indicates no first-order autocorrelation in this model.

Model C should be used in the case that PPP is not valid. Nevertheless, results show that only significant variable is prices differential in the tradable sector in this model. Based on these results, Model B seems to be the best. Based on the empirical results, we can say that percentage point increase of relative productivity growth is associated with an increase of 0.24 % of relative inflation rate in the Czech Republic (compared to euro area).

5 Discussion and Conclusions

The aim of this paper was to assess the influence of BS effect in the Czech Republic. Six econometric models were made. We can divide them into three groups based on their assumptions. The Model B seems to provide best results. It claims that growth of inflation rate differential is influenced by nominal depreciation of currency, faster growth in the domestic relative (tradable to non-tradable sector) productivity growth and faster growth in the domestic relative (tradable to non-tradable sector) wages. According to [2], the role of wages and imperfect internal mobility is important.

Based on this model, we can assume that there is a positive relationship between BS effect and relative prices. This founding is consistent with [5], [7] and [12], nevertheless, the size of the BS effect estimated in this paper is lower than in these papers ([12] estimated coefficients of BS effect in size over 2).

Acknowledgements

This paper was financially supported within the VŠB - Technical University SGS grant project No. SP2016/101 “Monetary, Fiscal and Institutional Aspects of Economic Policy in Selected Countries“.

References

[1] Balassa, B.: The Purchasing Power Parity Doctrine: A Reappraisal. *Journal of Political Economy* **72** (1964), 584–596.
[2] Cardi, O., and Restout, R.: Imperfect mobility of labor across sectors: a reappraisal of the Balas-sa-Samuelson effect. *Pôle européen de gestion et d'économie (PEGE). Working paper 16.* 2014
[3] CNB: *ARAD – Data Series System*. [online database]. Praha: Česká národní banka. Available at: http://www.cnb.cz/docs/ARADY/HTML/index.htm. 2015.
[4] ECB: *Key ECB interest rates*. [online database]. Available at: http://www.ecb.europa.eu/stats/monetary/rates/html/index.en.html. 2015.
[5] Égert, B.: Estimating the impact of Balassa-Samuelson effect on inflation and real exchange rate during transition. *Economic Systems* **26** (2002), 1-16.
[6] Égert, B.: Catching-up and inflation in Europe: Balassa–Samuelson, Engel's Law and other culprits. *Economic Systems* **35** (2011), 208-229.
[7] Égert, B., Drinec, I., and Lommatzschd, K.: The Balassa–Samuelson effect in Central and Eastern Europe: myth or reality? *Journal of Comparative Economics* **31** (2003), 552-572.
[8] EUROSTAT: *Eurostat - Database*. [online database]. Luxembourg: European Commision. Available at: http://ec.europa.eu/eurostat/data/database. 2015.
[9] Fisher, C.: Real currency appreciation in accession countries: Balassa-Samuelson and investment demand. *Review of World Economics* **140** (2004), 179-210.
[10] Jazbec, B.: Real Exchange Rates in Transition Economies. *William Davidson Working Paper, 482*. 2002.
[11] Kim, B., and Korhonen, I.: Equilibrium exchange rates in transition countries: Evidence from dy-namic heterogeneous panel models. *Economic Systems* **29** (2005), 144-162.
[12] Lojschová, A.: Estimating the impact of the Balassa-Samuelson effect in transition economies. *Reihe Ökonomie/Economic Series, 140*. 2003
[13] Mandel, M., and Tomšík, V.: *Monetární ekonomie v malé otevřené ekonomice*. 2. vyd. Praha: Management Press. 2008.
[14] Samuelson, P.: Theoretical Notes on Trade Problems. *Review of Economics and Statistics* **46** (1964), 145–154.
[15] Sadoveanu, D., and Ghiba, N.: Purchasing Power Parity: Evidence from Four CEE Countries. *Jurnal of Academic Research in Economics* **4** (2011), 80-89.

Italian Saving Banks efficiency, is unity strength? Bank groups versus stand-alone

Simona Alfiero[1], Filippo Elba[2], Alfredo Esposito[3], Giuliano Resce[4]

Abstract. This study investigates the Italian SBs sector efficiency over the 2012–2013 period. The measure of efficiency is estimated via SBM Data Envelopment Analysis. In the first stage, we evaluate the SBs efficiency while in the second we compare the performance of SBs belonging to bank groups with those stand-alone. To evaluate the impact of be part of a bank group we use Policy Evaluation tools, performing an impact evaluation with the controlled by a group considered as the *"treatment"* variable and checking for relevant banking ratios. To deal with self-selection bias (heterogeneity in treatment propensity related to variables), we use the PS Matching estimating the average treatment effects with the Ichino-Becker propensity scores. The research novelty resides in the combined application of DEA and Policy Evaluation tools for the specific field. Results show that when comparing SBs belonging to a banks group with stand-alone SBs, although a positive but not significant ATT, we find no relevant differences between the SBs part of group and the stand-alone. However, with reference to Technical Efficiency the stand-alone SBs experience the worst performance while after an insight into the inefficiency decomposition is clear that difficulties are due to managerial inefficiency. Finally, we present speculation, linked to real circumstances, with respect to the Italian SBs sector.

Keywords: Saving Banks, Efficiency, Data Envelopment Analysis, Policy Evaluation, Average Treatment Effect on the Treated

JEL Classification: G32 - G21 - C61 - C21 - O12
AMS Classification: 30C40 - 03B48 - 32S60

1 Introduction

The banking sector is still playing a crucial role in countries economics and Saving Banks (SBs), part of it, are fundamental in the economics developments as in [24]. The global crisis and the weak economic reestablishment belong to what Hansen [26] defined during the Great Depression as the economic *"secular"* stagnation being, in such periods, the efficiency pursue vital in supporting local economies mainly composed of SMEs and families. During the years, SBs evolved into full-service banks and JSCs (joint stock companies) and currently they are full commercial competitors that maximize profits being indistinguishable from other competitors as per [24]. Focusing on Italy, local SBs had a great influence on small territories while being the driver of local economies development.

With respect to the present framework, the aim of our study is to investigate whether, and to what extent, belong to a banking group improves the performance in terms of efficiency.

2 Theoretical background and a brief note on the Italian SBs sector

Studies on the Italian banking sector efficiency are few such as the Favero-Papi [21] and Resti [33] about the determinants as scale inefficiencies and regional disparities and one on Italian cooperative banks as per [6]. Casu and Girardone [14] compared results obtained in different countries focusing on commercial banks. Since the work of Sherman and Gold [37], considered the first on banking industry, scholars refers to the DEA technique as a useful tool to measure the relative efficiency of banks as in [13] and [22]. Berger and DeYoung [12] and Williams [42] shows that a low levels of efficiency lead to lack of credit monitoring and inefficient control of operating expenses. With respect to the role of management and efficiency, from a managerial perspective, the *"agency theory"* as in [20] and [29] is relevant when referring to the separation of ownership and control thus implying conflicts of interests between managers and shareholders. On the other hand, the *"stewardship theory"*

[1] University of Torino, Department of Management, simona.alfiero@unito.it
[2] University of Firenze, Department of Science for Economics and Business, filippo.elba@unifi.it
[3] University of Torino, Department of Management, alfredo.esposito@unito.it
[4] University of Roma 3, Department of Economics, giuliano.resce@uniroma3.it

argues that managers are reliable and there are no agency costs as in [18] and [19]. This latest viewpoint is typically a pre-financial scandals one within a paradigm of 10/15 years ago. Currently, the banking efficiency is evaluated via financial ratios, parametric and non-parametric approach as in [8] with different results by using divers methods as shown by Cubbin and Tzanidakis [16] where the mean DEA efficiency is higher than OLS analysis. Thanassoulis [38] found that DEA outperforms regression analysis on accuracy of estimates, but regression analysis offers a greater stability of accuracy even if Resti [32] shows that, both scores do not differ substantially. The Basel Committee [7] states that frontier efficiency measures are better representative in capturing the concepts of *"economic efficiency"*, *"overall economic performance"*, and *"operating efficiency"* providing a superior understanding on corporate governance topics. The main reason for the use of DEA technique is the smallest number of observations requirement and because it takes into account multiple inputs and outputs simultaneously, compared to ratios where one input relates to one output at each time as per [38].

Italian SBs started operating in the nineteenth century as institutions in which the credit and social aspects were living together rooted in the social and economic goals. However, at the end of the twentieth century due to sectoral regulatory developments they turn into full joint stock companies operating under complete commercial principle taking into consideration both stakeholder values and shareholder value. The Bank of Italy Report for 2014 [3]) states that in 2008-2014 period the bank employees and branches decreased by about 17,900 (-5.6%) and 3,400 (-9%) units because of distribution channels. As contribution to sector, Italian SBs at the end of 2014 are forty and holds 4,345 branches, more than 36 thousand employees, total assets of 206.2 billion and 144.4 billion in direct deposits. Currently, groups (which partly belongs to the local territorial Foundations) own SBs to cover the link between the ancient SBs and local communities. The SBs nature most of the time is still reflecting s the local political powers variable influence.

Recently, the Italian Government [26] after some financial scandals issued the Law Decree 22 November 2015, n. 183 (a.k.a. *"decreto salva banche"*) and chose to terminate four banks while setting up four new substitute entities. It was the *de facto* transfer on local territories of the high socio-economic cost impacts and families and countryside people lost all their money because of the lack of financial skills to understand the nature of the business and risk level (a real fraud via the sale of a huge amount of subordinated debt). Interestingly, two recipient of the decree are SBs and part of our sample as another well-known troublesome (CR Ferrara, CR Chieti and Carige). Currently, the SBs sector is mainly hold by bank groups as result of the last 10/12 years mergers and takeovers, which led SBs to the loss of independency and links to territories.

3 Data specifications and methodology

The analysis covers a 36 Italian SBs sample during the 2012-2013 and the data source is Bankscope [4]. Four SBs were not included due to lacks of data. We have 20 SBs part of a bank group (representing the 55% of the sample) and 16 stand-alone SBs (45%). In the first part of the analysis, we evaluate the efficiency of SBs as per the Tone [38] SBM model while in the second part we compare the performances of the 20 SBs part of a bank group with the performances of 16 stand-alone SBs via policy evaluation tools. Table 1 provides the variables descriptive statistics.

Variables	Type	2012 Min	Max	Std. Dev.	2013 Min	Max	Std. Dev.
Total Assets	Input	799,70	49.322,00	10249,91	709,70	50.162,70	9740
Operating Expenses	Input	17,20	1.187,40	237,193	19,80	1.009,90	195,1604
Impaired Loans	Input	45,70	3.141,00	815,513	60,90	3.895,00	1062,465
Loans	Output	703,40	35.128,10	7082,896	639,10	36.391,90	6844,524
Pre-tax Profit	Output	-205,10	163,20	50,3527	-2.205,80	251,90	380,1767
Customer Deposits	Output	410,50	22.018,80	4241,049	482,40	23.251,80	4396,648

Table 1 The 2012 and 2013 Data in Mln. of € – Source: Bankscope - Bureau Van Dijk (2015)

The most used approach in banking efficiency measurement is the intermediation approach as per [36], which considers the bank as intermediators by transforming the production factors into outputs. The alternative approach is the production approach that uses traditional factors of production such as capital and labour to produce the number of accounts of loan and deposit services. In a mix-approach, deposits are likely to be take into account as both inputs and outputs. Referring to the banking sector, Anouze [2] presented a table showing that the intermediation approach is the most applied approach while the mix-approach represents 1/5 of the studies. The chosen variables able to measure the relative efficiency of the Italian SBs lies within the mix approach and the intermediation one. Total assets, operating expenses and impaired loans are inputs variables. Total assets is a proxy for the bank size, operating expenses (sum of personnel other operating expenses) represents the labour input. Although many take it into the NPLs, we chose to refer to Impaired Loans as a risk proxy as in [33] that need, in any case, a reduction and envelops the potential risks that are likely to hit the banks and the stakeholders. With reference to the outputs side we chose the total loans, customer deposits (considered by many as an input and by other as an output) and pre-tax profits. The deposit variable is widely used as input or output as per [40] with a prevalence for an outputs consideration as in our case, and the pre-tax profits output variable is able to encompass both traditional operational and extraordinary activities that may, economically, impacts on financial statements.

To measure efficiency we use DEA as in [5], a non-parametric technique in a SBM as per the Tone [41] version; hence, t multiplies all the variables. The linear program is the following:

$$\begin{gathered}\min_{t,\lambda,S,Z} \tau = t - P_c Z \\ t + P_y S = 1 \\ x_0 = X\lambda + Z \\ y_0 = Y\lambda - S \\ \lambda \geq 0, S \geq 0, z \geq 0\end{gathered} \tag{1}$$

where x_0 is the vector (3x1) of actual inputs under evaluation, X is the costs matrix (3 x 72) of banks sample (36 times 2 years), Z is the vector (3 x 1) for inputs excesses, y_0 is the vector (3 x 1) of firms outputs under evaluation, Y is the matrix (3 x 72) with the outputs of all the firms sample, S is the vector (3 x 1) with the output slacks, λ is the vector (72 x 1) of intensity and P_c and P_y are the vectors (1 x 3) that contains weights.

The result of the linear program solution (1) is τ, the relative efficiency of the Bank, where (1 - τ) is Technical Inefficiency: the averaged distance from the Constant Returns to Scale frontier, it includes both Managerial and Scale Inefficiency. The first depends directly on the management while the second is due to the dimension of the Bank. The SB is efficient and has an optimal dimension $\tau = 1$. By changing the specification of the problem (1) it allows the measure of Managerial Efficiency τ_{VRS} (adding $e\lambda = t$ to the program) (1) as suggested by [5] and Scale Inefficiency ($\tau_{\text{vrs}} - \tau$). The Managerial Inefficiency (1 - τ_{vrs}) is due to the management, being the averaged distance from the Variable Returns to Scale frontier. The Scale Inefficiency is due to the dimension, so generally in the short run the management does not have enough discretion to fill this gap.

Uncovering the effects of being part of a group on efficiency is arduous because of the merge of a bank does not happens randomly. In order to evaluate this kind of influence, we perform an impact evaluation considering SBs part of a group as the *"treatment"* variable. Our check relies on a set of banking sector ratios measures such as, Impaired Loans to Gross Loans [33], Loans to Customer Deposits [42] and [1], and Interest Income on Loans to Average Gross Loans [23]. Moreover, we add the Cost to Income Ratio [9], [27] and [30]) and the Tier 1 Regulatory Capital [31] and [11]. The last two ratios are significant because of their being a proxy for cost efficiency and a benchmark for the operational efficiency while a high Tier 1 shown by banks during the subprime crisis led to higher returns; hence to a substantial resilience.

To address self-selection bias (heterogeneity in treatment propensity that relates to the variables of outcomes), we refer to the PS Matching as per Rosenbaum and Rubin [34]:

$$P(X) = \Pr(D = 1|X) = E(D|X) \tag{2}$$

where $P(X)$ is the PS: the conditional probability of receiving the treatment that in our case is being part of group)$(D = 1)$ given certain characteristics (X). The variables taken in to account for the PS estimates (X) are the above-mentioned ratios. Taking the PS, the results of the Average Effect of Treatment on Treated (ATT) is the estimation of the Becker-Ichino [10] as follows:

$$ATT = E_{P(X)|D=1}\{E[Y(1)|D = 1, P(X)] - E[Y(0)|D = 0, P(X)]\} \tag{3}$$

where Y (0) is the performance (efficiency) of non-treated and Y (1) is the performance of treated, with the same PS *P(X)*. To derive(3), given(2), we need two hypothesis: the Balancing Hypothesis and Unconfoundedness Hypothesis. The Balancing Hypothesis states that observations with the same PS (*P(X)*) must have the same distribution of observable and unobservable characteristics *(X)*, independent of treatment status (*D*):

$$D \perp X \mid P(X) \tag{4}$$

The Unconfoundedness Hypothesis says that the assignment to treatment is unconfounded given the PS:

$$Y(0), Y(1) \perp X \mid P(X) \tag{5}$$

with $Y(0)$ and $Y(1)$ representing the potential outcomes in the two counterfactual situations of non-treatment and treatment. In our case, the Technical Efficiency and the Managerial Efficiency outcome are the measure of

(1). As robustness checks, to estimate ATT, we use four matching methods: Stratification as per Rosenbaum and Rubin [34], Nearest Neighbour as per Rubin [35], Radius as per Dehejia and Wahba, [17] and Kernel [25].

4 Italian SBs efficiency results and Group versus Stand-alone

The figure 1 highlights the level of the SBM relative efficiency of the SBs part of a group and the stand-alone.

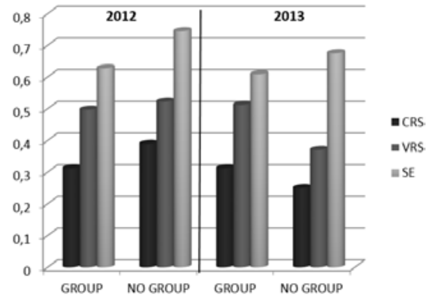


Figure 1 Efficiency of Italian SBs over the 2012 – 2013 period

The figure 2 and table 3 show the Technical Efficiency and inefficiency decomposition. With reference to Technical Efficiency, the stand-alone SBs are experiencing the worst performance while an insight into the inefficiency decomposition clarifies that difficulties are mainly due to managerial inefficiency. In terms of inefficiency the dimensions (Scale) of the stand-alone SBs allows for a better performance.

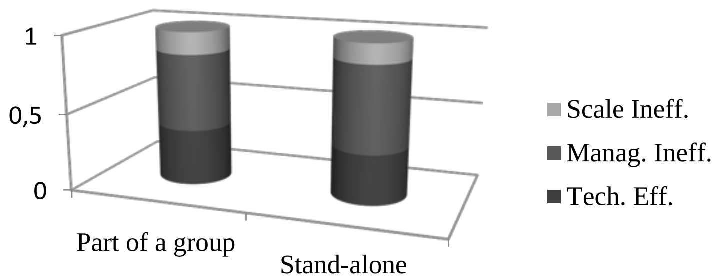


Figure 2 The Technical efficiency and Inefficiencies of Italian SBs over the combined 2012 – 2013 period

SBs efficiency decomposition	Technical Efficiency	Managerial inefficiency	Scale Inefficiency
Part of a group	0,359470231	0,495401225	0,145128544
Stand-alone	0,319707487	0,55351431	0,126778201

Table 2 The SBs efficiency decomposition over the 2012-2013 period

The figure 3 refers to the slacks issues. The SBs part of group outperform respect the others when considering Loans and Customer deposit while the stand-alone SBs outperform with respect to the pre-tax profit and total asset.

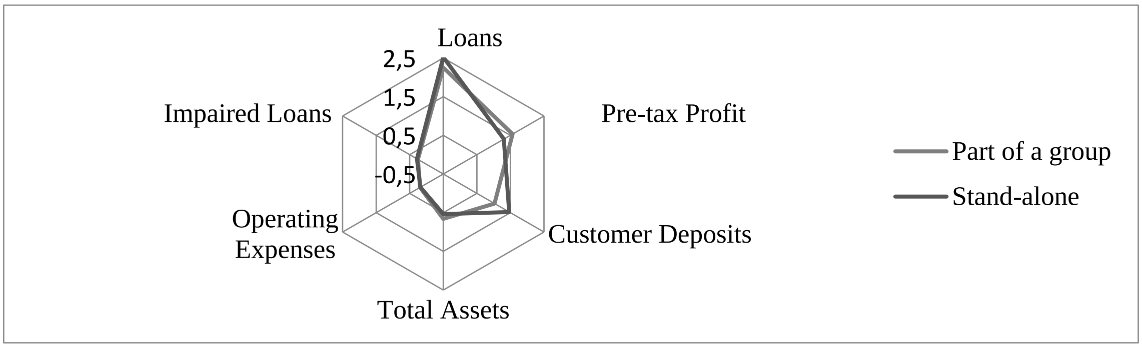


Figure 3 The Slacks (refers to the whole 2012 – 2013 period)

Yet we face the first self-selection problem: the performance difference is due to the Groups policies or to the self stand-alone performance? A merger or a takeover is never a casualty. Hence, we choose to compare the last performances by controlling for a set of variables using the PS Matching approach. Having estimated the PS and following the Becker-Ichino [10] algorithm, we detect the ATT, where the treatment is to be part of a group. As table 3 shows, although the ATT is slightly positive but not particularly significant revealing that the economic side controls shows no significant differences between the 20 SBs part of a bank group and the 16 stand-alone.

Method	Technical Efficiency				Managerial Efficiency			
	ATT	Std. Err.	95% Conf. Interval		ATT	Std. Err.	95% Conf. Interval	
Stratification	0,167	0,132	0,196	0,377	0,271	0,144	0,068	0,616
Nearest neighbour	0,152	0,097	-0,053	0,315	0,286	0,146	-0,147	0,473
Radius	0,089	0,090	-0,076	0,271	0,114	0,089	-0,040	0,293
Kernel	0,123	0,106	-0,087	0,271	0,176	0,136	-0,204	0,342

Table 3 The ATT Results, Table 3 Stand-alone (0) vs Group (1) - Bootstrap statistics100 replications

5 Conclusions

Our study shows the level of relative efficiency of Italian SBs and when consider their belonging to a group or not we observe no significant differences. However, with reference to Technical Efficiency the stand-alone SBs are experiencing the worst performance while after an insight into inefficiency decomposition it becomes clear that difficulties are due to managerial inefficiency. In addition, via policy evaluation tools, we find no relevant dichotomies among SBs belonging to groups with those not part of it. It is important to consider that the driver that leads to the dismissal of power from local influencing policy-maker politicians' people, which leads the Foundations that formally owns the SBs, usually (in the Italian framework), is the results of a financial scandal. According to the mentioned results, statuses are in line with real circumstances such as the Italian Law Decree 2183/2015 - the first application of the bail-in rules, even before the European Bank Recovery and Resolution Directive. Indeed, given the fact that ATT reveals no great differences, currently SBs seems to belong to a bank group or not on the base of their previous financial conditions or because of financial crime scandals that are always exceptional and not ordinary events. Not only a common bad management behaviour drives merger and takeovers but, actually, unrevealed and covered financial crimes, concretely, comparable to exogenous and unpredictable shocks. This is also an indirect confirmation that the mergers and takeovers of last 10/12 by large bank groups were, mainly, hiding rescue actions (widely known in Italy), being illogical the fact that the loss of independency by local political influencers and business groups was due only to low management standards.

5.1 References

[1] Altunbaş, Y., Marqués, D.: Mergers and acquisitions and bank performance in Europe: The role of strategic similarities. *Journal of Economics and Business*, 60(3) (2008), 204–222.

[2] Anouze, A.: Evaluating productive efficiency: comparative study of commercial banks in gulf countries, *Doctoral thesis*, University of Ashton, (2010).

[3] Banca di Italia: Annual Report, https://www.bancaditalia.it/pubblicazioni/relazione-annuale/2014/, (2014).

[4] Bankscope - Bureau Van Dijk, (2015).

[5] Banker, R., Charnes, A., Cooper, W.: Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis, Management science, 30, (1984), 1078–1092.

[6] Barra, C., Destefanis, S., Lubrano Lavadera, G.: Regulation and the crisis: the efficiency of Italian cooperative banks, *Centre for studies in economics and finance*, working paper 38, (2013).

[7] Basel Committee on Banking Supervision: Enhancing corporate governance for banking organisations. February, available at http://www.bis.org/publ/bcbs122.pdf, (2006).

[8] Bauer, P., Berger, A., Ferrier, G., Humphrey, D.: Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods. *Journal of economics and business*, 50(2) (1998), 85–114.

[9] Beccalli, E., Casu, B., Girardone, C.: Efficiency and stock performance in European banking. *Journal of Business Finance & Accounting*, 33.1-2 (2006), 245–262.

[10] Becker, S.O., Ichino, A.: Estimation of average treatment effects based on propensity score. *The Stata Journal* vol., 2,4 (2002), 358–377.

[11] Beltratti, A, Stulz, R.; Why did some banks perform better during the credit crisis? A cross-country study of the impact of governance and regulation. *Charles A. Dice Cent.* Work. Pap. 2009-12 (2009).

[12] Berger, A.N., De Young, R.: Problem loans and cost efficiency in commercial banking. *Journal of banking and finance* 21, (1997), 849-870.

[13] Berger A.N., Humphrey, D.B.: Efficiency of financial institutions international survey and directions for future research. *European journal of operational research*, 98 (1997), 175–212.

[14] Casu, B., Girardone, C.: A comparative study of the cost efficiency of Italian bank conglomerates. *Managerial finance*, 28(9) (2002), 3–23.
[15] Charnes, A., Cooper, W., Rhodes, E.: Measuring the efficiency of decision-making units. *European Journal of Operational Research*, vol., 2 (1978), 429–444.
[16] Cubbin, J., Tzanidakis, G.: Regression versus data envelopment analysis for efficiency measurement: an application to the England and wales regulated water industry. *Utilities policy*, 7(2) (1998), 75–85.
[17] Dehejia, R.H., Wahba, S.: Propensity Score Matching Methods for Nonexperimental Causal Studies. *The Review of Economics and Statistics*, 84,1 (2002), pp.151-161.
[18] Donaldson, L.: The ethereal hand: organizational economics and management theory. *Academy of management review*, 15 (1990), 369–381.
[19] Donaldson, T., Preston, L.E.: The stakeholder theory of the corporation: concepts evidence, and implications. *Academy of management journal*, 20 (1995), 65–91.
[20] Eisenhardt, K.: Agency theory: an assessment and review. *Academy of management review*, (1989) 57–74.
[21] Favero, C.A., Papi, L.: Technical efficiency and scale efficiency in the Italian banking sector: a non-parametric approach. *Applied economics*, 27.4 (1995), 385–395.
[22] Fethi, M.D., Pasiouras, F.: Assessing bank efficiency and performance with operational research and artificial intelligence techniques:a survey. *European journal of operational research*, n. 2 (2010), 189–98.
[23] Foos, D., Norden, L., Weber, M.: Loan growth and riskiness of banks. *Journal of Banking & Finance*, 34(12) (2010), 2929–2940.
[24] Gardener, E.P.M., Molyneux, P., Williams, J., Carbo, S.: European savings banks: facing up the new environment. *International journal of banking marketing*, vol. 15(7) (1997), 243–254.
[25] Heckman, J.J., Smith, J., Clements, N.: Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts. *The Review of Economic Studies*, 64 (1997), 487–535.
[26] Hansen, A.H.: Economic progress and declining population growth. *The American economic review*, 29, 1, (1939), pp. 1-15.
[27] Hess, K., Francis, G.: Cost income ratio benchmarking in banking: a case study. *Benchmarking: An International Journal*, 11(3), (2004), 303–319.
[28] Italian Government, http://www.gazzettaufficiale.it/eli/id/2015/11/23/15g00200/sg, (2015).
[29] Jensen, M.C., Meckling, W.: Theory of the firm: Managerial behavior, agency costs and capital structure. *Journal of Financial Economics*, 3, 4, (1976) pp. 305–360.
[30] Mathuva, D.M.,: Capital adequacy, cost income ratio and the performance of commercial banks: the Kenyan scenario. *The International Journal of Applied Economics and Finance*, 3.2 (2009), 35–47.
[31] Merle, J.: An Examination of the Relationship Between Board Characteristics and Capital Adequacy Risk Taking at Bank Holding Companies. *Academy of Banking Studies Journal*, 12.1/2 (2013):3A.
[32] Resti, A.: Evaluating the cost efficiency of the Italian banking system: what can be learned from the joint application of parametric and nonparametric techniques. *Journal of banking & finance*, 21 (1997), 221–250.
[33] Resti, A., Brunella, B., Nocera, G.: The credibility of European banks' risk-weighted capital: structural differences or national segmentations? *BAFFI CAREFIN Centre Research Paper*, 2015, 2015-9.
[34] Rosenbaum, P.R., Rubin, D.B.: The central role of the propensity score in observational studies for casual effects. *Biometrika* vol. 70, 1 (1983), 41–55.
[35] Rubin, D.B.: Matching to remove bias in observational studies. *Biometrics*, Vol. 29, n.1 (1973), pp.159-184.
[36] Sealey Jr., C.W., Lindley, J.T.: Inputs, outputs, and a theory of production and cost at depository financial institutions, *The journal of fin*ance, (1977), 1251–1266.
[37] Sherman, H.D., Gold, F.: Branch operating efficiency: evaluation with data envelopment analysis. *Journal of banking and finance*, no. 2, 9 (1985), 297–315.
[38] Thanassoulis, E.: A comparison of regression analysis and Data Envelopment Analysis as alternative methods for performance assessments. *Journal of the operational research society*, 44,11 (1993), 1129–1144.
[39] Thanassoulis, E., Boussofiane A., Dyson, R.G.: A comparison of Data Envelopment Analysis and ratio analysis as tools for performance measurement. *Omega, international journal of management science*, 24(3) (1996), 229–244.
[40] Toloo, M., Barat, M., Masoumzadeh, A.: Selective measures in data envelopment analysis, *Annals of operations research*, 226.1 (2015), 623–642.
[41] Tone, K.: A Slacks-based Measure of Efficiency in Data Envelopment Analysis. *European Journal of Operational Research* vol., 130 (2001), 498–509.
[42] Williams, J., Gardener, E.: The efficiency of European regional banking. *Regional Studies*, 37(4) (2003), 321–330.
[43] Williams, J.: Determining management behaviour in European banking. *Journal of banking and finance*, 28 (2004), 2427–2460.

Are remittances important for the aggregate consumption growth?

Ondřej Badura[1], Jana Závacká[2]

Abstract. Remittances, though often overlooked, can be an important source of income and thus an important determinant of consumption function even at the aggregate level. It is evident that the importance of remittances for the receiving economy is given not only by the relative amount of emigrants, but also by its economic level. However, the most important factor is the share of remittances on the aggregate income of the receiving country. The aim of this study is to analyze the growth of the aggregate consumption in terms of remittance flows.

To achieve this objective, there is primarily used panel data regression. To avoid great differences in economic conditions across the countries with high share of remittances on their national income, the regression was driven on chosen Asian economies only. The results confirm the statistical significance of the growth of remittances on the aggregate consumption growth in panel regression, the same result was obtained from the simple regression with robust estimations. These findings concurrently support the idea that remittance flows should be more considered.

Keywords: remittances, consumption, Asian economics.

JEL Classification: E21, F24
AMS Classification: 91B42

1 Introduction

Remittances are usually defined as income received by family members abroad. Though it doesn't have to include just family members itself, it is still the most common example. But the really important thing here is the income nature of remittances. There is no doubt in professional economic circles that income is the key determinant of consumption function for a simple household as well as at aggregate level. But income itself it's not or at least it doesn't have to be homogenous, which also implies that different sources of household financial budget may also influence consumption spending differently. By dividing income into few smaller and more uniform parts and by studying this pieces of household budget separately we can possibly find new relationships and patterns and thus to significantly enrich our knowledge of consumer behavior. And this is exactly, what this paper is about. Particularly the aim of this study is to analyze the growth of the aggregate consumption in terms of remittance flows.

When we analyze the macroeconomic aspects of received remittances, there are some factors determining the importance or rather their potential to significantly influence aggregate consumption spending. But all of these factors like relative amount of emigrants (relative to country's population) or economic level of considered country lead to one thing - the share of remittances on the aggregate income of the receiving country. If the share is not high enough, the possibly found pattern may be very interesting and breathtaking, but still couldn't make any significant changes in aggregate consumption behavior. That's why we choose only countries with this share of about 10 % and more. In further regression analysis we focus only on Asian economies just to incorporate their specific macroeconomics characteristics and to enable us to better compare results from individual regressions.

When we mention economies with high share of received remittances on their national income, it's not surprising that we talk mainly about developing countries. For the huge remittances inflows relative to their gross domestic product (GDP) it can be expected that also their effect on the characteristics of income or consumption behavior of local households is enormous, even at aggregate level. This logical assumption is then confirmed by many empirical studies. Of those which aim was to analyze the impact of this type of international transfers on income we can mention for example publications dealing with particular countries such as Hobbs and Jameson [5], who deal with the influence of remittances on the distribution of income and poverty in Nicaragua, Lázaro et

[1] VŠB - Technical University Ostrava, Department of Economics, Sokolská třída 33, 701 21 Ostrava, Czech Republic, e-mail: ondrej.badura@vsb.cz

[2] VŠB - Technical University Ostrava, Department of Systems Engineering, Sokolská třída 33, 701 21 Ostrava, Czech Republic, e-mail: jana.zavacka1@vsb.cz

al [8], who analyze the same issue using data from Bolivia or more general study of Koechlin and Gianmarco [7], where authors prove a clear link between international remittances and income inequality using analyzes of panel data from 78 countries worldwide. Further it can be mentioned Gibson, Mc Kenzie and Stillman [4], who provide another complex view of impact of emigration and remittance flows on income and poverty in developing world discussing primarily difficulties in estimating those effects. The influence of remittances directly on consumption or other characteristics of consumer behavior is then a major goal of papers like Incaltarau Maha [6] analyzing so for Romania's macroeconomic environment, Yousafzai [9] or Ahmad, Kahn and Atif [1] discussing the same problem on the example of Pakistan or Zhu [10] for the data representing the rural areas of China, all finally prove the expected influence (weaker or stronger) of remittances on consumption characteristics. Gallegos and Mora [3] then come up with the long term relationship between remittances, consumption and human development in Mexico. Again, we can mention studies with a range far wider (international), which may be, for example, analysis of panel data of Combes and Sebek [2], which, among other things, points to the fact that the inflow of remittances significantly reduces consumer instability.

All of those studies mentioned above show and prove how remittances can influence aggregate consumption spending (and income distribution). The main goal of this paper is then to bring another little contribution to this field of knowledge doing so for environment of Asian developing world.

The paper is organized as follows. The data description and the model are introduced in Section 2. The estimation results are presented in Section 3. Section 4 concludes.

2 The model

2.1 The data

We have tested the significance of remittances on the consumption growth of the yearly data of Armenia (1996-2014), Bangladesh (1996-2014), Georgia (1997-2014), Jordan (1996-2014), Kyrgyz Republic (1996-2014), Nepal (1996-2014), Pakistan (1996-2014), Philippines (1996-2014), and Sri Lanka (1996-2010). These countries were chosen according to their high share of remittance inflow to the country on GDP and the availability of the data. The share of remittances on GDP expressed in real prices together with the estimated mean value could be seen in Figure 1.[3]

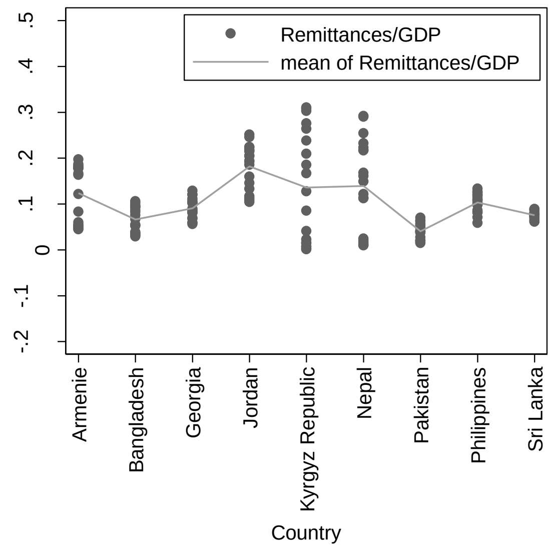


Figure 1 The share of remittances on GDP and its mean value (expressed in real prices) across countries

The data were downloaded from the World bank database. Because of the non-stationarity the data were transformed into their growth rates. All time series used for regression together with the plot of share of consumption on GDP is presented in Figure 3 – 6 in Appendix.

[3] In some cases (for example Kyrgyz Republic) the share of remittances on GDP was quite close to zero in the first years of observations, but its mean value is still relatively high, so even these countries are still suitable for our analysis.

The regression is based on following time series data:

- growth of cpi (*growth_cpi*) – the time series was counted as an annual growth rate of the Consumer price index (expressed to the baseline 2010 = 100);
- growth of consumption (*growth_c*) – the household final consumption expenditure (current US$) were firstly expressed in real prices of year 2010 dividing by consumer price index (expressed to the baseline 2010 = 100), than the annual growth was counted;
- growth of GDP (*growth_gdp*) – the annual growth of real GDP was counted from GDP (current US$) divided by consumer price index (expressed to the baseline 2010 = 100);
- growth of remittances (*growth_remittances*) – the annual growth of received personal remittances (current US$) divided by consumer price index[4] (expressed to the baseline 2010 = 100).

The regression was also controlled for the influence from the changes in short term interest rate, the rate of unemployment and the lagged values of all explanatory variables, however their presence in the model was rejected on the 10% level of significance. As for the unemployment rate, this result was quite expected. It is no secret that there is a strong interdependence of unemployment and development of GDP. Thanks to this correlation it was highly probable that the rate of explained variability in the model would not be significantly increased by adding the unemployment rate. Yet even this step was tested just to be sure, as well as the inclusion of lagged variables which influence was also not expected due to the annual nature of data. In other words, even the delay of one year is probably too much for a consumption function. Short-term interest rate surely should have an impact on the development of household consumption, but this "interest rate channel" is primarily significant in advanced economies. That’s why we attribute the statistical insignificance of the interest rate in our model to the developing nature of the selected countries and thus not fully developed market mechanisms.

2.2 The model

The consumption growth expressed in real prices together with its estimated mean value in initial countries could be seen in Figure 2. We can see that the mean value as well as the variance of the growth rate of consumption differs across the countries, thus the estimation by panel regression was done.

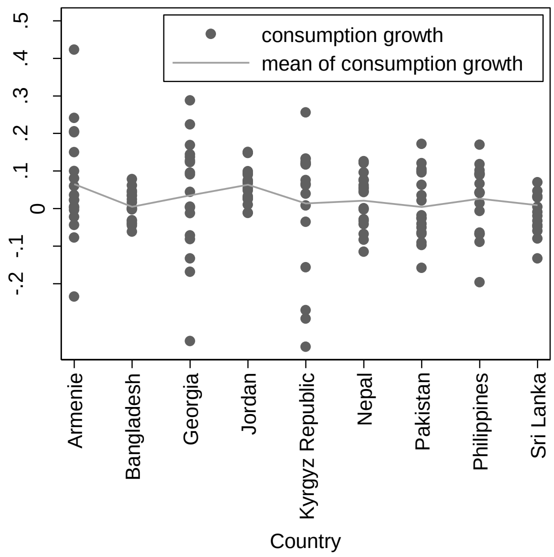


Figure 2 Consumption growth and its mean value across countries

[4] At first it is necessary to convert indicators of consumption and remittances from nominal to real values deflating by the CPI. First, the meaning of this work is to observe the impact of remittances on real household consumption, hence its explanatory variable must be in real terms and secondly there is already contained a development of a price level (through CPI) in remittances in nominal terms, therefore when we include CPI itself among the regressors, it is necessary to deflate the indicator of remittances, otherwise the development of CPI is included twice among the explanatory variables.

Hausmann test rejected the fixed-effect in panel regression on the 5% level of significance, hence the random effects were considered:

$$growth\_c_{it} = \beta_0 + \beta_1 growth\_gdp_{it} + \beta_2 growth\_remittances_{it} + \beta_3 growth\_cpi_{it} + u_{it}^{RE} + \varepsilon_{it}^{RE} \tag{1}$$

$\beta_0, \beta_1, \beta_2, \beta_3$ are model parameters, u_{it}^{RE} the between-entity error and ε_{it}^{RE} within-entity error. Breusch-Pagan Lagrange multiplier test did not reject at 5% level of significance the hypothesis that variances across countries are zero, thus the simple ordinary least square regression (OLS) was done.

$$growth\_c_{it} = \beta_0 + \beta_1 growth\_gdp_{it} + \beta_2 growth\_remittances_{it} + \beta_3 growth\_cpi_{it} + \varepsilon_{it}^{OLS} \tag{2}$$

ε_{it}^{OLS} is the error term from OLS regression. The residuals from the OLS model were not autocorrelated but the homoscedasticity was at the 5% level of significance rejected, hence the robust estimations were used.

3 Estimation results

The estimation results are presented in Table 1 (the parameter estimations are in bold, under the parameter estimations there are standard deviations, the p-value is presented in the braclets). However the random effects in panel regression were rejected, we are presented also the estimations from panel regression for comparison only. We can see that the results are quite stable and the new estimations of standard errors are similar.

Variable	Random-effects GLS regression	OLS regression
growth_gdp	**0.968**	**0.968**
	0.041	0.054
	(0.000)	(0.000)
growth_remittances	**0.028**	**0.028**
	0.006	0.005
	(0.000)	(0.000)
growth_cpi	**0.172**	**0.172**
	0.097	0.093
	(0.076)	(0.066)
const.	**-0.016**	**-0.016**
	0.008	0.007
	(0.045)	(0.027)
Number of observations	157	157
R^2		0.796
Adjusted R^2		0.792

Table 1 Estimation results

The estimation results confirmed at 5% level of singificance the statistical significance of the growth rate of GDP and remittances in the model. The high value of the parameter by GDP growth signalizes the high sensitivity of the consumption growth on the growth of GDP (the increse of GDP growth per one percent point is connected with the increase of consumption growth of 0.968 percent point ceteris paribus). Considering that all countires from the panel are developing countries and the short term interest rate in the model was not significant and hence dropped from the regression, the high value of this parameter only confirms that in these countries the immediate consumption is mainly prefered to savings. This is perfectly in line with the significance of the growth of remittances in the model. However the estimation of parameter of the growth of remittances is only 0.028, its significance prooved the importance of remittances' inflow into these countries and its direct contribution to the consumption growth.

The statistical significance of the inflation (growth of cpi) was confirmed on the 10% level of significance. The positive value of the parameter estimation (0.172) is in accordance with the expected impact of the inflation, e.g. that with the increasing price level consumers rather prefer immediate consumption to savings.

4 Conclusion

The aim of this study was to verify the initial hypothesis that international remittances have an impact on aggregate consumption of households using a growth model. This assumption has been actually confirmed in the regression model for environment of Asia described above. As mentioned in the introduction, from the perspective of macroeconomic impacts it does not make much sense to examine the impact of the inflow of remittances in countries where these flows aren't sufficiently large relative to the production output of the recipient country. That's why the model was driven only on Asian countries with high share of remittances on their GDP, which de facto represents the developing economies. It is necessary to take into account that this fact substantially influences the explanatory ability of the model itself.

The results are not fully general, but valid only for specific economic environment, where international transfers like these play an significant role. First, in developing world remittances are often an important part of household income and through consumption spending they can regularly increase GDP of the economy and so to contribute to real economic convergence. In some case where remittances are also used for investment and savings (which leads to increase in capital) this effect of long term economics growth and development is even raised to power. But it doesn't have to be the rule. If receiving household gets used to this transfers to much, it can lower its nature need to procure the resources itself. In other words, it can lower their labor supply. Another problem lies in potential real appreciation of country's currency, when inflows of remittances are too high. For export oriented economies, which developing countries usually are, this effect could be very dangerous. All of these aspects could be an important and interesting topic for the further scientific research.

In the end it is useful to emphasize the fact, that despite the arguable effect of remittances on economy in long term, there is no doubt, that in short term it reduces poverty significantly. And because this happens mainly through consumption spending (most of remittances received are spend for necessary goods), it is essential to explore this impact channel at the first place. The significance of remittances on consumption growth in this developing world is now confirmed, but the question of its precise nature and deeper understanding of patterns that stand behind, still remains open a thus a potential motive for extension of this work.

Acknowledgements

This paper has been funded by a grant from Students Grant Project EkF, VŠB-TU Ostrava within the project SP2016/112 "Remittances and their influence on macroeconomic and microeconomic indicators" and the Operational Programme Education for Competitiveness – Project CZ.1.07/2.3.00/20.0296.

References

[1] Ahmad, N., Khan, Z. U., and Atif, M.: Econometric Analysis of Income, Consumption and Remittances in Pakistan: Two Stage Least Square Method. *Journal of Commerce* **5** (2013), 1-10.

[2] Combes, J. L., and Ebeke, Ch.: Remittances and Household Consumption Instability in Developing Countries. *World Development* **39** (2011), 1076-1089.

[3] Gallegos, J. L. C., and Mora, J. A. N.: Remittances, Consumption and Human Development: Evidence of Mexico's Dependence. *Análisis Económico* **28** (2013).

[4] Gibson, J., Mc Kenzie, D., and Stillman, S.: Accounting for Selectivity and Duration-Dependent Heterogeneity When Estimating the Impact of Emigration on Incomes and Poverty in Sending Areas. *Economic Development and Cultural Change* **61** (2013), 247-280.

[5] Hobbs, A. W., and Jameson, K. P.: Measuring the Effect of Bi-directional Migration Remittances on Poverty and Inequality in Nicaragua. *Applied Economics* **44** (2012), 2451-2460.

[6] Incaltarau, C., and Maha, L. G.: The Impact of Remittances on Consumption and Investment in Romania. *Journal of European Studies* **3** (2012), 61-86.

[7] Koechlin, V., and Gianmarco, L.: International Remittances and Income Inequality: An Empirical Investigation. *Journal of Economic Policy Reform* **10** (2007), 123-141.

[8] Lazarte, A., et al.: Remittances and Income Diversification in Bolivia's Rural Sector. *Applied Economics* **46** (2014), 848-858.

[9] Yousafzai, T. K.: The Economic Impact of International Remittances on Household Consupmtion and Investment in Pakistan. *Journal of Developing Areas* **49** (2015), 157-172.

[10] Zhu, Y., et al.: Where Did All the Remittances Go? Understanding the Impact of Remittances on Consumption Patterns in Rural China. *Applied Economics* **46** (2014), 1312-1322.

Appendix

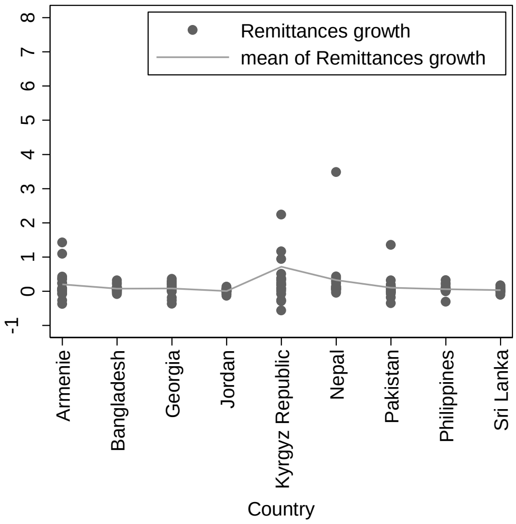


Figure 3 Remittances growth and its mean value across countries

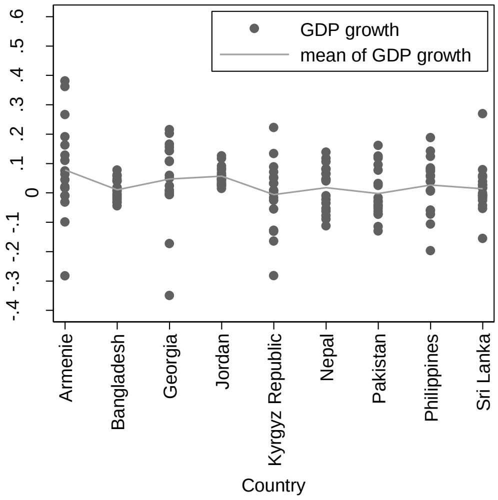


Figure 4 GDP growth and its mean value across countries

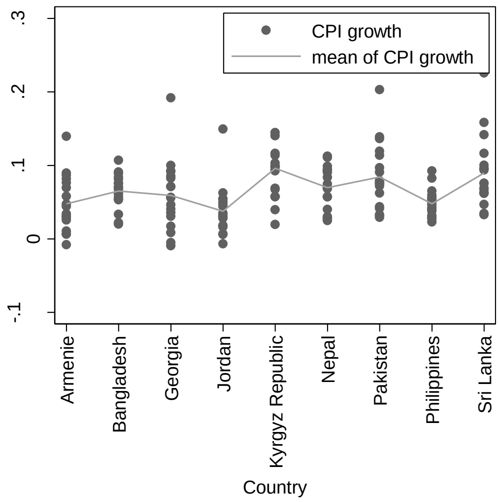


Figure 5 CPI growth and its mean value across countries

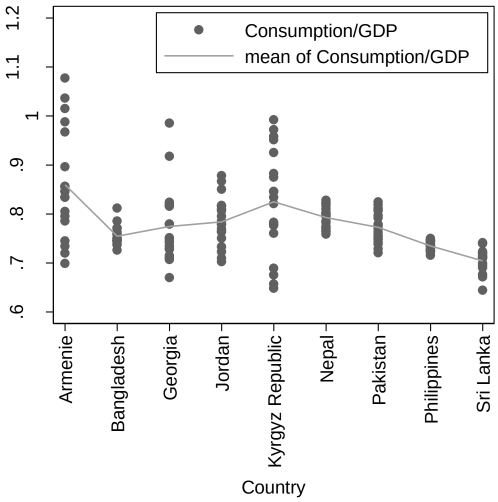


Figure 6 The ratio of consumption on GDP and its mean value across countries

Semantic Model of Organizational Project Structure

Jan Bartoška[1]

Abstract. The article discusses the use of semantic networks and Analytic Network Process (ANP) methods for quantifying a “soft” project structure. International standards of project management describe and set organizational, process and knowledge areas of project management for common practice, which can be designated as the so-called soft project structure. Although the prioritization of project roles in communication or project documentation in project management is a key to the success of a project, a unique approach to quantify the soft structure of a project has not yet been introduced. The semantic networks can be used in project management to illustrate and quantify links between internal and external objects of an environment of a project. Using the ANP method, it is possible to quantify elements even in a soft project structure, i.e. e.g. project roles, project documentation, and project constraints. The author of the paper loosely follows previous research in the application of semantic networks in the case of soft structures. This article aims to propose a procedure for displaying and quantifying the soft structure of a project.

Keywords: Project Management, Semantic Networks, Stakeholder Management, Analytic Network Process.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Project management is a collection of many professional disciplines and skills when, besides traditional quantitative approach, qualitative approach which is the content of international standards is gaining on importance. International standards of project management, i.e. PMBOK® Guide [9], PRINCE2 [6], ISO 21500:2012 [3] and IPMA ICB [1] describe a common practice to set organizational, instrumentation, process and knowledge management aspects of a project, giving rise to the so-called "soft" structure of the project, i.e. a partially ordered set of documents, roles, workflows and tools with variable influence of internal and external environment of the project. Although the prioritization of project roles in communication and project documentation in project management is a key for the success of the project, a unique approach to quantifying the "soft" structure of the project has not yet been introduced. "Soft" structures of the project are usually just displayed in the form of knowledge maps or semantic networks. For potential quantification it is appropriate to focus particularly on the use of semantic networks.

Semantic networks are conventionally utilized in ICT, especially in developing and maintaining software and web applications, e.g., Kim et al. [4] uses a semantic model for viewing and extracting structure tags for creating and managing web sites, or Zhu and Li [18] focus on developing a semantic approach in IT for the resolution of an application and contextual level. The use of semantic networks can lead to the improvement of the systems studied, such an example is a case study in the field of IT by Shing-Han, Shi-Ming, David and Jui-Chang [15]. Semantic networks are also used for expressing the relationships between users of information systems and content web servers, where significance quantification of elements is subsequently conducted with the help of the PageRank method [7]. In a similar manner, but using the ANP method (Analytic Network Process) [13], it is possible to quantify the elements of the "soft" project structure, i.e. e.g. project roles, project documentation and project constraints.

The use of semantic networks in project management is not yet prevalent. For the environment of IT project models, Schatten [14] proposes semantic social networks for the extraction and transfer of knowledge within the project team and stakeholders of the project. Whereas Williams [17] uses a semantic network for the display and analysis of success factors in the structure of mutual expectations and limitations of the project stakeholders. While El-Gohary, Osman and El-Diraby [2] for example indicate that insufficient or inadequate stakeholders' involvement in the management of projects, programs and portfolios is the main reason for failure, and they further propose the semantic model as a tool to visualize and manage relationships and knowledge towards stakeholders in a multi-project environment. Following the example of their use in IT, it is therefore appropriate

[1]Czech University of Life Sciences in Prague, Department of Systems Engineering, Kamýcká 129, Prague, Czech Republic, email: bartoska@pef.czu.cz

to continue developing semantic models to meet the needs of project management, i.e. a management structure of roles, documentation, and knowledge constraints in a project.

This article aims to propose a procedure for displaying and quantifying the soft structure of the project. The specific outcome of the article will be the draft of the general structure of project management according to international standards, using the semantic network, its quantification by means of the ANP method and a subsequent interpretation of feedback to project management practice.

2 Materials and methods

2.1 Project Management Institute – PMBOK® Guide

Project Management Institute (PMI) ranks among the largest non-profit associations in the world, dealing with project management. The association was founded in the United States in 1969. Since the 80s of the 20th century, in a broad discussion among experts and professional public, it has been developing a standard entitled "A Guide to the Project Management Body of Knowledge" (PMBOK® Guide). Since its fourth version, the PMBOK® Guide has been recognized as the standard of the Institute of Electrical and Electronics Engineers (IEEE) in the USA.

The PMBOK® Guide [9] is process-oriented and in its knowledge areas it presents the set of qualitative (e.g. team management, team motivation, etc.) or quantitative approaches (e.g. CPM, EVM, etc.) necessary for project management. The PMBOK® Guide is divided into four parts: introduction (18 pages), environment and project life cycle (28 pages), project management processes (16 pages), knowledge areas (353 pages).The standard defines five process groups (47 in total) for project management and their direct link to the fields of knowledge, 10 in total. Knowledge areas of project management for the PMBOK® Guide are the following[9]: *Project Integration Management*; *Project Scope Management*; *Project Time Management*; *Project Cost Management*; *Project Quality Management*; *Project Human Resource Management*; *Project Communications Management*; *Project Risk Management*; *Project Procurement Management*; *Project Stakeholder Management*. Knowledge areas of the PMBOK® Guide standard [9] represent the substantive content of the project lifecycle processes. The standard thus dissolves project management into atomic transformations where the entry is secured, necessary tasks (expert, professional, managerial, etc.) are performed and output is created. Outputs from one process become inputs of the second.

2.2 PRINCE2 – Projects in Controlled Environments

International project management methodology PRINCE2 originated at Simpact Systems Ltd. in 1975 as a methodology of PROMPT projects management. In 1979, the UK government has accepted PROMPT II methodology as the official methodology for managing IT projects in public administration. In 1989, the methodology was expanded and published under the new name PRINCE - "Projects in Controlled Environments". In 1996, based on the results of international studies in project management (with the participation of more than 120 organizations from around the world), the methodology was developed and released again in the second version as PRINCE2. Although the methodology has since undergone a series of amendments, the name has remained unchanged. For a long time, the owner of the methodology was the UK government (Office of Government Commerce). At present, it is again in the hands of a private company (Axelos Ltd.).

Using the topics [6], the methodology indirectly defines its own knowledge areas, i.e. the areas to be managed across processes, with the help of specific expertise and with a specific goal. The topics of PRINCE2 methodology are as follows [6]: *Business Case*; *Organization*; *Quality*; *Plans*; *Risk*; *Change*; *Progress*. The seven processes which PRINCE2 methodology provides describe the life cycle of a project from the position of roles and responsibilities. All seven PRINCE2 processes can be continuously connected. The linkage shows a full sequence of operations from start to finish of the entire project life cycle. According to PRINCE2, the life cycle processes of the project are as follows [6]: *Starting up a Project*; *Directing a Project*; *Initiating a Project; Controlling a Stage*; *Managing Product Delivery*; *Managing and Stage Boundary*; *Closing a Project*. PRINCE2 methodology processes [6] are based on the interaction of project roles and stakeholders in the organizational structure of the project. Project roles and roles of stakeholders in PRINCE2 are as follows [6]: *Steering Committee*; *Sponsor*; *Senior User*; *Senior Supplier*; *Project Manager*; *Team Manager*; *Change Authority*; *Project Support*; *Project Assurance*.

2.3 Semantic model

A semantic model consists of a semantic network that Mařík et al. [5] define as "natural graph representation", where each node of a graph corresponds to a specific object and each edge corresponds to a binary relation. Mařík et al. [5] further state that "semantic networks can conveniently express the relations of set inclusion and membership in a set, and unique as well as general terms can be represented there". Semantic networks emerged at the end of the 60s of the 20th century. The term "semantic network" was for the first time used by Quillian [10] in his dissertation on the representation of English words. According to Sow [16], semantic networks are used for their ability to provide an easy-to-use system of information representation. The semantic networks are suitable for displaying and expressing vast information resources, management structures and processes.

2.4 Analytic Network Process (ANP)

The ANP (Analytic Network Process) method is the generalization of the AHP (Analytical Hierarchy Process) method. The ANP model reflects and explores the increasing complexity of network structure, where the network is made up of different groups of elements. Each group (cluster) comprises a homogeneous set of elements. Linkages can exist among clusters as well as among the elements. In the ANP model the hierarchy 12], [13] is removed.

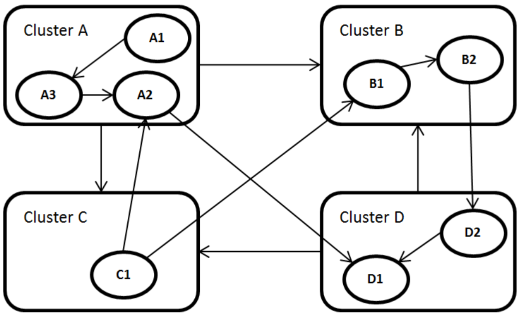


Figure 1 The **s**tructure of elements and clusters in ANP model

The benefit of the ANP method is the ability to express different preferences of links between elements and clusters. To express preferences the method of pair comparison is used. Preferences always occur precisely in assessing the importance of the two elements in terms of the element which refers to them - there rises a question "which of the elements is more important, and by how much". The resulting values for the sub-clusters are then combined into a super matrix where the normalisation of columns is performed [12], [13].

$$W = \begin{array}{c} \\ C_1 \\ C_2 \\ \vdots \\ C_N \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & \dots & C_N \end{array} \\ \begin{bmatrix} W_{11} & W_{12} & \dots & W_{1n} \\ W_{21} & W_{22} & \cdots & W_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ W_{n1} & W_{n2} & \dots & W_{nn} \end{bmatrix} \end{array} \tag{1}$$

Where each block of the super matrix consists of:

$$W_{ij} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix} \tag{2}$$

Under condition:

$$\sum_{i}^{n} w_{ij} = 1, j \in \langle 1, n \rangle \tag{3}$$

For a weighted super matrix (1) for which the relation (3) is valid, a calculation to obtain limit weights of the elements can be performed. The calculation is performed by squaring the weighted super matrix to a sufficiently large number. Since the super matrix has *N*x*N* size, the squaring is always feasible in a trivial manner (matrix multiplication).The result is the approximation of the weighted matrix to the limit matrix. Limit scales can be found in any column of the super matrix. The limit weight of each element expresses the strength of the effect on the overall structure of elements, i.e. it answers the question of how strongly an element affects the other elements [12], [13].

3 Results and Discussion

3.1 Semantic Model of Organizational Project Structure

The Semantic Model of Organizational Project Structure [5] which would be based on the definitions and procedures of international project management standards can be viewed with the help of individual project roles, processes, management practices, parameters (constraints) and documentation. The Organizational Project Structure is comprised of the clusters: Project Roles, Project Constraint, Process Groups, Knowledge Area, and Project Documentation.

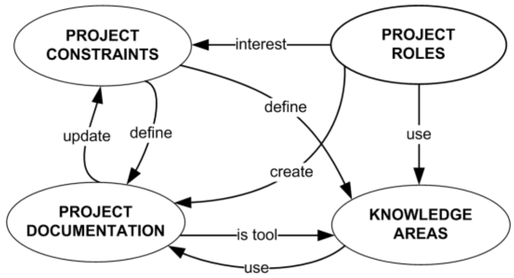


Figure 2 Semantic Model of Project Structure in the ANP model

The representation of the project management structure in the form of a semantic network in the form of Project Roles, Knowledge Areas, Constraints and Project Documentation is based on the processes of a project life cycle. In the process of the project life cycle, the roles are an active part, procedures a substantive part and documentation an input and output level. The processes define and present in detail both major international standards, PMBOK® Guide [9] and PRINCE2 [6]. The semantic network in Figure2 displays an aggregated view of the project structure in the form of clusters. The roles are interested in parameters (their value); they use management practices (Knowledge Areas) and generate documentation. Parameters (project constraints) define the scope and manner of procedures for the management of the project and the content and complexity of documentation. The documentation, which captures changes in a project, retroactively updates parameters and is a tool in sub-management practices. Clusters and their relations can be further processed in detail in sub-semantic networks. For example, the cluster of project documentation will look as follows:

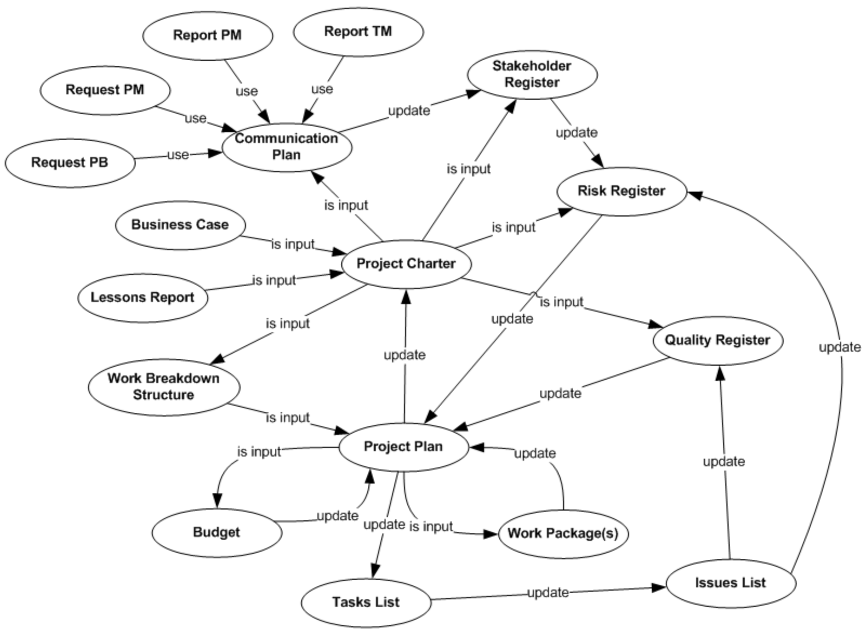


Figure 3 Semantic Model of Project Documentation in the ANP model

3.2 Semantic model quantification of the project structure

The semantic model quantification of the project structure can be performed by the ANP method. The advantage of this method is the possibility of bias preferences among elements. For example, with the help of the Saaty scale [13], the addressed roles of the project can differentiate their different attitude toward documentation. Creating the calculation of the ANP model (Figure 2) can be performed e.g. in a software tool Super Decisions Software 2.2 (http://www.superdecisions.com/). By calculating Calculus Type, a super matrix with limit weights can be obtained. The resulting weights of the individual elements reflect the influence of the element on the structure of the whole model even within individual clusters:

Project Documentation	**Weight in Model**	**Weight in Cluster**	**Weight cumulatively**	**Population cumulatively**
Project Plan	**0.09552**	**0.216**	**21.6%**	**5.9%**
Risk Register	**0.07752**	**0.176**	**39.2%**	**11.8%**
Stakeholder Register	**0.06071**	**0.138**	**53.0%**	**17.6%**
Project Charter	**0.04394**	**0.100**	**62.9%**	**23.5%**
Quality Register	0.03036	0.069	69.8%	29.4%
Work Package(s)	0.02925	0.066	76.4%	35.3%
Issues List	0.02809	0.064	82.8%	41.2%
Tasks List	0.02424	0.055	88.3%	47.1%
Communication Plan	0.0172	0.039	92.2%	52.9%
Business Case	0.01626	0.037	95.8%	58.8%
Budget	0.01309	0.030	98.8%	64.7%
Work Breakdown Structure	0.00369	0.008	99.6%	70.6%
Report TM	0.00123	0.003	99.9%	76.5%
Request PB	0.00028	0.000634	99.9864%	82.4%
Lessons Report	0.00002	0.000045	99.9909%	88.2%
Report PM	0.00002	0.000045	99.9955%	94.1%
Request PM	0.00002	0.000045	100%	100.0%

Table 1 Limit weights of elements in the ANP model

The resulting weights from Table 1 can be used for example in the case of project documentation to prioritize its creation and sharing in the project. The most important element in the cluster is the Project Plan.

3.3 Interpretation of the results for the semantic model of a project structure

Elements within the clusters can be assessed in terms of their weights, i.e. their relevance to the project management structure. The impact and abundances of elements in the cluster can be further assessed e.g. from the perspective of the existence of the Pareto principle [8] where 20% elements (e.g. project documentation) have an 80% impact on the management structure of the project within the same cluster:

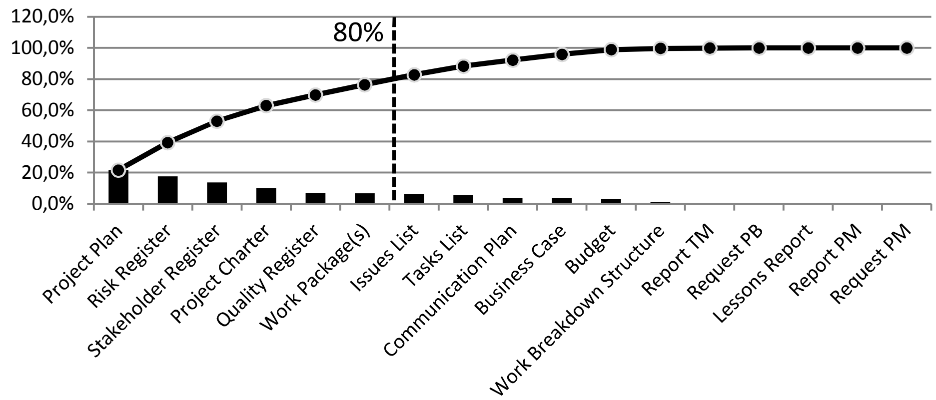


Figure 4 Pareto chart - Interpretation of the results with project documentation

For the project documentation (Figure 4) the Pareto principle did not occur. However, it can be concluded that the first seven documents (41.2%) have an impact on the management structure of the project within their cluster from 82.8%. In case of a real project environment of a company, this finding may lead to greater efficiency of project management - if, for example, this result is used in communications or management responsibilities for prioritizing in the creation and management of project documentation. By differentiating between e.g. significant and less significant project documentation it is possible to achieve the saving of work effort and cost of the project - a bureaucratic burden is thus reduced in a project.

4 Conclusion

In the article a semantic model was used for expressing general project structure under international standards PMBOK® Guide and PRINCE2 where elements of a semantic network represent specific project roles, documents, constraints or knowledge areas. Using the ANP method, the quantification of the model was achieved. The differentiation of preferences among elements allowed for an additional expression of the general model of reality from the current practice (e.g. the relation of the roles and the documents).The outputs of the calculation were the limit weights of the elements expressing the influence of the element on the structure of project management. The results elicited in the article were further interpreted against the practice, in particular the possible use of the Pareto chart was outlined (the search for the Pareto principle).The benefit of the described procedure is the ability to prioritize the elements that exist in a project environment and project culture of a company, based on their influence in project management.

References:

[1] Doležal, J. et al.: *Projektový management podle IPMA (Project management based on IPMA)*. Grada Publishing a. s., Prague, 2012.

[2] El-Gohary, N. M., Osman, H. and El-Diraby, T.E.: Stakeholder management for public private partnerships. *International Journal of Project Management* **24** (2006), 595-604.

[3] *ISO 21500:2012 Guidance on project management*. International Organization for Standardization, Geneva, 2012.

[4] Kim, H. L., Decker, S. and Breslin, J. G.: Representing and sharing folksonomies with semantics. *Journal of Information Science* **36** (2010), 57–72.

[5] Mařík, V. et. al.: *Umělá inteligence I. (Artificial Intelligence I.).* Academia, Prague, 1993.

[6] Murray, A. et al.: *Managing Successful Projects with PRINCE2®*. Fifth edition. AXELOS Ltd., 17 Rochester Row, London, 2013.

[7] Page, L., Brin, S., Motwani, R. and Winograd, T.: *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report, Stanford InfoLab, Stanford, 1999.

[8] Pareto, V.: *Manuel d'économie politique*. Fourth edition. Libraire Droz, Genève, 1966.

[9] Project Management Institute: *A Guide to the Project Management Body of Knowledge (PMBOK® Guide).* Fifth Edition. Project Management Institute Inc., 14 Campus Boulevard, Newtown Square, PA, 2013.

[10] Quillian, R.: *A recognition procedure for transformational grammars*. Doctoral dissertation. Massachusetts Institute of Technology, Cambridge (MA), 1968.

[11] Rydval, J., Bartoška, J., Brožová, H.: Semantic Network in Information Processing for the Pork Market. *AGRIS on-line Papers in Economics and Informatics* **6** (2014), 59-67.

[12] Saaty, T. L. and Vargas, L. G.: *Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Application with Benefits, Opportunities, Costs and Risks*. Springer, NY, 2006.

[13] Saaty, T. L.: *Decision Making with Dependence and Feedback: The Analytic Network Process*. RWS Publisher, Pittsburgh, 1996.

[14] Schatten, M.: Knowledge management in semantic social networks. *Computational and Mathematical Organization Theory* **19** (2013), 538-568.

[15] Shing-han, L., Shi-ming, H., David and C. Y., Jui-chang, S.: Semantic-based transaction model for web service. *Information Systems Frontiers* **15** (2013), 249-268.

[16] Sowa, J. F.: *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks/Cole Publishing Co., Pacific Grove (CA), 2000.

[17] Williams, T.: Identifying Success Factors in Construction Projects: A Case study. *Project Management Journal* **47** (2015), 97-112.

[18] Zhu, Y. and Li, X.: Representations of semantic mappings: A step towards a dichotomy of application semantics and contextual semantics. *International Journal of Project Management* **25** (2007), 121-127.

Modification of the EVM by Work Effort and Student Syndrome phenomenon

Jan Bartoška[1], Petr Kučera[2], Tomáš Šubrt[3]

Abstract. Any project as a unique sequence of tasks is a subject to the human agent impact. This article aims to analyze the design and use of nonlinear modification of BCWS (Budgeted Cost for Work Scheduled) parameter within EVM (Earned Value Management) for monitoring project planned progress respecting human factor impact on the actual work effort. Duration of project activities is very often measured by the EVM tools, where the BCWS is a key parameter. Because BCWS is derived from the work effort in a project, work effort has a hidden influence on EVM. Consequently, in the real world of human resources, the BCWS is not usually linear. To express the BCWS parameters, Non-Linear mathematical models are used in this paper, describing the course of different work contours incorporating the Student Syndrome phenomenon. The human agent impact is apparent especially in the work contours of the resources allocated to the project. At the same time, the human agent impact manifests itself in the form of the Student Syndrome phenomenon. Hence, planning and progress of the project is burdened by the variability of the resources work effort affected by both different work contours and the Student Syndrome phenomenon. The article builds on the previous work of the authors.

Key words: Project management, Earned Value Analysis, Non-Linear mathematical model, Work contour, Work effort.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Duration of project activities is very often measured by the EVM tools, where the BCWS is a key parameter. Because BCWS is derived from the work effort in a project, work effort has a hidden influence on EVM. Consequently, in the real world of human resources, the BCWS is not usually linear. The human agent impact is apparent especially in the work contours of the resource work. At the same time, the human agent impact manifests itself in the form of the Student Syndrome phenomenon. Hence, planning and progress of the project is burdened by the variability of the resources work effort affected by both different work contours and the Student Syndrome phenomenon.

Earned Value Management [11, 9 or 5] is based on comparing the Baseline and Actual plan of the project realization. Baseline and Actual plan is determined by the partial work of single resources in the particular project activities. Baseline is always based on expected resource work contours and the impact of human agent is usually not included. The impact of human agent is usually expected only in the actual course of the project. The versatility of the human agent in projects can be described also by the first “Parkinson’s first law” [8]. It is natural for people to distribute work effort irregularly to the whole time sequence which was determined by the deadline of the task completion. The questions of “Parkinson’s first law” in project management are further dealt with in e.g. [4].

Work effort of an allocated resource has very often been researched in projects from the area of information technologies and software development, as these projects contain a high level of indefiniteness, and even common and routine tasks are unique. At the same time, it concerns the area where it is possible to find a great number of approaches to estimate how work-intensive the project will be or how long the tasks will take, and also case studies. The proposal for mathematical apparatus for planning the course of tasks within a case study is dealt with for instance in [7], or Barry et al. [1]. The authors Özdamar and Alanya [7] propose a particular pseu-

[1] Czech University of Life Sciences in Prague, Department of Systems Engineering, Kamýcká 129, Prague, Czech Republic, email: bartoska@pef.czu.cz
[2] Czech University of Life Sciences in Prague, Department of Systems Engineering, Kamýcká 129, Prague, Czech Republic, email: kucera@pef.czu.cz
[3] Czech University of Life Sciences in Prague, Department of Systems Engineering, Kamýcká 129, Prague, Czech Republic, email: subrt@pef.czu.cz

do-heuristic approach to estimate the tasks course where the indefiniteness in the project is expressed by fuzzy sets. Barry et al. [1] concentrate on the existence and expression of the relation between project duration and total effort and in their theoretical starting points they point out the dynamics of the relation between the effort and project duration when a self-strengthening loop can be expected. The work effort can be described also using system dynamic models as presented e.g. in a study from project management teaching by Svirakova [10]. The others who research the project complexity and work effort are for instance Clift and Vandenbosh [3], who point out a connection between the length of life cycle and project management structure where a key factor is again a human agent.

The aim of this paper is to propose and analyze an application of a non-linear modification of the BCWS parameter within EVM for the observation of the planned project duration according to the actual work effort of the resource considering at the same time the human resource impact. The paper builds on the previous authors' works.

2 Materials and methods

2.1 Student Syndrome phenomenon

If there is a deadline determined for the completion of a task and a resource is a human agent, the resource makes its effort during the activity realization unevenly and with a variable intensity. Delay during activity realization with human resource participation leads to stress or to tension aimed at the resource or the tension of the resource him/herself. The development and growth of the tension evokes the increase in the work effort of the human agent allocated as a resource. For more details see previous paper [2].

2.2 Mathematical model of the Student Syndrome

Authors in previous paper [2] propose a mathematical expression of the Student Syndrome. Its brief description follows: First, a function expressing the proper Student Syndrome denoted by p_1 is introduced. It has three minima $p_1(t) = 0$ in $t = 0$, $t = 0.5$, and $t = 1$; and two maxima: former one close to the begin and latter one close to the end of the task realization. Beside this, functions denoted by p_2 expressing the resource allocation according to single standard work contours of flat, back loaded, front loaded, double peak, bell and turtle are proposed. All these functions are in the form of 4th degree polynomial. To express the strength of the Student Syndrome manifestation during the realization of a task the rate r of the Student Syndrome is introduced. It acquires values between 0 and 1 ($r = 0$ represents a situation when the Student Syndrome does not occur at all and the resource keep the work contour exactly; $r = 1$ means that the Student Syndrome manifests in its all strength and the resource absolutely ignore the work contour). As a result, the resource work effort during a real task realization can be modeled using function $p = rp_1+(1–r)p_2$. For more details see previous paper [2].

2.3 Modification of the BCWS by Work Effort

The EVM extension in the form of *BCWS* parameter modification for different work contours (turtle, bell, double peak, back loaded, front loaded, etc.) which is described below and applied in a case study is based on previous work of the authors of this paper [2, 6].

This approach can be applied when computing the *BCWS* of an activity in the project. It is computed in the classical way using the formula:

$$BCWS = \%PC \, . \, BAC \tag{1}$$

where *%PC* is the percentage of the work planned by the work calendar (planned completion percentage) and *BAC* is Budget at Completion of the project. The share of the whole work effort as a part of task duration requires can be calculated for a single resource as:

$$\int_0^a p(t)\,dt = 1 \tag{2}$$

Let there are n resources, indexed by 1, 2, …, n, allocated at the task. Let r_k, p_{1k}, p_{2k} denotes r, p_1, p_2 for k-th resource. Then the *BCWS* can be calculated:

$$BCWS = \left(\sum_{k=1}^{n} \int_{0}^{a} (r_k \, p_{1k}(t) + (1 - r_k) \, p_{2k}(t)) \, dt \right) \cdot BAC \qquad (3)$$

The resource work effort affects the growth of *BCWS*. It is not possible to expect uniform increase of *BCWS* in general and in case of all project tasks. In case of changing *BCWS*, EVM may provide fundamentally different results for the assessment of the state and development of the project.

3 Results and Discussion

The Student Syndrome phenomenon influences resource work effort, changes the work contour shape and transfers itself into the whole project EVM parameters. The work contours of the resources and the Student Syndrome impact may have a significant effect on both the actual and planned course of the project. As far as calculations in EVM are based on unrealistically expectations of *BCWS*, EVM may be unsuccessful.

3.1 Case Study

An illustrative fictitious case is described in Table 1. It is an IT project of a small extent. Planned tasks are differentiated by the type of work contour and by the human factor impact (Student Syndrome rate). IT projects are typical by a higher degree of procrastination, in particular in the last part of the project duration. Analytical work at the beginning is usually realized with the active participation of the customer side of the project.

	Student Syndrome Rate	**Work Contour**	**Start**	**Man-day**
Business Requirements	0.3	Front Loaded	0	5
End-User Requirements	0.3	Front Loaded	0	7
Conceptual Design	0.3	Front Loaded	7	5
Architectural Design	0.3	Front Loaded	7	8
Database Components	0.5	Back Loaded	13	10
Code/Logic Components	0.5	Back Loaded	13	17
GUI Interface Components	0.5	Back Loaded	13	15
User Acceptance Test	0.5	Back Loaded	30	5
Performance Test	0.5	Back Loaded	32	6
Product Release	0.5	Back Loaded	35	5
			Total Scope	83

Table 1 Case Study – The Project Plan

The project is planned for 40 working days. The total scope of the planned work is 83 man-days. As well, the project plan (see Figure 1) can be subjected to time analysis by Critical Path Method, etc.), however, this paper does not further follow up these results.

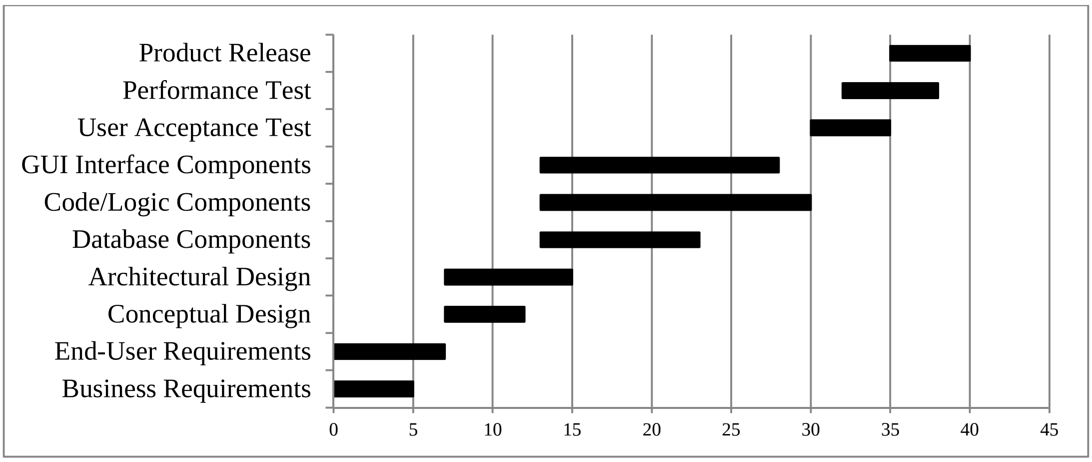


Figure 1 Case Study – Gantt Chart

In case of the first 4 tasks we expect the work accumulation at the beginning and a lower Student Syndrome phenomenon impact; therefore these tasks work contour is front loaded (see Figure 2) and the Student Syndrome rate value r = 0.3 (see Table 1). The reason may be a higher level of cooperation with stakeholders in the project, i.e.e.g. the analysis and collection of information with the participation of the contracting authority and end users (the meetings are planned in advance and their terms are fulfilled, which has a positive impact on tasks duration). In case of the remaining 6 tasks, it is possible to expect the work accumulation at the end and a higher impact of procrastination; therefore these tasks work is back loaded (see Figure 3) and the value r = 0.5 (see table 1). The reason may be more technologically demanding nature of these tasks (programming and expert work).

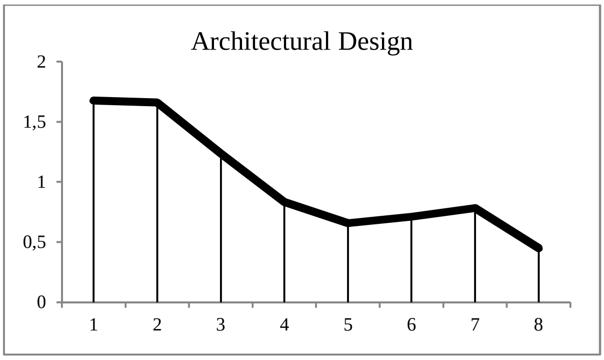


Figure 2 Case Study – a task with front loaded work contour and *r*=0,3

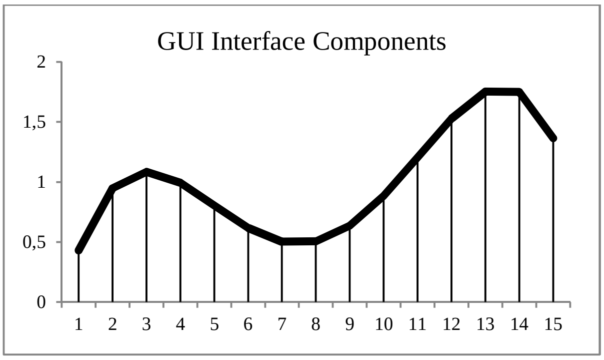


Figure 3 Case Study – a task with back loaded work contour and *r*=0,5

The application of the work contour of front loaded at the project beginning leads to a higher work effort rate of resources planned. The application of the back loaded work contour and a higher human factor impact causes significantly lower work effort rate than planned. The resulting difference may be a reason of EVM failures.

3.2 Linear and Non-Linear BCWS

The *BCWS* curves (formula 3 and Figure 4) bring two parts of information: 1) the area below the curve expresses the expended work effort by the resource; 2) the function value determines the fulfilled project extent (e.g. on 40th day it is 83 man-days). The whole planned work effort course is higher than the identified course comprising the task nature and human resource impact.

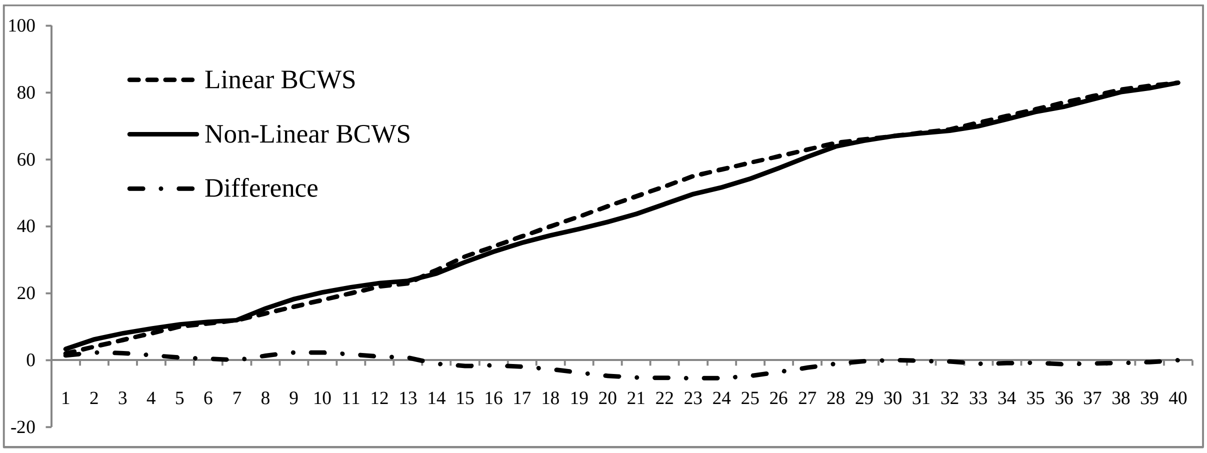


Figure 4 Case Study – difference between Linear and Non-Linear BCWS

From Figure 4 it can be evidently seen that since 13th day **the work is carried out with a lower work effort**. The overall difference between the Linear *BCWS* and the Non-Linear *BCWS* is –39.283 man-days (the difference between the areas below the curves). The observed difference is not an absent work in the project extent, but the work spent by the resources inefficiently due to the human factor impact. The reason lies in the nature of the tasks (back loaded versus front loaded) and the expected human factor impact (Student Syndrome phenomenon). Although the extent of the project (83 man-days) is completed through 40th day of duration, **the allocated resources do not make the expected work effort to the project during its realization**. The original work plan for single tasks is primarily based on the assumption of uniform work plan and work effort. When including the

Non-Linear course of work plan and work effort, a decline of the work effort occurs. At the same time, this work effort decline affects the Earned Value and increases the Actual Costs. What is more, the project completion date may be threatened. These conclusions should be subjected in further research.

3.3 Disillusionment in EVM by Linear BCWS

Including Actual Cost of Work Performed (ACWS) and the value of the actual work performed (% Work Performed per day) to the case study, the different progression of the values of the Linear BCWS and Non-Linear BCWS impacts significantly also other EVM parameters. The assessment of project status and development becomes inconsistent. Thus, the value of actual work performed may be affected by the Student Syndrome phenomenon e.g. in the following way (see Figure 5; for each day, an estimate of the work performed in relation to the plan is calculated).

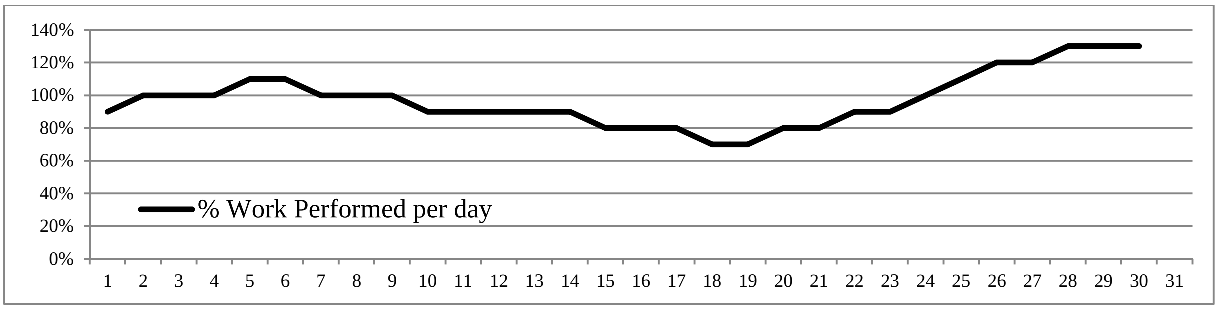


Figure 5 Case Study – % Work Performed per day

Further, to illustrate, let us expect that the project is in 30th day of its duration at this moment and the Actual Cost of Work Performed development is as follows (Figure 6, indicated in Man-days):

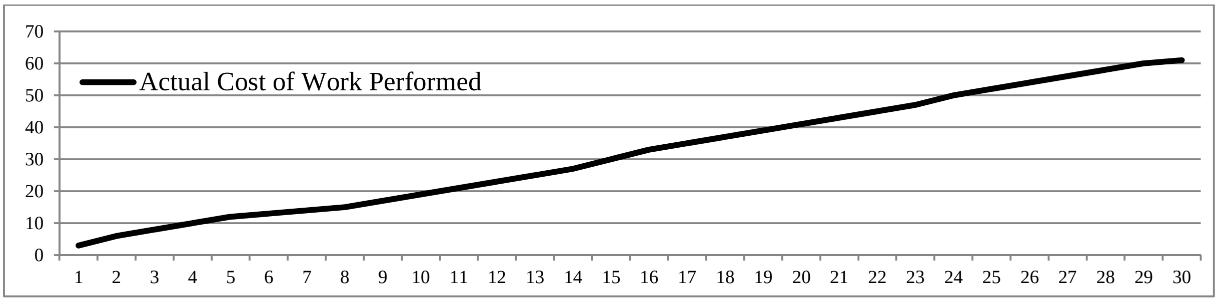


Figure 6 Case Study – Actual Cost of Work Performed (ACWP)

Other derived EVM parameters especially CV% (Cost Variance %) and SV% (Schedule Variance %), express the relative deviation of the project from its budget or plan (Figure 7 and 8):

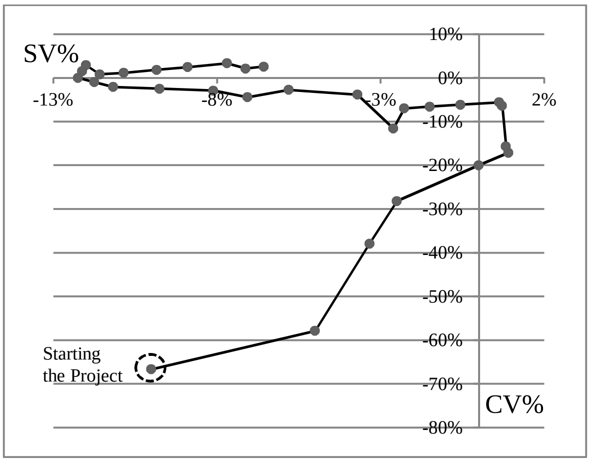


Figure 7 Case Study – Within using Linear BCWS

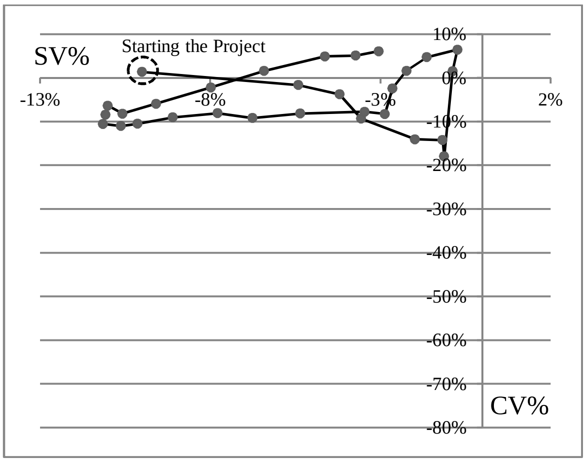


Figure 8 Case Study – Within using Non-Linear BCWS

From Figures 7 and 8 it is apparent that the observed project progress in case of the Linear BCWS values application differs completely form the identified progress in case of the Non-Linear BCWS values application. When including the human factor impact and various shapes of the work contour for single activities (Non-Linear BCWS, Figure 8), the project is assessed better at the beginning; however, during the course it often gets into a negative misalignment – it is more often in delay or over budget comparing to observing the duration with the Linear BCWS values (Figure 7). In case of the Linear BCWS values application the project would be assessed positively too: after the first five days mistakenly significant improvement of the state occurs (furthermore, it seems that from 6th to 9th day the project is ahead of time as the SV% values are greater than 0). The project is assessed completely differently. The inclusion of the human factor and work contours leads to discovery of a deeper crisis in the project.

4 Conclusion

The article discusses an application of a newly proposed *BCWS* calculation for monitoring the project duration within EVM. As the case study shows, as far as the task nature (work contour) and the human resource impact (Student Syndrome phenomenon) is comprised, the planned course of work effort (Non-Linear *BCWS*) may significantly differ from the commonly expected one (Linear *BCWS*). A possible decline or increase of resources work effort, which is not evident in case of the uniform work plan with uniform work effort, may manifest itself in an irregular increase or decrease of the Earned Value. This may result in malfunctioning EVM and project failure.

The difference in course between the Linear and Non-Linear BCWS leads within EVM to different conclusions in the assessment of the project status and development. Overmuch positive or negative project assessment may result in uneconomic decisions with fatal consequences. With the inclusion of the human factor impact and various work contours and the derivation of the Non-Linear BCWS it is possible to obtain a much more accurate (though possibly less pleasant) illustration of the project.

References

[1] Barry, E., Mukhopadhyay, T., and Slaughter, A. S.: Software project duration and effort: An empirical study. *Information Technology and Management* **3** (2002), 113–136.

[2] Bartoška, J., Kučera, P., and Šubrt, T.: Modification of the BCWS by Work Effort. In: *Proceedings of the 33rd International Conference Mathematical Methods in Economics MME 2015* (Martinčík, D., Ircingová, J., and Janeček, P., eds.). University of West Bohemia, Plzen, 2015, 19-24.

[3] Clift, T., and Vandenbosh, M.: Project complexity and efforts to reduce product development cycle time. *Journal of Business Research* **45** (1999), 187–198.

[4] Gutierrez, G.J., and Kouvelis, P.: Parkinson's law and its implications for project management. *Management Science* **37** (8) (1991), 990-1001.

[5] Kim, B., C., and Kim, H., J.: Sensitivity of earned value schedule forecasting to S-curve patterns. *Journal of Construction Engineering and Management* **140** (7) (2014), 14-23.

[6] Kučera, P., Bartoška, J., and Šubrt, T.: Mathematical models of the work contour in project management. In: *Proceedings of the 32nd International conference on Mathematical Methods in Economics* (Talašová, J., Stoklasa, J., and Talášek, T., eds.). Palacký University, Olomouc, 2014, 518-523.

[7] Özdamar, L., and Alanya, E.: Uncertainty modelling in software development projects (with case study). *Annals of Operations Research* **10** (2001), 157–178.

[8] Parkinson, C.N.: *Parkinson's Law and Other Selected Writings On Management*. 1st ed., Federal Publications (S) Pte Ltd, Singapure, 1991.

[9] Project Management Institute. *A Guide to the Project management body of knowledge (PMBOK® Guide).* Fifth edition. Newtown Square: Project Management Institute Inc., p. 589, 2013.

[10] Svirakova, E.: System dynamics methodology: application in project management education. In: *Proceedings of the 11th International Conference on Efficiency and Responsibility in Education* (Houška, M., Krejčí, I., and Flégl, M., eds.). Czech University of Life Sciences, Prague, 2014, 813-822.

[11] Vandevoorde, S., and Vanhoucke, M.: A comparison of different project duration forecasting methods using earned value metrics. *International Journal of Project Management* **24** (4) (2005), 289–302.

Equivalent scale and its influence on the evaluation of relative welfare of the household

Jitka Bartošová [1], Vladislav Bína [2]

Abstract. The quantitative description and income modeling usually employs equivalent income values based on the total income of households adjusted in order to provide a reasonable comparison of relative welfare of households differing in their extent and age structure. Equalized income of the household is perceived as a general indicator of economic sources available to each household member and suitable to perform comparisons across countries and regions. Unfortunately, there does not exist any generally accepted methodology of equivalent income calculation. In the presented paper authors discuss equivalent scales which in a different manner incorporate adults and children in the households according to their age in order to evaluate the number of consumption units. The influence of chosen method on basic summary characteristics of equalized incomes and on the chosen indicators of poverty and inequality is demonstrated on the data file from large European sample survey on incomes and living conditions of households EU-SILC. The paper focuses also on dependence of the relative welfare (represented by the scaling using equalized income) on the type of household. In certain types of households we can expect relatively strong dependence of the above mentioned characteristics on the type of household.

Keywords: consumption unit, equalized income, poverty and inequality indicators.

JEL Classification: C13, C15, C65
AMS Classification: 62H12

1 Equivalent scales

It is obvious that household expenditures grow with each additional household member, but they do not grow proportionally. Thanks to the possibility to share some of the expenditures, namely expenditures on housing, water and electricity, it is necessary to assign to each household a value (weight) in relation to its requirements. These values, so called consumption units, provide a possibility to compute equalized values of incomes which comprise income of standardized household or, in the other words, income per one equalized member of household. The choice of equivalence scale depends on the expected economies of scale in the consumption of household but also on the priorities assigned to the demands of particular individuals according to their age (adults and children). Equalized household income is perceived as a general indicator of economic resources available to each household member (OECD [8]). Thus, it can be used for analysis of incomes and their comparison across the regions and countries but also for the analysis of the risk of monetary poverty of individuals (see, e.g., Bühmann et al. [2], Flachaire and Nunez [3]).

The aim of this paper is to compare two different types of models used for the construction of consuming unit and present the influence of the choice of model type on the equivalent scales together with the influence of setting the model parameters on basic summary characteristics of incomes and indicators of monetary poverty and inequality. For this purpose we use the linear equivalent scales constructed according to the methodology of Organization for Economic Co-operation and Development (OECD scale) and its variant used for the computation of official Eurostat statistics (OECD modified scale) and we compare it with non-linear equivalent scale constructed for the calculation of official statistics on New Zealand. The impact of choice of equivalent scale on monetary incomes can be evaluated from different viewpoints. For the demonstration of sensitivity of selected characteristics of incomes and indicators of monetary poverty and inequality on the selection of equivalence scale model and its parameter settings we use the resulting data from sample survey EU-SILC 2012 for the Czech households. It is large representative survey on incomes and living conditions carried out annually according to unified methodology of Eurostat in all member countries of European Union, in Norway, Sweden and on Iceland.

[1,2] University of Economics in Prague, Faculty of Management, Department of Exact Methods, Jarošovská 1117/II, 37701 Jindřichův Hradec, {bartosov,bina}@fm.vse.cz.

2 Linear and non-linear equivalence scale

Evaluation of level, differentiation, non-uniformity of distribution or insufficiency of financial potential is influenced by the choice of scale (measure) in which the values of incomes are equalized. Two extreme possibilities of the scaling are on one hand the choice of total household incomes, i.e. the sum of all monetary means collected by the household members. In this case we completely ignore the size and age structure of household. On the other hand, the second extreme possibility is the assignment of the whole household income to all particular household members (Per Capita incomes). Larger households usually demand more resources but also have higher capacity for achieving higher incomes because they have more adult, economically active members. With the increase of number of household members there also appears a tendency to decrease relatively the amount of resources necessary to guarantee the household functioning. It can be expected that different age categories may have different requirements and among other relevant characteristics potentially influencing the financial needs counts also gender, job position and marital status. Integration of such differences to equivalence scale demands more sophisticated solution. Subsequently several types of equivalence scale definitions were created (e.g., OECD [8], Jorgensen and Slesnick [6], etc.) which aims to perform an optimal estimation of proportion of

- common households expenditures;
- individual household members' incomes.

It appears that unified and general approach for estimation does not exist since each country has different setting of price relations which, moreover, evolve in time.

2.1 Linear equivalence scale

Linear equivalence scale is constructed according to the methodology of the Organization for Economic Cooperation and Development (OECD scale). European Union employs the modified version (OECD modified scale) which is based on the different setting of model parameters. General linear model of equivalent consumption unit (*CU*) is defined as

$$CU(\alpha,\beta) = 1 + \alpha \cdot C + \beta \cdot (A-1) \tag{1}$$

where *A* is a number of adults in household, *C* is number of children in household younger than 14 years, α is the child adult equivalence parameter and β is the other persons adult equivalence parameter.

OECD scale has the model parameters set to the values $\alpha = 0.5$ and $\beta = 0.7$, whereas OECD modified scale uses lower values of parameters, namely $\alpha = 0.3$ and $\beta = 0.5$. Both limit possibilities, i.e. total incomes assigned to household without considering its size and age structure correspond to the choice of $\alpha = \beta = 0$, incomes Per Capita match to $\alpha = \beta = 1$. This means that in the first case we get $CU(0,0) = 1$ and Per Capita the consuming unit is defined as $CU(1,1) = A + C$.

The modified OECD scale has both parameters decreased by two tenth which lead to the emphasis of first component, namely the common expenditures on running of the household in contrary to the second component, i.e. individual expenditures of particular household members. This leads (in comparison to the original OECD scale) to the relative increase of household incomes with relatively higher number of members. Modification changes also the perspective of monetary poverty of households since relatively smaller households can now easily fall under the poverty threshold. The emphasis on common expenditures rather corresponds to the reality of countries of Western Europe where housing costs comprise higher percentage of total household expenditures. In post-communist countries still – in spite of the permanent growth of housing costs – persists the situation corresponding rather to the original OECD scale.

2.2 Non-linear equivalence scale

Non-linear equivalence scale was used for the calculation of mutually comparable values of household incomes on New Zealand. The change in methodology of income distribution analyses on New Zealand was stimulated by the Royal Commission of Inquiry into Social Security in years 1969 and 1972. The creation of models and estimation of their parameters were in that time in focus of many statisticians, e.g., Jensen [5]. And they suggested for income equalization to mutually comparable values several different models or suitable values of their parameters.

The model used in this paper is based on power of the weighted sum of adult persons and children in the household. The model was constructed on the basis of econometrical analysis of household expenditures (with typical consumer basket). The coefficients react on different consumption of adult household members and children (parameter α) and on economies of scale which emerge thanks to the cohabitation of household members. This non-linear model is in general given by the formula

$$CU(\alpha,\gamma) = (A + \alpha \cdot C)^{\gamma} \tag{2}$$

where *A* is a number of adults in household, *C* is the number of children in household younger than 14 years, *α* is the child adult equivalence parameter and *β* is the household economies of scale parameter.

Parameter *α* has the same role in both models (linear and power). For $\gamma = 1$ the power model changes to the linear and for $\alpha = 1$ and $\gamma = 0.5$ to the square root model which was also use on New Zealand but was not successful since it does not react sufficiently on the age structure of household. Similarly as in the case of the linear model there exists a choice of parameters corresponding to both limits. Total incomes of household without consideration of size and age structure correspond to $\gamma = 0$, incomes Per Capita correspond to $\alpha = \gamma = 1$.

2.3 Characterization of income distribution, inequality and monetary poverty

For assessing of the sensitivity of income distribution characteristics on the setting of parameters in models of equivalent consumption unit we employ only basic measures of location – median and lower and upper quartiles which are robust estimates of the middle half of income distribution.

The sensitivity of income inequality on the setting of model parameters can be observed using the well-known Gini index. .The Gini index is elicited from the Lorenz curve and measures the deviation of the income distribution of individuals or households from the perfectly uniform distribution. Its value is given by the ratio of the area between the line of absolute equality (diagonal $y = p$) and the Lorenz curve *L(p)* and the entire area under the diagonal. According to the fact that the area under the diagonal is equal to the half of area of the unit square we can obtain the Gini coefficient using the numerical integration of estimated Lorenz curve

$$G = 1 - 2 \cdot \int_0^1 L(p)dp \tag{3}$$

The Gini index takes values from the interval $\langle 0; 1\rangle$ – value approaching to 0 indicates more egalitarian distribution of incomes in the considered society and vice versa.

For the expression of monetary poverty we used the measures of poverty employed in EU countries. These stem from the class of Foster-Greer-Thorbecke poverty measures (see [4]) in general given by formula

$$P_{\alpha}(y,z) = \frac{1}{n}\sum_{i=1}^{q}\left(\frac{z - y_i}{z}\right)^{\alpha}, \tag{4}$$

where $z > 0$ is a beforehand given poverty threshold, $\mathbf{y} = (y_1, y_2, \cdots, y_n)$ is a vector of household incomes sorted by size $(y_1 \le y_2 \le \cdots \le y_q \le z)$, *q* is the number of households belonging to the group under the poverty threshold and *n* is the total count of households. Parameter *α* conditions the measure of sensitivity of deprivation in case of households belonging below the poverty threshold (see [8]). For $\alpha > 1$ the value of P_{α} begin to be distributionaly sensitive and with growing value of *α* grows the sensitivity on measuring the poverty of the poorest. For $\alpha \to \infty$ P_{α} reflects the poverty of the poorest persons (see [7]).

The most commonly used measure of monetary poverty, so called *head count index* or *risk-of-poverty-rate* can be obtained by choosing $\alpha = 0$. By choice of $\alpha = 1$ we obtain another measure, describing the *depth of poverty* (or *poverty gap*) which is based on the summary evaluation of poverty according to the poverty threshold. The value of *PG* relates to the distance of poor from the poverty threshold. Thus we obtain information about the extent of poverty. But even this measure is not sensitive enough when the “poor person” becomes “very poor”. This lack of sensitivity will be removed by choice of $\alpha = 2$.

3 Sensitivity of selected characteristics on the choice of equivalent scale

The choice of equivalent scale model and change of its parameters in a way projects into the changes of income distribution and its basic and advanced characteristics. It changes not only the momentum and quantile measures of location, variability and concentration (skewness and kurtosis) of the distribution. But this change of scale projects itself automatically also into the advanced indicators of income inequality, monetary poverty, etc.

Reaction of selected indicators, i.e. median, both quartiles, Gini index and Foster-Greer-Thorbecke measures of poverty on the selection of models and change of its parameters is depicted in Figures 1 – 3. The influence of parameter setting is presented simultaneously for both models – in the left part we can see the linear model, in the right part the non-linear model is depicted. The horizontal axis shows the change od parameter α $(0 \geq \alpha \leq 1)$ whereas the vertical axis shows in case of linear model the change of parameter β $(0 \geq \beta \leq 1)$, in case of non-linear model the change of parameter γ $(0 \geq \gamma \leq 1)$.

The sensitivity is presented using isoquants, i.e. the curves with constant value of considered characteristics Perpendicular curves to isoquants are gradients of considered characteristics. Density of the curves is proportional to the measure of change of the characteristics and thus it represents the sensitivity on the change of parameters in the given model.

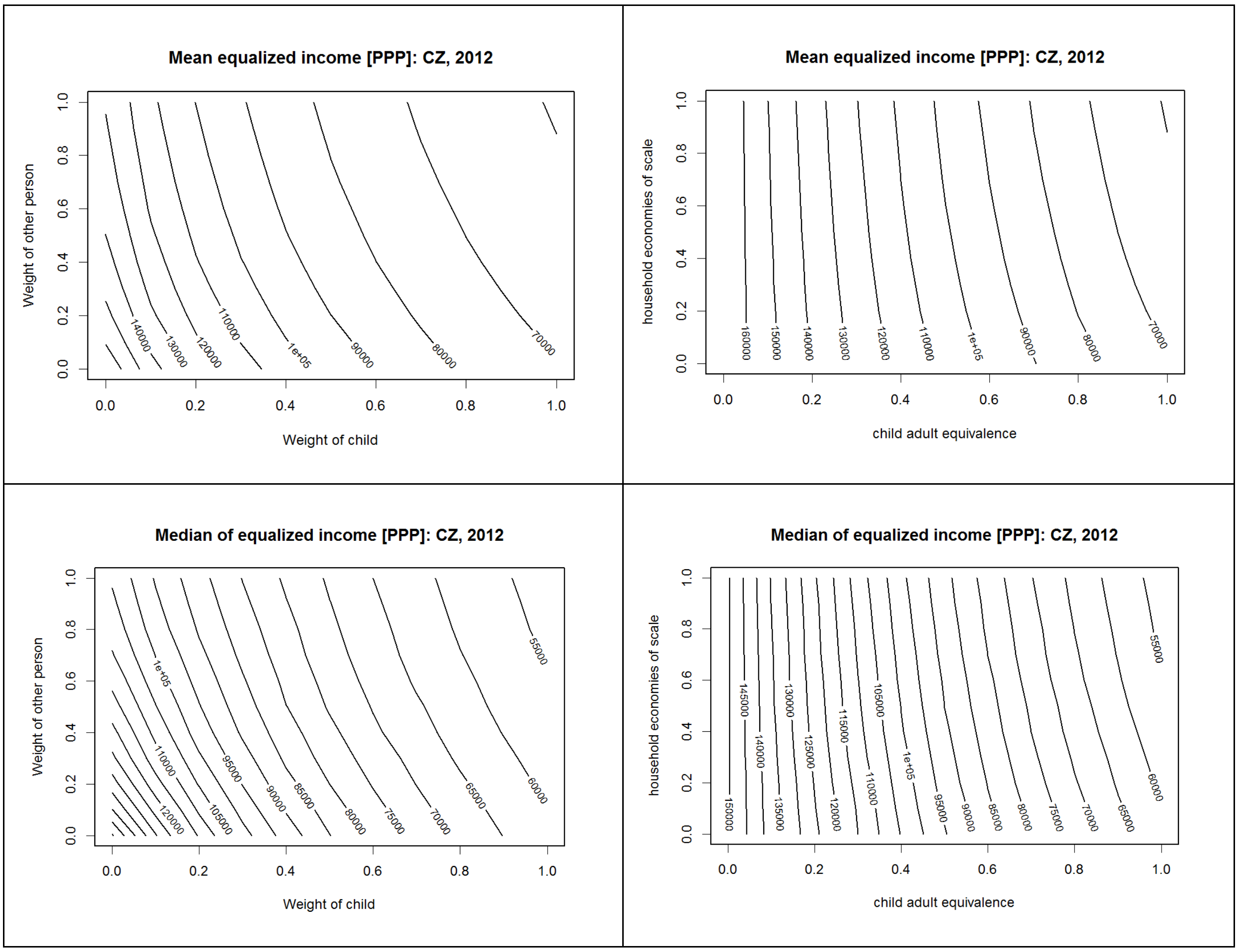


Figure 1 Reaction of mean (upper) and median (lower): linear model (left), power model (right)

Figure 1 presents the contour maps for mean and median. This allows us to visually depict the difference in reaction of robust and non-robust characteristics of "center" of the income distribution on change of parameters in both models. Different shapes and densities of depicted isoquants show simultaneously the similarity of reaction in the same type of model and dissimilarity of reaction in different model types. The choice of linear and non-linear model projects into the reaction on the change of particular parameters (direction and curvature of isoquants). The sensitivity of reaction – it can be observed that robust statistics, i.e. median, reacts more sensitively – their isoquants have higher density.

Figure 2 depicts the reaction of two basic measures of monetary poverty – risk of poverty and poverty gap. From the presented graphs we can infer that the choice of model is a fundamental decision in both assessing of the risk of poverty and the poverty gap. Even in this case we can see that the reaction is similar for both measures in one considered model. The sensitivity of reaction shows that the risk of poverty is more sensitive on the change of parameters in both models than the poverty gap. (Sensitivity analysis of risk of poverty, poverty gap and severity of poverty on the choice of parameters in linear model can be found also in Bartošová and Bína [1].)

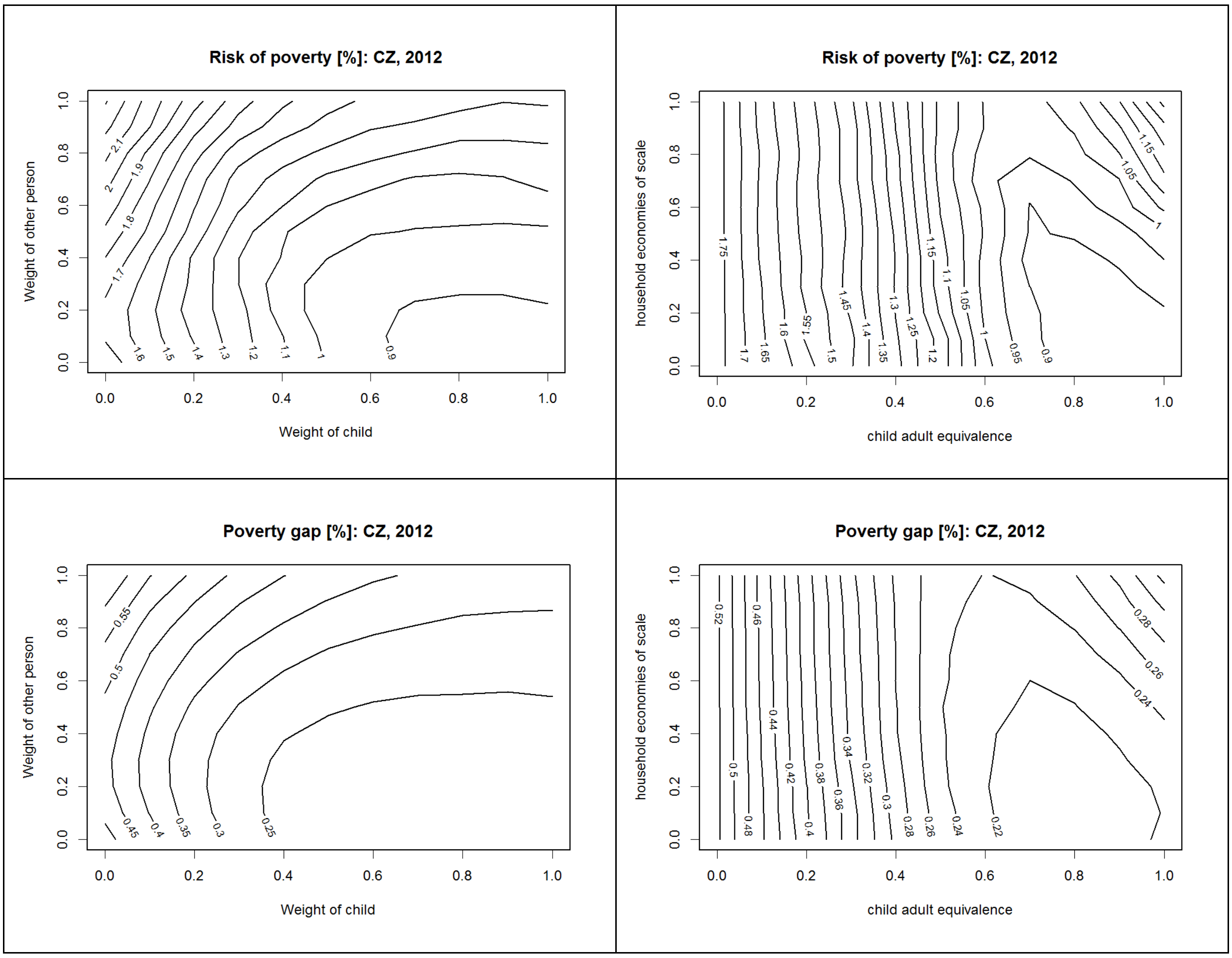


Figure 2 Reaction of risk of poverty (upper) and poverty gap (lower): linear model (left), power model (right)

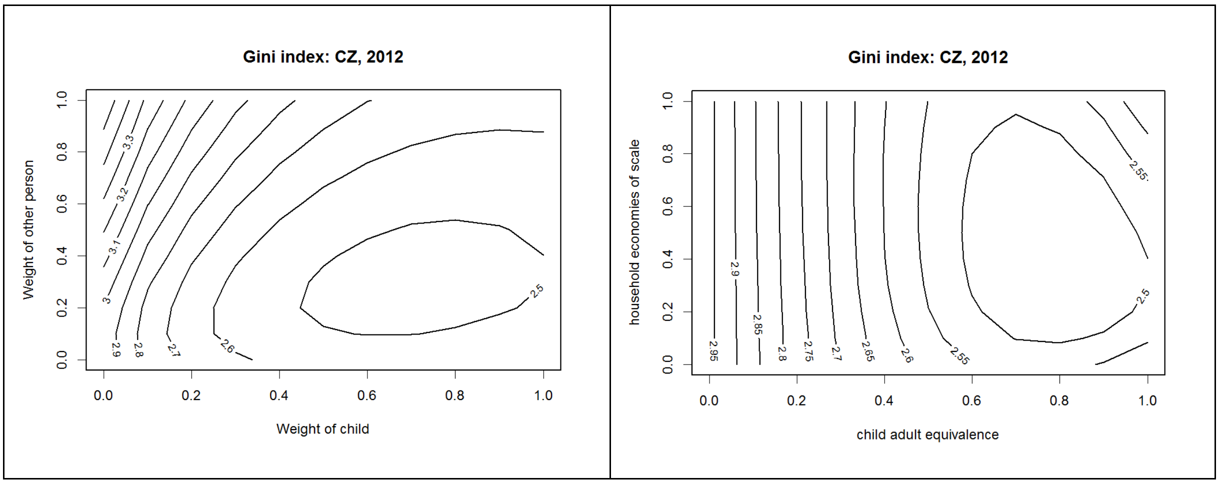


Figure 3 Reaction of Gini index: linear model (left), power model (right)

Figure 3 illustrates the response of Gini index. Different shapes and densities of isoquants show the dependence of reaction on both the choice of model and parameter setting. The perception of income inequality measured by Gini index is therefore affected by the choice of the model.

4 Conclusions

The empirical sensitivity analysis of selected statistics showed that in all cases we can observe significant influence of model type on the dependence of considered statistics with the given setting of model parameters. The selection of linear or non-linear expression of functional dependence has an important impact on the results. It implies the change of the relation of reaction on the setting of pairs of parameters in the models. We observed that this relation is for given type of model typical. Also the sensitivity, i.e. the measure of reaction on the change of parameters is in both types of models different and typical for the corresponding model.

The choice of model and its parameters radically influences the values of income indicators serving as a basis for the objective evaluation of financial situation of households (risk of monetary poverty and income inequality). The EU methodology which defines the equalized incomes using linear model with parameters α = 0.3 and β = 0.5 does not provide a possibility to describe properly national differences in economies of scale which are essential particularly in case of post-communist countries. Thus, the international comparison of income inequality and relative poverty does not necessarily reflect accurately true conditions and hence, the foundations of setting of social policy of particular countries appear to be disputable. We propose to focus on the proper choice of the model and its parameters in dependence on situation in particular country or region.

Acknowledgements

The research was supported by Internal Grant Agency of the University of Economics in Prague under project F6/34/2016.

References

[1] Bartošová, J. and Bína, V.: Sensitivity of monetary poverty measures on the setting of parameters concerning equalization of household size. In: *30th International Conference Mathematical Methods in Economics, MME 2012* (Ramík, J. and Stavárek, D., eds.). Silesian University in Opava, School of Business Administration, Karviná, 2012, 25–30.

[2] Bühmann, B., Rainwater, L., Schmauss, G. and Smeeding, T.: Equivalence Scales, Well-being, Inequality, and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study (LIS) Database. *Review of Income and Wealth* **34** (1988), 115–142.

[3] Flachaire, E. and Nunez, O.: Estimation of the income distribution and detection of subpopulations: An explanatory model. *Computational Statistics and Data Analysis* **5** (2007), 3368–3380.

[4] Foster, J., Greer, J. and Thorbecke, E. A Class of Decomposable Poverty Measures. *Econometrica* **52** (1984), 761–766.

[5] Jensen, J.: *Minimum Income levels and Income Equivalence Scales*, Department of Social Welfare, Wellington, 1978.

[6] Jorgenson, D. and Slesnick, D.: Aggregate Consumer Behavior and Household Equivalence Scales. *Journal of Business and Economic Statistics* **5** (1987), 219–232.

[7] Morduch, J. *Poverty Measures. Handbook on Poverty Statistics: Concepts, Methods and Policy Use*. New York: United Nations, Department of Economic and Social Affairs, 2005.

[8] OECD: *What Are Equivalence Scales?* Retrieved from http://www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf, 2016.

[9] Ravallion, M. *Poverty Comparisons: A Guide to Concepts and Methods*. Světová banka, Washington, DC, 1992.

Evaluation of the destructive test in medical devices manufacturing

Oldřich Beneš[1], David Hampel[2]

Abstract. Testing the quality of produced medical devices is undoubtedly a crucial issue. If a product of compromised quality is sold, it would first of all have an impact on the health of patients and, consequently, on the economic performance of the manufacturer. Destructive testing of medical devices is expensive and there are attempts to replace it by non-destructive tests. During the substitution of destructive tests with non-destructive ones, the strength of the test should be rationalized at the stage of process control planning. To do so, logit or probit models are the first choice. This article deals with an evaluation of typical destructive test results from both product and process verification and validation of medical devices. The differences between the estimated logit and probit models are discussed and arguments when to prefer which one are collected. Finally, replacement of destructive testing by non-destructive testing is evaluated numerically.

Keywords: destructive test, logit model, medical devices manufacturing, non-destructive test, probit model.

JEL Classification: L63
AMS Classification: 62P30

1 Introduction

Within highly regulated industries, like the food industry and pharmaceutical or medical device production, there are two main requirements in terms of manufacturing process definition, or maybe it would be better said duties. During the development phase, process validation is performed where the evidence of process capability is given. After the market launch, evidence that the processes are under the same conditions as during the process validation should be present during the product's life cycle.

At the same time, the strong demand for reducing the time and costs needed for any new product development or even the implementation of change should be reflected. One way of meeting these requirements is to replace expensive destructive testing performed during the validation phase with much cheaper non-destructive testing during the production phase. To do this without compromising the quality and reliability of the manufacturing processes, the strong correlation between the results of direct, typically destructive, testing and non-destructive testing of products during standard production could provide the solution. From a theoretical point of view, the way how to provide a system with this type of evidence should be described and rationalized before the planning phase.

So what do we have? Typically, we have information from the destructive test. During this test we directly measure the value that was required. This is usually called the structural value. If we find a depending value, which can be measured by a non-destructive test then we would have a so-called response variable. This sound like a standard linear model (e.g. a simple regression model) and can be thought of as having two 'parts'. These are the structural variable and the response variable. The structural component is typically more or less normally distributed, covered by a normally distributed error and is often the random component. The response variable cannot be normally distributed.

Replacing destructive tests with non-destructive ones is currently an issue in many fields. The use of statistical models for evaluating the reliability of non-destructive tests is described in the paper [10]. Probability models for the uncertainty in the flaw size determination and flaw detection are constructed. Flaw size models are based on the assumption that the measured size and true flaw size are related in a simple way: two models based on logarithmic and logit transformations of the flaw size are considered.

Authors of the paper [3] describe the discovery and comparison of empirical models for predicting corrosion damage from non-destructive test data derived from eddy current scans of USAF KC-135 aircraft. The results

[1] Mendel University in Brno, Department of Statistics and Operation Analysis, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail xbenes19@mendelu.cz.
[2] Mendel University in Brno, Department of Statistics and Operation Analysis, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail qqhampel@mendelu.cz.

also show that while a variety of modelling techniques can predict corrosion with reasonable accuracy, regression trees are particularly effective in modelling the complex relationships.

The aim of work [4] is to evaluate the reliability of non-destructive test techniques for the inspection of pipeline welds employed in the petroleum industry. The European methodology for qualification of non-destructive testing presented in [7] has been adopted as the basis of inspection qualifications for nuclear utilities in many European countries.

Quantitative non-destructive evaluation discussed in [1] provides techniques to assess the deterioration of a material or a structure, and to detect and characterize discrete flaws. Therefore, it plays an important role in the prevention of failure. The described techniques are used in processing, manufacturing and for in-service inspection; they are particularly important for the in-service inspection of high-cost and critical load-bearing structures whose failure could have tragic consequences.

The objective of this paper is to verify the possibility of replacing destructive testing based on tearing of the component with non-destructive testing involving only measurement of the component part.

2 Material and Methods

Our datasets come from destructive tests of a medical device. We focus on the part where two pipes of the same diameter are connected together. This procedure is illustrated in Fig. 1. Pipe A is expanded by tool C. Furthermore, pipes A and B are connected together under pressure. The overlap length of pipes A and B is measured (in mm). When the destructive test is performed, the strength for which the pipes are torn apart is measured (in N). Finally, we have a binomial variable that indicates whether the device passed (1) or failed (0) the test.

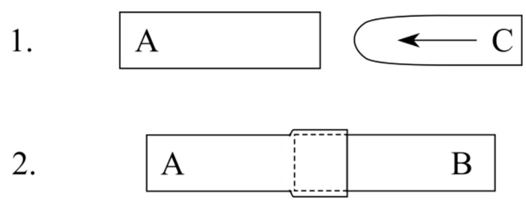


Figure 1 Illustration of pipe connections

We will work with the results of 6 destructive tests, in each 60 devices were tested. The first three sets we will use as training sets dedicated to estimation (we will use the notation O1, O2 and O3 for the particular datasets and O together). The next three sets (V1, V2, V3 and V together) will be used for verification of the estimated relationships.

The binary variable on passing the test is definitely not randomized because the data are collected in a predefined time or sequence. Nevertheless, a binary outcome is still thought to depend on a hidden, unknown Gaussian variable. The generalized linear model (GLiM) was developed to address such cases, and logit and probit models as special cases of GLiMs are appropriate for binary variables or multi-category response variables with some adaptations to the process [2, 6]. One of these two models should be the first choice solution for similar situations. A GLiM has three parts: a structural component, a link function, and a response distribution.

The way we think about the structural component here does not really differ from how we think about it in standard linear models; in fact, that is one of the great advantages of GLiMs. The link function is the key to GLiMs: since the distribution of the response variable is non-normal, it is what lets us connect the structural component to the response one i.e. it ‘links’ them. Since the logit and probit are links, the understanding link functions allow us to intelligently choose when to use which one. Although there can be many acceptable link functions, there is often one that is special. The link function that equates them is the canonical link function. The canonical link for binary response data (more specifically binomial distribution) is the logit.

Nevertheless, there are lots of functions that can map the structural component to the interval [0, 1], and thus be acceptable; the probit is also popular. The choice should be made based on a combination of knowledge of the response distribution, theoretical considerations and empirical fit to the data. To start with, if our response variable is the outcome of a Bernoulli trial (i.e. 0 or 1), our response distribution will be binomial, and what we are actually modelling is the probability of an observation being a 1. As a result, any function that maps the real number line to the interval [0, 1] will work. From the point of view of substantive theory, if we are thinking of our covariates as being directly connected to the probability of success, then we typically choose logistic regression because it is the canonical link.

However, the consideration is that both logit and probit are symmetrical, if you believe that the probability of success rises slowly from zero but then tapers off more quickly as it approaches one, then a cloglog is called for. A cloglog is slightly different from the others. It is based on a standard extreme value distribution and the curve approaches one more gradually than it approaches zero. The alternative of a ‘slower’ approach to zero can be achieved by redefining the response so as to apply the model to let us say unemployment rather than employment.

The logit and probit functions yield very similar outputs when given the same inputs, except that the logit is slightly further from the bounds when they ‘turn the corner’. In addition, we could add the cloglog function here and they would lay on top of each other. Notice that the cloglog is asymmetrical whereas the others are not; it starts pulling away from 0 earlier, but more slowly, and approaches close to 1 and then turns sharply.

Additional constraints could be based on the statement that probit results are not easy to interpret. This is not true, although interpretation of the betas is less intuitive. With logistic regression, one unit change in X is associated with a β change in the log odds of ‘success’ all else being equal. With a probit, this would be the change of β zβ z’s with two observations in a dataset with zz-scores of 1 and 2, for example. To convert these into predicted probabilities, you can pass them through the normal cumulative distribution function.

It is also worth noting that the usage of probit versus logit models is heavily influenced by disciplinary tradition. For instance, economists seem far more used to probit analysis while researchers in psychometrics rely mostly on logit models. So while both models are abstractions of real world situations, logit is usually faster to use on larger problems, which means multiple alternatives or large datasets. The multinomial logit functions are classically used to estimate spatial discrete choice problems, even though the actual phenomenon is better modelled by a probit. But in the choice situation the probit is more flexible, so more used today.

The major difference between logit and probit models lies in the assumption on the distribution of the error terms in the model. For the logit model, the errors are assumed to follow the standard logistic distribution while for the probit, the errors are assumed to follow the normal distribution. In principle, for general practice the model formalisms both work fine and often lead to the same conclusions regardless of the problem complexity. While the distributions differ in their theoretical rationale, they can for all practical purposes in my view be treated the same as it would take an impossibly large sample to distinguish them empirically.

Practically, employment of GLiM will be the last task for us. Before this, we need to verify the relationship between strength and a higher defined overlap length. For datasets O1, O2 and O3 we estimate the suitable regression function by both general OLS and robust regression comprising bisquare weights [8], because we want to eliminate the effects of possible outliers. The stability of these three estimated regressions is judged by the general linear hypothesis [5, 9], where the equality of all corresponding parameters is tested.

After this, the regression is estimated based on all training datasets O. Using this estimate, the MSE of all datasets is calculated; the MSE from verification datasets V1, V2, V3 is compared to the MSE of O1, O2 and O3. When the estimated regression function is verified, we use it for modelling the binomial variable indicating passing the test by an overlap of the pipes. For this purpose, we use both the logit and probit link. All of the calculations were performed in the computational system Matlab R2015b.

3 Results

Tearing strength was modelled by the overlap length of the pipes for all of the training datasets. The relationship is not symmetrical and we think it is well described by the cubic polynomial (all of the parameters are statistically significant, determination is higher than 0.95 for all of the datasets). There are no important differences between general and robust regression, see Fig. 2. Regressions for datasets O1, O2 and O3 can be considered to be identical, because the testing statistics value of the general linear hypothesis 1.004 is lower than the critical value 1.994. This means that the estimated relationship is stable at least over the training datasets.

The regression cubic polynomial was estimated based on all training datasets O. Using this estimate, the mean square error (MSE) was calculated for all of the particular datasets, see Tab. 1. It is clear that the regression estimated with dataset O is suitable also for of the verification datasets V1, V2 and V3. The highest MSE was calculated for dataset O1; this is caused by the presence of outliers evident from the visible difference of the OLS and a robust estimate of the regression function.

Sample	O1	O2	O3	O average	V1	V2	V3	V average
MSE	3.90	1.80	1.65	2.45	1.13	1.74	1.62	1.50

Table 1 MSE of all datasets when regression estimated for all training sets O is used

Furthermore, we modelled the binomial results of the destructive test by the pipe overlap length. Note that when we model this binomial variable by strength sufficient for tearing off the pipes, we obtain 100 % success. We used a function form describing the relationship between the tearing strength and the pipe overlap length i.e. a cubic polynomial. Logit and probit link functions are employed. The results are summarized in Tab. 2 and graphically illustrated in Fig. 3.

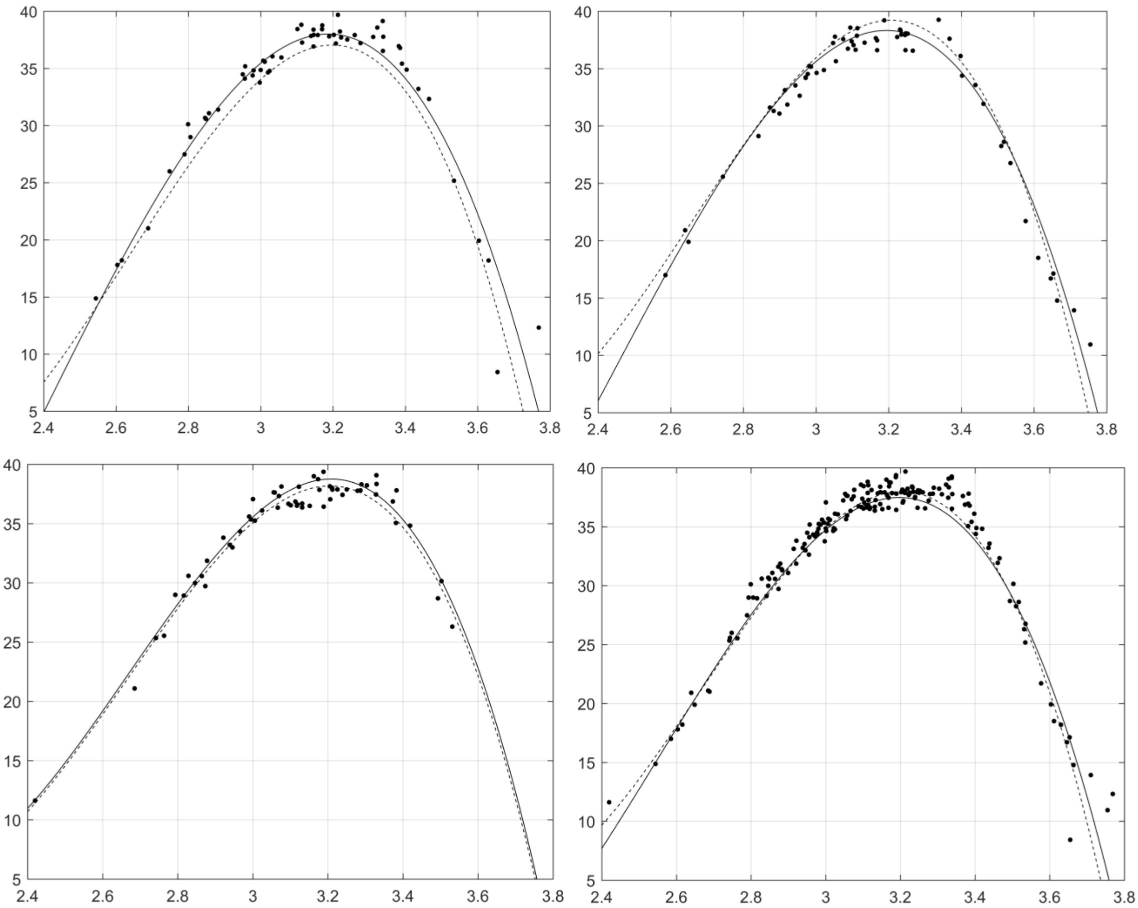


Figure 2 Estimated regression function – cubic polynomial (solid line OLS estimate, dashed line robust regression with bisquare weighting). X-axis pipe overlap length in mm, y-axis pipe tearing strength in N. Left top O1 dataset, left bottom O3, right top O2, right bottom all mentioned datasets together.

It is clear that there are no substantial differences between the results obtained with the logit and the probit links. It should be mentioned that for the probit link we obtain systematically lower p-values than for the logit link when testing the significance of particular regression parameters. There are only slight differences between the MSE for the logit and the probit links. The percentage of correctly classified cases is always the same. This percentage is relatively high, above 93.3 % for the particular datasets. In Fig. 3, it is clear that the transition from 0 to 1 is not symmetric to the transition from 1 to 0, which reflects the non-symmetry of the previously estimated cubic polynomial.

Sample	O1	O2	O3	all O	V1	V2	V3	all V
MSE logit	0.012	0.000	0.036	0.065	0.021	0.016	0.000	0.059
percentage logit	96.7	100.0	93.3	88.3	98.3	96.7	100.0	93.3
MSE probit	0.033	0.000	0.036	0.065	0.022	0.016	0.000	0.059
percentage probit	96.7	100.0	93.3	88.3	98.3	96.7	100.0	93.3

Table 2 MSE and percentage of correctly classified cases of GLiM estimate of binomial variable indicating passing the test by pipe overlap length with logit and probit link

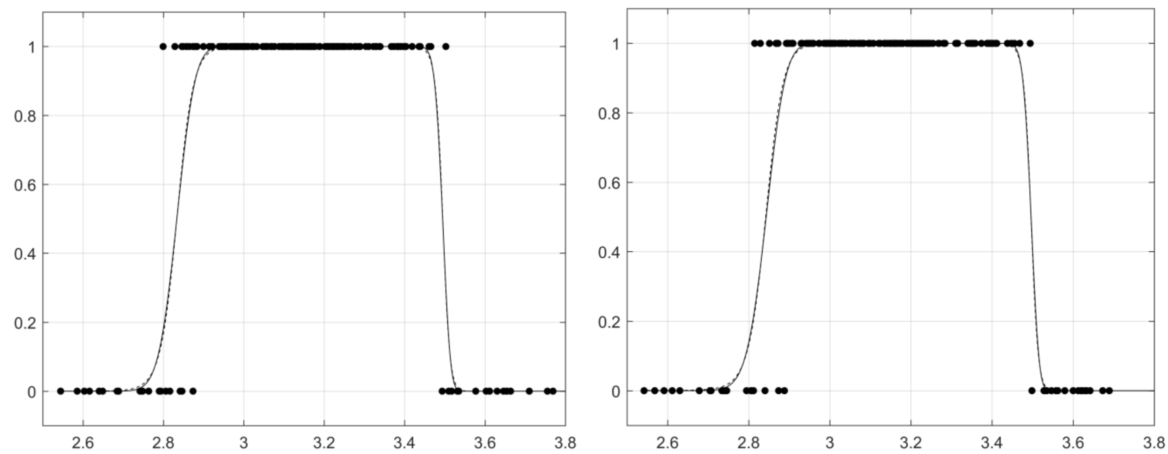


Figure 3 Estimated regression function – cubic polynomial (solid line GLiM estimate with probit link, dashed line GLiM estimate with logit link). X-axis pipe overlap length in mm, y-axis result of destructive test (1 passed, 0 failed). Left graph all training datasets together, right graph all verification datasets together.

4 Discussion and Conclusions

The relationship between the strength when the component is destructed and the length of the pipe overlap was determined based on the datasets dedicated to the estimation. An adequate description is given by the cubic polynomial, where the coefficients were estimated by the both OLS method and robust regression comprising bisquare weights. The stability of this relationship was judged by the test of equality of three regression functions based on the general linear hypothesis. Furthermore, the determined function form of the dependency was verified using control datasets.

The cubic polynomial enables a correct estimation of the asymmetrical relation. However, this representation has a limited range of applicability, because out of data we cannot assume proper behaviour. This shortage could be overcome by using a nonlinear Gauss function of order 2, but in this case we have 6 parameters instead of 4. It is also not easy to use this function in GLM.

In addition, we attempted to fit a Bernoulli-distributed variable, which means that the component either passed (1) or fails (0). This variable is fully described by the strength for which the component is destructed. When we use the length of pipe overlap instead of this strength, we obtain a percentage of correctly estimated cases higher than 93.3 % for both the training and verification datasets. Finally, the performed analysis shows that it is possible to replace the destructive test by a non-destructive one.

This fact has important economic consequences. The main reason for this alternative testing is to reduce the cost of process control during regular production without any compromise in the quality of product. The situation the test results are coming from is rather typical: high volume of annual production with a regular price of one piece. The figures for the mentioned production are the annual demand for the first year coming from a ramp up estimation i.e. 411,538 pieces with an annual increase of 1 %. The price of one piece in at the time of testing was 88.70 EUR. To be on the safe side the annual decrease in price is estimated as being 10 %.

For the same reason, the number of tests is estimated as 3 fails for one success. However, this test should be based on the strong results of an engineering study. The possibility of fail is still rather high and rules for risk management of this safety factor should be used. To obtain a better and more detailed picture, the return on investment is calculated with a time scale set at 6 months with a 5-year life cycle. In this situation, perfect for using of the alternative testing usage, a return on investment would take third months and the complete saving before the end of the life cycle is as high as 3,038,528 EUR. This is 2.03 % of the calculated turnover (see Fig. 4 Scenario 1). If we consider only one test attempt, the results will obviously be even more promising (see Fig. 4, Scenario 2).

But let us simulate a different situation. A real business case from a different company shows figures where the annual demand is 477 pieces only but with a high price of each unit of 324.38 EUR. The price is much more stable and the forecast is for an annual decrease of only 2 %. At the same time, this product is manufactured based on an individual order therefore standard process validation is not possible. In order to perform the verification two identical pieces need to be produced, whereby one of them will undergo destructive testing and the second will be sent to the customer. This situation still makes sense despite the small annual production. Despite the fact that the total amount of pieces produced during the estimated life cycle of five years will only be 2,434, the economic effect is still high due to the high price of a single product. Taking into account the safety factor mentioned above, this situation fits Scenario 3 in Fig. 4. The return on investment will be after two years and the total saving before the end of the life cycle will be 402,531 EUR (which is 53.60 % of the calculated turnover). Similarly, without the safety factor it would be even more promising. The impact is higher because of the higher price of one unit, which means a higher price for the alternative test validation at the beginning (see Fig. 4, Scenario 4).

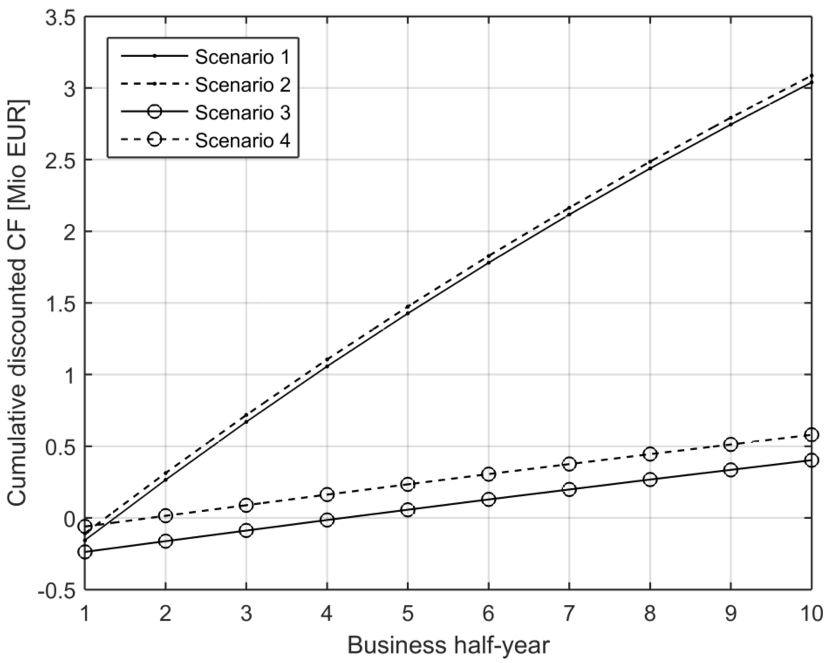


Figure 4 Effects of different production scenarios on company cumulative discounted cashflow

Finally, from the economic point of view, attempts to find an alternative test method make sense in the case of high volume production or with a high price of one unit together with verification or high volume process control sampling. In these situations, the initial cost of alternative test method validation will be paid back in an acceptable time and increase competitiveness on the market.

References

[1] Achenbach, J. D.: Quantitative nondestructive evaluation. *International Journal of Solids and Structures* **37**, 1–2 (2000), 13–27.

[2] Agresti, A.: *An Introduction to Categorical Data Analysis*. John Wiley & Sons Inc., New York, 2007.

[3] Brence, J. R. and Brown, D. E.: Data mining corrosion from eddy current non-destructive tests. *Computers & Industrial Engineering* **43**, 4 (2002), 821–840.

[4] Carvalho, A. A., Rebello, J. M. A., Souza, M. P. V., Sagrilo, L. V. S. and Soares, S. D.: Reliability of non-destructive test techniques in the inspection of pipelines used in the oil industry. *International Journal of Pressure Vessels and Piping* **85**, 11 (2008), 745–751.

[5] Doudová, L. and Hampel, D.: Testing equality of selected parameters in particular nonlinear models, In: *AIP Conference Proceedings* **1648**, 2015.

[6] Fitzmaurice, G. M., Laird, N. M. and Ware, J. H.: *Applied Longitudinal Analysis*. John Wiley & Sons Inc., New York, 2011.

[7] Gandossi, L. and Simola, K.: Framework for the quantitative modelling of the European methodology for qualification of non-destructive testing. *International Journal of Pressure Vessels and Piping* **82**, 11 (2005), 814–824.

[8] Huber, P. J. and Ronchetti, E. M.: *Robust Statistics*. John Wiley & Sons Inc., New York, 2009.

[9] Milliken, G. A. and Graybill, F. A.: Extensions of the general linear hypothesis model. *Journal of the American Statistical Association* **65**, 330 (1970), 797–807.

[10] Simola, K. and Pulkkinen, U.: Models for non-destructive inspection data. *Reliability Engineering and System Safety* **60**, 1 (1998), 1–12.

Sentiment Prediction on E-commerce Sites Based on Dempster-Shafer Theory

Ladislav Beranek[1]

Abstract. Predictions sentiment on e-commerce sites are usually based on assigning a numerical score that is followed eventually by free text review. These score and text review are written by people who bought certain products or services (especially hotel industry) and want to evaluate their purchases. Most of the works that try to evaluate customers' sentiment from these evaluations deal with textual analysis of the reviews. In this work, we will consider this review as a collection of sentences, each with its own orientation and with sentiment score. Certain score aggregation method is needed to combine sentence-level scores into an overall review rating. On the basis of the analysis of existing methods, we propose a new method that aggregate score on the base of the Dempster-Shafer theory of evidence. In the proposed method, we first determine the polarity of reviews with the use of machine learning and then consider a sentence scores as evidence for an overall assessment review. Results from online data files show the usefulness of our method in comparison with existing methods.

Keywords: E-commerce, Dempster-Shafer theory, Sentiment analysis, Customers' reviews, Recommendation.

JEL Classification: C49
AMS Classification: 28E15

1 Introduction

Online reviews can be good or bad, they are seldom neutral. These reviews are often provided in a free-text format by various websites that sell products and services. These reviews can help people decide better in their purchasing decisions. Mostly these text reviews are accompanied by data about the average review. The most common scheme of representing average review rating is five-star scores.
Sentiment prediction methods are also used to extract other useful information from review texts. For example, the text body of reviews may be used to extract reviewers' opinion toward different aspects of a product or service. For sellers or service providers (hotels, etc.), it is important to process these reviews, to find out sentiment of their customers and to check feedback.

However, full comprehension of natural language text contained in reviews remains beyond the power of algorithms used at the present time. On the other hand, the statistical analysis of relatively simple sentiment cues can provide a meaningful sense of how the online review impacts customers. There are essentially two ways to use the text body of reviews for predicting their overall scores: considering any review as a unit of text or treating it as a collection of sentences, each with its own sentiment orientation and score. The main drawback of the first viewpoint is that it ignores the fine-grained structural information of textual content [6]. Therefore, the second point of view is preferable. In this approach, the sentiment score of each sentence within a review is first computed. Then, a score aggregation method is used to combine sentence-level scores into an overall review score. In fact, score aggregation can be seen as a data fusion step in which sentence scores are multiple sources of information that should be taken to generate a single review score. Score aggregation is a widespread problem in sentiment analysis [5, 6, 10]. In this paper, we describe method based on the Dempster-Shafer (DS) theory of evidence [18]. We determine the consumers' sentiment from reviews from chosen e-shop, and also how this sentiment varies with time. We give example of our analysis. In this paper, we discuss several aspects of our model for sentiment analysis which include:

- Our model relies on tracking the reference frequencies of adjectives with positive and negative connotations. We present a method for creating a sentiment lexicon.
- We construct statistical indexes which meaningfully reflects the significance of sentiment of sentences and we apply the Dempster rule to integrate them.
- We provide evidence of the validity of our sentiment evaluation by correlating our index with evaluation of products on a chosen e-shop. Positive correlations prove that our model can measure public sentiment of consumers. Finally, we discuss possible applications and implications of our work.

[1] University of South Bohemia, Faculty of Economics, Department of Applied Mathematics and Informatics, Studentska 13, Ceske Budejovice, beranek@ef.jcu.cz.

The remainder of the paper is organized as follows. Section 2 and 3 reviews background and related work; Section 4 illustrates the materials and methods; Section 5 reports experimental results and presents a discussion of examined methods; finally, Section 6 sets out conclusion and future work.

2 Dempster-Shafer Theory

Information related to decision making about cyber situation is often uncertain and incomplete. Therefore, it is of vital importance to find a feasible way to make decisions about the appropriate response the situation under this uncertainty. Our model is a particular application of the Dempster-Shafer theory. The Dempster-Shafer theory [18] is designed to deal with the uncertainty and incompleteness of available information. It is a powerful tool for combining evidence and changing prior knowledge in the presence of new evidence. The Dempster-Shafer theory can be considered as a generalization of the Bayesian theory of subjective probability [8]. In this paper, we propose a unique trust model based on the Dempster-Shafer theory which combines evidence concerning reputation with evidence concerning possible illegal behavior on an Internet auction.

In the following paragraphs, we give a brief introduction to the basic notions of the Dempster-Shafer theory (frequently called theory of belief functions or theory of evidence).

Considering a finite set referred to as *the frame of discernment* Θ, a *basic belief assignment (BBA)* is a function $m\colon 2^{\Theta} \to [0,1]$ so that

$$\sum_{A \subseteq \Theta} m(A) = 1, \tag{1}$$

where $m(\emptyset) = 0$, see [18]. The subsets of 2^{Θ} which are associated with non-zero values of m are known as *focal elements* and the union of the focal elements is called *the core*. The value of $m(A)$ expresses the proportion of all relevant and available evidence that supports the claim that a particular element of Θ belongs to the set A but not to a particular subset of A. This value pertains only to the set A and makes no additional claims about any subsets of A. We denote this value also as a *degree of belief* (or *basic belief mass - BBM*).

Shafer further defined the concepts of *belief* and *plausibility* [18] as two measures over the subsets of Θ as follows:

$$Bel(A) = \sum_{B \subseteq A} m(B), \tag{2}$$

$$Pl(A) = \sum_{B \cap A \neq \phi} m(B) \ . \tag{3}$$

A *BBA* can also be viewed as determining a set of probability distributions P over Θ so that $Bel(A) \leq P(A) \leq Pl(A)$. It can be easily seen that these two measures are related to each other as $Pl(A) = 1 - Bel(\neg A)$. *Moreover, both of them are equivalent to m.* Thus one needs to know only one of the three functions m, Bel, or Pl to derive the other two. Hence we can speak about belief function using corresponding *BBA*s in fact.

Dempster's rule of combination can be used for pooling evidence represented by two belief functions Bel_1 and Bel_2 over the same frame of discernment coming from independent sources of information. The Dempster's rule of combination for combining two belief functions Bel_1 and Bel_2 defined by (equivalent to) *BBA*s m_1 and m_2 is defined as follows (the symbol $\otimes$ is used to denote this operation):

$$(m_1 \otimes m_2)(A) = \frac{1}{1-k} \sum_{B \cap C = A} m_1(B) \cdot m_2(C) \ , \tag{4}$$

where

$$k = \sum_{B \cap C = \emptyset} m_1(B) \cdot m_2(C) \ . \tag{5}$$

Here k is frequently considered to be a *conflict measure* between two belief functions m_1 and m_2 or a measure of conflict between m_1 and m_2 [18]. Unfortunately, this interpretation of k is not correct, as it includes also internal conflict of individual belief functions m_1 and m_2 [5]. Demspter's rule is not defined when $k = 1$, i.e. when cores of m_1 and m_2 are disjoint. This rule is commutative and associative; as the rule serves for the cumulation of beliefs, it is not idempotent.

When calculating contextual discounting [2] we also use the un-normalized (conjunctive) combination rule in the form [18] (we use the symbol ⊕ to denote this operation):

$$(m_1 \oplus m_2)(A) = \sum_{B \cap C = A} m_1(B) \cdot m_2(C)\,. \tag{6}$$

3 Related Work

Sentiment analysis has been receiving increasing attention at the present time. Some examples of sentiment analysis applications include predicting sales performance [11], ranking products and merchants, linking Twitter sentiment with public opinion polls [17], and identifying important product aspects [23] or services [3]. In addition to these traditional applications, sentiment analysis has presented new research opportunities for social sciences in recent years, for instance, characterizing social relations [7] or predicting the stock market using Twitter moods [4].

Sentiment analysis tasks can also be studied in terms of their sentiment prediction component. In this sense, existing approaches can be grouped into three main categories: machine learning approaches [13, 15], linguistic-based strategies, and lexicon-based methods [20, 21]. In the first category, input texts are first converted to feature vectors and then, using a machine learning algorithm, a classifier is trained on a human-coded corpus. Finally, the trained classifier is used for sentiment prediction. Linguistic approaches use simple rules based upon compositional semantics. They exploit the grammatical structure of text to predict its polarity. Lexicon-based approaches work primarily by identifying some predefined terms from a lexicon of known sentiment-bearing words. Lexicon-based techniques use a list of sentiment-bearing words and phrases that is called opinion lexicon [10]. In the current study we use a lexicon-based method. Lexicon-based techniques deal with the determining semantic orientation of words. Author in [20] evaluates adjectives for polarity as well as gradation classification. A statistical model groups adjectives into clusters, corresponding to their tone or orientation. The use of such gradable adjectives is an important factor in determining subjectivity. Statistical models are used to predict the enhancement of adjectives. Authors of the paper [12] evaluate the sentiment of an opinion holder (entity) using WordNet to generate lists of positive and negative words. They assume that synonyms (antonyms) of a word have the same (opposite) polarity. The percentage of a word's synonyms belonging to lists of either polarity was used as a measure of its polarity strength, while those below a threshold were considered neutral or ambiguous.

Several systems have been built which attempt to quantify opinion from product reviews. Pang, Lee and Vaithyanathan [16] perform sentiment analysis of movie reviews. Their results show that the machine learning techniques perform better than simple counting methods. They achieve an accuracy of polarity classification of roughly 83%. Author in [1] detect the polarity of reviews using a machine learning approach and then use integration based on belief function theory to the counting of overall review rating.

Nasukawa and Yi[14] identify local sentiment as being more reliable than global document sentiment, since human evaluators often fail to agree on the global sentiment of a document. They focus on identifying the orientation of sentiment expressions and determining the target of these sentiments. Shallow parsing identifies the target and the sentiment expression; the latter is evaluated and associated with the target. In [22], they follow up by employing a feature-term extractor. For a given item, the feature extractor identifies parts or attributes of that item. e.g., battery and lens are features of a camera.

4 Materials and methods

4.1 Sentiment lexicon generation

Two recent and widely used lexicon-based tools for sentiment strength detection are SO-CAL [20] and SentiStrength [21]. SO-CAL (Semantic Orientation CALculator) uses lexicons of terms coded on a single negative to positive range of −5 to +5 [30]. Its opinion lexicon was built by human coders tagging lemmatized noun and verbs as well as adjectives and adverbs for strength and polarity in 500 documents from several corpora (e.g., rating "hate" as −4 and "hilariously" as +4). In SO-CAL intensifying words have a percentage associated with them. For example, "slightly" and "very" have −50% and +25% modification impacts, respectively. Therefore, if "good" has a sentiment value of 3, then "very good" would have a value of 3 × (100% + 25%) = 3.75. Moreover, it amplifies the strength of negative expressions in texts and decreases the strength of frequent terms. Empirical tests showed that SO-CAL has robust performance for sentiment detection across different types of web texts [20].

4.2 The determination of belief functions for sentiment analysis

We denote $\Omega = \{positive, negative\}$ as a frame of discernment concerning the sentiment.
The power set of the set Ω (the set of all subsets) 2^Ω has three elements (we do not consider the empty set): $2^\Omega = \{\{positive\}, \{negative\}, \Omega\}$, where $\{positive\}$ represents positive for example evaluation of some product (positive sentiment), $\{negative\}$ means that this evaluation is negative (negative sentiment) and $\{\Omega\}$ denotes ignorance. It means that we cannot judge whether the respective evaluation is positive or negative, for example if evaluation is neutral (neither positive nor negative).

Polarity scores

We use the raw sentiment scores to track two trends over time:

- *Polarity:* Is the sentiment associated with the entity positive or negative?
- *Subjectivity:* How much sentiment (of any polarity) does the entity garner?

Subjectivity indicates proportion of sentiment to frequency of occurrence, while *polarity* indicates percentage of positive sentiment references among total sentiment references.

We focus first on polarity. We evaluate *world polarity* using sentiment data for all entities for the entire time period:

$$m_p(positive) = \frac{positive\_sentiment\_references}{total\_sentiment\_references},$$

$$m_p(negative) = \frac{negative\_sentiment\_references}{total\_sentiment\_references}, \tag{7}$$

$$m_p(\Omega) = 1 - m_p(positive) - m_p(negative).$$

We can evaluate *entity polarity$_i$* using these equations for that days (*day$_i$*) when we want to monitor changes in sentiment.

Subjectivity scores

The subjectivity time series reflects the amount of sentiment an entity is associated with, regardless of whether the sentiment is positive or negative. Reading all news text over a period of time and counting sentiment in it gives a measure of the average subjectivity levels of the world. We evaluate *world_subjectivity* using sentiment data for all entities for the entire time period:

$$m_s(\Omega) = \frac{total\_sentiment\_references}{total\_references},$$

$$m_s(positive) = \frac{1 - m_s(\Omega)}{2}, \tag{8}$$

$$m_s(negative) = 1 - m_s(positive) - m_s(\Omega).$$

We also can evaluate *entity subjectivity$_i$* using sentiment data for that day (*day$_i$*) only similar like in the previous case.

General index of sentiment

Once we obtain the belief functions, we combine them in a consistent manner to get a more complete assessment of what the whole group of signs indicates. The combination of belief functions is done with the help of the Dempster's combination rule, see equation (4).

$$(m_p \oplus m_s)(\{positive\}) = \frac{1}{K}\left[m_p(\{positive\}) \cdot m_s(\{positive\}) + m_p(\{positive\}) \cdot m_s(\Omega) + m_p(\Omega) \cdot m_s(\{positive\})\right]$$

$$(m_p \oplus m_s)(\{negative\}) = \frac{1}{K}\left[m_p(\{negative\}) \cdot m_s(\{negative\}) + m_p(\{negative\}) \cdot m_s(\Omega) + m_p(\Omega) \cdot m_s(\{negative\})\right]$$

$$(m_p \oplus m_s)(\{\Omega\}) = \frac{1}{K}\left[m_p(\Omega) \cdot m_s(\Omega)\right] \quad , \tag{9}$$

where K:

$$K = 1 - (m_p(\{negative\})m_s(\{positive\}) + m_p(\{positive\})m_s(\{negative\})) \,.$$

It should be emphasized that these scores are either positive or negative, or cannot decide. We cannot have a positive assessment while at the same time a negative. According this assumption, we also have defined appropriate frame of discernment.

5 Results and discussion

We evaluated the analysis system on a corpus of 158 hotel reviews crawled from the web. These reviews contained 642 reviews. For the evaluation, these reviews were manually classified with respect to their polarity, including the neutral polarity besides positive and negative ones. Also, we annotated the reviews whether they cover more than one topic. The distribution from this manual classification is shown in Table 1.

Reviews	Positive	Negative	Neutral	Multi-topic
642	305	142	123	72

Table 1. Manual corpus classification

The experiment data was formed by the following approach. First, the sentiment holder and target pair frequency is calculated. Then the most popular sentiment holders and targets are selected by identifying a threshold where the frequency drops significantly. This approach ensures that our selected data has enough common sentiment targets among the different sentiment holders for the evaluation.

For evaluation, we used the accuracy measure. The accuracy of a prediction method may be calculated according to the following equation:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \,, \tag{10}$$

where TP and TN are true positive and true negative, while FP and FN are false positive and false negative, respectively. In order to comparison our approach with others, we predict the sentiments using different algorithms: decision tree (J48 classification), K-Star, Naïve Bayes, and Support Vector Machines.

	Correct	False	Accuracy
J48	281	143	0.66
K-star	251	173	0.59
Naïve Bayes	263	161	0.62
SVM	277	147	0.65
Our model	268	156	0.63

Table 2. Comparison of various approaches with our model based on belief functions and their integration

Evaluated of different algorithms, the results in Table 2 were achieved. However, our model is relatively simple, and we expect improvement in accuracy after rebuilding and adjustment of our model.

6 Conclusion

In this paper we have studied sentiment prediction applied on data from hotel review sites. An evaluation of the techniques presented in the paper showed sufficient accuracy in sentiment prediction. Our experimental study and results showed that our suggested approach can be a useful tool for generating profiles of user sentiments.

There are many interesting directions that can be explored. In our future work, we want to explore how sentiment can vary by demographic group, news source or geographic location. By expanding our model with sentiment maps, we will be able to identify geographical regions of favorable or adverse opinions for given entities. We also want to analyze the degree to which our sentiment indices predict future changes in popularity or market behavior.

Reference

[1] Mohammad Ehsan Basiri, M.E., Naghsh-Nilchi, A.R., and Ghasem-Aghaee, N.: Sentiment Prediction Based on Dempster-Shafer Theory of Evidence. *Mathematical Problems in Engineering* **2014** (2014), Article ID 361201, 13 pages.

[2] Beranek, L., and Nydl, V.: The Use of Belief Functions for the Detection of Internet Auction Fraud. In: *Mathematical Methods in Economics 2013 (MME 2013)*, (Vojackova, H., ed.). Coll Polytechnics Jihlava, Jihlava, 2013, 31-36.

[3] Bjørkelund, E., Burnett, T.H., and Nørvåg, K.: A study of opinion mining and visualization of hotel reviews. In: *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services* (IIWAS '12). ACM, New York, NY, USA, 2012, 229–238.

[4] Bollen, J., Mao, H., and Zeng, X.: Twitter mood predicts the stock market. *Journal of Computational Science* **2** (2011), 1–8.

[5] Daniel, M.: Conflicts within and between belief functions. In: *Proceedings of the Computational Intelligence for Knowledge-based Systems Design, and 13th International Conference on Information Processing and Management of Uncertainty*. Dortmund, Germany, 2010, 696–705.

[6] Ganu, G., Kakodkar, Y., and Marian, A.: Improving the quality of predictions using textual information in online user reviews. *Information Systems* **38**, (2013), 1–15.

[7] Groh, G., and Hauffa,J.: Characterizing social relations via NLPbased sentiment analysis. In: *Proceedings of the International Conference onWeblogs and Social Media* (ICWSM '11), 2011.

[8] Haenni, R.: Shedding new light on Zadeh's criticism of Dempster's rule of combination. In: *Proceedings of the Eighth Int. Conf. on Information Fusion*. Philadelphia, IEEE, 2005, 879–884.

[9] Lappas, T.: Fake reviews: the malicious perspective. In: *Natural Language Processing and Information Systems*. Springer, 2012, 23–34.

[10] Liu, B.: *Sentiment Analysis and Opinion Mining*. *Synthesis Lectures on Human Language Technologies* vol. 5, Morgan & Claypool, 2012.

[11] Liu, B., Huang, X., An, A., and Yu, X.: ARSA: a sentiment-aware model for predicting sales performance using blogs . In: *Proceedings of the 30th Annual International ACMSIGIR Conference on Research and Development in Information Retrieval* (SIGIR '07). July 2007, 607–614.

[12] Liu, B.: Sentiment analysis: a multifaceted problem. *IEEE Intelligent Systems*, vol. 25, no. 3, 2010, 76–80.

[13] Minařík, M., and, Šťastný, J. Recognition of Randomly Deformed Objects. In: *MENDEL 2008, 14th International Conference on Soft Computing*. Brno, Czech Republic. 2008, pp. 275-280.

[14] Nasukawa, T., and, Yi, J.: Sentiment analysis: Capturing favorability using natural language processing. In: *The Second International Conferences on Knowledge Capture*. 2003, pp. 70–77.

[15] Paltoglou, G., and Thelwall, M.: Twitter, MySpace, Digg: unsupervised sentiment analysis in social media. *ACM Transactions on Intelligent Systems and Technology* (TIST) **3** (2012), article 66.

[16] Pang, B., and Lee, L.: Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, **2** (2008), 1–135.

[17] O'Connor, B., Balasubramanyan, R., Routledge, B.R., and Smith, N.A.: From tweets to polls: linking text sentiment to public opinion time series. In: *Proceedings of the International Conference on Weblogs and Social Media* (*ICWSM '10*). vol. 11, 2010, 122–129.

[18] Shafer, G.: *A mathematical theory of evidence*. Princeton University Press, Princeton, NJ, 1975.

[19] Stuckenschmidt, H., and Zirn, C.: Multi-dimensional analysis of political documents. In: *Natural Language Processing and Information Systems*. Springer, 2012, 11–22.

[20] Taboada, M., Brooke, J., Tofiloski,M., Voll, K., and Stede, M.: Lexicon-basedmethods for sentiment analysis. *Computational Linguistics* **37**, (2011), 267–307.

[21] Thelwall, M., Buckley, K., and Paltoglou, G.: Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology* **63** (2012), 163–173.

[22] Yi, J., Nasukawa, T., and Niblack, W.: Sentiment analyser: Extracting sentiments about a given topic using natural language processing techniques. In: *3rd IEEE Conf. on Data Mining* (ICDM'03), 2003, 423-434

[23] Yu, J., Zha,Z.J., Wang, M., and Chua, T.S.: Aspect ranking: Identifying important product aspects from online consumer reviews. In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (ACL-HLT '11). June 2011, 1496–1505.

OCA Index: Results for the EU's member states

Vendula Beránková[1]

Abstract. Probably the most famous monetary union in the world is the Eurozone, which now consists of nineteen European countries. The other EU member states are in the statute of candidate countries. The aim of the article is to evaluate an appropriateness of these nine EU countries standing out of the Eurozone for their membership according to the OCA index. This index, proposed by Bayoumi and Eichengreen [2], is based on the optimum currency area theory and its criteria.
The paper utilise the traditional approaches of the optimum currency area theory. The OCA index includes variables that are important in deciding the country's entry into the monetary union and it is estimated by using a panel regression method. First, the original equation of Bayoumi and Eichengreen [2], who estimated the OCA index first, is used. Then, the modified equation of Horváth and Komárek [7] is estimated. The sample covers the period from 2000 to 2014.
According to the results, there are countries which are more appropriate for the membership in the Eurozone because they reach satisfactory values of OCA index. But there are also countries with worse values and as such they are not appropriate candidates.

Keywords: Monetary Union, Optimum Currency Area, Eurozone, OCA Index.

JEL Classification: E42, E52
AMS Classification: 62M10

1 Introduction

The creation of monetary union is not something new in the world or not in the Europe. The reasons for creation of such a union can be different - historical, economic or political. There are few monetary unions such as the United States of America, the Eastern Caribbean Currency Area or the Central African Economic and Monetary Community in the world. But these days, probably the most famous monetary union is the monetary integration within the European Union – the Eurozone, which now consists of nineteen European countries[2]. The other nine EU member states are in the statute of candidate countries for their membership in the Eurozone. Assess, whether the country is a good candidate for membership in the monetary union, it is not as simple as it might seem at first glance. Undoubtedly, joining the monetary union brings the country certain benefits (removing of the transaction costs of exchange, removing of the exchange rate risks, or intensification of competition and pressure on production quality, etc.). But on the other hand, the membership in the monetary union also brings some disadvantages. And probably the biggest disadvantage is the loss of autonomous monetary policy – member country is losing the possibility to react to potential economic shocks. It is necessary, that these benefits outweigh the obvious disadvantages. The optimum currency area theory and its criteria are an appropriate instrument by which it is possible to assess the costs and benefits of membership in the monetary union.

The aim of this article is to evaluate an appropriateness of nine selected EU countries for their membership in the Eurozone according to the OCA index, which was created by Bayoumi and Eichengreen [2]. The OCA index is based on the optimum currency area theory and its criteria.

The remainder of the paper is organized as follows. Next section deals with the theoretical basis of this paper, methods and data are introduced in the next section. Then, results are discussed. The conclusion is offered in the final section.

2 Optimum Currency Area

Optimum currency area is a relatively young economic theory dating back to the 60s of the 20th century. In this period, the first landmark studies were written. These studies formed the basis of the optimum currency area theory and enable its further development. Namely it is the work of:

- Robert Mundell: *A Theory of Optimum Currency Areas* [12];

[1] VŠB-TU Ostrava, Faculty of Economics, Department of National Economy, Sokolská třída 33, Ostrava, Czech Republic, vendula.berankova@vsb.cz address.

[2] The Eurozone member states: Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxemburg, Malta, Netherlands, Portugal, Slovakia, Slovenia, Spain.

- Ronald McKinnon: *Optimum Currency Area* [10];
- Peter Kenen: *The Theory of Optimum Currency Areas: An Eclectic View* [8][3].

Based on these studies, three basic criteria for the optimum currency area were established. Concretely, these criteria are: Mundell criterion of workforce mobility, McKinnon criterion of openness of the economy and Kenen criterion of the diversification of production. These three criteria are very often accompanied by the criterion of alignment of economic cycles. When examining individual criteria, however we run into a problem, because empirical testing of fulfilling the OCA criteria does not give a clear answer to the question whether the ground for adoption of the single currency is advantageous or not. It is very difficult to assess what degree of fulfilment of the criteria and what combination of their performance is satisfactory, and what is not. The optimum currency area index is trying to overcome these problems (Hedija, [6]).

2.1 OCA Index

The OCA index assesses the costs and benefits of adopting the single currency. This index was first used by Tamim Bayoumi and Barry Eichengreen [2] in their article *Ever Closer to Heaven? An Optimum Currency Area Index for European Countries*. The authors wanted to develop a tool that would allow assessing on the basis of the OCA theory whether the country is suitable candidate for adopting the single currency. Hedija [6] suggests that the OCA index is constructed as a bilateral index, which assesses the appropriateness of introducing a single currency in the two countries. It is built on the optimum currency area theory and comes to the realization that the country's adoption of the single currency is the better, the smaller the tendency has nominal exchange rate to oscillate countries. These mutual fluctuations in nominal exchange rates of the currencies in the index are examined, depending on the level of fulfilment of the four criteria of optimum currency area:

1. the alignment of business cycles;
2. the similarity of economic structure;
3. the interdependence of trade;
4. the size of the economy.

It is clear that the OCA index is based on traditional approaches of the optimum currency area theory. It was originally compiled from a sample of 21 countries[4] during the period 1983 – 1992 using a panel regression method. The equation of Bayoumi and Eichengreen [2], who estimated the OCA index, is defined as follows:

$$SD(e_{ij}) = \beta_0 + \beta_1 SD(\Delta y_i - \Delta y_j) + \beta_2 DISSIM_{ij} + \beta_3 TRADE_{ij} + \beta_4 SIZE_{ij} \quad (1)$$

where $SD(e_{ij})$ is the standard deviation of the exchange in the logarithm of bilateral exchange rate at the end of the year between countries *i* and *j*, $SD(\Delta y_i - \Delta y_j)$ is the standard deviation of the difference in the logarithm of a real GDP between countries *i* and *j*, $DISSIM_{ij}$ is the sum of the absolute differences in the shares of agricultural, mineral, and manufacturing trade in a total merchandize trade, $TRADE_{ij}$ is the mean of the ratio of bilateral exports to domestic GDP for two countries *i* and *j*, $SIZE_{ij}$ is the mean of the logarithm of the two GDPs measured in U.S. dollars.

Horváth and Komárek [7] modify the equation of Bayoumi and Eichengreen when they calculate the OCA index – they use the variable $OPEN_{ij}$ (openness of the economy) instead of the variable $SIZE_{ij}$ (size of the economy). The reason was that Bayoumi and Eichengreen [2] found relatively little correlation between the size of the economy and trends in the transition to a fixed exchange rate regime. In contrast, the openness of the economy is one of the traditional optimum currency area criteria, and therefore Horváth and Komárek [7] include this variable in their regression model:

$$SD(e_{ij}) = \beta_0 + \beta_1 SD(\Delta y_i - \Delta y_j) + \beta_2 DISSIM_{ij} + \beta_3 TRADE_{ij} + \beta_4 OPEN_{ij} \quad (2)$$

where $SD(e_{ij})$ is the standard deviation of the exchange in the logarithm of bilateral exchange rate at the end of the year between countries *i* and *j*, $SD(\Delta y_i - \Delta y_j)$ is the standard deviation of the difference in the logarithm of a real GDP between countries *i* and *j*, $DISSIM_{ij}$ is the sum of the absolute differences in the shares of agricultural, mineral, and manufacturing trade in a total merchandize trade, $TRADE_{ij}$ is the mean of the ratio of bilateral

[3] The issue of optimum currency area is also engaged in a number of other, respectively recent studies. Their authors are e.g. Bayoumi and Eichengreen [2], Fidrmuc and Korhonen [5], Cincibuch and Vávra [3], Mongelli [13], Horváth and Komárek [7], Kučerová [9], Mink, Jacobs and Haan [11], Bachanová [1], Eichengreen [4], Skořepa [14] or Hedija [6].

[4] European Union countries that were members at the time the article was published, without Luxembourg (i.e. Belgium, Denmark, Finland, France, Ireland, Italy, Netherlands, Portugal, Austria, Greece, the United Kingdom, Spain, Sweden), followed by Norway, Switzerland, USA, Canada, Australia, New Zealand and Japan.

exports to domestic GDP for two countries i and j, $OPEN_{ij}$ is the mean of the ratio of the trade, i.e. export and import to their GDP.

Variable *SD*

The variable $SD(\Delta y_i - \Delta y_j)$ is the standard deviation of the difference in the logarithm of real GDP of two countries. The real GDP (in prices of year 2005) is measured in U.S. dollars. It is computed as the difference of annual real GDP for each country:

$$SD(\Delta y_i - \Delta y_j) = SD\left[\ln\frac{RGDP_{i(t)}}{RGDP_{i(t-1)}};\ln\frac{RGDP_{j(t)}}{RGDP_{j(t-1)}}\right] \tag{3}$$

where SD is standard deviation, $RGDP_{i(t)}$ is real GDP of country i in time t, $RGDP_{i(t-1)}$ is real GDP of country i in time $t-1$, $RGDP_{j(t)}$ is real GDP of country j in time t and $RGDP_{j(t-1)}$ is real GDP of country j in time $t-1$.

Variable *DISSIM*

The variable $DISSIM_{ij}$ is the sum of the absolute differences in the share of individual components[5] in total bilateral trade. It attains values from 0 to 2. The value 0 means the same structure of bilateral trade. The value 2 means that commodity structure of bilateral trade of two countries is absolutely different. In this case, lower value implies better conditions to adopt a common currency. The variable has the following specification:

$$DISSIM_{ij} = \sum_{A=1}^{N}\left|\frac{XA_{ij}}{X_{ij}} - \frac{XA_{ji}}{X_{ji}}\right| \tag{4}$$

where XA is the export of each economic category, X_{ij} is total export from country i to country j and X_{ji} is export from country j to country i.

Variable *TRADE*

The variable $TRADE_{ij}$ is the mean of the ratio of bilateral trade (import plus export) to nominal GDP of countries i and j. This nominal GDP is measured in U.S. dollars. The variable $TRADE_{ij}$ was computed as follows:

$$TRADE_{ij} = MEAN\left[\frac{X_{ij}}{GDP_i};\frac{X_{ji}}{GDP_j}\right] \tag{5}$$

where X_{ij} is nominal export from country i to country j, X_{ji} is nominal export from country j to country i, GDP_i is nominal GDP of country i and GDP_j is nominal GDP of country j. Higher value of this variable means better conditions to adopt a common currency because common currency is more convenient for countries which have higher level of bilateral trade.

Variable *SIZE*

The variable $SIZE_{ij}$ represents the size of the economies. It is computed as the mean of the logarithm of real GDP of countries i and j. Again, the real GDP (in prices of year 2005) is measured in U.S. dollars.

$$SIZE_{ij} = MEAN\left[\ln RGDP_i;\ln RGDP_j\right] \tag{6}$$

where $RGDP_i$ is the real GDP of country i and $RGDP_j$ is the real GDP of country j.

Variable *OPEN*

The variable $OPEN_{ij}$ measures the rate of openness of each economy. It is computed as the mean of the share of nominal trade (import plus export) on nominal GDP (measured in U.S. dollars) of two countries:

$$OPEN_{ij} = MEAN\left[\frac{X_i + M_i}{GDP_i};\frac{X_j + M_j}{GDP_j}\right] \tag{7}$$

[5] The variable $DISSIM_{ij}$ is computed using the UN`s Basic Economic Categories (BEC) classification which distinguishes economic categories into 7 parts: 1. Food and beverages; 2. Industrial supplies; 3. Fuels and lubricants; 4. Capital goods; 5. Transport equipment; 6. Consumption goods; 7. Others.

where X_i is total nominal export of country *i*, M_i is total nominal import of country *i*, GDP_i is nominal GDP of country *i*, X_j is total nominal export of country *j*, M_j is total nominal import of country *j*, GDP_j is nominal GDP of country *j*.

2.2 Data

Data for variables $SD(\Delta y_i - \Delta y_j)$, $SIZE_{ij}$ and $OPEN_{ij}$ is used from the World Bank Database. Variables $DISSIM_{ij}$ and $TRADE_{ij}$ are calculated with the use of the data from the United Nations Commodity Trade Statistics Database and the World Bank Database. Data for variables $SD(e_{ij})$ is used from Eurostat. For all calculations, an annual data from the period 2000 – 2014 are used.

3 Results

In this case, the OCA index is compiled from a sample of nine selected European countries during the period 1983 – 1992. These countries are: Bulgaria, Croatia, the Czech Republic, Denmark, Hungary, Poland, Romania, Sweden, and the United Kingdom. Of course, the panel regression method is used. The equations of the OCA index look like follows (standard errors of coefficient estimates are in parenthesis):

$$OCA\ index = 0{,}016 + 0{,}474\ SD(\Delta y_i - \Delta y_j) + 0{,}031\ DISSIM_{ij} + 0{,}145\ TRADE_{ij} - 0{,}12\ SIZE_{ij} \qquad (8)$$

(0,0518) (0,0045) (0,0731) (0,0113) (0,2702)

$n = 126$ $R^2 = 0{,}515$ $SE = 0{,}022$

$$OCA\ index = 0{,}014 + 0{,}397\ SD(\Delta y_i - \Delta y_j) + 0{,}031\ DISSIM_{ij} + 0{,}114\ TRADE_{ij} + 0{,}010\ OPEN_{ij} \qquad (9)$$

(0,1659) (0,0087) (0,1555) (0,0736) (0,6978)

$n = 126$ $R^2 = 0{,}507$ $SE = 0{,}022$

The equations for calculating the OCA index receives a specific form. Estimated values of coefficients β express the sensitivity of the index to the explanatory variable. When calculating the OCA index, both equations are used. The index is computed from annual data from the period 2000 – 2014. This selected time series are divided into two periods: pre-crisis period (2000 – 2007) and post-crisis period (2008 – 2014). Then, the OCA index for the whole period 2000 – 2014 is computed. This allows the appreciation of the OCA index over the time. The lower the OCA index is the more suitable candidate for adopting common currency the country is[6]. The empirical results are introduced in following Table 1 and Table 2.

Country	**2000 – 2007**	**2008 – 2014**	**2000 – 2014**
Bulgaria	0,0476	0,0385	0,0433
Croatia	0,0401	0,0316	0,0361
Czech Republic	0,0412	0,0429	0,0417
Denmark	0,0346	0,0337	0,0342
Hungary	0,0440	0,0428	0,0435
Poland	0,0427	0,0443	0,0435
Romania	0,0496	0,0443	0,0471
Sweden	0,0376	0,0394	0,0385
United Kingdom	0,0327	0,0378	0,0351

Table 1 OCA Index based on equation (8), own calculations

In the Table 1, there are presented results based on the equation (8). But the variable $SIZE_{ij}$ is not included in calculations, because this variable seems to be statistically insignificant. Therefore, the OCA index is calculated in this case only by significant variables $SD(\Delta y_i - \Delta y_j)$, $DISSIM_{ij}$ and $TRADE_{ij}$. The best values are observed

[6] Variables are computed for all nine EU countries outside the Eurozone, but there is a problem, because some data are not available for the Eurozone. So, when calculating variables $DISSIM_{ij}$ and $TRADE_{ij}$, Germany is used instead of the Euro Area.

in Croatia, Denmark, Sweden and the United Kingdom. Conversely, the worst value of the OCA index is in Romania. It should be also noted that since 2008 the OCA index has significantly improved in Bulgaria, Croatia and also Romania. On the other hand, since 2008 the OCA index has got worse in the Czech Republic, Hungary, Poland, Sweden and the United Kingdom.

In the Table 2, there are presented results based on the equation (9). The best values of the OCA index are in the case of Croatia, Denmark and the United Kingdom. The worst values are observed in Bulgaria, Hungary and Romania. Since 2008, the OCA index has significantly improved in Bulgaria, Croatia and also Romania. On the other hand, the OCA index has got worse in the Czech Republic, Hungary, Poland, Sweden and the United Kingdom after the crisis.

Country	**2000 – 2007**	**2008 – 2014**	**2000 – 2014**
Bulgaria	0,0516	0,0451	0,0486
Croatia	0,0445	0,0365	0,0408
Czech Republic	0,0442	0,0477	0,0456
Denmark	0,0393	0,0395	0,0394
Hungary	0,0483	0,0492	0,0487
Poland	0,0455	0,0474	0,0464
Romania	0,0526	0,0478	0,0504
Sweden	0,0421	0,0444	0,0432
United Kingdom	0,0359	0,0411	0,0383

Table 2 OCA Index based on equation (9), own calculations

However, neither the assessment of the appropriateness of adopting a common currency using reconstructed index is not entirely clear nor straight forward. Moreover, Bayoumi and Eichengreen [2] find relatively little correlation between the size of the economy and trends in the transition to a fixed exchange rate regime. This problem was solved by Horváth and Komárek [7] who use variable $OPEN_{ij}$ instead of the variable $SIZE_{ij}$ and their equation could have better informative value.

4 Conclusion

Probably the most famous monetary union is the Eurozone, which now consists of nineteen European countries. Nine EU member states are in the statute of candidate countries. The aim of this article was to evaluate an appropriateness of these nine EU states for a membership in the Eurozone according to the OCA criteria.

The research was based on traditional approaches of the optimum currency area theory. It was used the methodology of Bayoumi and Eichengreen [2] who estimated the OCA index. Then the modified equation of Horváth and Komárek [7] was used. The OCA index was computed for the period 2000 – 2014.

Using the equation of Bayoumi and Eichengreen, the best values of OCA index are observed in Croatia, Denmark, Sweden and the United Kingdom. Conversely, the worst value of the OCA index is in Romania. Note that after the crisis, the OCA index has significantly improved in Bulgaria, Croatia and Romania. On the other hand, it has got worse in the Czech Republic, Hungary, Poland, Sweden and the United Kingdom since 2008.

However, because Bayoumi and Eichengreen [2] find relatively little correlation between the size of the economy and trend in the transition to a fixed exchange rate regime, Horváth and Komárek [7] modified their equation. Using the equation of Horváth and Komárek, the best values of the OCA index are in Croatia, Denmark and the United Kingdom. The worst values are observed in Bulgaria, Hungary and Romania. Since 2008, the OCA index has significantly improved in Bulgaria, Croatia and Romania, while in the case of the Czech Republic, Hungary, Poland, Sweden and the United Kingdom it has got worse.

According to these results it may be concluded that there were some countries, which were more appropriate for the membership in the Eurozone. These countries were especially Croatia, Denmark and the United Kingdom. Conversely, Romania is not appropriate for the membership in the Eurozone. These nine countries can be divided into three groups. First group, the main candidates for the membership in the Eurozone, consists of the United Kingdom, Denmark and Croatia. These countries reach satisfactory degree of convergence. Second group consists of the Czech Republic, Poland and Sweden. These countries gradually converge to the Eurozone. And finally, the third group of countries, which has a small degree of convergence, includes Romania, Bulgaria and Hungary.

Of course, the candidate countries must fulfil all the convergence criteria if they want to join the Eurozone. But the OCA index can be used as a convenient additional instrument by which it is possible to determine the appropriateness of countries for the membership in the Eurozone.

Acknowledgements

This work was supported by the VSB-Technical University of Ostrava, Faculty of Economics under Grant SGS SP2016/101.

References

[1] Bachanová, V.: Index optimální měnové oblasti pro Českou republiku. *Ekonomická revue - Central European Review of Economic Issues* **11** (2008), 42–57.
[2] Bayoumi, T., and Eichengreen, B.: Ever Closed to Heaven? An Optimum-Currency-Area Index for European Countries. *European Economic Review* **41** (1997), 761–770.
[3] Cincibuch, M., and Vávra, D.: Na cestě k EMU: Potřebujeme flexibilní měnový kurz? *Finance a úvěr* **50** (2000), 361–384.
[4] Eichengreen, B.: Sui Generis EMU. *NBER Working Paper No. 13740*. University of California, Berkeley, 2008.
[5] Fidrmuc, J., and Korhonen, I.: Similarity of Supply and Demand Shocks Between the Euro Area and the CEECs. *Economic Systems* **27** (2003), 313–334.
[6] Hedija, V.: Index OCA – aplikace na země EU10. *Ekonomická revue - Central European Review of Economic Issues* **14** (2011), 85–93.
[7] Horváth, R., and Komárek, L.: Teorie optimálních měnových zón: rámec k diskuzím o monetární integraci. *Finance a úvěr* **52** (2002), 386–407.
[8] Kenen, P. B.: The Theory of Optimum Currency Areas: An Eclectic View. In: *Monetary Problems of the International Economy* (Mundell, R. A., and Swoboda, A. K., eds.). University of Chicago Press, Chicago, 1969.
[9] Kučerová, Z.: *Teorie optimální měnové oblasti a možnosti její aplikace na země střední a východní Evropy*. Národohospodářský ústav Josefa Hlávky, Praha, 2005.
[10] McKinnon, R. I.: Optimum Currency Areas. *American Economic Review* **53** (1963), 717–725.
[11] Mink, M., Jacobs, J., and Haan, J.: Measuring Synchronicity and Co-movement of Business Cycles with an Application on the Euro Area. *CESifo Working Paper No. 2112*. University of Groningen, Groningen, 2007.
[12] Mundell, R. A.: Theory of Optimum Currency Areas. *American Economic Review* **51** (1961), 657–665.
[13] Mongelli, F. P.: "New" Views on the Optimum Currency Area Theory: What is EMU Telling us? *European Central Bank Working Paper No. 138*. European Central Bank, Frankfurt, 2002.
[14] Skořepa, M.: A Convergence-Sensitive Optimum-Currency-Area Index. *IES Working Paper: 23/2011*. Charles University, Prague, 2011.

Trimmed L-moments in Modeling Income Distribution

Diana Bílková[1]

Abstract. This paper deals with the application of such robust method of point parameter estimation, as the method of TL-moments on economic data. The advantages of this highly robust parametric estimation method are aware when applied to small data sets, especially in the field of hydrology, meteorology and climatology, in particular considering extreme precipitation.
The main aim of this contribution is to use this method on large datasets, and comparison the accuracy of this method of parametric estimation with the accuracy of the other two methods, such as the method of L-moments and maximum likelihood method, especially in terms of efficiency of parametric estimation. The study is divided into a theoretical part, in which mathematical and statistical aspects are described, and an analytical part, during which the results of the use of three robust parametric estimation methods are presented. The basis for the research are slightly older data on household net annual income per capita (it failed to obtain some later data). Total 168 income distributions of the years from 1992 to 2007 in the Czech Republic (distribution of household net annual income per capita in CZK) were analyzed.

Keywords: L-moments and TL-moments of probability distribution, sample L-moments and TL-moments, probability density function, distribution function, quantile function, order statistics, models of income distribution.

JEL Classification: C13, C46
AMS Classification: 60E05

1 Introduction

Alternative robust version of L-moments, see for example [1], [3]–[4] or [7]–[8], will be now presented. This robust modification of L-moments is called „trimmed L-moments“, and labeled „TL-moments“. Three-parametric lognormal curves, see [5]–[6], represent the basic theoretical distribution whose parameters were simultaneously estimated by three methods of point parameter estimation (methods of TL-moments and L-moments and maximum likelihood method in combination with Cohen’s method) and accuracy of these methods was then evaluated.

This is a relatively new category of moment characteristics of the probability distribution. There are the characteristics of the level, variability, skewness and kurtosis of probability distributions constructed using TL-moments that are robust extending of L-moments. L-moments alone were introduced as a robust alternative to classical moments of probability distributions. However, L-moments and their estimations lack some robust properties that belong to the TL-moments.

The advantages of this highly robust parametric estimation method are aware when applied to small data sets, especially in the field of hydrology, meteorology and climatology. This paper shows the advantages of TL-moments when using to large data sets.

Sample TL-moments are linear combinations of sample order statistics, which assign zero weight to a predetermined number of sample outliers. Sample TL-moments are unbiased estimations of the corresponding TL-moments of probability distributions. Some theoretical and practical aspects of TL-moments are still under research or remain for future research. Efficiency of TL-statistics depends on the choice of α proportion, for example, the first sample TL-moments $l_1^{(0)}, l_1^{(1)}, l_1^{(2)}$ have the smallest variance (the highest efficiency) among other estimations from random samples from normal, logistic and double exponential distribution.

When constructing the TL-moments, the expected values of order statistics of random sample in the definition of L-moments of probability distributions are replaced by the expected values of order statistics of a larger random sample, where the sample size grows like this, so that it will correspond to the total size of modification, as shown below.

[1] University of Economics, Prague; Faculty of Informatics and Statistics; Department of Statistics and Probability; Sq. W. Churchill 1938/4; 130 67 Prague 3; Czech Republic; e-mail: diana.bilkova@vse.cz.

TL-moments have certain advantages over conventional L-moments and central moments. TL-moment of probability distribution may exist even if the corresponding L-moment or central moment of the probability distribution does not exist, as it is the case of Cauchy's distribution. Sample TL-moments are more resistant to existence of outliers in the data. The method of TL-moments is not intended to replace the existing robust methods, but rather as their supplement, especially in situations where we have outliers in the data.

2 Theory and Methods

2.1 Trimmed L-moments of Probability Distribution

In this alternative robust modification of L-moments, the expected value $E(X_{r-j:r})$ is replaced by the expected value $E(X_{r+t1-j:r+t1+t2})$. Thus, for each r we increase sample size of random sample from the original r to $r + t_1 + t_2$ and we work only with the expected values of these r treated order statistics $X_{t1+1:r+t1+t2}, X_{t1+2:r+t1+t2}, \ldots, X_{t1+r:r+t1+t2}$ by trimming the t_1 smallest and the t_2 largest from the conceptual sample. This modification is called the r-th trimmed L-moment (TL-moment) and is marked $\lambda_r^{(t_1,t_2)}$. Thus, TL-moment of the r-th order of random variable X is defined

$$\lambda_r^{(t_1,t_2)} = \frac{1}{r}\cdot\sum_{j=0}^{r-1}(-1)^j\cdot\binom{r-1}{j}\cdot E(X_{r+t_1-j:r+t_1+t_2}), \quad r = 1, 2, \ldots. \tag{1}$$

It is apparent that the TL-moments simplify to L-moments, when $t_1 = t_2 = 0$. Although we can also consider applications, where the values of trimming are not equal, i.e. $t_1 \neq t_2$, we focus here only on symmetric case $t_1 = t_2 = t$. Then equation (1) can be rewritten

$$\lambda_r^{(t)} = \frac{1}{r}\cdot\sum_{j=0}^{r-1}(-1)^j\cdot\binom{r-1}{j}\cdot E(X_{r+t-j:r+2t}), \quad r = 1, 2, \ldots. \tag{2}$$

Thus, for example, $\lambda_1^{(t)} = E(X_{1+t:1+2t})$ is the expected value of median from conceptual random sample of sample size $1 + 2t$. It is necessary here to note that $\lambda_1^{(t)}$ is equal to zero for distributions, which are symmetrical around zero.

First four TL-moments have the form for $t = 1$

$$\lambda_1^{(1)} = E(X_{2:3}), \tag{3}$$

$$\lambda_2^{(1)} = \frac{1}{2}E(X_{3:4} - X_{2:4}), \tag{4}$$

$$\lambda_3^{(1)} = \frac{1}{3}E(X_{4:5} - 2X_{3:5} + X_{2:5}), \tag{5}$$

$$\lambda_4^{(1)} = \frac{1}{4}E(X_{5:6} - 3X_{4:6} + 3X_{3:6} - X_{2:6}). \tag{6}$$

Note that the measures of location (level), variability, skewness and kurtosis of the probability distribution are based on $\lambda_1^{(1)}, \lambda_2^{(1)}, \lambda_3^{(1)}$ a $\lambda_4^{(1)}$.

Expected value $E(X_{r:n})$ can be written using the formula

$$E(X_{r:n}) = \frac{n!}{(r-1)!\cdot(n-r)!}\cdot\int_0^1 x(F)\cdot[F(x)]^{r-1}\cdot[1-F(x)]^{n-r}\,\mathrm{d}F(x). \tag{7}$$

Using equation (7) we can re-express the right side of equation (2)

$$\lambda_r^{(t)} = \frac{1}{r}\cdot\sum_{j=0}^{r-1}(-1)^j\cdot\binom{r-1}{j}\cdot\frac{(r+2t)!}{(r+t-j-1)!\cdot(t+j)!}\cdot\int_0^1 x(F)\cdot[F(x)]^{r+t-j-1}\cdot[1-F(x)]^{t+j}\,\mathrm{d}F(x), r = 1, 2, \ldots. \tag{8}$$

It is necessary to be noted here that $\lambda_r^{(0)} = \lambda_r$ is a normal the r-th L-moment without any trimming.

Expressions (3)−(6) for the first four TL-moments, where $t = 1$, can be written in an alternative manner

$$\lambda_1^{(1)} = 6\cdot\int_0^1 x(F)\cdot[F(x)]\cdot[1-F(x)]\,\mathrm{d}F(x), \tag{9}$$

$$\lambda_2^{(1)} = 6 \cdot \int_0^1 x(F) \cdot [F(x)] \cdot [1 - F(x)] \cdot [2F(x) - 1] \, \mathrm{d} F(x), \tag{10}$$

$$\lambda_3^{(1)} = \frac{20}{3} \cdot \int_0^1 x(F) \cdot [F(x)] \cdot [1 - F(x)] \cdot \{5[F(x)]^2 - 5F(x) + 1\} \, \mathrm{d} F(x), \tag{11}$$

$$\lambda_4^{(1)} = \frac{15}{2} \cdot \int_0^1 x(F) \cdot [F(x)] \cdot [1 - F(x)] \cdot \{14[F(x)]^3 - 21[F(x)]^2 + 9[F(x)] - 1] \, \mathrm{d} F(x). \tag{12}$$

Distribution may be identified by its TL-moments, although some of its L-moments or conventional central moments do not exit; for example $\lambda_1^{(1)}$ (expected value of median of conceptual random sample of sample size three) exists for Cauchy's distribution, although the first L-moment λ_1 does not exist.

TL-skewness $\tau_3^{(t)}$ and TL-kurtosis $\tau_4^{(t)}$ are defined analogously as L-skewness and L-kurtosis

$$\tau_3^{(t)} = \frac{\lambda_3^{(t)}}{\lambda_2^{(t)}}, \tag{13}$$

$$\tau_4^{(t)} = \frac{\lambda_4^{(t)}}{\lambda_2^{(t)}}. \tag{14}$$

2.2 Sample Trimmed L-moments

Let $x_1, x_2, \dots, x_n$ is a sample and $x_{1:n} \le x_{2:n} \le \dots \le x_{n:n}$ is an ordered sample. Expression

$$\hat{E}(X_{j+1:j+l+1}) = \frac{1}{\binom{n}{j+l+1}} \cdot \sum_{i=1}^{n} \binom{i-1}{j} \cdot \binom{n-i}{l} \cdot x_{i:n} \tag{15}$$

is considered to be an unbiased estimation of expected value of the $(j + 1)$-th order statistic $X_{j+1:j+l+1}$ in conceptual random sample of sample size $(j + l + 1)$. Now we will assume that we replace the expression $E(X_{r+t-j:r+2t}$) by its unbiased estimation in the definition of the r-th TL-moment $\lambda_r^{(t)}$ in (2)

$$\hat{E}(X_{r+t-j:r+2t}) = \frac{1}{\binom{n}{r+2t}} \cdot \sum_{i=1}^{n} \binom{i-1}{r+t-j-1} \cdot \binom{n-i}{t+j} \cdot x_{i:n}, \tag{16}$$

which we gain by assigning $j \to r + t - j - 1$ a $l \to t + j$ in (15). Now we obtain the r-th sample TL-moment

$$l_r^{(t)} = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot \hat{E}(X_{r+t-j:r+2t}), \quad r = 1, 2, \dots, n - 2t, \tag{17}$$

i.e.

$$l_r^{(t)} = \frac{1}{r} \cdot \sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \cdot \frac{1}{\binom{n}{r+2t}} \cdot \sum_{i=1}^{n} \binom{i-1}{r+t-j-1} \cdot \binom{n-i}{t+j} \cdot x_{i:n}, \quad r = 1, 2, \dots, n - 2t, \tag{18}$$

which is unbiased estimation of the r-th TL-moment $\lambda_r^{(t)}$. Note that for each $j = 0, 1, \dots, r - 1$, values $x_{i:n}$ in (18) are nonzero only for $r + t - j \le i \le n - t - j$ due to the combinatorial numbers. Simple adjustment of the equation (18) provides an alternative linear for

$$l_r^{(t)} = \frac{1}{r} \cdot \sum_{i=r+t}^{n-t} \left[\frac{\sum_{j=0}^{r-1} (-1)^j \cdot \binom{r-1}{j} \binom{i-1}{r+t-j-1} \cdot \binom{n-i}{t+j}}{\binom{n}{r+2t}} \right] \cdot x_{i:n}. \tag{19}$$

For example, we obtain for $r = 1$ for the first sample TL-moment

$$l_1^{(t)} = \sum_{i=t+1}^{n-t} w_{i:n}^{(t)} \cdot x_{i:n}, \tag{20}$$

where the weights are given by

$$w_{i:n}^{(t)} = \frac{\binom{i-1}{t} \cdot \binom{n-i}{t}}{\binom{n}{2t+1}}. \tag{21}$$

The above results can be used to estimate TL-skewness and TL-kurtosis by simple ratios

$$t_3^{(t)} = \frac{l_3^{(t)}}{l_2^{(t)}}, \tag{22}$$

$$t_4^{(t)} = \frac{l_4^{(t)}}{l_2^{(t)}}. \tag{23}$$

We can choose $t = n\alpha$ representing the amount of the adjustment from each end of the sample, where α is a certain proportion, where $0 \le \alpha < 0{,}5$. More on the TL-moments is for example in [2].

3 Results and Discussion

The researched variable is the net annual household income per capita (in CZK) within the Czech Republic (nominal income). The data obtained come from a statistical survey Microcensus – years 1992, 1996, 2002, and statistical survey EU-SILC (The European Union Statistics on Income and Living Conditions) – the period 2004-2007, from the Czech Statistical Office. Total 168 income distributions were analyzed this way, both for all households of the Czech Republic together and also broken down by gender, country (Bohemia and Moravia), social groups, municipality size, age and the highest educational attainment, while households are classified into different subsets according to the head of household.

Method of TL-moments provided the most accurate results in almost all cases, with the negligible exceptions. Method of L-moments results as the second in more than half of the cases, although the differences between the method of L-moments and maximum likelihood method are not distinctive enough to turn in the number of cases where the method of TL-moments came out better than the method of L-moments. Table 1 is a typical representative of the results for all 168 income distributions. This table provides the results for the total household sets in the Czech Republic. It contains the value of known test criterion χ^2. This is evident from the values of the criterion that the method of L-moments brought more accurate results than maximum likelihood method in four of seven cases. The most accurate results were obtained using the method of TL-moments in all seven cases. Figures 1–3 allow the comparison of these methods in terms of model probability density functions in choosing years (1992, 2004 and 2007) for the total set of households throughout the Czech Republic together. It should be noted at this point that other scale is on the vertical axis in Figure 1 than in Figures 2 and 3 for better legibility. It is clear from these three figures that the methods of TL-moments and L-moments bring the very similar results, while the probability density function with the parameters estimated by maximum likelihood method is very different from model probability density functions constructed using the method of TL-moments and the method of L-moments. Figures 4–6 only confirm this statement.

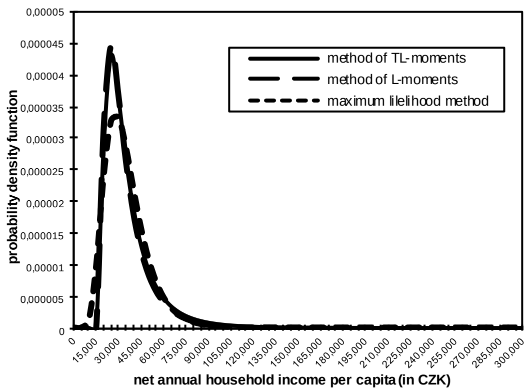


Figure 1 Model probability density functions of three-parametric lognormal curves in 1992 with parameters estimated using three various robust methods of point parameter estimation

Source: Own research

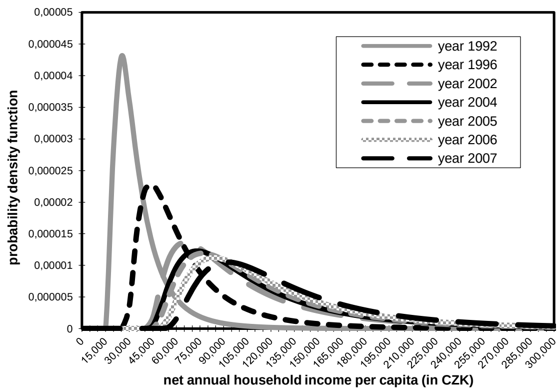


Figure 4 Development of probability density function of three-parameter lognormal curves with parameters estimated using the method of TL-moments

Source: Own research

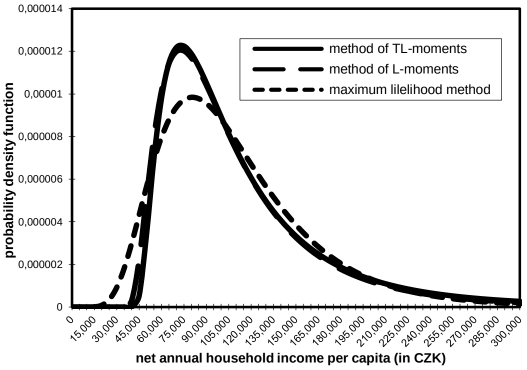


Figure 2 Model probability density functions of three-parametric lognormal curves in 2004 with parameters estimated using three various robust methods of point parameter estimation

Source: Own research

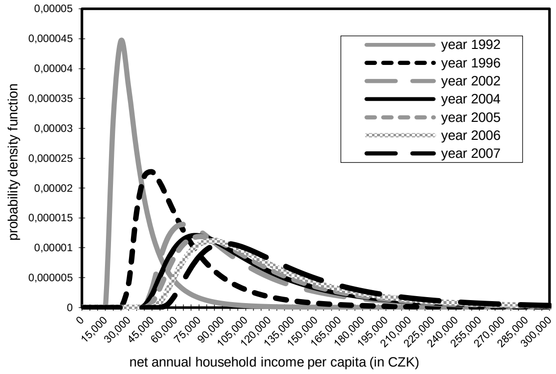


Figure 5 Development of probability density function of three-parameter lognormal curves with parameters estimated using the method of L-moments

Source: Own research

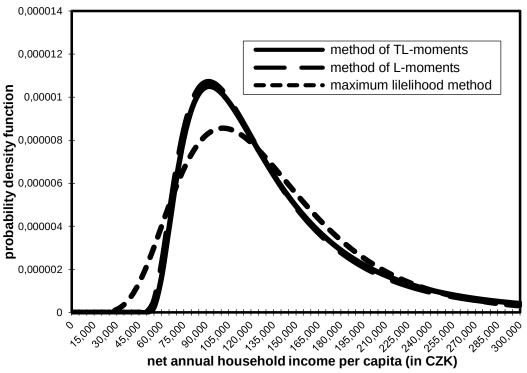


Figure 3 Model probability density functions of three-parametric lognormal curves in 2007 with parameters estimated using three various robust methods of point parameter estimation

Source: Own research

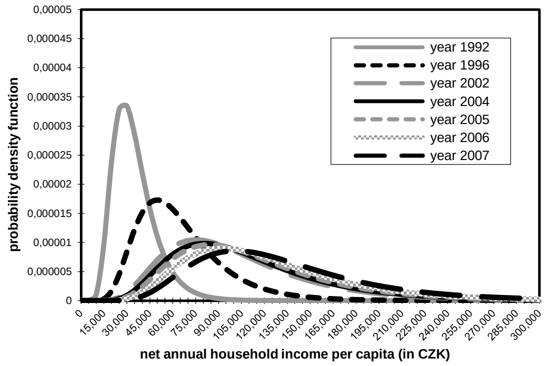


Figure 6 Development of probability density function of three-parameter lognormal curves with parameters estimated using the maximum likelihood method

Source: Own research

Microcensus				EU-SILC			
Year	TL-moments	L-moments	Maximum likelihood	Year	TL-moments	L-moments	Maximum likelihood
1992	739.512	811.007	1,227.325	2004	494.441	866.279	524.478
1996	1,503.878	1,742.631	2,197.251	2005	731.225	899.245	995.855
2002	998.325	1,535.557	1,060.891	2006	831.667	959.902	1,067.789
				2007	1,050.105	1,220.478	1,199.035

Table 1 Value of χ^2 criterion of parameter estimations of three-parametric lognormal curves obtained using three various robust methods of point parameter estimation

Source: Own research

Figures 7–9 then represent the model relative frequencies (in %) of employees by the band of net annual household income per capita in 2007 obtained using three-parametric lognormal curves with parameters estimated by the method of TL-moments, method of L-moments and maximum likelihood method. These figures also allow some comparison of the accuracy of the researched methods of point parameter estimation compared with Figure 10, where are the really observed relative frequencies in individual sample income intervals.

4 Conclusion

Relatively new class of moment characteristics of probability distributions were here introduced. There are the characteristics of location (level), variability, skewness and kurtosis of probability distributions constructed using TL-moments. The accuracy of TL-moments method was compared to that of L-moments and the maximum likelihood method. Higher accuracy of the former approach in comparison to that of the latter two methods has been proved by 168 income distribution data sets. Advantages of L-moments over the maximum likelihood method have been demonstrated by the present study as well.

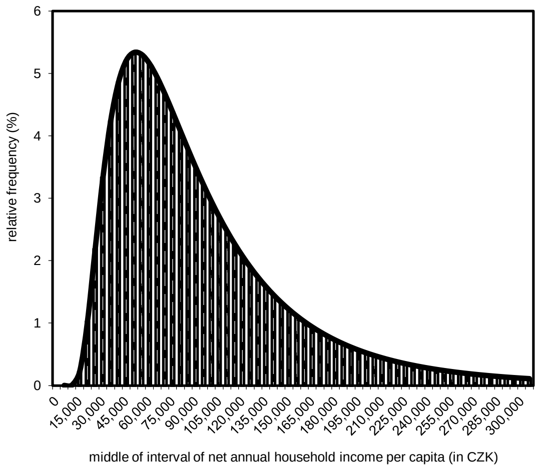


Figure 7 Method of TL-moments

Source: Own research

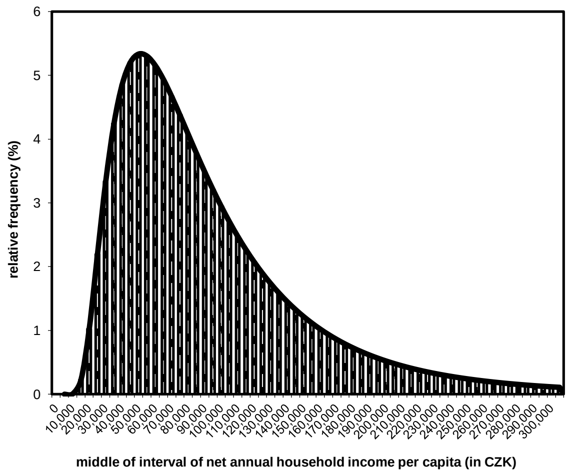


Figure 9 Maximum likelihood method

Source: Own research

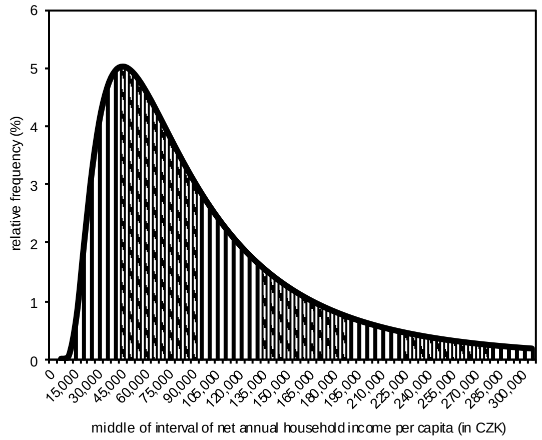


Figure 8 Method of L-moments

Source: Own research

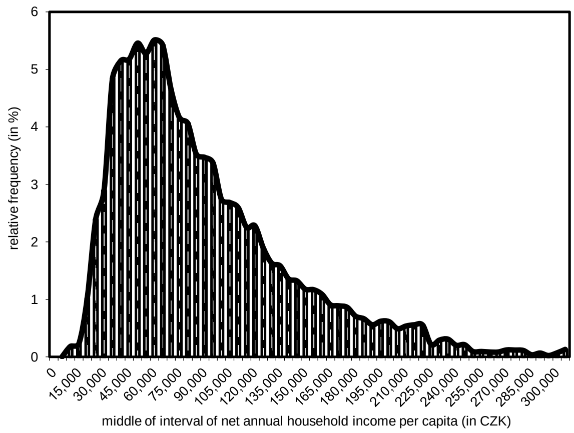


Figure 10 Sample ratios of employees

Source: Own research

Acknowledgements

Paper was processed with contribution of long term institutional support of research activities at Faculty of Informatics and Statistics, University of Economics, Prague (IP 400040).

References

[1] Adamowski, K.: Regional Analysis of Annual Maximum and Partial Duration Flood Data by Nonparametric and L-moment Methods. *Journal of Hydrology* **229** (2000), 219−231.

[2] Elamir, E. A. H., and Seheult, A. H.: Trimmed L-moments. *Computational Statistics & Data Analysis* **43** (2003), 299–314.

[3] Hosking, J. R. M.: L-moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *Journal of the Royal Statistical Society (Series B)* **52** (1990), 105–124.

[4] Hosking, J. R. M., and Wakes, J. R.: *Regional Frequency Analysis: An Approach Based on L-moments*. Cambridge University Press, Cambridge, 1997.

[5] Johnson, N. L., Kotz, S., and Balakrishnan, N.: *Continuous Univariate Distributions*. Wiley-Interscience, New York, 1994.

[6] Kleiber, C., and Kotz, S.: *Statistical Size Distributions in Economics and Actuarial Sciences*. Wiley-Interscience, New York, 2003.

[7] Kyselý, J., and Picek, J.: Regional Growth Curves and Improved Design value Estimates of Extreme Precipitation Events in the Czech Republic. *Climate research* **33** (2007), 243−255.

[8] Ulrych, T. J., Velis, D. R. Woodbury, A. D., and Sacchi, M. D.: L-moments and C-moments. *Stochastic Environmental Research and Risk Assessment* **14** (2000), 50–68.

Selection of the bank and investment products via stochastic multicriteria evaluation method

Adam Borovička[1]

Abstract. The main aim of the article is to choose the suitable bank products and investment instruments for a (partial) financial old-age security. For this purpose, a multicriteria evaluation of the selected alternatives is performed. Three main criteria are reflected – return, risk and cost. Return (or another criterion) can be described as a random variable with some probability distribution which is stated via statistical test. For evaluation of alternatives the stochastic multicriteria evaluation method is proposed. It is based on a combination of the principles of ELECTRE I and III methods. Further, the algorithms must be modified and improved (not only) in order to accept an input data with stochastic character. In a practical part, bank products (bank deposit, building savings and supplementary pension insurance) and investments products (open unit trusts) of the Česká spořitelna are chosen. These alternatives are evaluated by the proposed multicriteria approach for two types of strategies – risk-averse and risk-seeking. The results are analyzed and compared.

Keywords: financial old-age security, multicriteria evaluation, stochastic return.

JEL Classification: C44, G11
AMS Classification: 90B50, 62C86

1 Introduction

Every human being should think about a financial old-age security. The pension is often insufficient for paying all costs in this period of life. To eliminate such a problem, people usually try to save some money during their working age. They have a few possibilities how to do it. They can choose from various bank and investment products. But the question is which products are the most suitable and of course how to choose them. The answers are the main aim of this article.

Imagine the following real situation. If you are a long-term client of Česká spořitelna, you want to use selected products of this company in order to save free financial resources for future pension age. Two types of products are available – bank products (bank deposit, building savings, supplementary pension insurance) and investment products (open unit trusts). To make a complex decision I recommend evaluating the products by more than one criterion. The three main criteria are reflected – return, risk and cost. The value of return (or another criterion) can be very variable in time. Then such a criterion can be expressed as a random variable with some probability distribution. For a selection of the suitable products the stochastic multicriteria evaluation method is proposed. This method uses some fragments of the principles of ELECTRE I and III methods. Some principles are adopted, some are modified and improved to solve precisely a real financial situation. The proposed method is complex. It can make a quantitative analysis from more than one angle. Moreover, an importance of criteria can be different. Then a decision can be made for various strategies. Further, input information can have a stochastic character. These facts are the advantages of the proposed method compared to other concepts (e.g. fundamental or technical analysis). In practical situation, the strategies are differentiated – risk averse and risk-seeking strategy. For both types of the clients of Česká spořitelna the proposed method is applied to choose the suitable products. The results are analyzed and compared.

The article has the following structure. After the Introduction (Section 1), the proposed stochastic multicriteria evaluation method is described in Section 2. The original parts of the algorithm are described in more detail. A brief overview of the current concepts is also not missing. In Section 3, two real decision making situations about financial old-age security are introduced. Section 4 summarizes all the most important aspects of this article and provides some possible ideas for future research.

2 Stochastic multicriteria evaluation method

We know quite a number of multicriteria evaluation methods. One group of method is based on the evaluation principle of utility function – WSA [5], AHP [11], ANP [12] and the others. I think that the most important

[1] University of Economics, Department of Econometrics, W. Churchill Sq. 4, 130 67 Prague 3, adam.borovicka@vse.cz.

drawback of these methods can be just a utility function whose construction can be so difficult for a decision maker. But it depends on the particular method how the utility function is constructed. Another methods use a concept minimizing distance from the ideal solution for an evaluation of alternatives. Maybe the best known method from this group is TOPSIS [6]. These methods differentiate in a metric measuring a distance from the ideal solution. The methods of both mentioned groups usually provide the ranking of alternatives. But in my real decision making situation the ranking is not primarily required. Further, we know the methods working with the aspiration levels, for example conjunctive and disjunctive methods [6]. The aspiration levels must be specified by a decision maker which could be difficult for him/her. The next big group contents the methods using the concept of preference scoring relation. The best known are AGREPREF [8], ELECTRE I [9] or ELECTRE III [10]. The most methods from this group require information in the form of threshold values from a decision maker. A determination of these values (this value) could be a problem for a decision maker. It is obvious, that we know other multicriteria evaluation methods. Of course, the basic methods mentioned above were variously modified and improved (normalization techniques, importance of criteria, fuzzification, stochastic input data etc.).

As briefly indicated above, I see some drawbacks in the current approaches. Moreover, in order to solve the particular real situation the original method cannot be sometimes applied. Then the current methods must be modified, or new concept must be proposed. It is also my case. Firstly, in my practical application some data are variable in time, some are invariable. So the multicriteria evaluation method must accept both forms of the input data. The most methods do not enable this combination. Secondly, it is the best if any additional information (threshold values, specification of utility function) is not required from a decision maker because it could be very problematic for him/her. Further, the algorithm should take into account the differences in criteria values to make a representative analysis of alternatives. Many approaches do not make that (e.g. ELECTRE III, conjunctive or disjunctive method). Finally, for easy applicability the method should be user-friendly, the algorithm should be comprehensible for a decision maker. To fulfill all mentioned basic demands, a new stochastic multicriteria evaluation method is proposed in this paper.

2.1 Algorithm of the proposed method

The algorithm of the proposed method can be briefly described in the following several steps:

Step 1: We have n alternatives and k criteria. The evaluation of alternatives by each criterion is specified. This evaluation can be in strict form or in variable form in time. Then such an evaluation is formulated as random variable with some probability distribution. The probability distribution (with concrete parameters) is stated via a suitable statistical test (Chi-Square, Anderson-Darling or Kolmogorov-Smirnov test). Then for each probability distribution m random numbers are generated. Now m scenarios are specified. The following matrix of all criteria values for l-th scenarios $\mathbf{Y^l} = (y_{ij}^l)$ is specified where $y_{ij}^l \ (i = 1,2,...,n;\ j = 1,2,...,k;\ l = 1,2,...,m)$ is an evaluation of the i-th alternative by the j-th criterion. An importance of the j-th criterion is specified in the strict form as a weight $v_j \ (j = 1,2,...,k)$, when conditions $v_j \geq 0$ and $\sum_j v_j = 1$ hold.

Step 2: For each scenario the criteria values are compared. As in ELECTRE III, the following two sets of criteria indices are specified

$$I_{iPj}^l = \left\{ r \wedge s \mid y_{ir}^l > y_{jr}^l,\ y_{is}^l < y_{js}^l;\ r \in I^{\max}, s \in I^{\min} \right\} \qquad i, j = 1,2,...,n,\ i \neq j; l = 1,2,...,m$$
$$I_{jPi}^l = \left\{ r \wedge s \mid y_{jr}^l > y_{ir}^l,\ y_{js}^l < y_{is}^l;\ r \in I^{\max}, s \in I^{\min} \right\} \qquad i, j = 1,2,...,n,\ i \neq j; l = 1,2,...,m ,$$

where set $I^{\max}$, or $I^{\min}$ contains the indices of maximizing, or minimizing criteria.

Step 3: The matrices $\mathbf{S^l}$ and $\mathbf{R^l}$ are determined for l-th scenario. Their elements express the grades of preference. The matrix $\mathbf{S^l}$ is defined as in the ELECTRE III technique. The concept of matrix $\mathbf{R}$ is derived from ELECTRE I method. However, it takes into account the criteria importance and works with the standardized criteria values. The element of the matrix $\mathbf{R^l}$ is formulated for each couple of alternatives i and j for l-th scenario in the following form

$$r_{ij}^l = \frac{\sum_{h \in I_{iPj}^l} \left(v_h \mid {}^t y_{ih}^l - {}^t y_{jh}^l \mid \right)}{\sum_{h=1}^{k} v_h \mid {}^t y_{ih}^l - {}^t y_{jh}^l \mid} \qquad I_{iPj}^l \neq \varnothing, \qquad r_{ij}^l = - \quad i = j, \qquad r_{ij}^l = 0 \quad \text{else}$$

where ${}^{t}y_{ih}^{l}$, or ${}^{t}y_{jh}^{l}$ $(i, j = 1,2,...,n;\ h = 1,2,...,k;\ l = 1,2,...,m)$ is the standardized (normalized) criteria value. The standardized values are computed as follows

$${}^{t}y_{ij}^{l} = \frac{y_{ij}^{l}}{H_{j}^{l}} \qquad i = 1,2,...,n;\ j = 1,2,...,k;\ l = 1,2,...,m\ ,$$

where $H_{j}^{l} = \max_{i}(y_{ij}^{l})\ (j = 1,2,...,k;\ l = 1,2,...,m)$. For this formulation, all values must be positive. If they are not positive, the appropriate constant must be added.

Step 4: The aggregate preference of the *i*-th alternative in face of the *j*-th alternative in terms of the *l*-th scenario is set by a proposed rule

$$s_{ij}^{l} > s_{ji}^{l} \wedge r_{ij}^{l} > r_{ji}^{l}\ .$$

The rule is modification of the ELECTRE III technique. The thresholds are eliminated. Then for the *i*-th alternative in the *l*-th scenario the following indicator is specified

$$p_{i}^{l} = d_{i}^{l+} - d_{i}^{l-} \qquad i = 1,2,...,n;\ l = 1,2,...,m\ ,$$

where d_{i}^{l+}, or d_{i}^{l-} is the number of alternatives in face of which *i*-th alternative is preferred, or the number of alternatives that are preferred in face of the *i*-th alternative in terms of the *l*-th scenario. Now the set of *subeffective* alternatives is proposed as follows

$$E^{l} = \left\{a_{i};\ p_{i}^{l} = \max_{i}(p_{i}^{l})\right\} \qquad l = 1,2,...,m\ ,$$

where a_{i} denotes the *i*-th alternative. It is possible to say, that these alternatives are effective for *l*-th scenario. The concept of effective alternative for *l*-th scenario is inspired by ELECTRE I approach. However, it is quite not suitable (namely for low number of scenarios) because the effective alternative does not have to exist in the set of alternatives. To eliminate this drawback some ideas of ELECTRE III method are used for a specification of the effective alternative. Finally, the *effective* alternatives are selected over all scenarios. The effective alternative is such an alternative that is mostly times evaluated as subeffective. In addition, according to diminishing value of this indicator (how many times indicated as subeffective alternative) the alternatives can be ranked.

3 Selection of suitable product(s) for financial old-age security

In this part of the article, a very common situation of a real life is described. The most people in working age think about financial old-age security. I mean that this thinking is just very strong in the Czech Republic. The pension system in the Czech Republic is not stable, the rules are still changing. The age of retirement is still extended. Then many people think how to save some money for future usage during their pension age. They have a few options.

Many people are the clients of Česká spořitelna. So let us focus on the case of long-term clients who have a current account in Česká spořitelna. It is meaningful that such a person appeals to “home” bank to advise what to do it in the described situation. A worker of the bank offers two types of products – bank and investment products. The bank products are bank deposit, building savings and supplementary pension insurance. The investment products are the open unit trusts. This investment instrument is also suitable for a smaller investor. They can invest in stocks or bonds of big companies around the world via these funds.

3.1 Criteria and strategies

To make a decision as complex as possible, several characteristics of the bank and investment instruments are taken into account. Three main criteria are specified – return, risk and cost. According to these criteria all selected instruments are evaluated. Then a worker of the bank presents this information to the client. Another characteristics should be reflected, namely in the case of open unit trusts. For instance, a currency in which allotment certificates are traded. Locality of the capital market with the open unit trusts, approach of the management of the fund or mood in the capital market could be included. Another criterion could be a liquidity of the saved money. All these mentioned criteria are not explicitly included. But they can be taken into account in a selection of the bank or investment instruments which are offered to the client. Then these selected instruments are analyzed from the perspective of the most important three criteria in terms of the multicriteria evaluation process.

In order to involve a wider scale of real decision making cases, two client's strategies are specified – risk-averse and risk-seeking. *Risk-averse* client fears a loss of the investment so much. He/she is willing to undergo only low level of risk that the saved money loses a part of their value. He/she wants to have a certain confidence that the money will be available in his/her pension age. Therefore he/she is able to sacrifice some part of return for a lower risk. A criterion risk is the most important characteristic. The cost connected with the bank or investment products does not play a fundamental role in a decision making process. *Risk-seeking* client is more oriented to the return. He/she is able to undergo a greater risk for a greater level of return. His/her goal is to valorize his/her spared money as much as possible under a condition of greater level of risk. In contrast to risk-averse client, the return is more important criterion than risk. Therefore the cost connected with investment is a little bit more important for a risk-seeking client than for a risk-averse client.

The client should specify his/her preferences about an importance of these criteria. It is obvious that the preferences are based on the client's strategy. The preferences can be quantitatively expressed on some scale. So the client assigns the points (e.g.) from the interval $\langle 0,10 \rangle$ to each criterion, where 0 means the smallest (actually zero) importance, conversely 10 means the highest importance. To calculate the weights of criteria required by the proposed multicriteria evaluation method, the scoring method is applied [5]. This method is very simple for an application. It just uses the quantitative information about the importance of criteria. According to two client's strategies mentioned above, the points are assigned to the chosen three criteria. The following table shows the assigned points, the weights of criteria for both types of clients as well (Table 1).

	Risk-averse client		***Risk-seeking client***	
Criterion	**Points**	**Weight**	**Points**	**Weight**
Return	7	0,368	10	0,526
Risk	10	0,526	6	0,316
Cost	2	0.105	3	0,158

Table 1 Points and weights of criteria for both types of clients

According to expectation, the criterion risk has the greatest weight for a risk-averse client. On the contrary for a risk-seeking strategy, the greatest weight is calculated to the criterion return. The weight of cost for a risk-seeking client is a little bit greater than for a risk-averse client. This fact was also expected.

3.2 Data

We keep at disposition 3 bank products (bank deposit, building savings, supplementary pension insurance) and 8 open unit trusts (5 bond and 3 stock funds). The bond open unit trusts are High Yield dluhopisový, Korporátní dluhopisový, Sporobond, Sporoinvest and Trendbond. The stock open unit trusts are Global Stocks, Sporotrend and Top Stocks. The bank products are offered by Česká spořitelna and the open unit trusts are managed and offered by Česká spořitelna investment company.

All instruments are evaluated by the selected criteria. The value of return is deterministic for the bank products. It is on the annual basis and expressed in percentages. Return of the bank deposit is in actual value. This value is actually historically stable. Return of the building savings has two components. One part is a contribution from the state and second part is interest from a deposit. Return of the building savings reflects actual conditions specified by Česká spořitelna. Return of the supplementary pension insurance is also composed of two parts. One component is a state contribution and second component is a return of the Česká spořitelna pension fund. Return of the pension fund is average annual for last 20 years. It is computed as a strict value because of lower number of observations. Return of the open unit trusts is variable in time. Then return of each fund is described as a random variable with some continuous probability distribution. The distribution is set on the basis of non-parametric statistical Anderson-Darling test [1]. The monthly returns from the period 2010-2015 are used. This period is chosen because a development in the capital market was calmer. Then it can better describe a long-term development which is so important for a practical application. The generated returns from the particular distributions are converted to the annual base. Risk connected with the bank deposit is considered as zero because the interest rate is stable in time. Risk of the building savings is also zero. The conditions are given by the appropriate contract, so that a change of return at least during 6 years is not possible. Risk of the supplementary pension insurance is connected with a possible deviation of interest rate from a deposit. It is calculated as average absolute negative deviation of returns. This concept is proposed in [2]. The risk of the open unit trusts is also calculated as average absolute negative deviation of their returns. Cost connected with the bank deposit reflects the tax on interest. There are no other costs. The same situation is in the case of supplementary pension insurance. Cost connected with the building savings includes charge for an account management, conclusion of a contract and tax on interest. Cost connected with the open unit trust represents various charges, namely initial charge and management fee. Costs are also calculated in the percentage form as well as other two criteria.

For building savings, the most profitable alternative is expected. It means that monthly deposited amount is 1700 CZK. The criteria values for supplementary pension insurance are calculated for the most profitable situation when the monthly deposited amount is 1000 CZK. It is obvious that the return of building savings and supplementary pension insurance is based on the amount of invested money. In our analysis the best alternative is searched for a person who can periodically tuck away lower amount of money. In case of building savings and supplementary pension insurance he/she accepts mentioned amount of money. The sock inserted in other products will be similar. If he/she would want to save more money, another product could be chosen via the additional procedure mentioned below.

Table 2 shows evaluation of each product by each criterion. All data is expressed in the percentage form. According to statistical test the return of the open unit trusts can be described as a random variable with logistic probability distribution. Both parameters of distribution for each fund are also written in the following table.

Product	Return	Risk	Cost
Bank deposit	0.4	0	0.06
Building savings	6	0	2.84
Supplementary pension insurance	26.62	0.21	0.54
High Yield dluhopisový	Logistic (6.52, 15.35)	2.19	2.49
Korporátní dluhopisový	Logistic (4.42, 9.76)	1.23	2.75
Sporobond	Logistic (5.15, 5.99)	0.69	2.20
Sporoinvest	Logistic (0.82, 0.79)	0.18	1.05
Trendbond	Logistic (3.77, 10.79)	1.37	2.77
Global stocks	Logistic (11.29, 21.66)	2.21	6.33
Sporotrend	Logistic (-4.77, 44.05)	5.36	5.43
Tops Stocks	Logistic (18.7, 42.76)	3.93	5.82

Table 2 Return, risk and cost of bank and investment products (in percentages)

The market prices and cost of the open unit trusts are taken over from [7]. The needed information on the bank products is found in [3] and the historical returns of Česká spořitelna pension fund are taken over from [12]. On the basis of this information, the criteria values are calculated.

3.3 Results

Now the suitable products are selected for both types of strategies via the proposed multicriteria evaluation method. For representative results, 50 scenarios of returns of the open unit trusts are made. Thus, 50 values from particular probability distribution for each open unit trust are generated. The following table shows how many times each product is classified as a subeffective alternative for both types of clients (Table 3).

Risk-averse client		***Risk-seeking client***	
Product	**Frequency**	**Product**	**Frequency**
Building savings	36	Supplementary pension insurance	24
Bank deposit	25	Top Stocks	19
Supplementary pension insurance	14	Global Stocks	10
High Yield dluhopisový	0	Sporotrend	6
Korporátní dluhopisový	0	Trendbond	3
Sporobond	0	High yield dluhopisový	2
Sporoinvest	0	Sporobond	2
Trendbond	0	Korporátní dluhopisový	0
Global stocks	0	Sporoinvest	0
Sporotrend	0	Bank deposit	0
Tops Stocks	0	Building savings	0

Table 3 Frequency of incidence of each product as subeffective alternative

For a risk-averse investor, the effective alternative is the bank product building savings. None of the open unit trusts is specified as subeffective alternative because they have too great risk. The building savings is effective alternative thanks to a zero risk and a greater level or return than bank deposit. The worst position from the bank products belongs to the supplementary pension insurance thanks its nonzero risk. A better position of this product cannot be even achieved by significantly the greatest return from the bank products because the return has smaller weight than risk. For a risk-seeking investor, the effective alternative is the supplementary pension insurance. But according to expectation, the stock open unit trusts compete with it thanks a possible greater return. A decisive factor is a stable return of supplementary pension insurance compared to the stock open unit trusts. Because the return is the most important criterion for a risk-seeking investor, the other open unit trusts with lower level of return had not a chance to be an effective alternative. The same situation occurs in the case of bank deposit and building savings in spite of they have a zero level of risk.

According to proposed multicriteria analysis, a decision maker can see which products are “good” and which are “bad”. In case of need the ranking of alternatives can also be made by a frequency of incidence of alternatives as subeffective alternative. An appropriate portfolio of more products can be made by other methods.

4 Conclusion

The article deals with a selection of the suitable products for financial old-age security. Two strategies of saving free financial resources during a working age are specified. For both cases the suitable products of Česká spořitelna are chosen via the proposed stochastic multicriteria evaluation method whose algorithm is described. The results of both strategies are analyzed and compared.

The proposed method provides a complex quantitative multicriteria analysis. All stated criteria have a quantitative character. But some qualitative characteristics should be also explicitly taken into account in the analysis (e.g. mood in the capital market of the open unit trusts, some feeling or intuition of a decision maker etc.). Then they must be transformed to the quantitative form. One possible way should be via the fuzzy sets (numbers). So the proposed method would be fuzzified. Then it would become yet more complex for real decision making situations.

Acknowledgements

The research project was supported by Grant No. IGA F4/54/2015 of the Internal Grant Agency, Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] Anderson, T. W., Darling, D. A.: Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Annals of Mathematical Statistics* **23** (1952), 193–212.

[2] Borovička, A.: *Vytváření investičního portfolia podílových fondů pomocí fuzzy metod vícekriteriálního rozhodování*. Ph.D. Thesis. University of Economics, Prague, 2015.

[3] Česká spořitelna – products and services [online], available at: *http://www.csas.cz/banka/nav/osobni-finance/produkty-a-sluzby-d00019523*, [cit. 01-03-2016].

[4] Development of pension funds returns [online], available at: *http://www.penzfondy.cz/*, [cit. 01-03-2016].

[5] Fiala, P.: *Modely a metody rozhodování*. Oeconomica, Praha, 2013.

[6] Hwang, C. L., Yoon, K.: *Multiple Attribute Decision Making – Methods and Applications, A State-of-the-Art Survey*. Springer-Verlag, New York, 1981.

[7] Investment centrum of Česká spořitelna [online], available at: *https://cz.products.erstegroup.com/Retail/*, [cit. 01-03-2016].

[8] Lagrèze, E. J.: *La modélisation des préférences: Préordres, quasiordres et relations floues*. Ph.D. Thesis, Université de Paris V, 1975.

[9] Roy, B.: Classement et choix en présence de points de vue multiples (la méthode ELECTRE). *La Revue d'Informatique et de Recherche Opérationelle (RIRO)* **2** (1968), 57–75.

[10] Roy, B.: ELECTRE III: algorithme de classement basé sur une représentation floue des préférences en présence des critères multiples. *Cahiers du CERO* **20** (1978), 3–24.

[11] Saaty, T. L.: *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. McGraw-Hill, New York, 1980.

[12] Saaty, T. L.: *Decision Making with Dependence and Feedback: The Analytic Network Process*. RWS Publication, New York, 1996.

A chance constrained investment problem with portfolio variance and skewness criteria - solution technique based on the Successive Iterative Regularization

Martin Branda [1]

Abstract. We deal with an investment problem, where the variance of a portfolio is minimized and at the same time the skewness is maximized. Moreover, we impose a chance (probabilistic) constraint on the portfolio return which must be fulfilled with a high probability. This leads to a difficult nonconvex multiobjective stochastic programming problem. Under discretely distributed returns, this problem can be solved using the CCP-SIR solver (Chance Constrained Problems: Successive Iterative Regularization) which has been recently introduced by Adam and Branda [1]. This algorithm relies on a relaxed nonlinear programming problem and its regularized version obtained by enlarging the set of feasible solutions using regularizing functions. These both formulations as well as the solution technique are discussed in details. We report the results for a real life portfolio problem of a small investor. We compare the CCP-SIR solver with BONMIN applied to the deterministic mixed-integer reformulation.

Keywords: Investment problem, chance constraints, skewness, successive iterative regularization

JEL classification: C44
AMS classification: 90C15

1 Introduction

Markowitz [14] was the first, who considered moment conditions on portfolio composed from various assets available on a stock market. In particular, he maximized the expected return of a portfolio and at the same time minimized the variance of portfolio return. The variance served here to quantify the risk. The optimal portfolios then correspond to the efficient frontier in the mean-variance space. Skewness together with the mean and variance criteria was considered in the models for accessing portfolio efficiency, see, e.g., Joro and Na [12], Briec et al. [9]. Kerstens et al. [13] derived a geometric representation of the mean–variance–skewness portfolio frontier, and Briec et al. [10] generalized the well-known one fund theorem. Alternatively to the higher order moment criteria, many researchers focused on axiomatic definitions of risk measure classes with desirable properties, e.g. convex risk measures [11], coherent risk measures [3], and general deviation measures [16]. The resulting models can be supported by the relations to the utility functions and stochastic dominance efficiency, see [6, 7, 8].

In this paper, we focus on a multiobjective portfolio problem with the variance and skewness criteria. The set of feasible portfolios is defined using a chance (probabilistic) constraint which imposes a high probability on the events when portfolio return exceeds a given value. Due to the presence of the nonconvex cubic term in the objective function, we were not able to use algorithms which rely on convexity of the underlying problem. Therefore we select an alternative approach which has been recently introduced by Adam and Branda [1]. They proposed an algorithm based on a relaxed problem which is a nonlinear

[1] Charles University in Prague
Faculty of Mathematics and Physics
Department of Probability and Mathematical Statistics
Sokolovská 83, Prague 186 75, Czech Republic
tel.: +420 221 913 404, fax: +420 222 323 316, branda@karlin.mff.cuni.cz

programming problem, and its regularized version that is obtained by enlarging the set of feasible solutions using regularizing functions. The resulting CCP-SIR solver[1] (Chance Constrained Problems: Successive Iterative Regularization) is then able to obtain a stationary point of the original problem. We apply the CCP-SIR algorithm to the investment problem in the empirical part and compare its performance with the BONMIN [5] solver available in GAMS modelling system.

The paper is organized as follows. In Section 2, we propose the notation and the variance-skewness model with a chance constraint on portfolio return. In Section 3, we formulate the relaxed problem and derive the optimality conditions. We also discuss the CCP-SIR algorithm which is then employed in the empirical study in Section 4. Section 5 concludes the paper.

2 Problem formulation

In this section, we formulate the multiobjective stochastic portfolio optimization problem with minimization of the portfolio variance and maximization of the skewness. The chance constraint provides a probabilistic lower bound for random portfolio rate of return.

We consider n assets with random rates of return denoted by R_j, $j = 1, \dots, n$, with $\mathbb{E}|R_j|^3 < \infty$ and define the corresponding covariance matrix C and skewness tensor S elementwise as

$$\begin{aligned} C_{jk} &:= \mathbb{E}(R_j - \mathbb{E}R_j)(R_k - \mathbb{E}R_k), \\ S_{jkl} &:= \mathbb{E}(R_j - \mathbb{E}R_j)(R_k - \mathbb{E}R_k)(R_l - \mathbb{E}R_l). \end{aligned}$$

If we employ the aggregate function approach of multiobjective optimization with aggregation parameter $c > 0$, we obtain

$$\begin{aligned} \text{minimize } & \sum_{j=1}^{n} \sum_{k=1}^{n} C_{jk} x_j x_k - c \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} S_{jkl} x_j x_k x_l \\ \text{subject to } & P\left(\sum_{j=1}^{n} R_j x_j \geq b \right) \geq 1 - \varepsilon, \\ & \sum_{j=1}^{n} x_j = 1, \\ & lb_j \leq x_j \leq ub_j, \end{aligned} \tag{1}$$

where b is a minimal rate of return acceptable with probability $1 - \varepsilon$. We allow various restrictions on portfolio weights using lb_j, ub_j. In particular, if $lb_j < 0$, then short selling of the asset j is allowed.

3 Optimality conditions and algorithm

We consider discretely distributed returns with scenarios R_{ji} (index j identifies an asset, i a scenario) and general probabilities $p_i > 0$, $\sum_{i=1}^{S} p_i = 1$. Such distribution can be obtained as the observations of historical returns or by simulation from a multivariate continuous distribution. Standard approaches use mixed-integer programming reformulations, cf. [2, 15]. However, the employed algorithm relies on the

[1] Matlab source codes are freely available at http://staff.utia.cas.cz/adam/research.html

relaxed problem that can be formulated as a nonlinear programming problem:

$$\begin{aligned}
\text{minimize } & \sum_{j=1}^{n}\sum_{k=1}^{n} C_{jk}x_jx_k - c\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} S_{jkl}x_jx_kx_l \\
\text{subject to } & \sum_{i=1}^{S} p_iy_i \geq 1-\varepsilon, \\
& y_i\left(b - \sum_{j=1}^{n} R_{ji}x_j\right) \leq 0,\ i=1,\dots,S, \\
& \sum_{j=1}^{n} x_j = 1,\ lb_j \leq x_j \leq ub_j, \\
& 0 \leq y_i \leq 1.
\end{aligned} \tag{2}$$

According to [1, Theorem 2.2], the optimality condition for the problem at a point x can be stated as:

$$\begin{aligned}
2\sum_{k=1}^{n} C_{jk}x_k - 3c\sum_{j=1}^{n}\sum_{k=1}^{n} S_{jkl}x_kx_l - \sum_{i\in I_0(x)} \lambda_iR_{ji} + \mu + \mu_j = 0,\ j=1,\dots,n,& \\
\lambda_i = 0,\ i\in I_{00}(x,y),& \\
\lambda_i \geq 0,\ i\in I_{0+}(x,y)\cup I_{01}(x,y),& \\
\mu_j \leq 0,\ j\in J^-(x),& \\
\mu_j \geq 0,\ j\in J^+(x),& \\
\mu_j = 0 \text{ otherwise},& \\
\mu\in\mathbb{R},&
\end{aligned} \tag{3}$$

where the index sets are defined as follows

$$\begin{aligned}
J^-(x) &:= \{j : x_j = lb_j\},\ J^+(x) := \{j : x_j = ub_j\}, \\
I_0(x) &:= \{i : \sum_{j=1}^{n} R_{ji}x_j = b\}, \\
I_{00}(x,y) &:= \{i\in I_0(x) :\ y_i = 0\}, \\
I_{0+}(x,y) &:= \{i\in I_0(x) :\ y_i \in (0,1)\}, \\
I_{01}(x,y) &:= \{i\in I_0(x) :\ y_i = 1\}.
\end{aligned} \tag{4}$$

The algorithm is based on a regularizing function imposed on the random constraints, which enlarges the set of feasible solutions. We assume that $\phi_t : \mathbb{R}\to\mathbb{R}$ are continuously differentiable decreasing functions which depend on a parameter $t > 0$ and which satisfy the following properties:

$$\phi_t(z) > 0,\quad \text{for } z\in\mathbb{R},\ \phi_t(0) = 1, \tag{5}$$

$$\phi_t(z^t) \to 0,\quad \text{whenever } z^t \overset{t\to\infty}{\to} \bar{z} > 0, \tag{6}$$

$$\frac{\phi_t'(z^t)}{\phi_t'(\tilde{z}^t)} \to 0,\quad \text{whenever } \phi_t(z^t) \searrow 0 \text{ and } \phi_t(\tilde{z}^t) \to \bar{z} > 0. \tag{7}$$

We will consider $\phi_t(z) = e^{-tz}$ as a regularizing function. Using these functions we obtain the following regularized problem:

$$\begin{aligned} \text{minimize} \ & \sum_{j=1}^{n}\sum_{k=1}^{n} C_{jk}x_jx_k - c\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} S_{jkl}x_jx_kx_l \\ \text{subject to} \ & \sum_{i=1}^{S} p_iy_i \geq 1-\varepsilon, \\ & \phi_t\left(b - \sum_{j=1}^{n} R_{ji}x_j\right) \geq y_i \ i = 1,\dots,S, \\ & \sum_{j=1}^{n} x_j = 1, \ lb_j \leq x_j \leq ub_j, \\ & 0 \leq y_i \leq 1. \end{aligned} \tag{8}$$

The Successive Iterative Regularization algorithm then iteratively increases the parameter t leading to tighter approximation of the feasibility set. The problem is solved by a nonlinear programming solver and the obtained solution is then used as a starting point for the next iteration. It can be shown that the sequence of the solutions converge to a stationary point of the relaxed problem (2). If the obtained point is not a stationary point of the original chance constrained problem, an additional local search procedure can be performed to find such point, see [1] for details.

Finally, note that the problem (1) can be also reformulated using additional binary variables z_i and a large constant M:

$$\begin{aligned} \text{minimize} \ & \sum_{j=1}^{n}\sum_{k=1}^{n} C_{jk}x_jx_k - c\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} S_{jkl}x_jx_kx_l \\ \text{subject to} \ & \sum_{i=1}^{S} p_iz_i \geq 1-\varepsilon, \\ & b - \sum_{j=1}^{n} R_{ji}x_j \leq M(1-z_i), \ i = 1,\dots,S, \\ & \sum_{j=1}^{n} x_j = 1, \ lb_j \leq x_j \leq ub_j, \ z_i \in \{0,1\}. \end{aligned} \tag{9}$$

This leads to a nonlinear mixed-integer programming problem which will be employed in the following section.

4 Numerical study

We consider the following $n = 8$ US assets observed monthly from January 2011 to December 2015. In particular, we consider Boeing (BA), Coca-Cola (KO), JPMorgan Chase (JPM), Oracle (ORCL), Microsoft (MSFT), Nike (NKE), Intel (INTC), Apple (AAPL). Since the problem is highly demanding, we have chosen $S = 100$ scenarios. The other parameters of the problem were set as follows: minimal acceptable returns $b \in \{-4, -2, 0, 2\}$, probabilistic level $\varepsilon = 0.1$, and aggregation (risk aversion) parameter $c = 0.1$. We employed the multivariate skew normal distribution, cf. [4], for modelling the random rates of return. Emphasize that this distribution is able to fit the skewness of the real data. For each value of b, we have rerun the whole optimization 10 times.

In [1], we have also discussed the influence of the choice of M in problem (9) on the algorithm performance. It is highly desirable to choose this parameter as tight as possible. Since the maximal random value of R_{ji} was approximately 35% and since $|x_j| \leq 10$, we have decided to choose $M = 500$. As the maximal constraint violation from all generated solutions was 80, this M showed to be a very reasonable bound.

We compare three possible methods for the problem solution. The first one is based on the regularization and employs the CCP-SIR algorithm. The second one uses the solver BONMIN [5] available in the modelling system GAMS for solving the deterministic equivalent formulation which corresponds to a mixed-integer nonlinear programming problem. The last method uses their combination and employs the result of BONMIN as a starting point for the CCP-SIR algorithm. For the regularized problem, the average solution time was approximately 4.5 seconds while we stopped BONMIN after 10 minutes. Thus, the solutions obtained by BONMIN were always suboptimal.

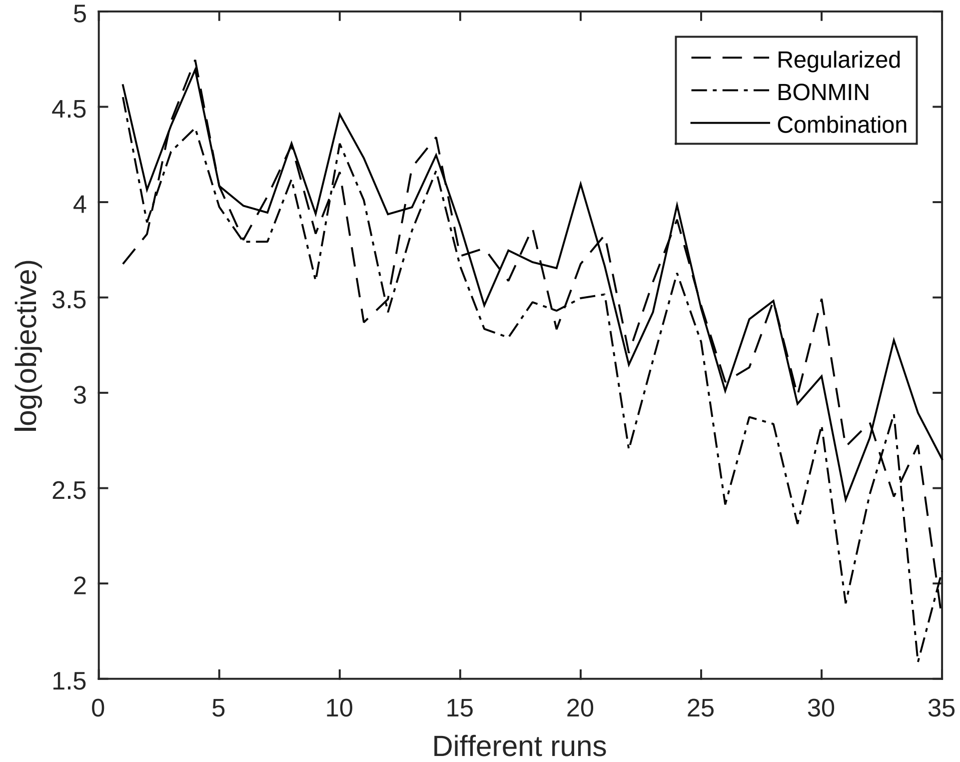


Figure 1 Comparison of different algorithms for problem (1).

The results of these three solvers can be seen in Figure 1. On the horizontal axis, there are different sets of simulated return realizations R_{ji} and different values of b. The first 10 values corresponds to $b = -4$, the next 10 for $b = -2$, the next 10 for $b = 0$ and finally, the last 5 (some infeasible problems were omitted) for $b = 2$.

Since the mean value of R_{ji} was 5%, for small b, the feasible set was huge and the relaxed problem contained many local minima. Because the regularized method uses only first order information (and no globalization in the form of cuts), BONMIN performed better than the regularized method in this case. However, with increasing b, the regularized method started to prevail. This can be especially seen for the last five values with $b = 2$. However, almost in all cases, when we used the regularized method on the BONMIN solution, we were able to improve the solution significantly. We can conclude that the employed methods are able to solve problems in a very good way. Moreover, if they for some reason fail, they are still able to quickly improve suboptimal solutions obtained by other solvers.

5 Conclusions

We have proposed a portfolio optimization problem with the variance–skewness objective and a chance constraint on the portfolio return. Under the discrete distribution of returns, we have relaxed the problem and provided the optimality conditions. Algorithmic issues have been discussed based on the regularized problem obtained by enlarging the set of feasible solutions using regularizing functions. The good performance of the algorithm has been demonstrated using the real data and compared with the solver BONMIN available in GAMS. The combination of CCP-SIR with BONMIN seems to be also promising.

Acknowledgements

This work was supported by the Czech Science Foundation under the Grant P402/12/G097 (DYME – Dynamic Models in Economics).

References

[1] Adam, L., Branda, M.: Nonlinear chance constrained problems: optimality conditions, regularization and solvers, *Journal of Optimization Theory and Applications* (2016), to appear. DOI: 10.1007/s10957-016-0943-9

[2] Adam, L., Branda, M.: Sparse optimization for inverse problems in atmospheric modelling. *Environmental Modelling & Software* **79** (2016), 256–266.

[3] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. *Mathematical Finance* **9** (1999), 203–228.

[4] Azzalini, A., Dalla Valle, A.: The multivariate skew-normal distribution. *Biometrika* **83** (4) (1996), 715–726.

[5] Bonami, P., Biegler, L.T., Conn, A.R., Cornuejols, G., Grossmann, I.E., Laird, C.D., Lee, J., Lodi, A., Margot, F., Sawaya, N., and Waechter, A.: An algorithmic framework for convex mixed integer nonlinear programs, *Discrete Optimization* **5** (2) (2008), 186–204.

[6] Branda, M.: Diversification-consistent data envelopment analysis based on directional-distance measures. *Omega* **52** (2015), 65–76.

[7] Branda, M., Kopa, M.: On relations between DEA-risk models and stochastic dominance efficiency tests. *Central European Journal of Operations Research* **22** (1) (2014), 13–35.

[8] Branda, M., Kopa, M.: DEA models equivalent to general N-th order stochastic dominance efficiency tests. *Operations Research Letters* **44** (2) (2016), 285–289.

[9] Briec, W., Kerstens, K., Jokung, O.: Mean-variance-skewness portfolio performance gauging: a general shortage function and dual approach. *Management Science* **53** (2007), 135–149.

[10] Briec, W., Kerstens, K., Van de Woestyne, I.: Portfolio selection with skewness: a comparison of methods and a generalized one fund result. *European Journal of Operational Research* **230** (2) (2013), 412–421.

[11] Follmer, H., Schied, A.: *Stochastic Finance: An Introduction In Discrete Time*. Walter de Gruyter, Berlin, 2002.

[12] Joro, T., Na, P.: Portfolio performance evaluation in a mean–variance–skewness framework. *European Journal of Operational Research* **175** (2006), 446–461.

[13] Kerstens, K., Mounir, A., Van de Woestyne, I.: Geometric representation of the mean–variance–skewness portfolio frontier based upon the shortage function. *European Journal of Operational Research* **210** (1) (2011), 81–94.

[14] Markowitz, H.M.: Portfolio selection. *Journal of Finance* **7** (1) (1952), 77–91.

[15] Raike, W.M.: Dissection methods for solutions in chance constrained programming problems under discrete distributions. *Management Science* **16** (11) (1970), 708–715.

[16] Rockafellar, R.T., Uryasev, S., Zabarankin, M.: Generalized deviations in risk analysis. *Finance and Stochastics* **10** (2006), 51–74.

Post-crisis development in the agricultural sector in the CR – Evidence by Malmquist index and its decomposition

Helena Brožová[1], Ivana Boháčková[2]

Abstract. This contribution aims to use Data Envelopment Analysis models, Malmquist index and its decomposition for analysis of the efficiency of the agricultural sector in the regions of the Czech Republic and its development from the year 2008 to 2013.
The regions are characterized by three inputs (agricultural land, annual work units and gross fixed capital formation) and one output (gross value added).
Malmquist index shows significant increase of efficiency in 2011 although in other years rather negative development or no change can be seen. The same shape can be seen for technological change. Pure and scale efficiency changes are similar and both show different behaviour than Malmquist index and technological change. Pure and scale efficiency are significantly increasing between years 2008 and 2009 and, further, slowly decreasing or without change.

Keywords: Crisis, agricultural sector, efficiency, CCR model, BCC model, Malmquist index decomposition.

JEL Classification: C44, C61, D24
AMS Classification: 90B50, 90C05, 90C90

1 Introduction

The productivity and efficiency of the agricultural sector is provided for long time. The first studies on this topic have appeared already in the 40s and 50s of the last century. Starting from Clark [6] and Bhattacharjee [3] analysis, a number of studies tried to analyse the differences in agricultural productivity of analysed units and its development (for instance, Yu et al. [14], Trueblood and Coggins [13]). The Data Envelopment Analysis (DEA) is the mostly used one from the non-parametric methods.

The contribution of inputs to outputs is evaluated using the DEA models firstly developed by Charnes, Cooper and Rhodes [5] and Banker, Charnes and Cooper [2], which applied the Farrel's approach (Farrel [9]) to measurement of productivity efficiency. The dynamic context of the relation of inputs and outputs, the growth of productivity efficiency was studied by Malmquist [11]. Measuring productivity change and Malmquist index (MI) decomposing into its sources is important, because enhancement of the productivity growth requires knowledge of the relative importance of its sources. In this regard the Malmquist productivity index (MI) is particularly used and decomposed into the technical efficiency change index and the technological change index and, further, these indices can be decomposed again (Yu et al.[14], Lovell [10], Ray [12], Zbranek [15]).

The objective of this paper is to show the relative technical efficiency and its changes of the agricultural sector in the Czech regions in the crisis period 2008-2013. Efficiency progress according to the Malmquist index will be compared with changes of selected indices receiving by Malmquist index decomposition.

The paper is organized as follows. Section 2 defines shortly basic DEA models, Malmquist index and its decomposition. Section 3 contains an analysis of the efficiency of the agricultural sector of the Czech regions and changes of the elements of the MI decomposition in the years 2008-2013. Section 4 summarizes the findings.

2 Malmquist index and its decomposition

The Malmquist index is usually applied to the measurement of productivity change over time, and can be multiplicatively decomposed into a technical efficiency change index and a technological change index. Its calculation can be based on decision-making units (DMU) efficiency according to the CCR and BCC models combining inputs and outputs from different time periods.

[1] Czech University of Life Sciences, Faculty of Economics, Department of Systems Engineering, 165 00 Praha 6 – Suchdol, e-mail: brozova@pef.czu.cz.
[2] Czech University of Life Sciences, Faculty of Economics, Department of Economics, 165 00 Praha 6 – Suchdol, e-mail: bohackiv@pef.czu.cz.

The model CCR (according to authors Charnes, Cooper and Rhodes [5] is evaluating the efficiency of a set of DMU, which convert multiple inputs into multiple outputs with the constant returns to scale (CRS) assumption. The outputs and inputs can be of various characteristics and of variety of forms that can be also difficult to measure. If the variable returns to scale (VRS) is supposed, so-called BCC model (according to authors Banker, Charnes and Cooper [2]) should be used. DEA models with super efficiency (Andersen, Petersen [1]) can be used for the distinction of the efficient DMUs. In this type of the DEA model a DMU under evaluation is not included in the reference set of the original DEA models.

Model CCR

The input oriented CCR DEA model (Charnes et al. [5]) measuring efficiency of DMU in time t is obtained as the maximum of a ratio of weighted outputs to weighted inputs under the constraints that this ratio for all DMUs is less than or equal to 1and the efficiency of the efficient DMU is equal to 1. The efficiency Φ_H^t explicitly shows the necessary decreasing of inputs with the same amount of inputs of non-efficient DMU H in time t. Linearization of this model follows

$$\Phi_H^t(\mathbf{x}^t,\mathbf{y}^t)=\sum_{j=1}^{n}u_{jH}\,y_{jH}^t \rightarrow MAX \qquad \text{subject to} \qquad \begin{aligned}&\sum_{i=1}^{m}v_{iH}\,x_{iH}^t=1\\&-\sum_{i=1}^{m}v_{iH}\,x_{ik}^t+\sum_{j=1}^{n}u_{jH}\,y_{jk}^t\le 0,\ k=1,\ 2,\dots,\ p\\&u_{jH}\ge 0,\ \ j=1,\ 2,\dots,\ n\\&v_{iH}\ge 0,\ \ i=1,\ 2,\dots,\ m\end{aligned} \tag{1}$$

where u_{jk} and v_{ik} are weights of outputs and inputs, y_{jk}^t is the amount of j^{th} output from unit k, and x_{ik}^t is the amount of i^{th} input to k^{th} unit in time t, H is index of the evaluated DMU (this notation is used through the whole paper) .

Output orientation of this model for all DMUs supposes the minimization model. The efficiency of all DMUs is greater than or equal to 1and the efficiency of the efficient DMU is equal to 1. The efficiency shows the necessary increasing of outputs with the same amount of inputs of non-efficient DMU.

For the decomposition of the MI, we have to compute following models with different combination of inputs and output from different time periods (Zbranek [15]). We need to compute efficiency of the DMU inputs and outputs in time t+1 in the conditions of the time t. For this we use the model

$$\Phi_H^t(\mathbf{x}^{t+1},\mathbf{y}^{t+1})=\sum_{j=1}^{n}u_{jH}\,y_{jH}^{t+1} \rightarrow MAX \qquad \text{subject to} \qquad \begin{aligned}&\sum_{i=1}^{m}v_{iH}\,x_{iH}^{t+1}=1\\&-\sum_{i=1}^{m}v_{iH}\,x_{ik}^t+\sum_{j=1}^{n}u_{jH}\,y_{jk}^t\le 0,\ k=1,\ 2,\dots,\ p\\&u_{jH}\ge 0,\ \ j=1,\ 2,\dots,\ n\\&v_{iH}\ge 0,\ \ i=1,\ 2,\dots,\ m\end{aligned} \tag{2}$$

The efficiency of the DMU inputs and outputs in time t in the conditions of the time t+1 is calculated using the model

$$\Phi_H^{t+1}(\mathbf{x}^t,\mathbf{y}^t)=\sum_{j=1}^{n}u_{jH}\,y_{jH}^t \rightarrow MAX \qquad \text{subject to} \qquad \begin{aligned}&\sum_{i=1}^{m}v_{iH}\,x_{iH}^t=1\\&-\sum_{i=1}^{m}v_{iH}\,x_{ik}^{t+1}+\sum_{j=1}^{n}u_{jH}\,y_{jk}^{t+1}\le 0,\ k=1,\ 2,\dots,\ p\\&u_{jH}\ge 0,\ \ j=1,\ 2,\dots,\ n\\&v_{iH}\ge 0,\ \ i=1,\ 2,\dots,\ m\end{aligned} \tag{3}$$

Last two models are intended for efficiency calculation in case of combination of inputs and outputs in different time with conditions in different time. We use the following models:

Inputs in time t+1, outputs in time t, and production conditions in time t

$$\Phi_H^t(\mathbf{x}^{t+1},\mathbf{y}^t)=\sum_{j=1}^{n}u_{jH}\,y_{jH}^t \rightarrow MAX \qquad \text{subject to} \qquad \begin{aligned}&\sum_{i=1}^{m}v_{iH}\,x_{iH}^{t+1}=1\\&-\sum_{i=1}^{m}v_{iH}\,x_{ik}^t+\sum_{j=1}^{n}u_{jH}\,y_{jk}^t\le 0,\ k=1,\ 2,\dots,\ p,\\&u_{jH}\ge 0,\ \ j=1,\ 2,\dots,\ n,\\&v_{iH}\ge 0,\ \ i=1,\ 2,\dots,\ m\end{aligned} \tag{4}$$

and inputs in time t+1, outputs in time t, and production conditions in time t+1

$$\Phi_H^{t+1}(\mathbf{x}^{t+1},\mathbf{y}^{t}) = \sum_{j=1}^{n} u_{jH}\, y_{jH}^{t} \rightarrow MAX \qquad \text{subject to} \qquad \begin{array}{l} \sum_{i=1}^{m} v_{iH}\, x_{iH}^{t+1} = 1 \\ -\sum_{i=1}^{m} v_{iH}\, x_{ik}^{t+1} + \sum_{j=1}^{n} u_{jH}\, y_{jk}^{t+1} \le 0,\ k = 1,\ 2,\dots,\ p \\ u_{jH} \ge 0,\ \ j = 1,\ 2,\dots,\ n \\ v_{iH} \ge 0,\ \ i = 1,\ 2,\dots,\ m \end{array} \tag{5}$$

Model BCC

The model BCC DEA model (Banker et al. [2]) supposes variable returns to scale. The input oriented BCC model uses the following linearization of the original DEA model:

$$\Theta_H^{t}(\mathbf{x}^{t},\mathbf{y}^{t}) = \sum_{j=1}^{n} u_{jH}\, y_{jH}^{t} + q_H \rightarrow MAX \qquad \text{subject to} \qquad \begin{array}{l} \sum_{i=1}^{m} v_{iH}\, x_{iH}^{t} = 1 \\ -\sum_{i=1}^{m} v_{iH}\, x_{ik}^{t} + \sum_{j=1}^{n} u_{jH}\, y_{jk}^{t} + q_H \le 0 \\ \qquad k = 1,\ 2,\dots,\ p \\ u_{jH} \ge 0,\ \ j = 1,\ 2,\dots,\ n \\ v_{iH} \ge 0,\ \ i = 1,\ 2,\dots,\ m \\ q_H \ \text{without limitation} \end{array} \tag{6}$$

where Θ_H^t is efficiency of DMU H, u_{jk} and v_{ik} are weights of outputs and inputs, y_{jk}^t is the amount of the j^{th} output from unit k, and x_{ik}^t is the amount of the i^{th} input to the k^{th} unit in time t, H is index of the evaluated DMU, and q_H is variable allowing variable returns to scale.

Malmquist index

Malmquist productivity index was introduced into the literature by Caves, Christensen, and Diewert [4]. This index quantifies the change in multi-factor productivity of the DMU with the CRS assumption at two different points in time. A decade later, Färe, Grosskopf, Lindgren, and Roos [7] identified technological change and technical efficiency change as two distinct components of productivity change. Subsequently, Färe, Grosskopf, Norris, and Zhang [8] relaxed the CRS assumption allowing VRS at different points on the production frontier and offered a new decomposition of the Malmquist productivity index incorporating a new component representing the contribution of changes in scale efficiency.

Malmquist index measures the efficiency change between two time periods by geometric mean of the ratios of the efficiency of each data point relatively to a common technological frontier in the first and second period.

$$M_H(\mathbf{x}^{t+1},\mathbf{y}^{t+1},\mathbf{x}^{t},\mathbf{y}^{t}) = \left(\frac{\Phi_H^{t}(\mathbf{x}^{t+1},\mathbf{y}^{t+1})}{\Phi_H^{t}(\mathbf{x}^{t},\mathbf{y}^{t})} \times \frac{\Phi_H^{t+1}(\mathbf{x}^{t+1},\mathbf{y}^{t+1})}{\Phi_H^{t+1}(\mathbf{x}^{t},\mathbf{y}^{t})} \right)^{1/2} \tag{7}$$

where x_{ik} is the amount of i^{th} input to k^{th} unit and H is index of the evaluated DMU.

Malmquist index is often decomposed into two values.

$$M_H(\mathbf{x}^{t+1},\mathbf{y}^{t+1},\mathbf{x}^{t},\mathbf{y}^{t}) = E_H T_H = \frac{\Phi_H^{t+1}(\mathbf{x}^{t+1},\mathbf{y}^{t+1})}{\Phi_H^{t}(\mathbf{x}^{t},\mathbf{y}^{t})} \times \left(\frac{\Phi_H^{t}(\mathbf{x}^{t+1},\mathbf{y}^{t+1})\Phi_H^{t}(\mathbf{x}^{t},\mathbf{y}^{t})}{\Phi_H^{t+1}(\mathbf{x}^{t+1},\mathbf{y}^{t+1})\Phi_H^{t+1}(\mathbf{x}^{t},\mathbf{y}^{t})} \right)^{1/2} \tag{8}$$

The first ratio - efficiency change E_H - represents the DMU technical efficiency change between two periods, and the second one - technological change T_H - represents the efficiency frontier shift between two periods (Färe et al. [7]). Considering VRS, the technical efficiency change E_H can be calculated using enhanced formula

$$E_H = \frac{\dfrac{\Phi_H^{t+1}(\mathbf{x}^{t+1},\mathbf{y}^{t+1})}{\Theta_H^{t+1}(\mathbf{x}^{t+1},\mathbf{y}^{t+1})}}{\dfrac{\Phi_H^{t}(\mathbf{x}^{t},\mathbf{y}^{t})}{\Theta_H^{t}(\mathbf{x}^{t},\mathbf{y}^{t})}} \times \frac{\Theta_H^{t+1}(\mathbf{x}^{t+1},\mathbf{y}^{t+1})}{\Theta_H^{t}(\mathbf{x}^{t},\mathbf{y}^{t})} = SE_H \times PE_H \tag{9}$$

where the first ratio represents scale efficiency change SE_H and the second one pure efficiency change PE_H of evaluated DMU.

Also technological change T_H can be further decomposed into the magnitude index and bias index, decomposed into the output bias index and input bias index. Since the values of these indices are not analysed in this study, this decomposition is not explain here.

Malmquist index correction

Malmquist index and its elements values are greater than 0. A value greater than 1 means efficiency increasing and contrary, while one less than 1 means decreasing of the efficiency. For greater clarity Malmquist index can be corrected or transformed by the following formulas (Zbranek [15])

$$\begin{aligned} &\text{If } M_H < 1 \text{ than } \bar{M}_H = 1 - \tfrac{1}{M_H} \\ &\text{If } M_H = 1 \text{ than } \bar{M}_H = 0 \\ &\text{If } M_H > 1 \text{ than } \bar{M}_H = M_H - 1 \end{aligned} \tag{10}$$

If any value of corrected Malmquist index is equal to 0, there is no efficiency change. If any corrected Malmquist index value is greater than 0, there is efficiency increasing, and if it is lower than 0, there is efficiency decreasing. Also other elements of Malmquist index decomposition can be corrected by the same way with the same meaning. For clarity of their values corrected form of indices will be used in this analysis.

3 Efficiency of the agricultural sector in the regions of the Czech Republic

For the analysis of the efficiency of the agricultural sector in the Czech regions we used data from Eurostat databases and yearbooks of the Czech regions. Comparable basic factors of production, namely the area of agricultural land (ha), labour as annual work units and gross fixed capital (CZK mil) were chosen as inputs, while gross value added (CZK mil) created in the agricultural sector was chosen as output (**Table 1**). The gross fixed capital and gross value added were recalculated using interest rates and inflation index.

	NUTS III	2008	2009	2010	2011	2012	2013
Area of agricultural land (ha)	Min	25,537	25,458	25,432	16,521	16,674	16,841
	Median	395,365	394,516	392,785	392,031	394,143	392,509
	Max	970,279	969,861	969,426	859,074	857,859	857,797
Overall work (AWU)	Min	2,227	2,777	3,017	3,357	3,757	3,338
	Median	12,988	13,131	12,973	13,300	12,988	12,886
	Max	26,959	26,228	24,582	26,186	25,626	26,132
Gross fixed capital (CZK mil.)	Min	211	334	406	582	742	749
	Median	2,381	1,741	2,044	1,909	3,060	2,634
	Max	5,014	3,469	3,904	4,449	5,828	5,257
Gross value added (CZK mil.)	Min	1,584	1,577	1,731	2,265	2,676	2,528
	Median	5,839	4,900	4,154	6,034	6,279	6,364
	Max	11,876	8,683	9,076	12,810	14,155	13,725
Interest rates (MF, according to Eurostat)		*4.6*	*4.8*	*4.2*	*3.7*	*3.9*	*3.9*
Inflation (MF, according to Eurostat)		*6.3*	*0.6*	*1.4*	*2.3*	*1.7*	*1.7*

Table 1 Basic information about the data

Starting calculations of CCR model (1) were performed for all 14 regions of the Czech Republic. Since Prague region is rated by the model DEA significantly better than other regions in all periods and the other regions generally do not achieved efficiency and therefore could not be sufficiently analysed, Prague region was excluded from further evaluation. Then CCR models (1) with the CRS assumption and also BCC model (6) assuming VRS were used for the global evaluation of 13 Czech regions excluding Prague region (**Table 2**).

Firstly we compare the overall technical efficiency in each period of each region under the CRS and VRS assumptions. Five, six or eight (more than one half) regions are shown as technical efficient according to the BCC model, only two, three or four (less than one third) regions are effective according to the CCR model.

Input orientation	CCR efficiency (%)						BCC efficiency (%)					
	2008	2009	2010	2011	2012	2013	2008	2009	2010	2011	2012	2013
Min	54.15	79.22	81.98	72.29	72.59	73.07	57.13	80.54	84.79	82.03	74.98	83.88
Median	77.42	90.53	87.49	91.67	90.87	90.97	95.46	100.00	98.44	97.82	98.21	94.23
Average	79.53	91.05	90.13	87.77	90.84	89.01	87.71	95.58	95.38	94.36	93.45	94.90
Standard Deviation	15.86	7.51	6.02	8.59	7.41	8.40	14.70	6.30	5.33	6.73	7.76	4.94

Table 2 Efficiency of the regions according to the models CCR and BCC (without Prague)

3.1 Changes of the elements of MI

Subsequently, DEA models (1) to (6) were computed for all 13 regions and all periods and MI and selected MI elements were calculated and corrected for cleared explanation. The development of MI has expected behaviour after crisis. The agricultural sector shows efficiency decreasing in periods 2008/9 and 2009/10, efficiency increasing between 2010 and 2011 after crisis followed by little or no efficiency changes during next three years. The efficiency development in all regions are nearly equal (**Table 3**, **Figure 1**).

	2008/09	2009/10	2010/11	2011/12	2012/13
Min	-57%	-28%	15%	-15%	-15%
Median	-8%	-13%	38%	0%	4%
Average	-17%	-12%	37%	1%	2%
Max	18%	13%	57%	28%	11%
Standard Deviation	22%	11%	11%	11%	8%

Table 3 Corrected Malmquist indices – The annual change of the regions efficiency

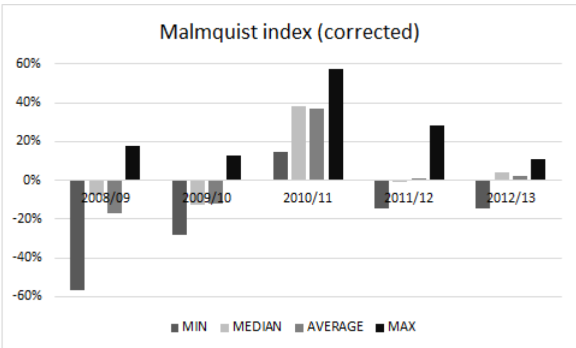


Figure 1 Corrected Malmquist index - The annual change of the regions efficiency

The analysis of components of MI indicates that this development is driven by the development of technological changes, changes of the production possibility frontier during crisis and after it. Technological change has the similar characteristics as MI in contrast to the technical efficiency change (**Figure 1**, **Figure 2**).

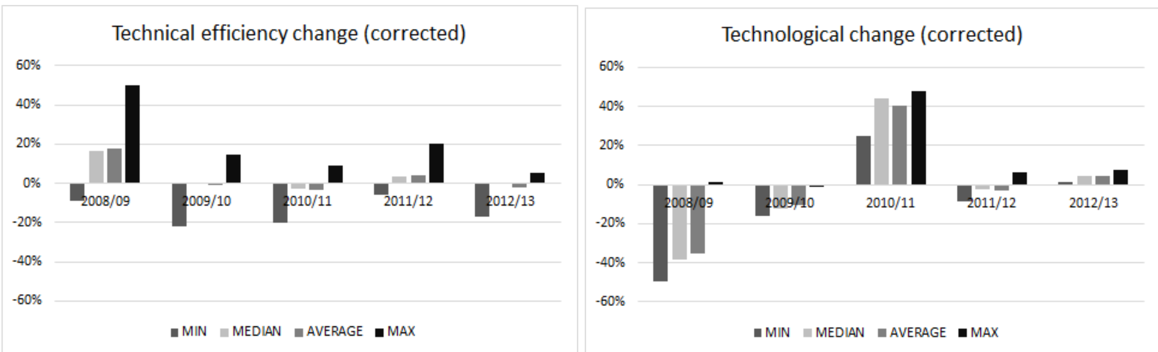


Figure 2 The annual technical efficiency change and technological change

The technical efficiency change index and its components pure efficiency change and scale efficiency change indices have similar values in all analyzed years (**Figure 2**, **Figure 3**). These indices show slight improvement from year 2008 to 2009 and show little or no changes during next five years. Main differences are in minimal and maximal values of these year to year indices. It means, that agricultural sector does not change its production system after overcoming the crisis. Small differences are caused by the change of production volume.

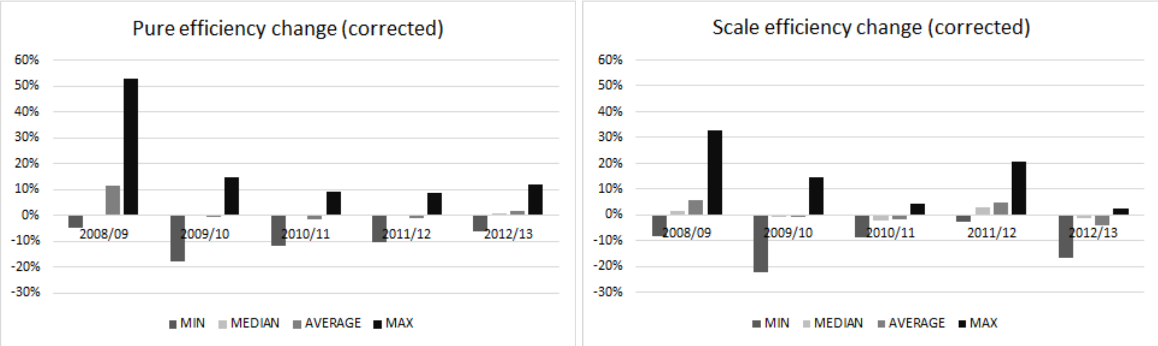


Figure 3 Pure and scale annual efficiency change of the regions in five periods

4 Conclusion

In 2008 the world economy faced its own most dangerous (financial) crisis. Malmquist index identifies expected development of efficiency in years 2008/2009 and following years; an economic collapse in 2008, followed by economic recovery till 2011 and the development in next years without fluctuations. Also the technological change index shows similar development as the global MI. On the other hand the efficiency change (technical, pure or scale efficiency change indices) shows very small positive as well as negative changes in all years. It can be explained by the relative permanence of the agricultural sector, by the effort of efficient productivity frontier to return to the previous level and also by permanent level, of technical efficiency.

References

[1] Andersen, P. and Petersen N. C.: A Procedure for Ranking Efficient Units in Data Envelopment Analysis. *Management Science* **39** (1993), 1261-1264.

[2] Banker, R. D., Charnes, R. F. and Cooper, W. W.: Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science* **30** (1984), 1078–1092.

[3] Bhattacharjee, J.: Resource Use and Productivity in World Agriculture. *Journal of Farm Economics* **37** (1955), 57-71

[4] Caves, D. W., Christensen, L. R. and Diewert, W. E.: The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity. *Econometrica* **50** (1982), 1393-1414.

[5] Charnes, A., Copper, W. and Rhodes, E.: Measuring the efficiency of decision-making units. *European Journal of Operational Research* **2** (1978), 429–444.

[6] Clark, C.: *The Conditions of Economic Progress.* Macmillan & Co., London, 1st edition, 1940.

[7] Färe, R., Grosskopf, S., Lindgren, B. and Ross P.: Productivity changes in Swedish pharmacies 1980-1989: a non-parametric Malmquist approach. *Journal of Productivity Analysis* **3** (1992), 85-101.

[8] Färe, R., Grosskopf, S., Norris, M. and Zhang Z.: Productivity Growth, Technical Progress and Efficiency Changes in Industrialised Countries. *American Economic Review* **84** (1994), 66-83.

[9] Farrel, M.: The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A* **120** (1957), 253-281.

[10] Lovell, C. A. K.: The Decomposition of Malmquist Productivity Indexes. *Journal of Productivity Analysis* **20** (2003), 437-458.

[11] Malmquist, S.: Index numbers and indifference surfaces. *Trabajos de Estadistica* **4** (1953), 209−242.

[12] Ray, S. C.: On An Extended Decomposition of the Malmquist Productivity Index. In: *Seventh European Workshop on Efficiency & Productivity Analysis*, Informs, 2001.

[13] Trueblood, M.A. and Coggins, J.: *Intercountry Agricultural Efficiency and Productivity: A Malmquist Index Approach*. Washington DC, U.S. Dept. of Agriculture, Economic Research Service, 2001.

[14] Yu, B., Liao, X. and Shen, H.: Parametric Decomposition of the Malmquist Index in Output-Oriented Distance Function: Productivity in Chinese Agriculture. *Modern Economy* **5** (2014), 70-85.

[15] Zbranek, P.: *Spoločná polnohospodárská politika a jej dopad na výkonnosť polnohospodárských podnikov: Aplikácia Malmquistových indewxov a Luenbergerových indikátorov*, Doctoral thesis, Slovak University of Agriculture in Nitra, 300, 2015.

Modelling Effects of Oil Price Fluctuations in a Monetary DSGE Model

Jan Brůha[1], Jaromír Tonner[2], Osvald Vašíček[3]

Abstract. The significant drop in oil prices after the end of 2014 has caused a positive supply shock for oil importing economies and at the same time it has further fuelled deflationary tendencies that already had been presented in many advanced countries. Because of the combination of these two effects in the environment of low inflation and low interest rates, the evaluation of the overall effect is not trivial. Therefore, it is not surprising that the drop in oil prices has become increasingly discussed also in the area of monetary policy. In this paper, we propose an extension of a standard DSGE model used for monetary policy analyses by the oil sector. Unlike existing extensions, we address issues not yet covered by Czech monetary policy models, such as the incidence of motor fuel excise taxes. Using the extended model, we empirically analyse the effects of the oil price drop on the Czech inflation and on the real economy.

Keywords: oil prices, fuel taxes, DSGE.

JEL classification: C44
AMS classification: 90C15

1 Motivation

The goal of the paper is to look at the impact of the changes in oil prices on the Czech economy through the lens of a monetary dynamic stochastic general equilibrium (DSGE) model. A reader may ask why we want to look at the impact of oil prices through the lens of a *monetary* structural model. Oil and energy prices are more often investigated in the area of energy, environmental, or transport economics and policy rather than in the area of monetary economics and probably. Indeed, the most widely used quantitative tools in this area are static or quasi-dynamic computable general equilibrium (CGE) models rather than DSGE models with forward looking agents and policy. This can be seemingly supported by the fact that the change in the oil price is after all a change in the relative price of one input and monetary policy deals with the overall price level rather than with relative prices of various goods.

This naive position can be easily challenged if we take into the account the fact that because of price stickiness, the change in a relative price is usually also accompanied by a change in the overall price level. This is even more true for commodities that have a significant share in the consumption basket and therefore the corresponding price changes affect the aggregate price indexes in a non-negligible way. And this is precisely the case of motor fuels, which in advanced countries exhibit some 4% share of consumption basket and which are very oil-intensive. Therefore, even if it is true that in the very long-run the change in the oil price is just a change in the relative price of one of the production inputs to which the production and consumption structure of the economy would adapt in the respective way, from the medium-term perceptive it is natural that even stabilization policy in general, and therefore also monetary policy, should monitor the fluctuations in the oil price and their impacts on the economy.

After the significant drop in oil prices after the end of 2014, this reasoning has gained even more prominence. On the one hand, the drop of oil price is a positive supply shock for oil importing countries, on the other hand, the resulting drop in motor fuel prices may have fueled deflationary tendencies that already had been presented in many advanced countries. The magnified deflationary tendencies may decrease the supply-shock benefits since in the environments of the low interest rates, the monetary

[1]Czech National Bank, Economic Research Department, Na Příkopě 28, Prague, Czech Republic, Jan.Bruha@cnb.cz
[2]Czech National Bank, Monetary Department, Na Příkopě 28, Prague, Czech Republic, Jaromir.Tonner@cnb.cz
[3]Masaryk University, Deparment of Economics, Lipová 41a, Brno, Czech Republic, osvald@econ.muni.cz

policy is constrained in its reaction. Moreover, the resulting drop in the consumption price index might negatively affect the overall inflation expectations and only a credible central bank with transparent policy can avoid this implausible side effect of the shock.

These arguments show that it is important to analyze the effects of fluctuations in oil prices not only for the purpose of energy or environmental policy, but also for the purpose of monetary policy. This is the motivation of this paper to look at the aggregate implications of the oil price changes on the Czech economic growth and inflation through the lens of the monetary DSGE model. There is another contribution of this paper which is that unlike existing extensions of DSGE models, we address issues not yet covered by Czech monetary policy models, such as the incidence of motor-fuel excise taxes. therefore, the paper may be interesting not only for those readers who wish to learn about the effects of the recent drop in the oil price, but for DSGE modelers in general.

The rest of this paper is organized as follows. The next section describes the extension and discusses its properties. Section **3** describes the results. The last section **4** concludes.

2 Data and Model

The model we are using for the analysis is a modification to the g3 model of the Czech National Bank. We refer to the publication (see [**1**]) for the description of the baseline version of the g3 model.

2.1 Data

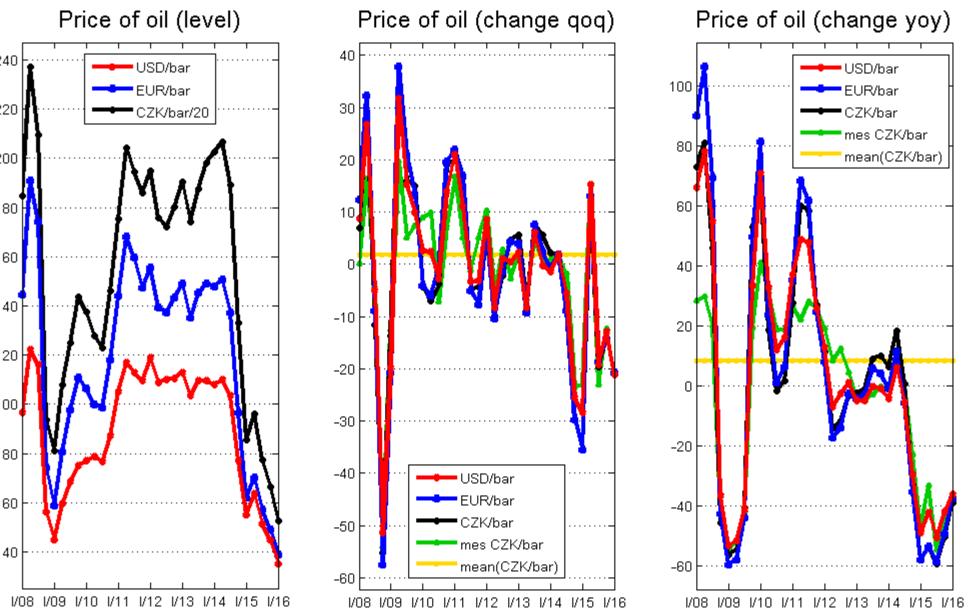


Figure 1 Price of oil

2.2 Model

There are four major modifications of the original model.

First, we assume that the oil is one of the imported goods and that it is domestically transformed to motor fuels (we ignore other usage of oil, for example for the production of plastic materials, as this is minority role). Given the fact motor fuels are very tradable goods, it is immaterial whether they are produced domestically or not. The motor fuel price is then affected not only by the oil price, but also by excise tax and –for households– by the value added tax. Following the decomposition analysis by Pisa ([**4**]), the elasticity of the motor fuel price with respect to the oil price is about 0.4. The inclusion of the

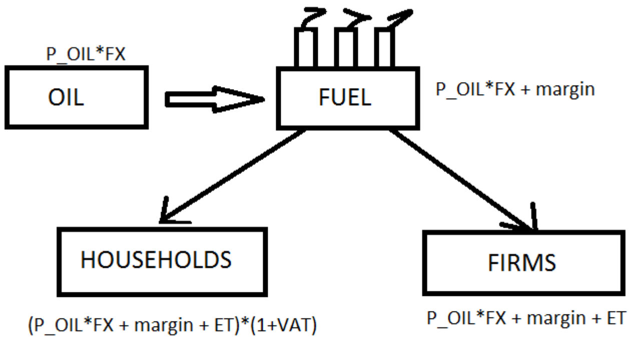


Figure 2 Structure of the model

effect of excise taxes on the final price of the motor fuel is often neglected in monetary models (unlike in their CGE counterparts) and presents one of the contributions of this paper for DSGE modelers. This motor-fuel production chain chain is illustrated on Figure **2**.

Second, we assume that the production function for the intermediate goods is extended as follows:

$$y_t = A_t K_t^{\alpha_K} L_t^{\alpha_L} \mathcal{F}_t^{\alpha_F}, \tag{1}$$

where y_t is the production of the intermediate good, A_t is the productivity, K_t is the capital used, L_t is the labour input, $\mathcal{F}_t$ are motor fuels used for the production of the intermediate good, and the non-negative shares obey the restriction: $\alpha_K + \alpha_L + \alpha_F = 1$.

The production function (1) together with the usual cost-minimization assumption imply the following reduced form for the production of the intermediate goods:

$$y_t = A_t^{1/(1-\alpha_F)} K_t^{\alpha_K/(1-\alpha_F)} L_t^{\alpha_L/(1-\alpha_F)} \left(P_t^{MF}/P_t^y\right)^{-\alpha_F/(1-\alpha_F)}, \tag{2}$$

where P_t^{MF}/P_t^y is the relative price of the motor fuel to the price of the intermediate good. This basically means that the decrease in the relative price of motor fuels is homeomorphic to the productivity increase with the elasticity $alpha_F/(1-alpha_F)$.

Third, we extend the utility function of the households as follows:

$$U = \sum_{t>0} \beta^t \log\left(\frac{C_t - \chi\overline{C}_t}{1-\chi}\right) + \kappa_{MF} \log\left(\frac{C_t^{MF} - \chi C_t^{MF}}{1-\chi}\right), \tag{3}$$

where C_t is the private consumption, $\overline{C}$ is the external habit, C_t^{MF} is the motor fuel consumption, and $\overline{C}_t^{MF}$ is the habit of the external motor fuel consumption. The intratemporal optimal consumption pattern is given as:

$$P_t^C(C_t - \chi\overline{C}_t) = \kappa_{MF} P_t^{MF}(C_t^{MF} - \chi\overline{C}_t^{MF}), \tag{4}$$

where P_t^C is the ex-fuel CPI and P_t^{MF} is the end-user price of motor fuels. The inter-temporal substitution is given by the standard Euler equation ([**1**]).

Fourth, as oil is imported, the net foreign asset equation should be modified accordingly. The rest of the model is unchanged.

3 Empirical analysis

To assess the plausibility of the model with oil prices, standard tests were performed on historical data, mainly impulse response analysis and decompositions to shocks.

3.1 Impulse responses

The oil price shock is a representant of typical supply-side shocks, thus an increase of oil prices should lead to higher inflation, interest rates, more depreciated exchange rate (via negative net foreign assets) and to a decrease of the real economy. The calibration assumes that the growth of oil prices by ten p.p implies decrease of domestic real gdp by 0.15 p.p. The negative effect to real variables is only partially compensated by the export growth via exchange rate depreciation.

The excise tax shock has similar effect to the main variables as the oil price shock except the exchange rate. Because of the fact that higher taxes represent competitive disadvantage, higher taxes imply more appreciated exchange rate.

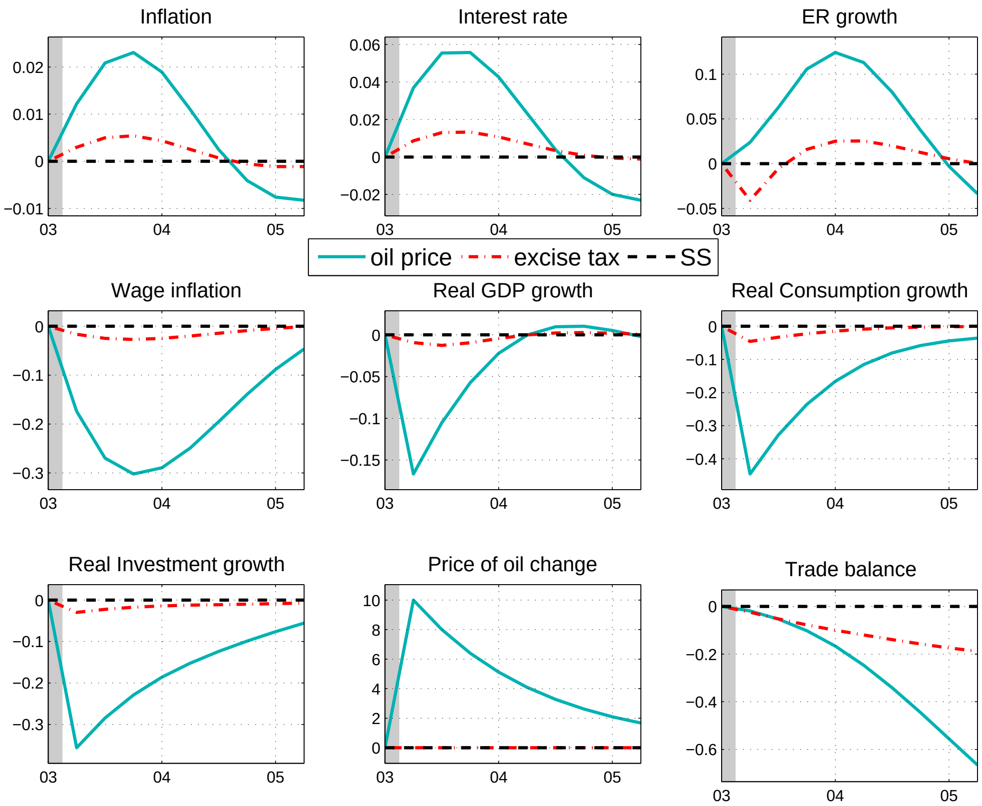


Figure 3 Oil price, exice tax and motor fuel shocks

3.2 Shock decomposition

Shock decomposition of the extended model has also revealed the contribution of oil price changes to the GDP growth. The overall effect is approximately 0.7 p.b. to GDP growth and -0.5 p.p to the CPI inflation for the years 2014 and 2015.

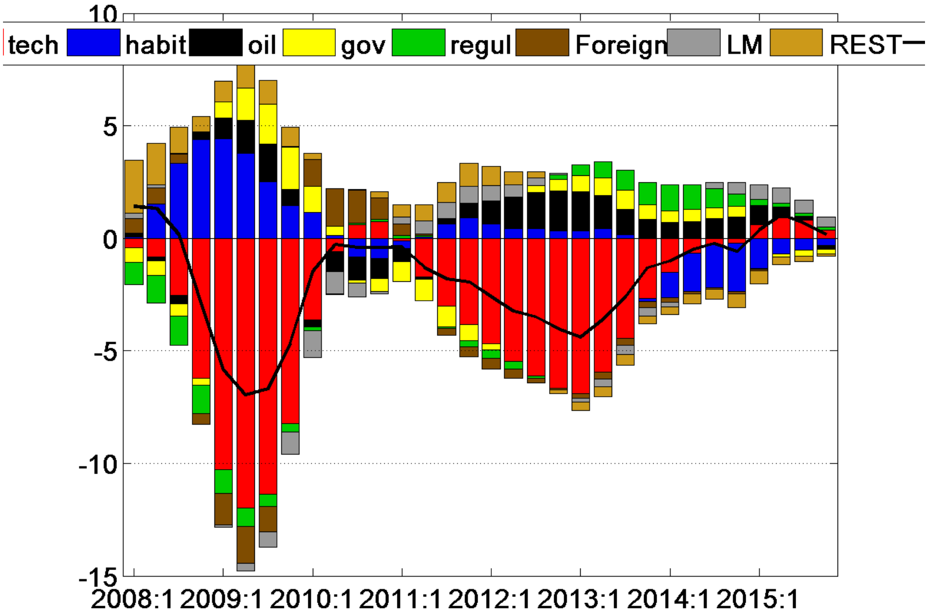


Figure 4 Real GDP growth shock decomposition

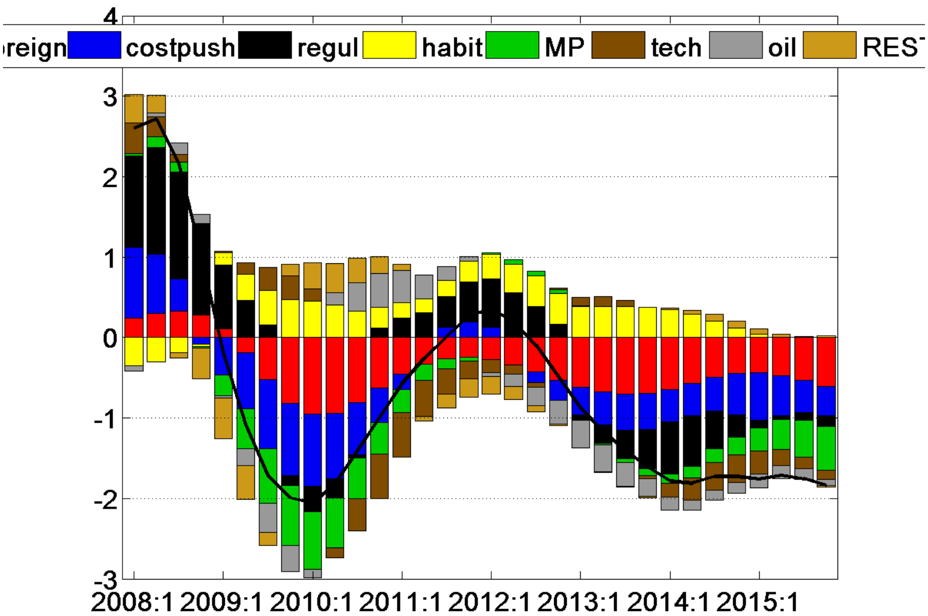


Figure 5 CPI inflation shock decomposition

The *tech* label in the figures denotes technology shocks as labour-augmented technology shocks and TFP shocks. The *oil* label denotes the oil price shocks (excise tax shock included). The *cost-push* label denotes costpush shocks in consumption, investment, government, export, import and intermediate sectors. The *Foreign* label denotes shocks to foreign variables, i.e. foreign demand, foreign interest rates and foreign prices. The *gov* label denotes government shocks. The *habit, euler, inv, UIP, MP*, and *regul* labels denote habit in consumption, wedge in the Euler equation, investment specific, uncovered interest rate parity, monetary policy and regulated prices shocks respectively. The *LM* label denotes labour market shocks. The *REST* label comprises the effects of those shocks which do not appear in the legend (including initial conditions).

4 Summary

The significant drop in oil prices after the end of 2014 has caused a positive supply shock for oil importing economies and at the same time it has further fuelled deflationary tendencies that already had been presented in many advanced countries. The proposed extension of a standard DSGE model has revealed that the effect of the oil price drop on the real economy had been about 0.7 p.p. to GDP growth and about -0.5 p.p to the CPI inflation for the years 2014 and 2015.

Acknowledgements

This work was supported by funding of specific research at Faculty of Economics and Administration, project MUNI/A/1040/2015. This support is gratefully acknowledged.

References

[1] Andrle, M., Hlédik, T., Kameník, O., and Vlček, J.: Implementing the New Structural Model of the Czech National Bank. Czech National Bank, Working Paper Series 2, 2009.

[2] Baumeister, C., Kilian, L.: What Central Bankers Need To Know About Forecasting Oil Prices. *International Economic Review* **55** (2014), 869–889.

[3] Blanchard, O., Galí, J.: The Macroeconomic Effects of Oil Shocks: Why are the 2000s so Different from the 1970s?. C.E.P.R. Discussion Papers, https://ideas.repec.org/p/cpr/ceprdp/6631.html, 2008.

[4] Píša, V.: The Demand for Motor Fuels in the Central European Region and its Impacts on Indirect Tax Revenues. Proceedings of the Conference Technical Computing Bratislava, 2012. Available at https://dsp.vscht.cz/konference_matlab/MATLAB12/full_paper/064_Pisa.pdf

Information Support for Solving the Order Priority Problem in Business Logistic Systems

Robert Bucki[1], Petr Suchánek[2]

Abstract. The ordering system diagram plays an important role in many manufacturing plants. There are cases in which orders cannot be made at the moment they are received so it is necessary to change the sequence of making them to increase production efficiency. Setting the priority of orders plays an important role as it lets manufacturers organize the course of production. Analytical methods are often confronted with their limits and, therefore, the use of computer simulation increases progressively. The simulation method as a tool is more and more often used to define the appropriate procedure for the selection of orders and organizing a manufacturing process. They are based on mathematical models which are subsequently used as a source for simulation programs. These programs are especially useful for strategic production planning. The main aim of the article is to design an adequate mathematical model based on the heuristic approach. The model represents one of the alternatives for the definition of the proper scheme of the sequence of orders with a possibility of changing their priorities.

Keywords: information, mathematical modeling, simulation, ordering system, production system, business logistics, organization of production, strategic planning.

JEL Classification: C02, C63, L21
AMS Classification: 93A30, 00A72, 90B06

1 Introduction

Each product which is produced in a company and subsequently sold comes into being by means of a certain sequence of actions. Production can be fully automatized or organized using a combination of machines and human labour. Production is realized on the basis of orders. In case when a company produces only one product type which requires modifications to the manufacturing process, the number of output products is given by the manufacturing capacity of a manufacturing line. If a company is engaged in manufacturing of a variety of different products which require changes in the production process (e.g. adjustments of production lines, the use of other machines, correcting dislocation of machines, etc.), there is a need to tackle optimization problems, respectively the relationship between the management of orders and production itself. Orders come independently throughout the manufacturing process and to achieve the highest possible efficiency of the enterprise as a whole it is necessary to have in place an adequate system of order control which is part of the production process. To optimize the presented problem there is a need to base it on the theory of process management which is currently a prerequisite for the use of advanced control methods (including quality control) and cannot be done without computer support [1], [3], [8], [11], [15].

Advanced software for decision support and management is equipped with increasingly integrated tools to support the simulations required to define the outputs of the company under the given circumstances [2], [12]. Simulation models are based on a variety of analytical and mathematical methods which are the fundamental basis for creation of simulation models and algorithms. To create simulation models there is a need to implement, for example, multi-agents approach [10], Petri nets [9], [16], hybrid approaches [7], heuristic approaches [14], [4] and others obviously with the support of mathematical and statistical methods. In this article, the authors use a heuristic approach accordingly to all possible variants of solving the problem (see algorithms in Tables 1-3). Finding all possible solutions to the problem is just one of the areas where heuristic algorithms are effectively implemented. Specifically, the paper highlights the problem of finding the sequence of orders to be made in the hypothetical manufacturing system. Orders can be chosen on the basis of the decision made by manufacturing algorithms including orders with the maximal priority. The second solution to this kind of problem can be effective in case of implementing a big number of simulation courses.

[1] Institute of Management and Information Technology, ul. Legionów 81, 43-300 Bielsko-Biala, Poland, rbucki@wsi.net.pl.
[2] Silesian University in Opava, School of Business Administration in Karviná, Department of Informatics and Mathematics, Univerzitní náměstí 1934/3, 73340 Karviná, Czech Republic, suchanek@opf.slu.cz.

2 Mathematical model

Let us assume there is a matrix of orders with elements which have adjusted priorities (1):

$$Z(\omega)^k = \left[z(\omega)_{m,n}^k \right] \ , \ m = 1,\ldots, M \ , \ n = 1,\ldots, N \ , \ k = 1,\ldots, K \ , \ \omega = 1,\ldots, \Omega \tag{1}$$

where: $z(\omega)_{m,n}^k$ - the n-th order for the m-th customer at the k-th stage with the ω-th priority. Orders are subject to change after each decision about manufacturing as follows (2):

$$Z(\omega)^0 \rightarrow Z(\omega)^{1k} \rightarrow \ldots \rightarrow Z(\omega)^k \rightarrow \ldots \rightarrow Z(\omega)^K \tag{2}$$

At the same time $z(\omega)_{m,n}^k \geq 0$ (the n-th order can be made for the m-th customer); $z(\omega)_{m,n}^k = -1$ (otherwise). The order matrix elements take the following values after each decision about manufacturing: $z(\omega)_{m,n}^k \leq z(\omega)_{m,n}^{k-1}$ (in case there is a manufacturing decision at the stage $k-1$; $z(\omega)_{m,n}^k = z(\omega)_{m,n}^{k-1}$ (otherwise). For simplicity reasons the following grades of the priority are introduced $\omega = 0$ (if there is no priority); $\omega = \min$ (in case of the minimal priority); $\omega = \max$ (in case of the maximal priority). It is assumed that if orders are not made in accordance with the demand, there are fines imposed. Let us introduce the matrix of fines for not making the nth order for the m-th customer (3):

$$C^k = \left[c_{m,n}^k \right] \ , \ m = 1,\ldots, M \ , \ n = 1,\ldots, N \ , \ k = 1,\ldots, K \tag{3}$$

where: $c_{m,n}^k$ - the fine to be paid to the m-th customer for not making the agreed n-th order in time till the k-th stage of the manufacturing process (expressed in conventional monetary units). At the same time if $c_{m,n}^k = -1$ the m-th customer does not have any priority for the n-th product, otherwise $c_{m,n}^k > 0$. To simplify the highly complex problem it is assumed that orders are directed to the manufacturing system which can consist of more subsystems. Such systems were analysed in detail in [5], [6] and [13]. To make orders there is a need to make use of the control methods - heuristic algorithms. They choose orders according to the requirements specified in them. The number of algorithms can be immense and it seems that it is worth testing all of them as they can deliver a solution which leads to minimizing order making time. There are algorithms taking into account only the amount of the order without considering its kind and a customer it is made for (Alg_1 – Alg_4) shown in Table 1, algorithms considering the customer - order correlation (Alg_5 – Alg_20) shown in Table 2 and algorithms based on the order - customer correlation (Alg_21 – Alg_36) shown in Table 3. It is assumed that all considered algorithms have the same priority ω. Moreover, if the heuristic approach either fails to deliver a satisfactory solution or it is decided to implement methods “at random” focusing on either algorithms or orders or customers is unavoidable. There is also a possibility to combine all mentioned methods e.g. mixing the random choice of all the above.

2.1 Algorithms_α_$_{max/min}$

It is assumed that orders for manufacturing are chosen in accordance with the amount of the order itself. Either the maximal or minimal orders are searched for. Moreover, there is an additional need to decide which order matrix element should be made in case there are more identical orders (see Table 1).

No. of algorithm	Choice of the order	Decision in case of more identical orders
Alg_1	$z(\omega)_{\mu,\eta}^k = \max\limits_{\substack{1\le\alpha\le M \\ 1\le\beta\le N}} z(\omega)_{\alpha,\beta}^k$	$z(\omega)_{\mu',\eta'}^k = \min\limits_{\substack{1\le\mu\le M \\ 1\le\eta\le N}} z(\omega)_{\mu,\eta}^k$
Alg_2	$z(\omega)_{\mu,\eta}^k = \max\limits_{\substack{1\le\alpha\le M \\ 1\le\beta\le N}} z(\omega)_{\alpha,\beta}^k$	$z(\omega)_{\mu',\eta'}^k = \max\limits_{\substack{1\le\mu\le M \\ 1\le\eta\le N}} z(\omega)_{\mu,\eta}^k$
Alg_3	$z(\omega)_{\mu,\eta}^k = \min\limits_{\substack{1\le\alpha\le M \\ 1\le\beta\le N}} z(\omega)_{\alpha,\beta}^k$	$z(\omega)_{\mu',\eta'}^k = \min\limits_{\substack{1\le\mu\le M \\ 1\le\eta\le N}} z(\omega)_{\mu,\eta}^k$
Alg_4	$z(\omega)_{\mu,\eta}^k = \min\limits_{\substack{1\le\alpha\le M \\ 1\le\beta\le N}} z(\omega)_{\alpha,\beta}^k$	$z(\omega)_{\mu',\eta'}^k = \max\limits_{\substack{1\le\mu\le M \\ 1\le\eta\le N}} z(\omega)_{\mu,\eta}^k$

Table 1 Choosing orders characterized by the maximal priority

First, the maximal priority order $z(\omega)^k_{\mu,\eta}$, $1 \le \mu \le M$, $1 \le \eta \le N$, $k = 0,1,..., K$ is chosen and then, if there are two or more such orders, the algorithm chooses the order $z(\omega)^k_{\mu',\eta'}$, $1 \le \mu' \le M$, $1 \le \eta' \le N$, $k = 0,1,..., K$.

2.2 Algorithms_α_custumer

It is assumed that orders for manufacturing are chosen in accordance with the detailed customer analysis. The μ-th customer, $\mu = 1,..., M$ characterized by the specific total order is searched for (4):

$$\mu : \sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \max \sum_{n=1}^{N} z(\omega)^k_{m,n} \ ; \ \mu : \sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \min \sum_{n=1}^{N} z(\omega)^k_{m,n} \ . \tag{4}$$

If there are more customers with identical number of orders, the one with the number μ' is taken into account: $\mu' = \max_{1\le\mu\le M} \mu$ or $\mu' = \min_{1\le\mu\le M} \mu$. Having found the μ'-th customer there is a need to detect the η-th order, $\eta = 1,..., N$; η: $z(\omega)^k_{\mu',\eta} = \max_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$; η : $z(\omega)^k_{\mu',\eta} = \min_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$. If there are more identical orders for the set μ'-th customer, the one with the number η' is taken into account $\eta' = \max_{1\le\eta\le N} \eta$, $\eta \in \langle 1; N \rangle$; $\eta' = \min_{1\le\eta\le N} \eta$, $\eta \in \langle 1; N \rangle$. Finally, the order $z(\omega)^k_{\mu',\eta'}$ becomes subject to manufacturing at the k' stage (see Table 2).

Alg. no.	μ	μ'	η	η'
Alg_5	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \max \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \max_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \max_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \max_{1\le\eta\le N} \eta$
Alg_6	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \max \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \max_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \max_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \min_{1\le\eta\le N} \eta$
Alg_7	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \max \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \max_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \min_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \max_{1\le\eta\le N} \eta$
Alg_8	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \max \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \max_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \min_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \min_{1\le\eta\le N} \eta$
Alg_9	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \max \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \min_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \max_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \max_{1\le\eta\le N} \eta$
Alg_10	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \max \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \min_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \max_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \min_{1\le\eta\le N} \eta$
Alg_11	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \max \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \min_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \min_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \max_{1\le\eta\le N} \eta$
Alg_12	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \max \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \min_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \min_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \min_{1\le\eta\le N} \eta$
Alg_13	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \min \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \max_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \max_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \max_{1\le\eta\le N} \eta$
Alg_14	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \min \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \max_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \max_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \min_{1\le\eta\le N} \eta$
Alg_15	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \min \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \max_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \min_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \max_{1\le\eta\le N} \eta$
Alg_16	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \min \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \max_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \min_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \min_{1\le\eta\le N} \eta$
Alg_17	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \min \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \min_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \max_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \max_{1\le\eta\le N} \eta$
Alg_18	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \min \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \min_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \max_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \min_{1\le\eta\le N} \eta$
Alg_19	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \min \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \min_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \min_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \max_{1\le\eta\le N} \eta$
Alg_20	$\sum_{n=1}^{N} z(\omega)^k_{\mu,n} = \min \sum_{n=1}^{N} z(\omega)^k_{m,n}$	$\mu' = \min_{1\le\mu\le M} \mu$	$z(\omega)^k_{\mu',\eta} = \min_{1\le\eta\le N} z(\omega)^k_{\mu',\eta}$	$\eta' = \min_{1\le\eta\le N} \eta$

Table 2 Choosing an order based on the specific customer

2.3 Algorithm_α_{order}

It is assumed that orders for manufacturing are chosen in accordance with the detailed order matrix analysis. The η-th order, $\eta = 1,..., N$ associated by the specific customer is searched for η (5):

$$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \max \sum_{m=1}^{M} z(\omega)_{m,n}^{k} \; ; \; \eta : \; \sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \min \sum_{m=1}^{M} z(\omega)_{m,n}^{k} \, . \tag{5}$$

If there are more identical orders, the one with the number η' is taken into account: $\eta' = \max_{1 \le \eta \le M} \eta$ or $\eta' = \min_{1 \le \eta \le M} \eta$, $\eta \in \langle 1; N \rangle$. Having found the η'-th order there is a need to detect the m-th customer, $m = 1,..., M$; μ : $z(\omega)_{\mu,\eta'}^{k} = \max z(\omega)_{m,\eta'}^{k}$; μ : $z(\omega)_{\mu,\eta'}^{k} = \min z(\omega)_{m,\eta'}^{k}$. If there are more customers where η'-th order, the one with the number μ' is taken into account $\mu' = \max_{1 \le \mu \le M} \mu$, $\mu \in \langle 1; M \rangle$; $\mu' = \min_{1 \le \mu \le M} \mu$, $\mu \in \langle 1; M \rangle$. Finally, the order $z(\omega)_{\mu',\eta'}^{k}$ becomes subject to manufacturing at the k' stage (see Table 3).

Alg. no.	η	η'	μ	μ'
Alg_21	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \max \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \max_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \max z(\omega)_{m,\eta'}^{k}$	$\mu' = \max_{1\le\mu\le M} \mu$
Alg_22	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \max \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \max_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \max z(\omega)_{m,\eta'}^{k}$	$\mu' = \min_{1\le\mu\le M} \mu$
Alg_23	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \max \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \max_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \min z(\omega)_{m,\eta'}^{k}$	$\mu' = \max_{1\le\mu\le M} \mu$
Alg_24	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \max \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \max_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \min z(\omega)_{m,\eta'}^{k}$	$\mu' = \min_{1\le\mu\le M} \mu$
Alg_25	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \max \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \min_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \max z(\omega)_{m,\eta'}^{k}$	$\mu' = \max_{1\le\mu\le M} \mu$
Alg_26	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \max \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \min_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \max z(\omega)_{m,\eta'}^{k}$	$\mu' = \min_{1\le\mu\le M} \mu$
Alg_27	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \max \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \min_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \min z(\omega)_{m,\eta'}^{k}$	$\mu' = \max_{1\le\mu\le M} \mu$
Alg_28	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \max \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \min_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \min z(\omega)_{m,\eta'}^{k}$	$\mu' = \min_{1\le\mu\le M} \mu$
Alg_29	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \min \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \max_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \max z(\omega)_{m,\eta'}^{k}$	$\mu' = \max_{1\le\mu\le M} \mu$
Alg_30	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \min \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \max_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \max z(\omega)_{m,\eta'}^{k}$	$\mu' = \min_{1\le\mu\le M} \mu$
Alg_31	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \min \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \max_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \min z(\omega)_{m,\eta'}^{k}$	$\mu' = \max_{1\le\mu\le M} \mu$
Alg_32	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \min \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \max_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \min z(\omega)_{m,\eta'}^{k}$	$\mu' = \min_{1\le\mu\le M} \mu$
Alg_33	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \min \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \min_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \max z(\omega)_{m,\eta'}^{k}$	$\mu' = \max_{1\le\mu\le M} \mu$
Alg_34	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \min \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \min_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \max z(\omega)_{m,\eta'}^{k}$	$\mu' = \min_{1\le\mu\le M} \mu$
Alg_35	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \min \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \min_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \min z(\omega)_{m,\eta'}^{k}$	$\mu' = \max_{1\le\mu\le M} \mu$
Alg_36	$\sum_{m=1}^{M} z(\omega)_{m,\eta}^{k} = \min \sum_{m=1}^{M} z(\omega)_{m,n}^{k}$	$\eta' = \min_{1\le\eta\le M} \eta$	$z(\omega)_{\mu,\eta'}^{k} = \min z(\omega)_{m,\eta'}^{k}$	$\mu' = \min_{1\le\mu\le M} \mu$

Table 3 Decisions made on the basis of the specific order

3 Complex search for the satisfactory choice of products

To present the pseudo-code representing the model responsible for searching for the satisfactory solution (see Table 4) there is a need to assume the following:

W - the number of simulations for the random choice of algorithms;
Y - the number of simulations for the random choice of orders;
A - the set of algorithms;
T_α - the total manufacturing time with the use of the α-th algorithm, $\alpha = 1,\dots,A$;

$T_{\alpha\_\min}$ - the minimal total manufacturing time with the use of the α-th algorithm, $\alpha = 1,\dots,A$;

T_A^w - the total manufacturing time with the use of the combination of algorithms in the w-th simulation, $w = 1,\dots,W$;

$T_{A\_\min}^w$ - the minimal total manufacturing time with the use of the combination of algorithms chosen at random in the w-th simulation, $w = 1,\dots,W$;

T_Z^y - the total manufacturing time with the use of the combination of orders chosen at random in the w-th simulation, $w = 1,\dots,W$;

$T_{Z\_\min}^y$ - the minimal total manufacturing time with the use of the combination of orders chosen at random in the w-th simulation, $w = 1,\dots,W$;

$T_{\min}$ - the satisfactory manufacturing time.

i)	Introduce W, Y, $Z(\omega)^k$, A, S^k, etc.	xv)	If w = 1, then $T_A^1 = T_{A\_\min}$. Otherwise, go to (xvi).
ii)	$w = 1$	xvi)	If $T_A^w < T_{A\_\min}$, then $T_A^w = T_{A\_\min}$. Otherwise, go to (xvii).
iii)	$y = 1$	xvii)	If $w = W$, then go to (xviii). Otherwise, go to (xix).
iv)	$\alpha = 1$	xviii)	Send $T_{A\_\min}$ to (xxviii) and then go to (xx).
v)	Simulate the production process for α.	xix)	$w =: w + 1$ and go to (xii).
vi)	$T_{\alpha\_\min} = T_\alpha$	xx)	Draw $z(\omega)_{m,n}^k$ at random.
vii)	$\alpha =: \alpha + 1$	xxi)	Simulate the manufacturing process till the system is brought to a standstill.
viii)	Simulate the production process for α.	xxii)	$\underset{1\le m\le M}{\forall}\ \underset{1\le n\le N}{\forall}\ z(\omega)_{m,n}^{k+1} \neq 0$? If *Yes*, go to (xx). If *No*, go to (xiii).
ix)	If $T_\alpha < T_{a\_\min}$, then $T_\alpha = T_{a\_\min}$. Otherwise, go to (x).	xxiii)	If $y = 1$, then $T_Z^1 = T_{Z\_\min}$. Otherwise, go to (xxiv).
x)	$\alpha = A$? If *Yes*, go to (xi). If *No*, go to (vii).	xxiv)	If $T_Z^y < T_{Z\_\min}$, then $T_Z^y = T_{Z\_\min}$. Otherwise, go to (xxv).
xi)	Send $T_{a\_\min}$ to (xxviii) and then go to (xii).	xxv)	If $y = Y$, then go to (xxvii). Otherwise, go to (xxvi).
xii)	Draw α at random.	xxvi)	$y =: y + 1$ and go to (xx).
xiii)	Simulate the manufacturing process till the system is brought to a standstill.	xxvii)	Send $T_{Z\_\min}$ to (xxviii).
xiv)	$\underset{1\le m\le M}{\forall}\ \underset{1\le n\le N}{\forall}\ z(\omega)_{m,n}^{k} \neq 0$? If *Yes*, go to (xii). If *No*, go to (xv).	xxviii)	$T_{\min} = \min(T_{\alpha\_\min}, T_{A\_\min}, T_{Z\_\min})$

Table 4 Pseudo-code representing searching for the satisfactory solution

4 Conclusion

Optimization of production and business processes is currently one of the major competitiveness prerequisites of enterprises. Optimization is performed with the use of information technology which is the primary tool for simulating the behaviour of the system under the given conditions. The starting point for computer simulations are simulation models developed using a range of approaches and mathematical and statistical methods. The paper presents a model how to search for a satisfactory solution with the use of heuristic algorithms based on the maximal or minimal total order, specific customers, specific orders, the random choice of algorithms, the random choice of orders. Realizing that the control of making orders with the use of heuristic algorithms only may not deliver a satisfactory solution there is a strong need to simulate making orders using the biggest possible number of simulations. Simulation replaces wasting resources as it does not require the real manufacturing environment. Further steps require carrying out a big number of simulations which have to result from introducing adequate initial data corresponding with the real manufacturing system. It seems to be obvious that the proper specification and projecting course should result in building a simulation tool to be implemented in the real system. The model presented in the paper is to be the basis for creating a dedicated simulator.

Acknowledgement

This paper was supported by the project SGS/19/2016 – Advanced Mining Methods and Simulation Techniques in the Business Process Domain.

References

[1] Alexe, C. G. and Alexe, C. M.: Computer solutions of management for the development process of new products. In: *Proceedings of the 8th International Conference Interdisciplinarity in Engineering, Inter-Eng 2014*. Elsevier Science Bv, Tirgu Mures, 2015, 1031-1037.

[2] Al-Sugair, F., Al-Suwailem, M., Al-Shedukhi, W. and Bin Sulaiman, A.: Using Simulation Software in the Classroom for the Implementation of Lean Principles in A Waterproof Manufacturing Company. In: *Proceedings of the 4th International Conference on Education and New Learning Technologies (EDULEARN)*. Iated-Int Assoc Technology Education A& Development, Barcelona, 2012, 7565-7572.

[3] Armillotta, A., Cavallaro, M. and Minnella, S.: A tool for computer-aided orientation selection in additive manufacturing processes. In: *Proceedings of the 6th International Conference on Advanced Research in Virtual and Physical Prototyping (VRatP), High Value Manufacturing: Advanced Research in Virtual and Rapid Prototyping*. Polytechn Inst Leiria, Leiria, 2014, 469-475.

[4] Arreola-Risa, A., Gimenez-Garcia, V. M. and Martinez-Parra, J. L.: Optimizing stochastic production-inventory systems: A heuristic based on simulation and regression analysis. *European Journal of Operational Research* **213**(1) (2011), 107-118.

[5] Bucki, R., Chramcov, B. and Suchánek, P.: Heuristic Algorithms for Manufacturing and Replacement Strategies of the Production System. *Journal of Universal Computer Science* **21**(4) (2015), 503-525.

[6] Bucki, R., Suchánek, P. and Chramcov, B.: Logistic optimization of the complex manufacturing system with parallel production lines. *Journal of Applied Economic Sciences* **8**(3) (2013), 270-285.

[7] del Conte, E. G., Schutzer, K. and Abackerli, A. J.: A hybrid monitoring-simulation system for contour error prediction on complex surfaces manufacturing. *International Journal of Advanced Manufacturing Technology* **77**(1-4) (2015), 321-329.

[8] Honchar, L. and Lendyuk, T.: Computer support of business-processes and multiperspective management as the basis of business operation. In: *Proceedings of the 2nd IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2003*. House of ScientistsLviv, 2003, 508-513.

[9] Kahloul, L., Bourekkache, S., Djouani, K., Chaoui, A. and Kazar, O.: Using High Level Petri Nets in the Modelling, Simulation and Verification of Reconfigurable Manufacturing Systems. *International Journal of Software Engineering and Knowledge Engineering* **24**(3) (2014), 419-443.

[10] Lin, J. and Long, Q. Q.: Development of a multi-agent-based distributed simulation platform for semiconductor manufacturing. *Expert Systems with Applications* **38**(5) (2011), 5231-5239.

[11] Modrak, V. and Marton, D.: Configuration complexity assessment of convergent supply chain systems. International Journal of General Systems **43**(5) (2014), 508-520.

[12] Ortega, N., Celaya, A., Plaza, S., Lamikiz, A., Pombo, I. and Sanchez, J. A.: Implementation of simulation software for better understanding of manufacturing processes. In: *New Frontiers in Materials Processing Training and Learning II* **692** (2011), 58-64.

[13] Suchánek, P., Bucki, R., Marecki, F. and Litavcová, E.: Information modelling of the manufacturing centre with the use of the heuristic algorithms. *International Journal of Mathematics and Computers in Simulation* **8**(1) (2014), 60-66.

[14] Varthanan, P. A., Murugan, N. K. and Mohan, G.: A simulation based heuristic discrete particle swarm algorithm for generating integrated production-distribution plan. *Applied Soft Computing* **12**(9) (2012), 3034-3050.

[15] Wang, K., Zhang, C., Su, J., Wang, B. and Hung, Y.: Optimisation of composite manufacturing processes with computer experiments and Kriging methods. *International Journal of Computer Integrated Manufacturing* **26**(3) (2013), 216-226.

[16] Xu, X. J., Xing, H. F. and Wang, S. G.: Research on stochastic manufacturing unit performance analysis based on simulation of Petri net. In: *Proceedings of the International Conference on Mechatronics Engineering and Computing Technology (ICMECT)*. Trans Tech Publications Ltd, Shanghai, 2014, 6414-6418.

Modelling Legal Merger Time to Completion with Burr Regression

Aleksander Buczek[1]

Abstract. In literature various parametrizations of Burr type III and XII distributions have been used to fit legal merger time to completion data. This approach was based on the assumption that the entire legal merger process is always executed in the same way and has not changed over time. In the following paper this assumption will be relaxed by considering procedure specific factors (its derogations and changes) while fitting statistical distributions to legal merger duration data.
In the statistical area of survival analysis this can be achieved with the use of accelerated failure time model in which explanatory variables have a direct impact on the scale parameter of the considered distribution (via exponential link function). Its special case known as Burr regression will be estimated for Polish merger market on the dataset derived from the obligatory announcements posted in *The Journal of the Ministry of Justice* between 1st January 2002 and 31st December 2013.

Keywords: Legal merger procedure; Duration analysis; Polish merger market; Distribution fitting; Burr distribution; Burr regression; Accelerated failure time model.

JEL Classification: C13, C41, G34, K20
AMS Classification: 62E15, 62J12, 62N02

1 Introduction

The legal process in a merger involves fulfilling all the regulatory requirements, in order to ensure an effective merger from the legal point of view. The companies taking part in a merger may originate from the same or different countries. In the former case, such a transaction is known as a domestic merger, in the latter one, as a cross-border merger.

The Code of Commercial Partnerships and Companies (CCPC, pol. *Kodeks Spółek Handlowych*) [1] in the **articles 491-527** defines a legal process in Poland for both types of merger transactions mentioned in the previous paragraph. According to the CCPC, legal merger consolidation may be executed as a **merger by takeover** or **merger by formation of a new company**. The former approach should be understood as transfer of all assets of a company or partnership (the target one) to another company (the bidding one) in exchange for the shares that the bidding company issues to the shareholders or partners of the target company or partnership. In turn, the latter one assumes the formation of a company to which the assets of all merging companies or partnerships devolve in exchange for shares of the new company. Both procedures differ in the way the assets are exchanged for the shares but they are still constructed in a similar way and consist of three main stages presented in the order of appearance in the fig. 1.

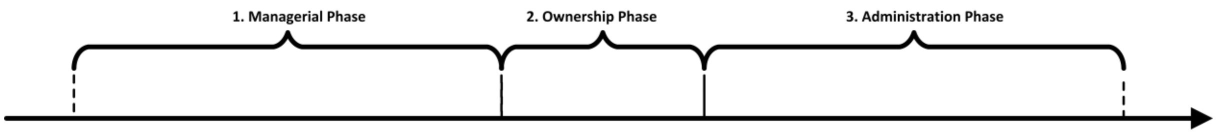


Figure 1 Phases of the legal merger process

Managerial phase should be understood as an introductory phase in which management boards of merging companies negotiate and stipulate the draft terms of merger. This document is then sent to the court register and announced in a public manner. Ownership phase starts and ends once the merger resolution is adopted during the general shareholder meeting. In turn, administration phase consists of merger registration, its public announcement and target company cessation (only in case when the companies merge by takeover). Detailed description of the consecutive process stages can be found in Buczek and Mercik [3] and Buczek [4].

Legal merger time to completion (merger duration) should be defined as the timeframe in which all the stages inherent in the legal process of consolidating a merger are successfully completed. In other words, it is the time needed to complete the three main stages of the legal procedure presented in the fig. 1. Buczek and Mercik [3]

[1] Wroclaw University of Science and Technology, Department of Computer Science and Management, Łukasiewicza 5, 50-370 Wrocław, Poland, e-mail: aleksander.buczek@pwr.edu.pl.

proposed four definitions of the legal merger duration. In this study definition of **public legal merger duration** will be adopted. It states that merger duration is calculated as a difference between the registration announcement and the draft terms of merger announcement.

The article is set up as follows. The next section introduces the legal merger process derogations and amendments. Section 3 outlines the concept of Burr regression while section 4 presents empirical regression model for Polish merger market. Finally, there are some conclusions and suggestions for future research.

2 Legal Merger Process Derogations and Amendments

According to the article 516 of CCPC the following three derogations in the merger consolidation procedure are possible:

- acquiring company general shareholder meeting is not obligatory;
- the reports of administrative or management bodies explaining the rationale of the draft terms of merger, its economic and legal justification are optional;
- the independent opinion about the draft terms of merger raised by the assessor nominated by the appropriate juridical or administrative authority is optional;

if the following prerequisites are met:

- the buyer company is not public and it possesses more than 90%, but not all, of the target company shares (CCPC Article 516 § 1);
- the buyer company is not public and is the only shareholder in the target company (CCPC Article 516 § 6).

What is more, if both merging companies are limited liability companies which are owned by at most 10 individual shareholders (CCPC Article 516 § 7), then the draft terms of merger announcement and notification obligations maybe skipped. Independent opinion raised by the assessor is not required as well in this case.

Across the years legal merger process has been amended several times. Very often these changes were driven by the European Union legislation. The CCPC has been enacted on 15th September 2000 and defined legal merger framework in Poland. On 12th December 2003 (effective on 26th December 2003) the first set of clarification amendments was introduced. The second amendment was announced on 25th April 2008 (effective on 28th May 2008) and brought the following significant changes:

- regulation of cross-border mergers has been introduced as a transposition of EU directive 2005/56/EC;
- the period between the announcement of the draft terms of a merger and the general shareholder meeting was shortened from 6 to 4 weeks;
- the period between the first notification of a merger and the general shareholder meeting was shortened from 6 to 4 weeks.

Not long after, the next amendment was enacted on 5th December 2008 (effective on 4th January 2009) and allowed that a written opinion on the merger by an assessor was not needed anymore if all the shareholders of the companies involved in the transaction have given their consent. The last amendment was introduced on 19th August 2011 (effective on 18th September 2011) and provided additional simplifications to the legal merger process as:

- announcement of the draft terms of a merger in the *Journal of the Ministry of Justice* became optional if this document was continuously available on the website of the merging company for at least one month before the general shareholder meeting;
- written opinion on the merger by the management board does not have to be prepared if all the shareholders of the companies involved in the transaction have given their consent;
- the management board is not obliged to publish relevant changes in a company's assets or liabilities noted after stipulating the draft terms of the merger, but before the shareholder meeting, under the same condition as in the previous point;
- the physical presence of the documents relevant to the merger at the companies' headquarters is not required anymore, if the merging company gives continuous access to these documents on its Internet page for at least one month before the general shareholder meeting.

3 Burr Regression

In 1942 Irving Burr published *„Cumulative Frequency Functions”* article [5] in which he presented the new system of twelve continuous distributions that are usually referred by the number. Among them, the most popular one is type XII distribution often called simply as the Burr distribution. Alternatively, it is also known under the name of Singh-Maddala distribution as it was rediscovered in 1976 by Singh and Maddala in the article „*A Function for Size Distributions of Incomes*” [9].

Its probability density function of the lifetime variable $T > 0$ contains two shape parameters $\varphi > 0$ and $\gamma > 0$ and is defined in the following way:

$$f(t;\varphi;\gamma) = \varphi\gamma t^{\gamma-1}\left(1+t^{\gamma}\right)^{-(\varphi+1)} \tag{1}$$

Three parameter Burr distribution

Probability density function from equation (1) can be extended easily with an additional scale parameter θ. The following parameterization is achieved [2, 6]:

$$f(t;\varphi;\gamma;\theta) = \frac{\varphi\gamma}{\theta^{\gamma}} t^{\gamma-1}\left(1+\left(\frac{t}{\theta}\right)^{\gamma}\right)^{-(\varphi+1)} \tag{2}$$

The associated survival function is then:

$$S(t;\varphi;\gamma;\theta) = \left(1+\left(\frac{t}{\theta}\right)^{\gamma}\right)^{-\varphi} \tag{3}$$

Three parameter log-Burr distribution

Let's assume that $Y=log(T)$, $\gamma = 1/\sigma$ and $\theta = exp(\mu)$. In this case the density function of Y can be defined as:

$$f(y;\varphi;\sigma;\mu) = \frac{\varphi}{\sigma}\exp\left(\frac{y-\mu}{\sigma}\right)\left[1+\exp\left(\frac{y-\mu}{\sigma}\right)\right]^{-(\varphi+1)} \tag{4}$$

where $\varphi > 0$, $\sigma > 0$ and $-\infty < \mu < \infty$. The distribution (4) is known as the log-Burr XII distribution [6]. The respective survival function is:

$$S(y;\varphi;\sigma;\mu) = \left(1+\exp\left(\frac{y-\mu}{\sigma}\right)\right)^{-\varphi} \tag{5}$$

Regression model

Parametrization used in equation (4) allows expressing random variable Y in terms of log-linear model:

$$Y = \mu + \sigma Z \tag{6}$$

where the random variable Z has the density function equal to:

$$f(z;\varphi) = \varphi\exp(z)\left[1+\exp(z)\right]^{-(\varphi+1)},\ -\infty < z < \infty \tag{7}$$

Log-Burr distribution conversion to the log-linear model allows scale parameter μ to vary with explanatory variables vector $x_i^T = (x_{i1},\dots, x_{in})$:

$$y_i = \mu_0 + \beta x_i^T + \sigma z_i \tag{8}$$

Thereby the log-Burr XII distribution given by (4) is extended with additional regression model parameters as now its scale parameter μ is equal to $\mu_0 + \beta x_i^T$. The same applies to the Burr XII distribution as its scale parameter θ is defined as $exp(\mu) = exp(\mu_0 + \beta x_i^T) = \theta_0 exp(\beta x_i^T)$.

The model specified in equation (8) is known as Burr XII regression and was initially proposed by Beirlant et al. [2] in 1998 and applied to portfolio segmentation for fire insurance. In 2012 Hashimoto et al. [6] used the same model for grouped survival data, but they named it as the log-Burr XII regression. The former authors also proposed an alternative model which assumes that scale parameter θ is constant and shape parameter γ is exponentially linked with explanatory variable vector x_i^T: $\gamma = \gamma_0 exp(\beta x_i^T)$.

Moreover, the possibility to express (log)-Burr regression in the form of the log-linear model in equation (8) causes that it should be treated as a special case of the accelerated time failure model which assumes that variable z_i follows extreme value distribution. Please refer to Kalbfleisch and Prentice [7] for detailed explanation.

4 Empirical Model for Polish Merger Market

Buczek and Mercik [3] decided to fit various parametrizations of Burr type III and XII distributions to legal merger time to completion data (with and without additional location parameter). They assumed that entire legal merger process is always executed in the same way and has not changed over time. The following section will be an attempt to relax this assumption and extend Burr distribution given by (2) and (3) with exogenous variables related to the legal merger procedure derogations and amendments.

Empirical Burr regression model will be estimated for Polish merger market. The required data was extracted from *The Journal of the Ministry of Justice*. This journal is the best available source of information about mergers which have occurred in Poland, as in the majority of cases it contains all the announcements required by the legal procedure of merger.

Constructed dataset contains **3570 completed domestic merger transactions** which took place between **1st January 2002 and 31st December 2013**. The following transaction specific information was collected: stipulation of the draft terms of a merger and announcement dates, date of registering a merger, as well as its announcement, number of companies taking part in the transaction. Additionally, for each transaction the following company specific information was extracted: KRS identification numbers[2], company names, cessation dates if applicable, legal form, initial capital and ownership structure. The dataset was enhanced with the information on whether the merging entities were listed on the Warsaw Stock Exchange (WSE) and NewConnect[3] at the time when the draft terms of merger were announced.

529 merger transactions, in which more than two companies took part, were excluded from further analysis. **320 transactions** were excluded, due to missing data about the initial capital of the target company and its ownership structure. Thus, the author of this article conducted his analysis on a subset consisting of **2721 transactions**. The definition of the **public** transaction time to completion was adopted after Buczek and Mercik [3].

Table 1 contains binary explanatory variables which were initially taken into consideration when estimating the log-linear model given by (8).

Variable	Value 1 condition	Observations
A1	The draft terms of merger are publicly announced after 25th December 2003 (1st amendment is effective).	2448
A2	The draft terms of merger are publicly announced after 27th May 2008 (1st and 2nd amendments are effective).	1468
A3	The draft terms of merger are publicly announced after 3rd January 2009 (1st, 2nd and 3rd amendments are effective).	1248
A4	The draft terms of merger are publicly announced after 17th September 2011 (all amendments are effective).	379
D1	The buyer company is not public and it possesses more than 90%, but not all, of the target company shares (CCPC Article 516 § 1).	100
D6	The buyer company is not public and is the only shareholder in the target company (CCPC Article 516 § 6).	1337
D7	Both merging companies are limited liability companies which are owned by at most 10 individual shareholders (CCPC Article 516 § 7).	41

Table 1 Explanatory binary variables

All the calculations are performed with the use of SAS software and its native procedure SEVERITY. Regression model specified in the equations (6) to (8) will be estimated with the use of Maximum Likelihood Method. Newton-Raphson Optimization with Line Search will be used to determine the regression parameter values for which the likelihood function reaches its maximum if it exists. More information about the parameter estimation and optimization methods can be found in the documentation [8]. Dummy coding of exogenous variables was adopted. Backward elimination method was applied to obtain the final model. The results of model

[2] National Court Register (*pol. Krajowy Rejestr Sądowy – KRS*).
[3] This is an alternative stock exchange in Poland designed for smaller companies with simplified entrance criteria and limited reporting requirements.

estimation are summarized in table 2. Scale parameter is defined as $\theta_0 exp(\beta x_i^T)$ where θ_0 is its base value. Model 0 (without any explanatory variables) is presented for comparative purposes.

Parameter	Model 0		Model 1		Model 2		Model 3	
	Estimate	T-Value	Estimate	T-Value	Estimate	T-Value	Estimate	T-Value
Shape (γ)	5.5491 (0.1955)	28.39***	5.9816 (0.2086)	28.67***	5.9801 (0.2086)	28.68***	5.9822 (0.2088)	28.65***
Shape (φ)	0.4326 (0.0248)	17.43***	0.4225 (0.0236)	17.88***	0.4226 (0.0236)	17.89***	0.4223 (0.0236)	17.88***
Base Scale (θ_0)	91.4858 (1.5699)	58.28***	118.4187 (3.6321)	32.60***	118.3238 (3.6218)	32.67***	118.2557 (3.6179)	32.69***
D1	--	--	-0.1032 (0.0401)	-2.57**	-0.1024 (0.0401)	-2.56**	-0.1031 (0.0401)	-2.57**
D6	--	--	-0.0780 (0.0159)	-4.90***	-0.0771 (0.0158)	-4.88***	-0.0760 (0.0157)	-4.83***
D7	--	--	-0.0304 (0.0686)	-0.44	--	--	--	--
A1	--	--	-0.1084 (0.0285)	-3.81***	-0.1084 (0.0285)	-3.81***	-0.1086 (0.0285)	-3.81***
A2	--	--	-0.1189 (0.0298)	-3.99***	-0.1190 (0.0298)	-3.99***	-0.1407 (0.0176)	-8.01***
A3	--	--	-0.0271 (0.0301)	-0.90	-0.0270 (0.0301)	-0.90	--	--
A4	--	--	-0.2049 (0.0240)	-8.53***	-0.2047 (0.0240)	-8.53***	-0.2099 (0.0233)	-9.02***

Table 2 Estimation results

Initially the regression model with 7 binary explanatory variables was estimated (model 1). Variable D7 related to the derogations of the CCPC article 516 § 7 appeared to be insignificant and was removed. This outcome is the result of the data collection methodology as the sample does not contain transactions for which the draft terms of merger announcement is not obligatory (thereby not present in *The Journal of the Ministry of Justice*). The sample contains only transactions for which this derogation was available, but at the end not executed.

After model re-estimation (model 2), all explanatory variables were significant except variable A3 representing the third process amendment from 3rd January 2009. Optionality of assessor's written opinion is not influencing legal merger time to completion. After removal of this variable, the final model was obtained (model 3).

All the parameter estimates are now significant and have negative sign meaning that they in general shorten the expected legal merger duration. The strongest effects are associated with the law amendments which shortened notification periods and introduced electronic circulation of the merger documents – variable A2 and A4 respectively. Derogations resulting from the CCPC articles 516 § 1 and 516 § 6 have similar influence on the time to completion. This result is concordant with the expectations as both derogations provide exactly the same procedure simplifications under different conditions. The cumulative effect of these explanatory variables on Burr distribution survival function is visible on the below fig. 2.

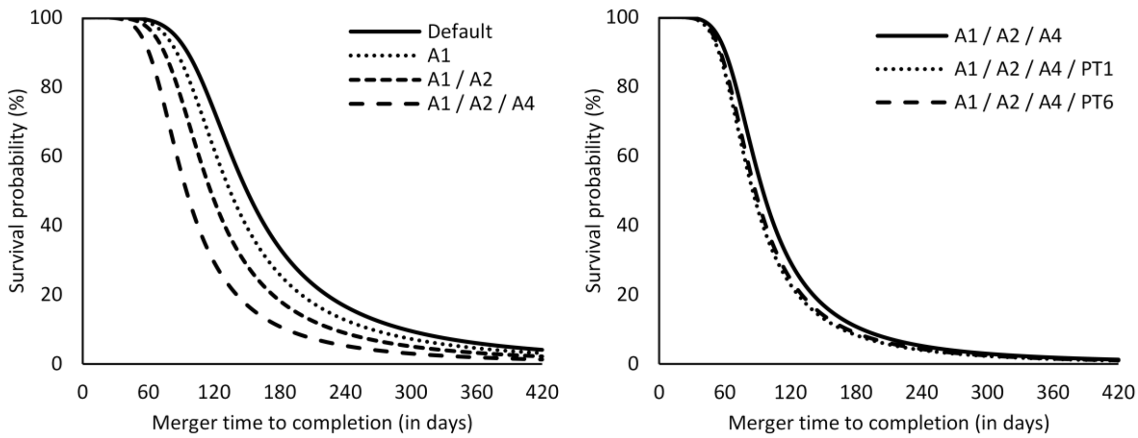


Figure 2 Survival distribution functions

Table 3 presents the goodness of fit statistics. All the statistics except Kolmogorov-Smirnov are the lowest for the final model number 3. They confirm that the introduction of exogenous variables to Burr distribution improved the quality of the fit.

Fit Statistic	Model 0	Model 1	Model 2	Model 3
-2 Log Likelihood	32658	29672	29672	29672
Akaike's Information Criterion	32666	29691	29689	29688
Schwarz's Bayesian Information Criterion	32690	29750	29742	29735
Kolmogorov-Smirnov Statistic	1.1722	1.2664	1.2638	1.2669
Anderson-Darling Statistic	28.8947	10.9735	10.9803	10.9652
Cramer-von Mises Statistic	0.4590	0.3445	0.3463	0.3432

Table 3 Goodness of fit statistics

-2 Log Likelihood statistics allows calculating likelihood-ratio test. It is used to compare the goodness of fit of two models, one of which (the null model) is a special case of the other (the alternative model). Assuming that model 0 is the null model and the model 3 is the alternative model, test value equal to 2986 is obtained. The probability distribution of this test statistic is approximately a chi-squared distribution with 5 degrees of freedom. Very high test statistics allows rejecting at 1% significance level the null hypothesis stating that the null model fits the data better than the alternative one.

5 Conclusion

For Polish merger market it appears that the time to completion of the legal merger process depends on its derogations and law changes. Shorter notification periods and electronic circulation of the merger documents have caused that the entire process is now completed in less days. The same applies to the derogations from the CCPC article 516 in the situation when the acquiring company possess 90% or more of the target company shares.

In the estimated regression model, these exogenous variables have direct impact on the scale parameter of Burr distribution (via exponential link function). Such parametrization has increased the quality of fit compared to the default model estimated without any explanatory variables. This fact supports the decision to treat the legal merger as a not homogenous process.

Some interesting extensions to this paper can be considered for future research. Estimated Burr regression model can be extended with additional explanatory variables related to the transaction and company specific characteristics. National economic indicators can be used here as well. Moreover, different parametrization of the regression model is possible. Please refer to Beirlant et al. [2] for the model specification in which exogenous variables are exponentially linked with the shape parameter γ.

Lastly, the approach presented is flexible and may be adapted to different legal procedures. The division of companies into separate entities is an obvious choice here – such a process is quite similar to legal merger consolidation, as it is initiated by a plan to divide and ends with the registration of the newly separated companies.

References

[1] Act of 15th of September 2000. The Code of Commercial Partnerships and Companies [in Polish].

[2] Beirlant, J., Goegebeur, Y., Verlaak, R., Vynckier, P.: Burr regression and portfolio segmentation. *Insurance: Mathematics and Economics* **23**, 3 (1998), 231-250.

[3] Buczek, A., Mercik, J.: On conformance of legal merger duration with Burr type III and XII distributions. *The Wroclaw School of Banking Research Journal* **15**, 5 (2015), 597–608.

[4] Buczek, A.: The time to completion of a legal merger: General concepts, statistical analysis and the case of Poland. *Operations Research and Decisions* **26**, 1 (2016), 19-44.

[5] Burr, I.W.: Cumulative frequency functions. *The Annals of Mathematical Statistics* **13**, 2 (1942), 215–232.

[6] Hashimoto, E.M., Ortega, E.M.M, Cordeiro, G.M., Barreto, M.L.: The Log-Burr XII Regression Model for Grouped Survival Data. *Journal of Biopharmaceutical Statistics* **22**, 1 (2012), 141-159.

[7] Kalbfleisch, J.D., Prentice, R.L.: *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, Hoboken, New Jersey, 2002.

[8] SAS Institute Inc.: *SAS/ETS 9.3 User's Guide*. SAS Institute Inc., Cary, 2011.

[9] Singh, S.K., Maddala, G.S.: A Function for Size Distribution of Incomes. *Econometrica* **44**, 5 (1976), 963-970.

Fiscal Policy for the 21st Century: Does Barack Obama Effect the Real Economic Policy?

Jakub Buček[1]

Abstract. In this paper we investigate whether behavior of politicians on social network can effect the real economic policy, namely the economic growth. We investigate Twitter account of the US president Barack Obama (@barackobama) from March 2012 to January 2016 using techniques of text mining. Time series of ratio of positive to all tweets, and ratio of negative to all tweets is constructed and afterwards these are compared with the growth rate of output represented by the growth rate of industrial production index. We found out via Granger causality analysis, that both, ratio of negative to all tweets and ratio of positive to all tweets, has predictive information about the growth rate of output. Lastly, we estimated several linear models that showed the effect of both variables on economic growth.

Keywords: Twitter, Barack Obama, Social media, Text mining, Economic growth.

JEL classification: O49
AMS classification: 91B62

1 Introduction

On the edge of the 21st century, developed countries have to face a new economical situation. The interest rates are on the lowest level possible which puts limits on monetary policy. In this situation, more researchers shift their interest to fiscal policy. Besides the new economical situation, a modern civilization has to deal with information revolution that brought not only new opportunities but also new threats to makers of fiscal policy.

With the arrival of social networks, politicians are watched every hour and every day. They have got a tool to communicate directly to ordinary citizens in any daytime without need of a mediator. This is new for both, politicians and citizens, none had this experience before. The power of social networks gives more responsibility to politicians, because even good fiscal policy can be ruined with reckles statements.

Twitter is microblogging service where users post messages (called *tweets*) with maximum limit of 140 characters, and on average 11 words per messages. O'Connor et al. [7] suggest that Twitter is suitable for public mood research because there are a very large number of messages, which majority is publicly available, and obtaining them is technically easier than searching through web. Social networks as indicator of public mood were investigated mainly in the context of financial market. Bollen et al. [3] used information gained from Twitter to find out a mood of a society in a certain time point and showed that it can significantly improve prediction power of stock market model. To the best of author's knowledge, no article that would connect social networks to fiscal policy was published.

In this paper we investigate whether behavior of politicians on social network can effect the real economic policy, namely the economic growth. We analyse tweets posted by Barack Obama's acount @barackobama which is, with more than 74 million followers in April 2016, one of the most followed acount on Twitter worldwide. The causality between economic growth and Barack Obama's tweets is investigated using Granger causality analysis.

Finally, we estimate the linear model similar to the one presented by Bergh and Karahalios [1]. In their specification, they use growth rate of GDP per capita in PPP as dependent variable and it is explained by following regressors: the growth rate of labor force, the investment rate as a share of GDP, and a government sector, measured as taxes, government revenue and government expenditure, all in relation to GDP. They also include additional control variables, specifically the unemployment rate, the annual inflation rate, and economic openness as sum of imports and exports over GDP. The estimation of the model is made on the annual data that covers 21 OECD-countries over the period 1970 to 2001.

Unfortunately, because of the character of the studied problem, we have to work with time series with higher frequency and some variables, most importantly GDP, are measured only once per a quarter. Therefore, we have to work with substitute time series with monthly frequency.

[1]Masaryk University, Department of Economics, Lipová 507/41a, Pisárky, Brno, jakubbucek@mail.muni.cz

2 Data and methods overview

We obtained a collection of public tweets from Barack Obama's account that starts on March 1, 2012 and ends on January 31, 2016 (11071 tweets)[2]. Furthermore, we collected the following macroeconomic series[3]: the growth rate of industrial production index [%], inflation as percentage change in consumer price index [%], the growth rate of import and export [%], the growth rate of labor force [%], and change of unemployment [%] from FRED [2], [8], [9], [10], [11], and [12]. The monthly data about tax revenue [%] and government spending [%] of the US government were obtained from Bureau of the Fiscal Service [4].

Afterwards, we try to investigate whether a tweet is positive, negative or neutral. Step by step we remove stop-words and punctuation, convert all letters to lower case, split the tweet into words and compare these words with lexicon of positive and negative opinion words. We use the lexicon[4] provided by Hu and Liu [6]. The lexicon contains over 2000 positive opinion word such as *progress* or *perfect*, and over 4500 negative opinion word, e.g. *tax* or *catastrophe*. We also add some extra positive and negative opinion word related to fiscal policy, e.g. *growth* or *invest* as positive opinion words, and *debt* or *loan* as negative opinion words.

Tweet	Date	Score
POTUS[5] on the auto industry *recovery*: "These jobs are *worth* more than just a paycheck. Theyre a source of *pride*."	Feb 28, 2012	3
FACT: Romneys corporate *tax* plan could *displace* U.S. jobs and create 800K jobs *overseas* by eliminating *taxes* on companies foreign profits.	Jul 17, 2012	-4
"We know that the *cost* of these events can be measured in *lost* lives and *lost* livelihoods, *lost* homes and *lost* businesses." -President Obama	Jun 25, 2013	-5
"*Health care* reform in this state was a *success*. That doesn't mean it was *perfect right* away." President Obama on *health* reform in Mass.	Oct 30, 2013	6

Table 1 Example of negative and positive opinion tweets. Identified lexicon words are emphasized.

For each tweet we determine whether it contains any number of positive or negative terms from the lexicon. We count a number of occurance of positive terms and subtract a number of occurance of negative terms and call it a score. We consider a tweet to be negative, when a score of the tweet is lesser than 0, and positive, when the score is greater than 0. If the score is equal to 0, then the tweet is said to be neutral. Example of the tweet scoring is in Table 1.

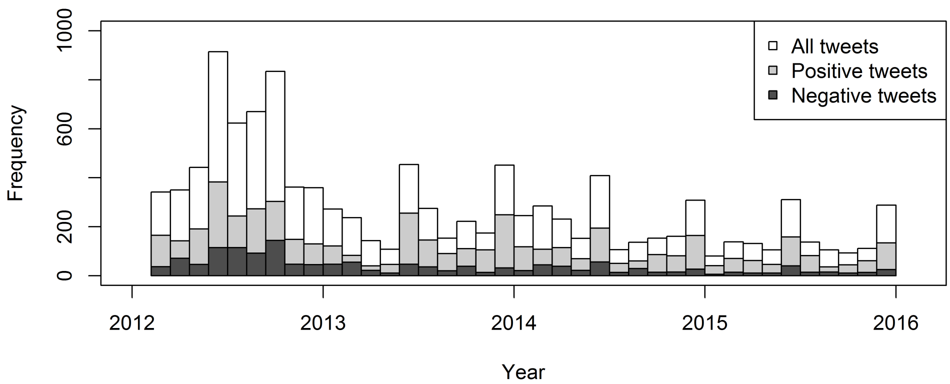


Figure 1 Distribution of tweets over the observed period. Bars do not represent cummulative frequency but the actual level.

A frequency and distribution of positive and negative tweets among all tweets is on Figure 1. From the figure, it is clear that positive tweets outnumber negative tweets in any given time during the observed period. The peak in

[2]The official API allows to download only the latest 3200 public tweets. These 3200 tweets cover only period starts in April 2014. We downloaded the data from website `http://greptweet.com/`, which cover much longer period.

[3]All variables are according to KPSS stationary on significant level 5 %.

[4]The lexicon can be downloaded from website `https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#lexicon`

[5]POTUS stands for President of the United States.

total number of tweets during the second half of 2012 is caused by presidental election that took place in November 6, 2012.

All tweets are aggregated into bins according to the month they were posted. In each bin, a ratio of positive to all tweets, and negative to all tweets are created. We get two time series, ratio of positive and negative tweets to all tweets, with monthly frequency. We compare these time series with a growth rate of industrial production index which substitute a growth rate of GDP that is measured only once per quarter. On Figure 2 are graphs of these time series. All variables are standardized to provide a common scale for comparisons of all time series.

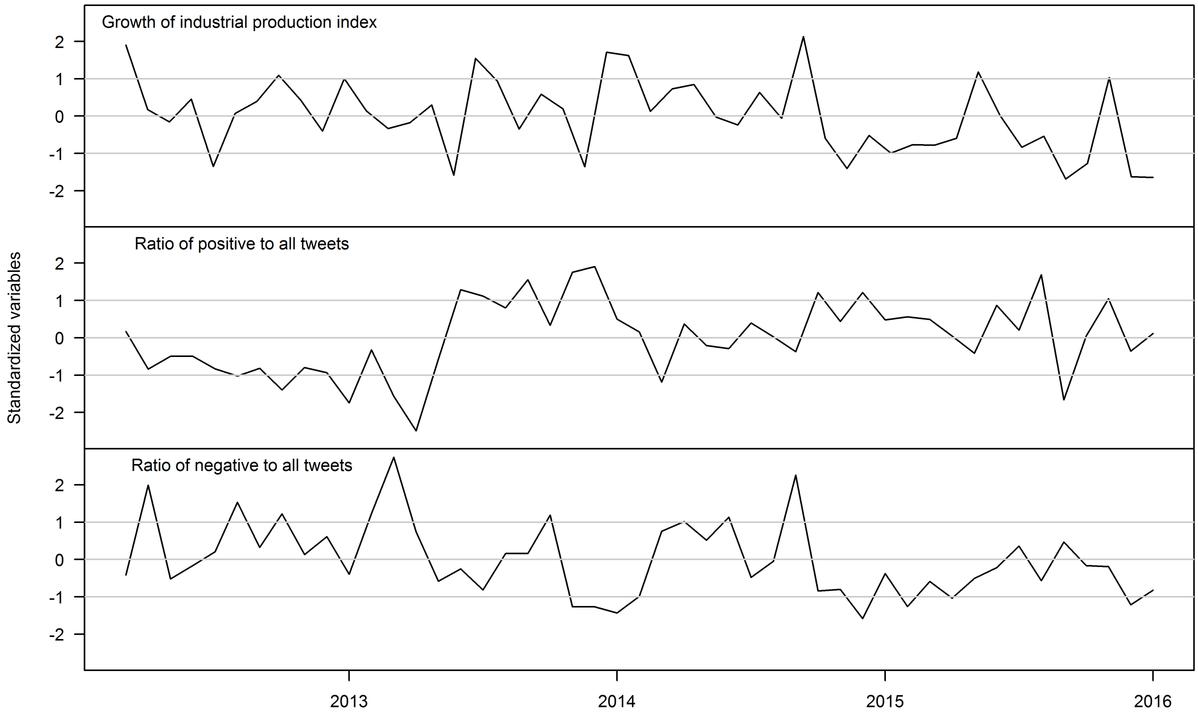


Figure 2 Time series of positive and negative tweets, and growth rate of industrial production index.

3 Granger causality

We are concerned with the question whether the negative and positive tweets correlate with changes in growth rate of industrial production index. To answer this question, we perform the Granger causality analysis according to Gilbert and Karahalios [5]. Although there is a "causality" in the name of this technique, we are not testing true causality but rather whether one time series has predictive information about the other or not.

We present two models, M_1 and M_2, where model M_1 is "nested" within model M_2.

$$M_1 : y_t = \alpha + \sum_{i=1}^{n} \beta_i y_{t-i} + \sum_{i=1}^{n} \gamma_i x_{t-i} + \epsilon_t \tag{1}$$

$$M_2 : y_t = \alpha + \sum_{i=1}^{n} \beta_i y_{t-i} + \sum_{i=1}^{n} \gamma_i x_{t-i} + \sum_{i=1}^{n} \gamma_i pos_{t-i} + \sum_{i=1}^{n} \delta_i neg_{t-i} + \epsilon_t, \tag{2}$$

where y_t denotes growth rate of industrial production index, x_t is a vector of regressors that contains: inflation, the growth rate of import and export, the growth rate of labor force, change of unemployment, the growth rate of tax revenue, and the growth rate of government spending. Explanatory variable pos_t is ratio of positive to all tweets, and neg_t is ratio of negative to all tweets. Parameters $\alpha, \beta_i, \gamma_i$ and δ_i are regression coefficients.

We test whether the model M_2 performs significantly better than the model M_1. The following hypotheses are tested:

- For all $i : \gamma_i = 0$ against alternative hypothesis that exists at least one $i : \gamma_i \neq 0$.
- For all $i : \delta_i = 0$ against alternative hypothesis that exists at least one $i : \delta_i \neq 0$.
- For all $i : \gamma_i = 0$ and $\delta_i = 0$ against alternative hypothesis that exists at least one $i : \gamma_i \neq 0$ or $\delta_i \neq 0$.

Results of these hypotheses for $n = 1, 2, 3$ are in Table 2. On significant level 5 %, the null hypothesis can be rejected in all three variants for all n. This finding implies that both of the variables of interest has predictive information about growth rate of industrial production index.

n	Only positive ($\gamma_i = 0$)	Only negative ($\delta_i = 0$)	Both ($\gamma_i = \delta_i = 0$)
1	$F_{1,35} = 7.42326$	$F_{1,35} = 4.68894$	$F_{2,35} = 3.92601$
	(0.0099811)	**(0.0372574)**	**(0.0289531)**
2	$F_{2,24} = 6.25666$	$F_{2,24} = 4.0714$	$F_{4,24} = 3.47281$
	(0.00650309)	**(0.0300284)**	**(0.0225117)**
3	$F_{3,13} = 6.867$	$F_{3,13} = 6.40596$	$F_{6,13} = 8.9846$
	(0.00516613)	**(0.00671488)**	**(0.000523181)**

Table 2 Granger causality analysis for $n = 1, 2, 3, 4$ and three types of hypoteses (p-values in brackets).

4 Linear model

The last step in our analysis is estimation of appropriate linear model of the economic growth. Our main interest is to investigate the effect of "Twitter" variables on the economic growth. We take in account effects of all variables presented in previous section.

Before the estimation, we compute correlation coefficients between the ratio of negative to all tweets and positive to all tweets, and others regressors so we can reject the hypothesis that the constructed variables are just subtitution for another important factor of the economic growth. The highest, and the only statistical significant, correlation coefficient for both variables of interest are the one between them (value -0,5501).

To check robustness of the results, we try several specifications of the model. Selected models are available in Table 3. Also, the model with total number of tweets instead of the two variables with only positive and negative tweets was taken in account. In this specification, the total number of tweets was not significant predictor.

From the results, we get a picture of the effects of regressors on the growth rate of industrial production index. The most significant variable is lagged government spending with positive effect on the dependent variable. The effect of tax revenue is negative but only for 2nd lag of the variable. The change of unemployment has negative effect on growth rate of industrial production index, while inflation and the growth rate of labor force is not significant on significant level 5 % in any specification. Strangely enough, the effect of the growth rate of export is positive in the same time point, but negative for the 1st lag. The effect of the growth rate of import is positive for the 1st lag. The lagged dependent variable is not significant on significant level 5 %. Finaly, the effect of the ratio of negative to all tweets on dependent variable is negative in time t and positive for the 1st lag. The ratio of positive to all tweets is not significant in time t, but has positive sign for the 1st lag and negative sign for the 2nd lag.

5 Conclusion

Although the presented result could imply that Barack Obama's behaviour on Twitter can effect the economic growth, we should adopt a cautious approach to this finding. During the analysis, we have to face several limitations. The first limitation is the fact that we only analyse tweets by Barack Obama and do not include his political opponents. Extension of analysed Twitter profiles could bring new information to the studied problem. Also, the classification of tweets as positive and negative is kind of arbitrary. The used approach can not identify whether the particular word is used in positive or negative context (e.g. "growth of unemployment" is considered as positive, even though it is in fact negative). The second limitation is availability of macroeconomics time series with suitable frequencies. In any case, this paper shows that there is some kind of relationship between economic growth and Twitter activity of Barack Obama. Any further explanation of this relationship is still necessary to be found.

Acknowledgements

This work was supported by funding of specific research at Faculty of Economics and Administration, project MUNI/A/1055/2015. This support is gratefully acknowledged.

Variable	Lag	Model 1	Model 2	Model 3
Export	0	0.395274 (**0.0154**)	0.467055 (**0.0044**)	0.490538 (**0.0035**)
	1	-0.327240 (**0.0063**)	-0.275352 (**0.0350**)	-0.410480 (**0.0022**)
	2		0.125605 (0.3066)	
Import	0	-0.102689 (0.3435)	-0.116599 (0.2652)	-0.132362 (0.1826)
	1	0.322633 (**0.0003**)	0.226793 (**0.0414**)	0.372151 (**0.0005**)
	2		0.0504183 (0.6297)	
Industry	1			0.128333 (0.2725)
	2			0.0976434 (0.4504)
Inflation	0	-0.583314 (0.0993)	-0.647656 (0.0540)	-0.667652 (0.0860)
Labor force	0	-0.344977 (0.1111)	-0.327675 (0.0665)	-0.230853 (0.3047)
Negative	0	-0.0338648 (**0.0048**)	-0.0488813 (**0.0004**)	-0.0566466 (**0.0011**)
	1	0.0447708 (**0.0366**)	0.0417752 (0.0610)	0.465798 (**0.0309**)
	2		-0.0180385 (0.3007)	-0.0235911 (0.1665)
Positive	0	-0.00553685 (0.5164)	-0.00773950 (0.2847)	-0.00660318 (0.4257)
	1	0.0105553 (**0.0478**)	0.0142311 (**0.0399**)	0.0126425 (**0.0182**)
	2		-0.0142408 (**0.0132**)	-0.0188846 (**0.0009**)
Spending	0		-0.00289649 (0.4676)	
	1	0.0157283 (<**0.001**)	0.0136061 (**0.0152**)	0.0183624 (<**0.001**)
	2	0.00880319 (**0.0478**)	0.00894177 (<**0.001**)	0.0110757 (**0.0336**)
Tax revenue	0		0.00309766 (0.6274)	
	1	-0.000818883 (0.8929)	0.00170134 (0.6952)	-0.00205625 (0.7155)
	2	-0.0119819 (0.0505)	-0.0124321 (**0.0019**)	-0.0144938 (**0.0323**)
Unemployment	0	-0.945134 (**0.0348**)	-1.08003 (**0.0021**)	-1.31261 (**0.0066**)
Constant	0	-0.162053 (0.8073)	0.918273 (0.2296)	1.20002 (0.1681)
R^2		0.570365	0.666668	0.625986
adj R^2		0.348140	0.362321	0.341735

Table 3 Linear models of relationship between economic growth and its factors. Dependent variable is the growth rate of industrial production index. Standard errors robust to autocorrelation and heteroskedasticity were used. Corresponding p-values in brackets.

References

[1] Bergh, A., and Öhrn, N.: Growth Effects of Fiscal Policies: A Critical Appraisal of Colombiers (2009) Study. *IFN Working Paper* **864** (2011)

[2] Board of Governors of the Federal Reserve System (US), *Industrial Production Index* [INDPRO], retrieved from FRED, Federal Reserve Bank of St. Louis https://research.stlouisfed.org/fred2/series/INDPRO, April 30, 2016.

[3] Bollen, J., Mao, H., and Zeng, X.: Twitter mood predicts the stock market. *Journal of Computational Science* **2** (2011), 1–8.

[4] Bureau of the Fiscal Service, *Monthly Treasury Statement (MTS)*, https://www.fiscal.treasury.gov/fsreports/rpt/mthTreasStmt/backissues.htm , April 30, 2016.

[5] Gilbert, E., and Karahalios, K.: Widespread worry and the stock market. In: *Fourth International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence, Washington, D.C., 2010, 58–65.

[6] Hu, M., and Liu, B.: Mining and summarizing customer reviews. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. The Association for Computing Machinery, Seattle, 2004, 168–177.

[7] O'Connor, B., Balasubramanyan, R., Routledge, B.R., and Smith, N.A. From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. In: *Fourth International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence, Washington, D.C., 2010, 122–129.

[8] US. Bureau of Labor Statistics, *Civilian Labor Force* [CLF16OV], retrieved from FRED, Federal Reserve

Bank of St. Louis https://research.stlouisfed.org/fred2/series/CLF16OV, April 30, 2016.

[9] US. Bureau of Labor Statistics, *Civilian Unemployment Rate* [UNRATE], retrieved from FRED, Federal Reserve Bank of St. Louis https://research.stlouisfed.org/fred2/series/UNRATE, April 30, 2016.

[10] US. Bureau of Labor Statistics, *Consumer Price Index for All Urban Consumers: All Items* [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis https://research.stlouisfed.org/fred2/series/CPIAUCSL, April 30, 2016.

[11] US. Bureau of Labor Statistics, *Export (End Use): All commodities* [IQ], retrieved from FRED, Federal Reserve Bank of St. Louis https://research.stlouisfed.org/fred2/series/IQ, April 30, 2016.

[12] US. Bureau of Labor Statistics, *Import (End Use): All commodities* [IR], retrieved from FRED, Federal Reserve Bank of St. Louis https://research.stlouisfed.org/fred2/series/IR, April 30, 2016.

Specifying the use of Taguchi's loss function in manufacturing and service sectors

Pavol Budaj[1], Miroslav Hrnčiar[2]

Abstract. An important aspect of quality management is the consideration of quality cost importance. A turning point when considering the quality costs was the introduction of the Taguchi loss function in perception disagreements. This paper focuses on a comparison of the approaches to the application of the Taguchi approach in manufacturing and services and highlights some differences that need to be respected to assess the full potential of the Taguchi method. The focus is not only on the parameters to which the Taguchi loss function is applied, but also measures to reduce possible deviations. The study shows how the Taguchi loss function can be used when determining the basic success factors for the selection of control elements and their tolerances for manufacturing as well as the service sector. The examples highlight the need not only for different approaches in the methods used, but also the selection of measures to reduce quality costs. In the service sector, a Service Level Agreement approach is used, which takes into account a service guarantee.

Keywords: Tachugi's loss function, service sector, quality costs, service level agreement.

JEL Classification: L91, M11
AMS Classification: 90C15

1 Interaction between the manufacturing and service sectors in the development of quality management

There has always been a large traffic in ideas on quality management between manufacturing and services. Initially, quality management grew out of statistical techniques intended to improve the management of the manufacturing of physical products. As comprehensive systems of quality management developed, attention came to focus increasingly on customers and their requirements. This was an area where services repaid study, because they benefit from closer contact with the customer than there is in manufacturing.

Systematic and institutional approaches to quality management do not currently distinguish between product types; the generic approaches under ISO 9001 are intended for both products and services. An analysis of the influences that manufacturing and services have had on each other's approaches to quality management is able to identify a number of vectors of knowledge transfer:

- Services have influenced quality management in the manufacturing sector
  - through the ability to individually tailor production, i.e. to satisfy a customer's specific requirements, which is now one of the competitive advantages of successful manufacturing enterprises;
  - through the flexibility that has always been a condition for success in the service sector – services had to find ways to manage resources and processes that were resilient to fluctuations in demand. Capacity planning approaches (queueing theory) in the service sector have certainly been a source of inspiration for the manufacturing sector;
  - through the provision of added value – the manufacturing sector increasingly looks to "services beyond the product" as a way to create added value for their material products.
- The manufacturing sector has influenced quality management in services
  - through the use of standardization as a way to achieve significant cost savings in the design and development of new products and services. Services are now profiting from standardization in the same way as industry, by using it to reduce variability as a fundamental distinguishing feature of services. Standardization has made the greatest progress in those types of services that are provided collectively and involve a high level of technical support (public transport, postal services, and telecommunications).
  - potentially also through the development of systems for monitoring and measuring performance, which is the basic success criterion for planning, managing and achieving an enterprise's objectives. The intangible nature of services means that the management of performance is generally a characteristic problem of the sector. The establishment and maintenance of systems for evaluating and improving performance

[1] Catholic University of Ružomberok, Faculty of Pedagogy, Department of Management, pavol.budaj@ku.sk.
[2] University of Žilina, Faculty of Management Science and Informatics, miroslav.hrnciar@fri.uniza.sk

through quality management instruments and methods is mainly oriented towards the manufacturing sector. The material character of the products of manufacturing makes it easier to determine key performance indicators, desired target values and a method for measuring and evaluating them.

One of the quality management methods for improving performance which were first developed in the manufacturing sector is the Taguchi loss function.

2 Applications of the Taguchi loss function

Genichi Taguchi revolutionized the traditional concept of quality according to which a product whose parameters fall within set tolerance limits is a good product, of acceptable quality, which will not cause any loss. Taguchi's statistical research led him to the conclusion that actual losses fall on a continuum and can be visualized using a function that generates a loss curve. Losses are lowest when the studied parameter of the product matches precisely to the required value. As the deviation from this ideal value increases, the loss increases exponentially. According to Taguchi's Quality Loss Function, quality loss increases exponentially as the deviation increases between the value of the monitored product characteristic and its target value, which is situated, as a rule, equidistantly from the upper and lower tolerance thresholds and represents the target value for all producers – customer satisfaction.

The modern holistic approach to the management of manufacturing processes requires a systematic analysis of the concept of the "customer". Customers are the most important stakeholder. A holistic approach must also take into account the requirements of other stakeholders (e.g. process owners, organizational units), who have other expectations than just offering or receiving a good-quality product, and whose expectations may be contradictory. The perspective on the quality of products and processes thus needs to be a multi-dimensional perspective as in the "4E" philosophy. The fundamental aim of this approach is to achieve the satisfaction of all customers and other interested parties in four dimensions: effectiveness – the satisfaction of a product's customers with its quality and other parameters; efficiency – reductions in inputs and resource use without adverse impacts on the other dimensions; economy – the avoidance of waste; and ethics – managing all of the above in accordance with moral and ethical standards [1]. Measuring quality using a Taguchi loss function is a way to monitor and promote a multidimensional approach in an enterprise. Its effectiveness is increased when it is implemented during the design and development of products and the pre-production planning process. Another way it helps to increase effectiveness is by basing quality improvements on reductions in variability in performance.

2.1 An example of use of the loss function in manufacturing

The Taguchi loss function has been applied successfully in mass production where there is a low level of variability in final production [7]. Every deviation from the ideal target value (T) results in an economic loss (L), which can be calculated according to the formula in (1).

$$L = k * (X - T)^2 \tag{1}$$

L – loss (€) – incremental loss
k – cost ratio

$$k = A \,.\, \Delta^{-2} \tag{2}$$

where A represents the cost of a change by value Δ
T – the target value of the monitored indicator,
X – the actual value of the monitored indicator.

Research carried out textile mass production focused on quality losses in fifteen products (pullovers, half-button shirts), with evaluation being carried out for all the size groups produced for each product. Quality loss was measured in terms of wastage of textile material or live work:

- higher material consumption within the acceptable "upper tolerance limit";
- losses resulting from repairs carried out when the tolerance limits for dimensions were exceeded;
- material losses or higher material costs incurred as a result of an irrational procedure for the technical preparation of products when modelling the "tuck-under" of sleeves and the bottom edge of products for the finishing of the products in an aesthetically and technically satisfactory manner on a 2-needle sewing machine.

This machine permits products to be finished with a tuck-under of 1.2 cm, 1.8 cm or 2.5 cm. The following loss calculation selects from the large quantity of measurements referred to above the measurements for the product short-sleeved half-button shirt and the product part – sleeve. The reason for this selection was that according to the technical requirements, the two-needle trimming of the sleeve required a 2.5 cm feed apparatus and therefore the length of the cut sleeve needed to be increased by 2.5 cm. A rational approach (that would have no adverse effect on technical or aesthetic quality) would be to use a feed apparatus set to one of the smaller sizes mentioned earlier. The measurements shown are for half-button shirts size "S". For this size, the tolerance range is on the "right-hand side" (size S products cannot be reallocated to a lower size group if the length actually achieved is below the lower tolerance threshold). Losses were studied for 140 half-button shirts [7], i.e. 280 sleeve pieces. The results of measurement are shown in table 1. The lengths measured for sleeve parts are given in column 1.

Product group	Length x_i (cm)	$x_i – T_1$	$(x_i – T_1)^2$	Number n_i (items)	c.3 x c.4	$x_i – T_2$	$(x_i – T_2)^2$	c.7 x c.4
r	1	2	3	4	5	6	7	8
1	23.5	0.7	0.49	0	0	1.3	1.69	0
2	23.6	0.8	0.64	8	5.12	1.4	1.96	15.68
3	23.7	0.9	0.81	14	11.34	1.5	2.25	31.50
4	23.8	1.0	1.00	26	26.00	1.6	2.56	66.56
5	23.9	1.1	1.21	56	67.76	1.7	2.89	161.84
6	24.0	1.2	1.44	66	95.04	1.8	3.24	213.84
7	24.1	1.3	1.69	52	87.88	1.9	3.61	187.72
8	24.2	1.4	1.96	28	54.88	2.0	4.00	112.00
9	24.3	1.5	2.25	18	40.50	2.1	4.41	79.38
10	24.4	1.6	2.56	12	30.72	2.2	4.84	58.08
Total	x	x	X	280	419.24	x	x	926.60

Table 1 Measured values and factors for the Taguchi loss function (n = 280) Source: own production

Interpretation of the results of measurement:
1. In order to minimize the risk of customer complaints, sleeves are cut with a positive tolerance (from 0.1 to 1.0 cm), which gives rise to the first loss through material costs. Since the parts are cut in layers of 20 items, the length of individual cuts is variable due to human factors. The resulting sleeve length is also affected by the operator's skill in assembling the garment.
2. The fact that the tuck-under adds 2.5 cm to the length where 1.8 cm would also be possible causes the loss of an additional 0.7 cm compared to the desired (minimum admissible) length T_1.

The material consumed in both cases is taken into account in column 2 of table 1. The required (minimum admissible) length of the cut sleeve (with a tuck-under of 1.8 cm) is T_1 = 22.8 cm.

Column 6 of table 1 shows the loss of material that would result if the tuck-under for the lower edge of the fabric were set using the option of just 1.2 cm for the feed apparatus. The required (minimum admissible) length of the cut sleeve (with a tuck-under of 1.2 cm) is T_2 = 22.2 cm.
Applying relationship (1) to the specific conditions found by the foregoing research, when the length of the measured parts is in the range $\mu \in (T - \varepsilon; T + \varepsilon)$, where

$$T - u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \le \mu \le T + u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \quad (3)$$

and assuming a normal distribution, the relationship can be transformed as shown below [3]:

$$L = k \sum_{i-1}^{r} (x_i - T)^2 * n_i \quad (4)$$

where r = the number of product classes (groups) and n_i is the number of products in class i
and, $L = \sum_{i=1}^{r} L_i$, i.e. L is the sum of the two causes of the qualitative losses described in (1) and (2).

The value of "k" is calculated from the price of 1 m^2 of fabric. In the given case, k = 0.0077. At this value of "k" costs are incurred as a result of the quality loss connected with increased fabric use. In variant 1 (changing the hem of the at the end of the sleeve from 2.5 cm to 1.8 cm), the value of the loss function L1 = 0.0077 x 419.24 = EUR 3.228. In variant 2 (changing the hem of the end of the sleeve from 2.5 cm to 1.2 cm), the value of the loss function L1 = 0.0077 x 926.60 = EUR 7.135.

These calculations show that with regard to the first variant, the cost incurred due to quality loss in the form of excess material consumption when producing 140 items (using 280 sleeves) is EUR 3.228 (or EUR 0.023 per item). In relation to the second variant, the cost is EUR 7.135 (or EUR 0.051 per item).

This calculation takes into consideration only the sleeves. It is highly likely that the tuck-under at the bottom of the front and rear parts of the product would produce similar losses. These findings regarding quality losses should motivate the organization to adopt effective measures in design and modelling and in the technological arrangements for the manufacture of similar products. Given annual production of around 150,000 items and the application of the proposed procedure to sleeves and the bottom hem of the front and back parts, savings in material costs in the range EUR 6,900 to EUR 15,300 should be realistically achievable.

2.2 An example of use of the loss function in the service sector

The literature also includes many applications of the Taguchi loss function in the service sector [4]. Tolga [8] integrated the relationship between cost and variability using the Taguchi loss function in the performance and parameters of the design of medical applications.

Nevertheless, the available literature includes practically no examples from the area of transport making use of a service level agreement (SLA). Customers evaluate public transport services according to various criteria, few of which are easily expressed by measurable indicators. This is one of the key differences from the manufacturing sector, which results from the intangible nature of services. Criteria that are harder to measure include the politeness of staff, cleanliness and the feeling of safety. The service level agreement between the provider of the service and the customer must, however, be based on clearly measurable indicators.

In the case public transport, such a criterion could be a specified guarantee level associated with an attribute such as fulfilment of the planned timetable, with a penalty function used to determine the impact of deviating from the guarantee level. The role of the service provider is to ensure that the service does not deviate from the timetable. Whether or not public transport runs on time is a major factor in customer satisfaction. Every deviation from planned arrival/departure times can be considered a cost in terms of reduced customer satisfaction and reduced confidence in the organization providing the given service [9]. The service level agreement between the contracting entity (usually a municipality) and the service provider (transport company) can be based on such a criterion, with set tolerance limits for deviation from the timetable. The usual approach is to ask whether services operated on time in the sense of arrival at the set time plus or minus a tolerated amount, or arrival outside the tolerated interval. A guarantee that applies only in terms of tolerance limits does not consider the size of deviation but only the fulfilment or non-fulfilment of the agreement. Applying a Taguchi loss function would be big benefit for customers in this case because it would create much stronger pressure for the transport operator (service provider) to conform to the timetable.

The service level agreement also includes measurement conditions. The research real measurements from 165 arrivals of town bus arrivals with timetabled arrivals HH:MM, with evaluation of 165 arrivals scheduled for HH:10. The results and calculations are shown in table 2.

time of arrival of bus	$x_i - T_1$	fulfilled with SLA 9.0-11.0	$(x_i - T_1)^2$	number of occurrenc-es of value	c.7 x c.5	$x_i - T_2$	fulfilled with SLA 9.5-10.5	$(x_i - T_2)^2$	c.9 x c.5
1	2	3	4	5	6	7	8	9	10
8.0	-1.0	no	1.00	0	0.00	-1.5	no	2.25	0.00
8.5	-0.5	no	0.25	1	0.25	-1.0	no	1.00	1.00
9.0	0.0	yes	0.00	4	0.00	-0.5	no	0.25	1.00
9.5	0.0	yes	0.00	6	0.00	0.0	yes	0.00	0.00
10.0	0.0	yes	0.00	21	0.00	0.0	yes	0.00	0.00
10.5	0.0	yes	0.00	32	0.00	0.0	yes	0.00	0.00
11.0	0.0	yes	0.00	36	0.00	0.5	no	0.25	9.00
11.5	0.5	no	0.25	18	4.50	1.0	no	1.00	18.00
12.0	1.0	no	1.00	20	20.00	1.5	no	2.25	45.00
12.5	1.5	no	2.25	9	20.25	2.0	no	4.00	36.00
13.0	2.0	no	4.00	5	20.00	2.5	no	6.25	31.25
13.5	2.5	no	6.25	6	37.50	3.0	no	9.00	54.00
14.0	3.0	no	9.00	2	18.00	3.5	no	12.25	24.50
14.5	3.5	no	12.25	0	0.00	4.0	no	16.00	0.00
				165	120.50				219.75

Table 2 Measured values and factors for the Taguchi loss function (n = 165) Source: own production

The service level agreement concluded between the service provider and the contracting entity stipulates that the service provider is obliged to return a part of the subsidy for providing the service equal to 0.05 times the percentage of operations that are found to operate outside the tolerance limits for punctuality [5]. The results of control measurements are shown columns 1 and 5. According to the results of measurement (overall, 36.9% of arrivals fell outside the tolerance limits), the amount that the service provider is obliged to repay to the contracting entity is 0.05 x 36.9% = 1.845%.

If the evaluation of deviations from the timetable were based on a Taguchi loss function (columns 2, 4 and 6), the penalties would be much higher. The calculation would be based on the cost factor (2) and would represent the value of customers' time lost as a result of non-conformity with the timetable. Assuming A = 0.5€ and the same measured values, the penalty rate would be around 3.766%. By way of illustration, table 2 shows a calculation with tolerance thresholds 9.5 and 10.5 (columns 8, 9 and 10), which are stricter. Without the adoption of any improvement measures, the service provider's costs would be on the level 15.063%. Implementing this function in the service level agreement would increase the pressure on the service provider to improve services by eliminating deviations from the timetable.

3 Summary – differences in views on the use of the Taguchi loss function

Many use cases for the Taguchi loss function are reported in the literature, including many sophisticated quality improvement instruments. In certain cases, there were differences in the way the method was used, which were mainly due to the special characteristics of the service sector and the Taguchi loss function's primary orientation towards industrial mass-production.

Industrial production has profited from the broad application of the Taguchi loss function as a way of providing a rationale for investments in quality improvements and reducing the losses caused by lack of quality. Research has shown that manufacturing's good experience in applying this method could be successfully replicated in services. The example from the service sector (public transport) also shows how the method could be applied to the relationship between a contracting entity and a service provider based on a service level agreement. The method has found many applications on the customer/producer interface and is changing the service sector from a managed to a guaranteed environment, which will be particularly beneficial for end customers.

References

[1] Budaj, P.: *Dimension and approaches for increasing the performance of the operations system*. Fribourg: S.É.C.T., 2013.

[2] Johnston, A., and Ozment, J.: A firm-specific analysis of service quality costs. *International Journal of Logistics Research and Applications* **18,** 5 (2015), 387– 401.

[3] Lysá, L.: *Štatistika pre manažérov*. Ružomberok: Verbum, 2010.

[4] Mishra, A.: Importance of Taguchi's Method in Optimization of Various Problems in Service Sector. *Industrial Engineering & Management* **4**, 5 (2015).

[5] Mitra, A., and Patankar, J.: Estimation of penalty cost in service industries. *Advances in Business and Management Forecasting* **9** (2013), 47–57.

[6] Sangbok, R., Young Hyun, P., and Hanjoo, Y.: A study on education quality using the Taguchi method. *Total Quality Management & Business Excellence* **25 (**2014),7–8.

[7] Smith, T. M.: Taguchi methods and sample surveys. *Total Quality Management & Business Excellence* **5**, 5 (1994), 243–254.

[8] Tolga, T., and Jiju, A.: Applying Taguchi methods to health care. *Leadership in Health Services* **19**, 1 (1997), 26–35.

[9] Hrnčiar, M.: Service Level Agreement – nástroj garancie kvality poskytovanej služby. *Kvalita – odborný časopis Slovenskej spoločnosti pre kvalitu* **2** (2007), 32–34.

Labour Market Institutions and Total Factor Productivity An Evidence in the European Union

Petra Čekmeová[1]

Abstract. In recent years, several economic studies relate the relatively unfavourable development of total factor productivity (TFP) in the European countries to inherent rigidities in labour markets. Therefore, the issue of labour market is in the centre of theoretical and empirical research. The aim of this paper is to determine the effects of five selected labour market institutions on the development of TFP in the member states of the European Union. In particular, by means of an endogenous growth model extended by institutions, we provide a verification of the following hypothesis: Labour market institutions have significant but, in average, more likely negative impact on the development of TFP in the European Union. To estimate the policy augmented productivity equation with a broad measure of TFP we use a panel data model with fixed effect. The analysis is executed on 19 member states of the European Union and covers the period of 1999-2013. Observations for productivity growth are calculated via growth accounting method. The results suggest a statistically significant effect of labour market institutions and confirm the validity of our hypothesis.

Keywords: Total factor productivity, Labour market institutions, European Union, Panel data model, Growth Accounting.

JEL Classification: C23, E24, J48
AMS Classification: 62J05

1 Introduction

In recent years, several empirical studies highlight a problem of inherent labour market rigidities in the European countries. Especially, the connection between these rigidities and unfavourable development of the European productivity is frequently investigated. It is not surprising given the significant contribution of productivity to economic performance. Particularly, the relatively low level of total factor productivity in the member states constitutes a serious problem for national economies and their competitiveness. However, it is not possible to improve the situation without an exact knowledge of factors that could induce this unsatisfactory development. One of the potential sources of the declining total factor productivity can be found in an unsuitable institutional framework on the European labour markets. As the economic theory provides an evidence of both positive and negative effect of labour market institutions, an empirical research is inevitable to make unambiguous conclusion and thereby to be able to design productivity enhancing institutional reforms on the European labour markets.

The aim of this paper is to determine the effects of five selected labour market institutions on the development of total factor productivity in the member states of the European Union. We are interested in the impact of (1) employment protection legislation, (2) unemployment benefits, (3) trade unions, (4) minimum wages and (5) active labour market policies. In particular, by means of an endogenous growth model extended by institutions, we provide a verification of the following hypothesis: *Labour market institutions have significant but, in average, more likely negative impact on the development of total factor productivity in the European Union.* To estimate the productivity equation with a catch-up specification augmented by institutions we apply a panel data model with fixed effect. The analysis is executed on 19 member states of the European Union[2] and covers the period from 1999 to 2013. Observations for the broad measure of total factor productivity growth are calculated via growth accounting method.

The structure of the paper is as follows. After a short introduction, the second section provides an overview of the theoretical relations between total factor productivity and labour market institutions. The third section is devoted to methodological issues, the first subsection contains a description of the empirical model and its estimation and the second one briefly presents the data used for the analysis. Consequently, the empirical results are presented in the section four. The last section summarises the main findings.

---

[1] Masaryk University, Department of Economics, Lipová 41a, Brno, Czech Republic, e-mail: 441117@gmail.com.

[2] Bulgaria, Croatia, Cyprus, Estonia, Latvia, Lithuania, Malta, Romania and Slovenia are excluded from the analysis regarding the unavailability of required data.

2 Theoretical background

Total factor productivity (TFP) reflects the ability of production factors to jointly generate output [17]. Regarding its computation, TFP growth is derived as a residual component of economic growth [30]. In the standard growth accounting approach, this residual is viewed as a proxy for exogenously given technological progress. According to new growth theories TFP is determined endogenously that allows for wider interpretation of TFP growth and investigation of its different sources, including labour market institutions and policies.[3]

Labour market institutions (LMI) represent a set of laws, norms and conventions, outcomes of collective choice mechanisms, which alter decisions of labour force by imposing constraints or incentives [13]. As Betcherman [12] pointed out, beside their main aims (protection of workers and income redistribution) LMI can also lead to efficiency gains as they have impact on the functioning of labour markets and on the productivity. But at the same time, an unsuitable institutional framework may impede productivity growth.

The recent economic research provides only an ambiguous answer on the role of LMI in determining TFP growth. Theoretically, we can identify various channels through which LMI influence productivity in both positive and negative ways. Even the empirical research does not offer an unequivocal conclusion as there is an evidence of both effects. Among the most important channels the followings can be included: investment in human capital, adoption of new technologies and innovations, labour reallocation, quality of job matches or workers´ moral.

The positive effect of trade unions can be explained by encouragement of training, labour reallocation and technological progress as firms have incentives to promote productivity enhancing measures [22] or improve organization and efficiency of production when labour costs rise [25]. According to Freeman and Medoff [18] TU may also increase productivity by lowering quit rate or improving workers´ moral. At the same time, management could be reluctant to introduce productivity enhancing technologies if regulations negotiated by trade unions are restrictive or job loss is expected [7].

Wage-setting institutions as minimum wages (MW) or TU may have negative impact on productivity as they create barriers for potential high-growth firms [21]. According to Bassanini and Duval [10] a reduction of wage differential between high-skill and low-skill job, due to higher MW, could reduce workers´ incentive to invest in training. On the contrary, MW increase average productivity via involvement of more skilled labour force into the production process [1] or by encouraging low skilled workers to invest more in human capital in order to avoid unemployment [4].

The favourable impact of unemployment benefits (UB) on productivity is viewed mainly in incentives to create more productive, high quality post-unemployment jobs [24] or to generate better matches and higher-productivity jobs [2]. However, too generous UBs likely increase the duration of unemployment leading to human capital depreciation and inefficient use of resources [26]. In addition, they may also reduce work effort of employees [29] or incentives to innovate [9]. Negative consequences of generous UB systems might be mitigated by suitable active labour market policies (ALMP) that are introduced with aim to make worker more employable by increasing their skills, thereby having positive effect on TFP [15]. Thus, the final effect of passive and active policies is given by the relative extent of concrete programmes.

Employment protection legislation (EPL), the most considered LMI in recent studies, may encourage workers´ commitment and their willingness to be involved in productivity enhancing activities [11] or firms to adjust by investing more in both physical and human capital [12]. Other theoretical channel through which EPL may influence productivity is a positive impact of less stringent regulations on the flexibility of high-risk entrepreneurial firms and their chance to expand and become high-growth firms [3]. As before, the theory considers not only positive but also negative implications of EPL for aggregate productivity. For example, increasing adjustment costs impedes the reallocation of resources from declining sectors to expanding ones [27]. Or as Ichino and Riphahn [23] claim EPL make workers more willing to put less effort.

3 Methodology and data

This section contains a description of the empirical model and methods used for its estimation as well as the data.

3.1 Model and methods of estimation

To estimate the impact of selected labour market institutions on total factor productivity we use a panel data model. The main assumptions behind our empirical specification are the followings. First, we assume that TFP is endogenously determined. Second, the impact of LMI may vary across countries with different technological level. The

[3] For more details on this topic see e.g. [5] or [8].

latter assumption is in compliance with a theoretical framework of Aghion and Howitt [6] about "distance-dependent" institutions. It means that countries at different stages of development require different institutional set-ups to maximize their productivity.

Given these assumptions, our baseline model will be derived using a catch-up specification of productivity augmented by institutions. The advantage of the proposed specification is that besides the analysis of the impacts of LMI on TFP, it allows to consider the differences in productivity levels among the European countries. In addition, the model is derived based on endogenous growth models, allowing to analyse the effect of various TFP drivers. Similar specifications were used in [28] for labour market regulations or in [14] for product market regulations. However, in our model we include different institutions to provide more complex study of LMI.

Under the described theoretical framework, TFP for a given country *i* in time *t* may be expressed using an autoregressive distributed lag ADL(1,1) process in which the level of TFP is co-integrated with the level of TFP at the technological frontier TFP_F. Hence, TFP can be formally modelled as follows:

$$lnTFP_{it} = \beta_1 lnTFP_{it-1} + \beta_2 lnTFP_{Ft} + \beta_3 lnTFP_{Ft-1} + \omega_{it} \tag{1}$$

where a subscript *F* stands for country with the highest level of TFP and ω represents all observable and unobservable factors influencing the level of TFP. Assuming long-run homogeneity ($\beta_1 + \beta_2 + \beta_3 = 1$) and rearranging the equation (8) we get:

$$\Delta lnTFP_{it} = \beta_2 \Delta lnTFP_{Ft} - (1-\beta_1) ln\left(\frac{TFP_{it}}{TFP_{Ft}}\right) + \omega_{it} \tag{2}$$

where the second right-hand side term represents the productivity gap between the follower country *i* and the frontier country *F*. The last term of this equation, which catches up all other determinants of the TFP growth, can be expressed in the following way:

$$\omega_{it} = \sum_k \gamma_k X_{kit-1} + \varepsilon_{it} \tag{3}$$

where X_{kit} is a vector of the TFP determinants and ε_{it} is an error term.

In the first step, we consider only the effect of five selected LMI as we are interested in these particular relations. Then, the formal model to be estimated is as follows:

$$\Delta lnTFP_{it} = \beta_2 \Delta lnTFP_{Ft} - (1-\beta_1) ln\left(\frac{TFP_{it}}{TFP_{Ft}}\right) + \sum_k \gamma_k LMI_{kit-1} + \varepsilon_{it} \tag{4}$$

The productivity equation augmented by institutions (4) suggests that TFP growth in follower countries is a function of TFP growth at the technological frontier, technological gap and set of LMI.

As we focus on the specific set of the European countries and the inference will be restricted on the behaviour of these countries, we assume the presence of unobserved heterogeneity. Therefore, the error structure of disturbance term can be decomposed into an individual time-invariant effect α_i and an iid error term μ_{it}, leading to the following form of our baseline model:

$$\Delta lnTFP_{it} = \beta_2 \Delta lnTFP_{Ft} - (1-\beta_1) ln\left(\frac{TFP_{it}}{TFP_{Ft}}\right) + \sum_k \gamma_k LMI_{kit-1} + \alpha_i + \mu_{it} \tag{4a}$$

In addition, we run a regression model based on a two-way error component specification with both country- and time-specific effects in order to control also for common aggregate shocks that could have impact on all the Europeans countries in a given year. In this case, the model to be estimated is as follows:

$$\Delta lnTFP_{it} = \beta_2 \Delta lnTFP_{Ft} - (1-\beta_1) ln\left(\frac{TFP_{it}}{TFP_{Ft}}\right) + \sum_k \gamma_k LMI_{kit-1} + \alpha_i + d_t + \mu_{it} \tag{4b}$$

where d_t stand for time dummies.

In the second step, the baseline model (4) is augmented by a set of control variable to check the robustness of our estimates for LMI. Then, we get:

$$\Delta lnTFP_{it} = \beta_2 \Delta lnTFP_{Ft} - (1-\beta_1) ln\left(\frac{TFP_{it}}{TFP_{Ft}}\right) + \sum_k \gamma_k LMI_{kit-1} + \sum_l \delta_l CV_{lit-1} + \varepsilon_{it} \tag{5}$$

where CV_{lit} is a vector of additional TFP determinants, namely R&D and human capital.[4] As before, the productivity equation (5) is estimated via fixed effect estimator in two separate regressions, with country-specific effects (5a) and with both country- and time-specific effects (5b).

The correctness of the estimation method is tested by Hausmann specification test according to [20]. In all regressions, a null hypothesis of common intercept is rejected at any reasonable significance level preferring the application of fixed effects estimator to random effects estimator. As the data indicate the presence of heteroscedasticity we run all regressions with HAC error terms (Arellano).

3.2 Data

The empirical analysis is conducted on a balanced panel data set. Observations on 19 member states of the European Union cover a period from 1999 to 2013. The choice of sample was determined by the availability of data for the given period. In total, we consider 9 explanatory variables and TFP growth as dependent variable. The list of variables with their description and reference to sources is presented in the Table 1.

variable	**source**	**description**
TFP_growth	Own calculation	Log difference of total factor productivity
TFP_F	Own calculation	Log difference of total factor productivity at the frontier
TFP_gap	Own calculation	Productivity gap between follower and frontier country
EPL	OECD.Stat	Strictness of employment protection, overall index (0-7)
ALMP	OECD.Stat	Public expenditures on active labour market policies (% of GDP)
UB	OECD.Stat	Public expenditures on unemployment (% of GDP)
TU	OECD.Stat	Trade union density (members to population)
MW	OECD.Stat	Minimum wages relative to median wages
R&D	OECD.Stat	Total patent applications
HC	Eurostat	Population with tertiary education (% of total)

Table 1 Description and sources of data

Growth rates of TFP were calculated via growth accounting method in our previous work [16] and represent the broadest measures of TFP growth. Beside disembodied technological progress, it includes the effects of technological progress embodied in physical capital, as well as, human capital accumulation. Notice that the productivity gap does not depend on growth rates, but on levels of TFP in follower and frontier country. TFP levels for the initial year are calculated according to [19] in the following form

$$TFP_{it} = \frac{Y_{it}}{\overline{Y}_t} \times \left(\frac{\overline{L}_t}{L_{it}}\right)^{a_{it}} \times \left(\frac{\overline{K}_t}{K_{it}}\right)^{b_{it}} \tag{6}$$

where a bar denotes an average over all countries for a year *t*. Then, TFP indices are extended over the sample period using our estimates of TFP growth.

4 Empirical results

The empirical results of panel regressions for baseline models (4a) and (4b) and their extensions (5a) and (5b) are reported in the Table 2. The results presented in the second column (4a) approve the importance of the catch-up specification of productivity as both productivity gap and TFP growth at the frontier are statistically significant. Regarding the institutional variables, three of the selected LMI have statistically significant effect on TFP growth, namely EPL with negative impact and ALMP and TU with positive one. However, the overall effect of LMI is seem to be negative given the sign and size of corresponding parameters.

After accounting for time-specific effects (4b), the results suggest that only two LMI, TU and MW, have statistically significant role in explaining TFP growth during the analysed period. But as before, the net effect of these LMI on the dependent variable is negative. By comparison the within R^2 of these regression models, it is clear that the latter one is able to explain bigger portions of variability between variables. Therefore, we rely on the findings

[4] R&D and human capital are explanatory variables of TFP commonly proposed by the economic theory.

of this model which are in line with our hypothesis about significant but in average negative impact of LMI on the development of TFP in the sample of selected EU member states.

The role of certain LMI in explaining TFP growth stay significant also after the inclusion of two additional variables, R&D and human capital, which are frequently considered as the most important determinants of TFP growth. In both specification (with and without time dummies) the estimates show statistically significant negative effect in the case of EPL and MW and positive one in the case of TU. The difference lies in the size of estimates that are lower in the case of (5b). Notice that also in the case of extended models the specification with time dummies fit better our data.

	(4a)	**(4b)**	**(5a)**	**(5b)**
const	0,061**	0,079**	-0,045	0,062
	(0,027)	(0,036)	(0,046)	(0,050)
ld_TFP_F	0,700***	0,180	0,768***	0,238
	(0,063)	(0,204)	(0,064)	(0,209)
TFP_gap	0,104***	0,185***	0,188***	0,189***
	(0,030)	(0,036)	(0,020)	(0,038)
EPL_1	-0,025**	-0,014	-0,022**	-0,016*
	(0,010)	(0,010)	(0,010)	(0,009)
ALMP_1	0,016*	0,000	0,015	-0,002
	(0,010)	(0,010)	(0,013)	(0,010)
UB_1	0,002	-0,001	-0,004	-0,002
	(0,007)	(0,005)	(0,008)	(0,005)
TU_1	0,001**	0,003***	0,003***	0,003***
	(0,001)	(0,001)	(0,001)	(0,001)
MW_1	-0,035	-0,033**	-0,047**	-0,037**
	(0,022)	(0,016)	(0,020)	(0,017)
R&D_1	-	-	2,729e-6***	2,836e-8
			(8,936e-7)	(1,161e-6)
HC_1	-	-	0,003***	0,001
			(0,001)	(0,001)
Time dummy	no	yes	no	yes
Observations	266	266	266	266
Within R^2	0,43	0,68	0,47	0,68
F Statistic	86,21	12,65	57,32	15,95
P-value (F)	6,72e-62	1,09e-13	1,18e-054	4,01e-20

Table 2 Regression results – Institutions-augmented productivity equations

5 Conclusion

In recent years, several empirical studies highlight a problem of inherent labour market rigidities in the European countries. Especially, the connection between these rigidities and unfavourable development of the European total factor productivity is frequently investigated. The aim of this paper was to determine the effect of five selected labour market institutions on the development of total factor productivity in the member states of the European Union. We considered the impact of (1) employment protection legislation, (2) unemployment benefits, (3) trade unions, (4) minimum wages and (5) active labour market policies during the period of 1999-2013 in 19 member states. In particular, by estimations of institutions-augmented productivity equations with catch-up specification, we provided a verification of the following hypothesis: *Labour market institutions have significant but, in average, more likely negative impact on the development of total factor productivity in the European Union.*

Our results do not reject the proposed hypothesis. The estimates suggest a statistically significant effect of certain labour market that is negative in average. More precisely, we found out a significant negative impact of employment protection legislation and minimum wages on the growth rate of total factor productivity and a significant positive effect of trade unions. In addition, these findings are robust to the inclusion of other important determinants of the total factor productivity, namely research and development and human capital. Based on these outcomes we can conclude that less stringent employment protection and lower minimum wages would improve, or at least not impede, the development of the European total factor productivity. However, to be able to make unambiguous conclusion about the role of these institutions and design suitable institutional reforms, consideration of additional factors, such as institutional interactions or industry context, will be inevitable.

Acknowledgements

Supported by the specific research project No. MUNI/A/1049/2015 at Masaryk University.

References

[1] Aarosnosn, D. and French, E.: Product Market Evidence on the Employment Effects of the Minimum Wage. *Journal of Labor Economics* **25** (2007), 167-200.
[2] Acemoglu, D. and Pischke, J.: *Minimum Wages and On-the-Job Training*. MIT Working papers **99-25** (1999).
[3] Acs, Z. J.: Foundations of high impact entrepreneurship. *Foundations and Trends in Entrepreneurship* **4** (2008), 535–620.
[4] Agell, J. and Lommerud, K. E.: Minimum wages and the incentives for skill formation. *Journal of Public Economics* **64** (1997), 25-40.
[5] Aghion, P. and Howitt, P.: A model of growth through creative destruction. *Econometrica*, **60** (1992), 323-351.
[6] Aghion, P. and Howitt, P.: *The Economics of Growth*. The MIT Press, Cambridge, 2009.
[7] Aidt, T. and Tzannatos, Z.: *Unions and Collective Bargaining: Economic Effects in a Global Environment.* The World Bank, Washington, DC, 2002.
[8] Barro, R. J.: *Notes on Growth Accounting.* Working Paper **6654** (1998).
[9] Bartelsman, E. et al.: Comparative Analysis of Firm Demographics and Survival: Evidence from Micro-level Sources in OECD countries. *Industrial and Corporate Change* **14** (2005), 365-391.
[10] Bassanini, A. and Duval, R.: *Employment Patterns in OECD Countries: Reassessing the Roles of Policies and Institutions*. OECD Economics Department Working Papers **486** (2010).
[11] Belot, M. et al.: Welfare-improving employment Protection. *Economica* **74** (2007), 381-396.
[12] Betcherman, G: *Labour Market Institutions: A review of the Literature.* WB Policy Research Working Paper **6276** (2012).
[13] Boeri, T. and van Ours, J.: *The Economics of Imperfect Labor Markets*. Princeton university Press, USA, 2013.
[14] Bourlès et al.: *Do product Market regulations in Upstream Sectors Curb productivity Growth? Panel Data Evidence for |OECD Countries*. NBER Working Paper **1650** (2010).
[15] Calmfors, L., et al.: *Does active labour market policy work? Lessons from the Swedish experiences.* IFAU Working Paper **4** (2002).
[16] Čekmeová, P.: Total Factor Productivity and its Determinants in the European Union. *Journal of International relations*, **14** (2016), 19-35.
[17] Compnet Task Force: *Compendium on the diagnostic toolkit for competitiveness.* ECB Occasional Paper Series **163** (2015).
[18] Freeman, R. B. and Medoff, J. L.: *What do unions do?* Basic Books, New York, 1984.
[19] Harrigan, J.: Estimation of Cross-Country Differences in Industry Production Functions. *Journal of International Economics* **47** (1999), 267-93.
[20] Hausman, J. A.: Specification Tests in Econometrics. *Econometrica* **46**, 1978, 1251-71.
[21] Henrekson, M. and Johansson, D.: Competencies and Institutions Fostering High-Growth Firms. *Foundations and Trends in Entrepreneurship* **5** (2009), 1-80.
[22] Heyes, J. and Rainbird, H.: Mobilising resources for union learning: a strategy for revitalisation? *Industrial Relations Journal* **42** (2011), 565-579.
[23] Ichino, A. and Riphahn, R.: The Effect of Employment Protection on Worker Effort: Absenteeism during and after Probation. *Journal of the European Economic Association* **3** (2005), 120-143.
[24] Lippman, S. A. McCall, J.: *Studies in the Economics of Search*. North-Holland, Amsterdam 1979.
[25] Machin, S. and Wandhwani, S.: The Effects of Unions of Investment and Innovation: Evidence from WIRS. *Economic Journal* **101** (1991), 324-30.
[26] OECD: *Employment Outlook 2006*. Paris: OECD, Paris, 2006.
[27] Saint-Paul, G.: Is Labour Rigidity Harming Europe´s Competitiveness? *European Economic Review* **41** (1997), 499-506.
[28] Scarpetta, T., Tressel, T: *Boosting Productivity via Innovation and Adoption of New Technologies: Any Role for Labor Market Institutions?* World Bank Policy Research Working Paper 3273 (2004).
[29] Shapiro, C. and Stiglitz, J. E.: Equilibrium Unemployment as a Worker Discipline Device. *The American Economic Review* **74** (1984), 433-444.
[30] Solow, R.: Technical Change and the Aggregate Production Function. *Review of Economics and Statistics* **39** (1957), 217-235.

A Note on Optimal Network of Tourist Paths Reflecting Customer Preferences

Anna Černá[1], Vladimír Přibyl[2], Jan Černý[3]

Abstract. The paper deals with leisure cycling (CY) and hiking (HI). A digraph whose arcs may be used or reconditioned for CY or HI (not for both) is a "candidate network". Each arc has a length [km] and duration of transit [min]. In addition, the list of tourist attractions, e.g. church, park or lake, is given for each vertex and arc. Each attraction is evaluated by a positive number called attractiveness, i.e. the degree of pleasure when a tourist sees it for the first time. In the theory of tourism networks it is a relatively new concept introduced by the authors in 2013 and explored in 2014 in cooperation with Italian colleagues. The attractiveness of a route is derived from them. It is not additive and less than the sum of individual values of attractiveness. Some vertices are terminal points of tourist trips and an OD matrix expresses the flows of tourist among them. Another matrix of the same dimension represents the duration limits. An original method of the column generation type is proposed for maximization of the total attractiveness, meeting the time and budget constraints.

Keywords: optimization, network, cycling, hiking, tourist paths, attractiveness.

JEL Classification: C61, O18, R58
AMS Classification: 90B06, 90B10

1 Introduction, Basic Problem

The article deals with the following **basic problem (BP)**: In an interesting tourist area, say 20 times 20 km, there is relatively dense “candidate” network of field and forest paths and roads of varying quality. From these, it would be necessary to choose a subnet that would be adapted or reconditioned for the needs of hiking (HI) or cycling (CY) but not for both simultaneously. This *candidate network* can be represented by an (undirected) connected graph $G = (V, E, d, c)$ with the *vertex set* $V = \{1, 2, \ldots, n\}$ where $n > 2$, $E \subset V \times V$ is the *edge set*, the *duration function* d: $E \rightarrow \langle 0; \infty)$ represents length or transit duration of edges, the *cost function* c: $E \rightarrow \langle 0; \infty)$ represents the costs necessary for adaptation of the track. The function d can be extended to the set $V \times V$. Moreover, there is defined a set of *terminals* $W = \{1, \ldots, |W|\} \subset V$, i.e. the nodes where the HI or CY trips can start and end, e.g. rail or bus stations, or well-located parking places with sufficient capacity (mainly for HI) etc. Finally, if $R(W)$ denotes the set of all routes connecting any pair of vertices v, w from W, even the pair v, v with equal ending points, then for each route $r \in R(W)$ there are defined numbers $d(r)$ – the length or transit duration of r and $a(r)$ – tourist *attractiveness* value (briefly attractiveness).

In Chapter 2, the current paper presents an original optimization method for the solution of BP.

The concept of "attraction" plays a very important role in the theory of tourism, especially in Destination Management. Swarbrooke [9] identifies four main types of attractions:

1. features within the natural environment,
2. human-made buildings, structures and sites that were designed for a purpose other than attracting visitors, such as religious worship, but which now attract substantial numbers of visitors who use them as leisure amenities,
3. human-made buildings, structures and sites that are designed to attract visitors and are purpose-built to accommodate their needs, such as theme parks,
4. special events.

Similar view on tourist attraction can be found in Leask [7] and in the book by Page [8] Chapter 8 (‘Visitor Attractions’).

[1] University of Economics/Faculty of Management, Jarošovská 1117, Jindřichův Hradec, Czech Republic cerna@fm.vse.cz.
[2] University of Economics/Faculty of Management, Jarošovská 1117, Jindřichův Hradec, Czech Republic pribyl@fm.vse.cz.
[3] University of Economics/Faculty of Management, Jarošovská 1117, Jindřichův Hradec, Czech Republic cerny@fm.vse.cz.

The common denominator of these views is dealing with such points of tourists' interest that do not influence the "operating quality" of tourism as good surface of routes, availability of appropriate accommodation, food or repair service (the last one for CY). On the contrary, they are interested in delight of tourists from visiting interesting attractions. However, these authors do not try to quantify the degree of delight. A quantitative approach can be seen only rarely, e.g. in [10], but this approach is applied to "operation quality" and not to the "delight" of tourists. Therefore, one can say that the approach of the authors is original, starting from the paper [1].

In the BP formulation, it is supposed that there are given: The graph $G = (V, E, d, c)$ with the duration function d and the cost function c, the set of terminals W, the total cost limit c_o, the $|W| \times |W|$ dimensional OD-matrix F where $f(v, w)$ means the flow of tourists from the terminal v to the terminal w and another matrix D_o of the same dimension where $d_o(v, w)$ means the duration limit for the trip through the route connecting the terminals v and w. Further, there are given the sets $R(v, w) \subset R(W)$ for each pair $(v, w) \in W \times W$ representing the c*andidate sets of routes*, i.e. suitable connections of the pairs, meeting the constraint $r = r(v, w) \in R(v, \mathrm{w}) \Rightarrow d(r) \leq d_o(v, w)$. The **problem** is to find a subgraph $G' = (V, E', d, c)$ of G such that

C1. For each pair $(v, w) \in W \times W$ there exist a route $r = r(v, w) \in R(v, w)$ such that $E(r) \subset E'$ where $E(r)$ means the set of edges passed by the route r.

C2. The edge set E' meets the constraint

$$\sum_{e \in E'} c(e) < c_o \tag{1}$$

C3. The total value of provided attractiveness is maximal, i.e.

$$A(E', F, R) = \sum_{(v,w) \in W \times W} f(v, w) \max\{a(r) : \big(r \in R(v, w)\big) \wedge \big(r \subset E'\big)\} \to \max \tag{2}$$

1.1 Notes to Cost and Duration Functions

The value of cost $c(e)$ for each $e \in E$ is usually determined by experts who are familiar with the situation in the field and with the cost of each of the necessary work.

Similarly, the value of transit duration $d(e)$ for each $e \in E$ is usually determined by experts who are familiar with the situation in the field, know the lengths of edges in km and are able to estimate the duration after the recondition. It is assumed that the duration is equal in both directions of movement through the edge e. This assumption is more or less practically applicable in most Czech tourist regions. Differences between altitudes of peaks from the set W there tend to 100-200 m, which in the duration of the transit route r (v, w) in both directions causes a difference within 20 min. If these differences have exceeded the acceptable limit, it is possible to replace the graph by a digraph, but it is not the subject of this paper.

1.2 Note to Attractiveness Functions and to sets $R(v, w)$

The attractiveness $a(r)$ for each $r \in R(W)$ can be also determined by experts who are familiar with the situation in the field. Another possibility is to use some of the known, much more objective approaches as described e.g. in [1], [4], [3]. However, this issue is not the subject of this paper.

Then candidate sets $R(v, w) \subset R(W)$ for each pair $(v, w) \in W \times W$ can be also determined by experts who are familiar with the situation in the field or designed by more exact and objective methods mentioned e.g. in [4] or [3]. The philosophy of this paper is similar to the one presented in [6] using a candidate set of urban transport routes R_0 and the operating subset $R \subset R_0$ is chosen by nonlinear programming and to [2] where a similar problem with different objective function is solved by linear programming (briefly LP). In both papers, the methods of design of the candidate set R_0 were not the subject of the papers.

In the present paper, the goal is a bit different from [6] or [2] where the optimal subset of candida route set was looked for. In this case, the LP does not choose the subset of the candidate set of routes, but the subset of edges passed by the candidate routes.

2 LP Solution of the Basic Problem

The LP model is a modification of the one from [5]. The goal is to find the optimal subset $E' \subset E$.

2.1 Variables and Denotations

The denotations from Chapter one are used here as well. Moreover, the complete set of candidate routes

$$R_0 = \bigcup_{(v,w)\in W\times W} R(v,w) \tag{3}$$

For each $e \in E$ the binary variable x_e is defined, where $x_e = 1$ when e is chosen into E' and $x_e = 0$ otherwise.

For each $(v, w) \in W \times W$ and each $r \in R(v, w)$ the binary variable y_r is defined, where $y_r = 1$ means that all edges passed by r belong to E' and therefore the tourists may be sure that all edges from r have been reconditioned for their use, while in the case of $y_r = 0$ it is not sure.

The following model uses all denotations defined until now.

2.2 LP Model

The problem is to find values of all variables x_e and y_r meeting the following constraints:

CM1
$$\sum_{r\in R(v,w)} y_r = 1 \text{ for each } (v, w) \in W \times W \tag{4}$$

CM1a
$$10^4 x_e \geq \sum_{r\in R_0, e\in E(r)} y_r \quad \text{for each e} \in \text{E} \tag{5}$$

CM2
$$\sum_{e\in E} c(e)x_e \leq c_o \tag{6}$$

CM3
$$\sum_{(v,w)\in W\times W}\ \sum_{r\in R(v,w)} f(v,w)a(r)y_r \to \max \tag{6}$$

Obviously, the constraints CM1-CM3 correspond to C1-C3.

3 Computational Experience

The above described LP model was tested on different sets of candidate routes using the Gurobi 6.5.1 LP solver. Although this paper does not deal with the problem of candidate routes selection it is necessary to mention several facts about the test network and our approach to routes selection for better understanding of the presented results.

The network of cycle-tourist candidate routes, attractiveness and costs data which describes a part of the Třeboň touristic region were used for testing purpose. It has more than 80 nodes and more than 140 edges. The network was originally used in [4] but it is very well suited for the purpose of this paper too. There are attractiveness values assigned to all the vertices and edges of the network available. The details can be seen in [4]. Figure 1 depicts the topology of this network. The edges are labeled by its average duration time determined by an expert closely familiar with the area.

First of all the set {1, 60, 68, 70, 80} of 5 terminal vertices W was chosen. The terminal vertices are emphasized in Figure 1 by shadowed background of vertex mark. These are municipalities with a bus or railway stop. Further, there were automatically generated the sets $R_p(v, w)$ for each pair $(v, w) \in W \times W$ representing the primary sets of candidate routes, meeting the constraint $r = r(v, w) \in R_p(v, \mathrm{w}) \Rightarrow d(r) \leq d_o(v, w)$. The duration limit $d_o(v, w) = 180$ minutes and flow $f(v, w) = 10$ were set for each pair. The total number of automatically generated primary candidate routes was more than 178 000. The total attractiveness $a(r)$ and estimated costs of each route r was computed too. The final sets (with lower cardinality) of candidate routes $R(v, w)$ were chosen from the appropriate $R_p(v, w)$ with these different strategies:

- **S1**: Only the most attractive routes were chosen preferably.
- **S2**: Several cheapest routes and several most attractive routes were chosen.
- **S3**: The routes with minimum ratio between price and attractiveness were chosen preferably.

The LP model was tested with the costs limit c_o varying from 1 to 10 million. Figure 2 shows the total value of provided attractiveness according to costs limit for the set of candidate routes R_0 with the relatively small cardinality $|R_0| = 150$ chosen with the strategies S1, S2 and S3. It can be seen, that in this case (relatively small cardinality of R_0) the selection strategy strongly affects the results. Strategy S2 allows the LP model to find solution for lower costs limit than in the case of strategy S1. It's clear that from a certain threshold value of costs limit the LP model chooses routes with the maximal attractiveness. In this test case the strategy S3 caused, that the candidate routes with highest attractiveness have not been chosen to R_0. This is the reason, that the maximal total value of provided attractiveness is lower in this case than in the case of strategies S1 and S2.

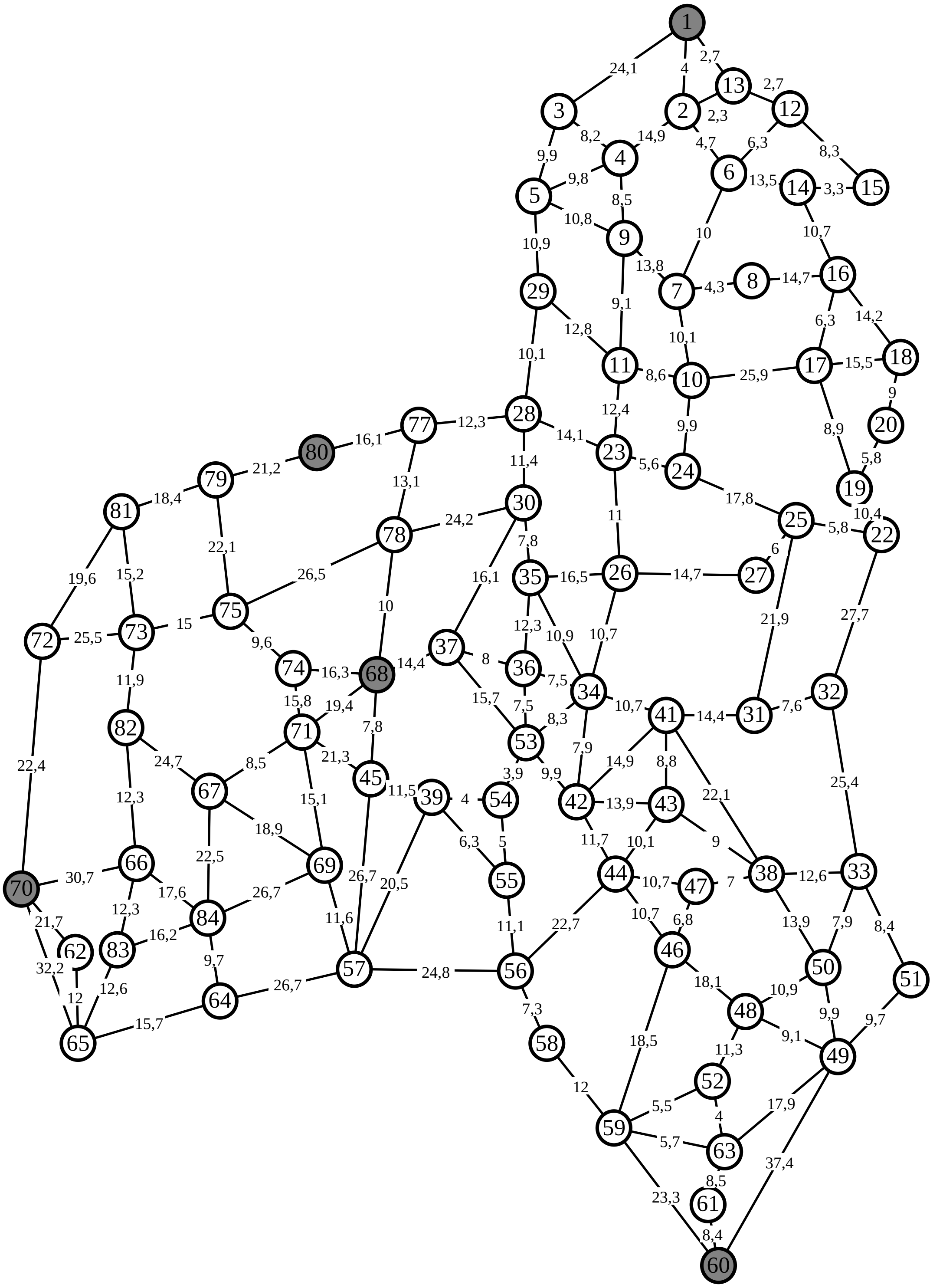


Figure 1 Test Network Topology

Tests showed that the model is well solvable. The computational times did not exceed several seconds even in case of cardinality $|R_0| > 10000$ (i.e. in case of more than 10000 variables in model). It's clear, that importance of the selection strategy decreases with increasing cardinality $|R_0|$. This fact is well documented in Figure 3, which shows the results in case of strategy S1 with cardinality $|R_0| = 150$ and $|R_0| = 11000$.

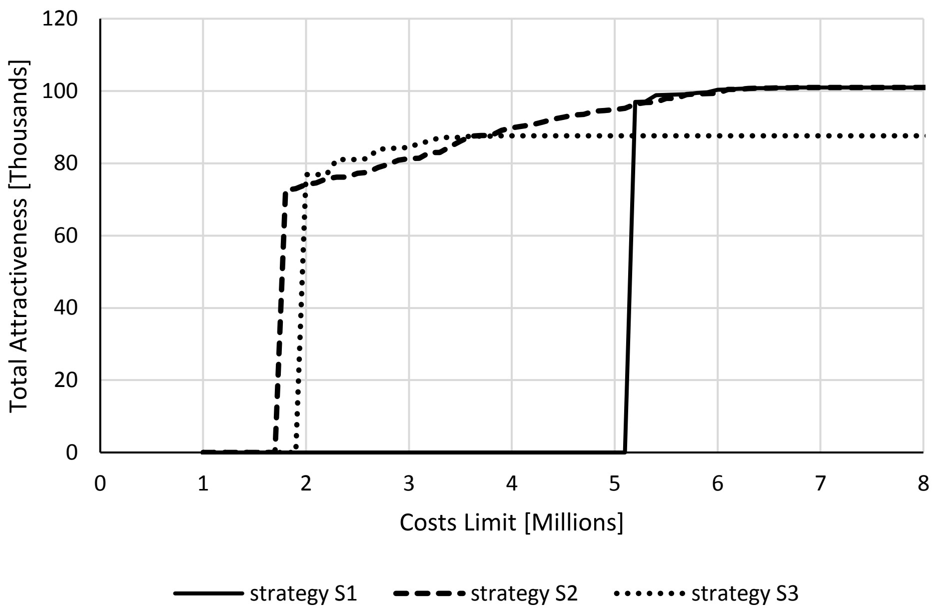


Figure 2 Total value of provided attractiveness (cardinality $|R_0|$ = 150, different strategies)

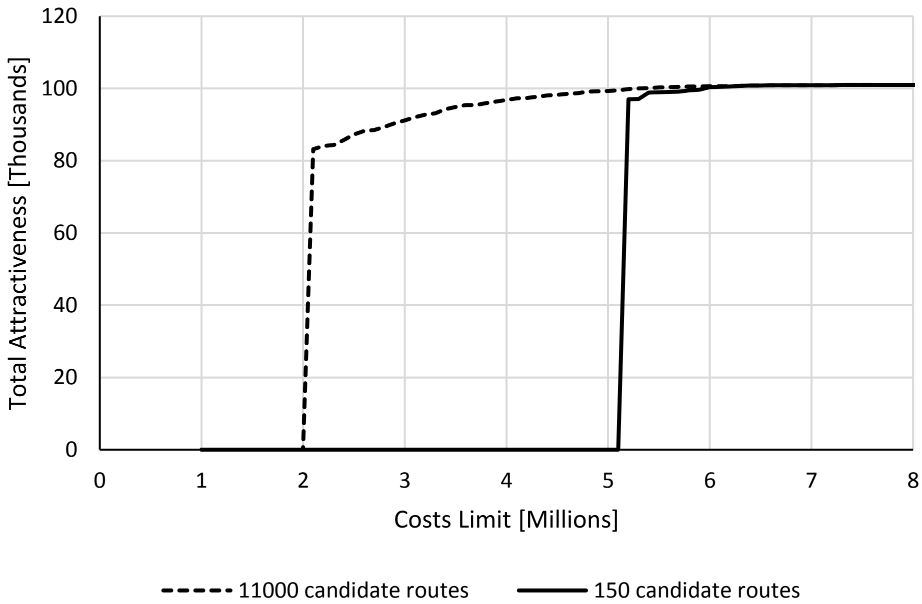


Figure 3 Total value of provided attractiveness (strategy S1, different cardinalities $|R_0|$)

4 Conclusion

The paper presents solution of the problem how to choose a subnet of a network of forest and field paths and roads in some tourist area for reconditioning them in order to let them serve for tourists, mainly hiking or cycling (one of them, not both). The tourists come to see main attractions located near the nodes and paths of the network. The objective is to maximize attractiveness for tourists meeting the time and budget constraints.

The original method of solution for the problem represents the main theoretical contribution of the paper. It is based on linear programming model maximizing the total attractiveness for the average number of tourist visitors, provided the budget constraint and trip duration limits are met. The LP model starts with the wide set of all feasible routes connecting the possible tourist terminals and one of the constrains insures that, for all pair of terminals, the tourists will have the connecting route as attractive as possible, chosen from the wide set of feasible routes.

The Chapter 3 shows that the computing time is small (in seconds) for an actual network from South Bohemia with more than 80 nodes and 140 edges.

Acknowledgements

The authors greatly appreciate the support of research, whose results are published in this paper, by the Faculty of Management, University of Economics in Prague within the resources for long-term institutional development. Project number IP600040.

References

[1] Černý, J. and Černá, A.: Note on Optimal Paths for Non-Motorized Transport on the Network. *Proceedings of the30th International Conference Mathematical Methods in Economics*, Karviná, Czech Republic, September 11-13, 2012, 91-94.

[2] Černý, J. and Černá, A.: Erlander Principle in Managerial Decision Making on Czech and Slovak Urban Transport Routes. *E&M Economics and Management* **16** (2013), 93-100.

[3] Černá, A., Černý, J. and Přibyl, V.: Extended Model of Tourist Routes Optimization. *Proceedings of the 33rd International Conference Mathematical Methods in Economics MME 2015*. Plzeň: Faculty of Economics University of West Bohemia, Plzeň, 2015, 92-98.

[4] Černá, A., Černý, J., Malucelli, F., Nonato, M., Polena, L. and Giovannini, A.: Designing Optimal Routes for Cycle-tourists. *Transportation Research Procedia* **3** (2014), 856–865.

[5] Černá, A., Černý, J., Palúch, S., Peško, Š. and Přibyl, V.: Note to Economically Optimal Road Subnetwork. In *Quantitative Methods in Economics (multiple Criteria Decision Making XVII)*. Bratislava: EKONÓM, 2014, 20-25.

[6] Erlander, S. and Schéele, S.: A mathematical programming model for bus traffic in a network. In D.J. Buckley (ed) *Transportation and Traffic Theory*, New York: Elsevier, 1974, 581-605.

[7] Leask, A. The Nature and Role of Visitor Attractions. In: Fyall, A., Garrod, B., Leask, A. and Wanhill, S (eds) *Managing Visitor Attractions: New Directions*. Elsevier Butterworth-Heinemann, Amsterdam, 2008.

[8] Page, S.J. *Tourism Management, An Introduction* (4th Edition) Elsevier Ltd. Amsterdam, 2011 (Chapter 8 - Visitor attractions)

[9] Swarbrooke, J. *The Development & Management of Visitor Attractions*, 2nd edn. Butterworth Heinemann, Oxford, UK, 2002.

[10] Tilahun, N., D. Levison, and K. Krizek. Trails, Lanes, or Traffic: The Value of Different Bicycle Facilities Using an Adaptive Stated Preference Survey. *Transportation Research:A Policy and Practice*, 41(2007), 287-301

Adaptive wavelet method for the Black-Scholes equation of European options

Dana Černá [1], Václav Finěk [2]

Abstract. We use the Black-Scholes model for calculating the price of European put and call options on a basket of assets. The explicit solution of the Black-Scholes equation is known only for some special cases, otherwise it has to be solved numerically. We present the numerical method based on wavelets for an adaptive solution of the Black-Scholes equation. We use several quadratic and cubic spline wavelet bases. Wavelets are very-well known for their compression property. It means that the representation of the solution in a wavelet basis requires a small number of coefficients and the computation of the solution with desired accuracy can be performed with the small number of degrees of freedom. Furthermore, this method enables high-order approximation, the system of linear algebraic equation arising from discretization is well-conditioned and the number of iterations for computing the solution is relatively small. A numerical example is presented for the two-dimensional Black-Scholes equation with real data.

Keywords: Black-Scholes model, European option, wavelets, adaptive method.

JEL classification: C63, G13
AMS classification: 35K20, 65M60, 65T60

1 Introduction

We consider European options on a basket of assets. Recall that a European put option gives its holder the right, but not the obligation, to sell a group of underlying assets at a specific price on a certain date. Similarly, a European call option gives its holder the right, but not the obligation, to buy a group of underlying assets at a specific price on a certain date. Several models have been proposed for computation of the market price of the option at a given time, see e.g [1].

We use the Black-Scholes model for calculating the price of options. The explicit solution of the Black-Scholes equation is known only for some special cases, otherwise it has to be solved numerically. The equation can be solved by the classical methods such as the binomial tree method, Monte Carlo method, the finite difference method or the finite element method, see e.g. [1] and the references therein. Recently also other approaches such as the discontinuous Galerkin method have been used for an efficient solution [9, 10]. Also methods based on wavelets were already used for solving the Black-Scholes equation [14]. We use a different approach and propose an adaptive wavelet method that is a modification of the method developed in [3] and uses other wavelet bases than bases previously used in the literature for solving this problem.

2 Black-Scholes equation

We focus on a basket containing two assests, whose prices are $S_1 > 0$ and $S_2 > 0$, but all that follows can be generalized for multi-asset cases. We assume that the variable t represents time to maturity, r is a risk-free rate, $\sigma_1 = \sigma_1 (S_1, S_2, t)$ and $\sigma_2 = \sigma_2 (S_1, S_2, t)$ are the corresponding volatilities of the assets

[1]Department of Mathematics and Didactics of Mathematics, Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic, dana.cerna@tul.cz
[2]Department of Mathematics and Didactics of Mathematics, Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic, vaclav.finek@tul.cz

and the parameter $\rho \in (-1,1)$ is the correlation factor. Then the market price $V(S_1,S_2,t)$ of the option at time t can be computed as the solution of the Black-Scholes equation [2]:

$$\frac{\partial V}{\partial t} - \mathcal{L}_{BS}(V) = 0, \quad t \in (0,T), \tag{1}$$

where the Black-Scholes operator $\mathcal{L}_{BS}$ is given by

$$\mathcal{L}_{BS}(V) = \frac{\sigma_1^2 S_1^2}{2}\frac{\partial^2 V}{\partial S_1^2} + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 V}{\partial S_1 \partial S_2} + \frac{\sigma_2^2 S_2^2}{2}\frac{\partial^2 V}{\partial S_2^2} + rS_1\frac{\partial V}{\partial S_1} + rS_2\frac{\partial V}{\partial S_1} - rV. \tag{2}$$

We choose maximal prices S_1^{max} and S_2^{max} large enough and approximate the unbounded domain $\mathbb{R}_+^2$ by a domain $\Omega = (0, S_1^{max}) \times (0, S_2^{max})$. We denote the parts of the boundary $\partial\Omega$ by

$$\Gamma_1 = \{[S_1, 0], S_1 \in (0, S_1^{max})\}, \quad \Gamma_2 = \{[0, S_2], S_2 \in (0, S_2^{max})\}, \tag{3}$$

and

$$\Gamma_3 = \{[S_1, S_2^{max}], S_1 \in (0, S_1^{max})\} \cup \{[S_1^{max}, S_2], S_2 \in (0, S_2^{max})\}. \tag{4}$$

It is clear that the value of a European call option at maturity is

$$V(S_1,S_2,0) = \max(\alpha_1 S_1 + \alpha_2 S_2 - K, 0), \quad (S_1,S_2) \in \Omega, \tag{5}$$

where α_1 and α_2 are the positive weights such that they sum is 1 and K is the strike price. We set the boundary conditions in the same way as in [10, 11]:

$$V(S_1,S_2,t) = \begin{cases} \alpha_1 S_1 \Phi(d_1) - Ke^{-rt}\Phi\left(d_1 - \sigma_1\sqrt{t}\right) & \text{on } \Gamma_1, \\ \alpha_2 S_2 \Phi(d_2) - Ke^{-rt}\Phi\left(d_2 - \sigma_2\sqrt{t}\right) & \text{on } \Gamma_2, \\ \alpha_1 S_1 + \alpha_2 S_2 - Ke^{-rt} & \text{on } \Gamma_3, \end{cases} \tag{6}$$

where Φ is the distribution function of the standard normal distribution and

$$d_i = \frac{\ln(\alpha_i S_i / K) + \left(r + \sigma_i^2/2\right)t}{\sigma_i\sqrt{t}}, \quad i = 1, 2. \tag{7}$$

Similarly, we set the initial and boundary conditions for a European put option. The value at maturity is:

$$V(S_1,S_2,0) = \max(K - \alpha_1 S_1 - \alpha_2 S_2, 0) \tag{8}$$

and boundary conditions are given by

$$V(S_1,S_2,t) = \begin{cases} Ke^{-rt}\Phi\left(-d_1 + \sigma_1\sqrt{t}\right) - \alpha_1 S_1 \Phi(-d_1) & \text{on } \Gamma_1, \\ Ke^{-rt}\Phi\left(-d_2 + \sigma_2\sqrt{t}\right) - \alpha_2 S_2 \Phi(-d_2) & \text{on } \Gamma_2, \\ 0 & \text{on } \Gamma_3. \end{cases} \tag{9}$$

We transform the given equation with non-homogeneous Dirichlet boundary conditions to the problem with homogeneous Dirichlet boundary conditions by the usual way. Let $\tilde{V} = V - W$, where V is the solution of the equation (1) satisfying the initial and boundary conditions defined above and W be a smooth enough function satisfying boundary conditions. In the case of a put option we can set $W(S_1,S_2,t) = W_1(S_1,t)\,W_2(S_2,t)\,e^{rt}/K$, where

$$W_i(S_i,t) = Ke^{-rt}\Phi\left(-d_i + \sigma_i\sqrt{t}\right) - \alpha_i S_i \Phi(-d_i), \quad i = 1, 2. \tag{10}$$

Then $\tilde{V}$ is the solution of the equation

$$\frac{\partial \tilde{V}}{\partial t} - \mathcal{L}_{BS}\left(\tilde{V}\right) = f(W), \quad f(W) = -\frac{\partial W}{\partial t} + \mathcal{L}_{BS}(W) \tag{11}$$

satisfying the initial condition

$$\tilde{V}(S_1,S_2,0) = V(S_1,S_2,0) - W(S_1,S_2,0) \tag{12}$$

and homogeneous Dirichlet boundary conditions on $\partial\Omega \times (0,T)$.

3 Semidiscretization in time

For discretization in time we use the θ–scheme. Let $M \in \mathbb{N}$, $\tau = M^{-1}$, $t_l = l\tau$, $l = 0, \ldots, M$, and denote $\tilde{V}_l(S_1, S_2) = \tilde{V}(S_1, S_2, t_l)$ and $W_l(S_1, S_2) = W(S_1, S_2, t_l)$. The θ–scheme has the form:

$$\frac{\tilde{V}_{l+1} - \tilde{V}_l}{\tau} - \theta \mathcal{L}_{BS}\left(\tilde{V}_{l+1}\right) - (1-\theta)\mathcal{L}_{BS}\left(\tilde{V}_l\right) = \theta f(W_{l+1}) + (1-\theta) f(W_l), \tag{13}$$

where $\theta \in [0,1]$ and $l = 0, \ldots, M-1$. The choice $\theta = 1$ corresponds to the backward Euler scheme and $\theta = 0.5$ corresponds to the Crank-Nicolson scheme. Both methods are implicit and stable, but the Crank-Nicolson method is more accurate, because it is the second order method with respect to the variable t, while the backward Euler method is only of the first order.

4 Wavelet bases

We briefly introduce the concept of a wavelet basis. Let H be a Sobolev space or the L^2–space, $\mathcal{J}$ be an index set such that each index $\lambda \in \mathcal{J}$ takes the form $\lambda = (j, k)$ and $|\lambda|$ denotes the level. A wavelet basis of the space H is defined as a family $\Psi = \{\psi_\lambda, \lambda \in \mathcal{J}\}$ such that

i) Ψ is a Riesz basis for H, i.e. the closure of the span of Ψ is H and there exist constants $c, C \in (0, \infty)$ such that

$$c\,\|\mathbf{b}\|_2 \leq \left\| \sum_{\lambda \in \mathcal{J}} b_\lambda \psi_\lambda \right\|_H \leq C\,\|\mathbf{b}\|_2, \tag{14}$$

for all $\mathbf{b} = \{b_\lambda\}_{\lambda \in \mathcal{J}}$ such that $\|\mathbf{b}\|_2^2 := \sum_{\lambda \in \mathcal{J}} b_\lambda^2 < \infty$.

ii) The functions are local in the sense that diam supp $\psi_\lambda \leq \mathrm{C}2^{-|\lambda|}$ for all $\lambda \in \mathcal{J}$, and at a given level j the supports of only finitely many wavelets overlap at any point x.

A wavelet basis on the interval I has typically the hierarchical structure:

$$\Psi^I = \Phi^I_{j_0} \cup \bigcup_{j=j_0}^{\infty} \Psi^I_j. \tag{15}$$

The functions from the set $\Phi^I_{j_0}$ are called scaling functions and the functions from the set Ψ^I_j are called wavelets on the level j. Wavelets ψ_λ, $\lambda = (j,k)$, in the inner part of the interval are typically translations and dilations of a function ψ also called wavelet, i.e. $\psi_{j,k}(x) = d_{j,k}\psi\left(2^j x - k\right)$, $d_{j,k} \in \mathbb{R}$, and the functions near the boundary are derived from functions called boundary wavelets constructed such that they satisfy the boundary conditions. In adaptive methods it is required that wavelets have vanishing moments. It means that

$$\int_I x^k \psi_{j,k}(x) = 0, \quad k = 0, \ldots L-1, \tag{16}$$

where L is dependent on the type of wavelet. We assume that $L \geq 1$. A wavelet basis Ψ on the rectangle $I \times J$ is constructed by a tensor product: $\Psi = \Psi^I \otimes \Psi^J$, for more details see e.g. [8]. Cubic spline wavelets from [8] are displayed in Figure 1.

5 Adaptive wavelet method

We use an adaptive wavelet method for discretization with respect to the variables S_1 and S_2. While the classical adaptive methods use refining a given mesh according to a-posteriori local error estimates, the wavelet approach is different. One starts with a variational formulation but instead of turning to a finite dimensional approximation, using the suitable wavelet basis the continuous problem is transformed into an infinite-dimensional problem. Then an iteration scheme is proposed for this problem. Finally, all infinite-dimensional quantities have to be replaced by finitely supported ones and the routine for an application of an infinite matrix approximately have to be designed.

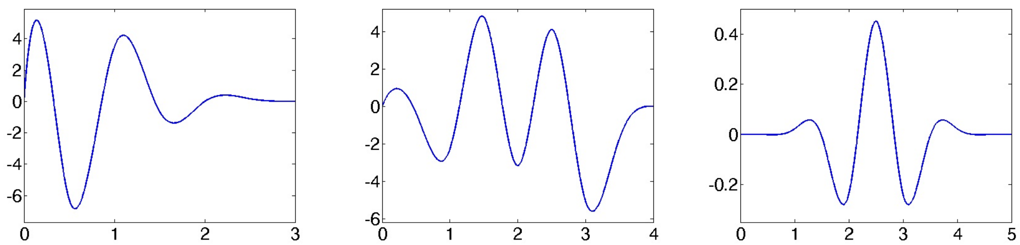


Figure 1 Two boundary cubic spline wavelets (left, middle) and one inner wavelet ψ (right) from [8].

The standard weak formulation of (13) has the form:

$$\frac{\left(\tilde{V}_{l+1}, v\right)}{\tau} - \theta a\left(\tilde{V}_{l+1}, v\right) - (1-\theta)\, a\left(\tilde{V}_l, v\right) = \frac{\left(\tilde{V}_l, v\right)}{\tau} + \theta\left(f\left(W_{l+1}\right), v\right) + (1-\theta)\left(f\left(W_l\right), v\right), \tag{17}$$

where $v \in H_0^1(\Omega)$, $a(u,v) = (\mathcal{L}_{BS}(u), v)$ and $(\cdot,\cdot)$ denotes $L^2(\Omega)$–inner product.

Let $\Psi = \{\psi_\lambda, \lambda \in \mathcal{J}\}$ be a wavelet basis for $H_0^1(\Omega)$ and $\mathbf{u} = \{u_\lambda\}_{\lambda\in\mathcal{I}}$ be the coefficients of the solution $\tilde{V}_{l+1}$ of the problem (17) in a basis Ψ, i.e.

$$\tilde{V}_{l+1} = \sum_{\lambda\in\mathcal{J}} u_\lambda \psi_\lambda. \tag{18}$$

Using (18) and setting $v = \psi_\mu$ we obtain the infinite matrix equation $\mathbf{Au} = \mathbf{f}$ with

$$\mathbf{A}_{\mu,\lambda} = \frac{(\psi_\lambda, \psi_\mu)}{\tau} - \theta a\left(\psi_\lambda, \psi_\mu\right), \tag{19}$$

and

$$\mathbf{f}_\mu = (1-\theta)\, a\left(\tilde{V}_l, \psi_\mu\right) + \frac{\left(\tilde{V}_l, \psi_\mu\right)}{\tau} + \theta\left(f\left(W_{l+1}\right), \psi_\mu\right) + (1-\theta)\left(f\left(W_l\right), \psi_\mu\right). \tag{20}$$

It is clear that $\mathbf{f}$ and $\mathbf{u}$ depend on the time level t_l, but for simplicity we omit the index l in the notations. The Richardson iterations are typically used for solving this infinite-dimensional problem. We choose a different approach and use the method of generalized residuals (GMRES), because in our numerical experiments it was significantly faster. For preconditioning we use the diagonal matrix $\mathbf{D}$, where the diagonal elements of $\mathbf{D}$ satisfy $\mathbf{D}_{\lambda,\lambda} = \sqrt{\mathbf{A}_{\lambda,\lambda}}$. We obtain the preconditioned system $\tilde{\mathbf{A}}\tilde{\mathbf{u}} = \tilde{\mathbf{f}}$ with

$$\tilde{\mathbf{A}} = \mathbf{D}^{-1}\mathbf{A}\mathbf{D}^{-1}, \quad \tilde{\mathbf{f}} = \mathbf{D}^{-1}\mathbf{f}, \quad \tilde{\mathbf{u}} = \mathbf{Du}. \tag{21}$$

For the given time level t_l the algorithm for solving the infinite-dimensional problem comprises the following steps:

1. Compute sparse representation $\tilde{\mathbf{f}}_j$ of the right-hand side $\tilde{\mathbf{f}}$ with the error smaller than given tolerance ϵ_j^1.

2. Compute K steps of GMRES iterations for solving the system $\tilde{\mathbf{A}}\mathbf{v} = \tilde{\mathbf{f}}_j$ with the initial vector $\mathbf{v}_j$. Each iteration of GMRES requires multiplication of the infinite-dimensional matrix with a vector. It is computed approximately with the given tolerance ϵ_j^2 by the method from [7]. We denote the resulting vector by $\mathbf{z}$.

3. Compute sparse representation $\mathbf{v}_{j+1}$ of $\mathbf{z}$ with the error smaller than ϵ_j^3.

We repeat the steps 1., 2., and 3. until the residual is not smaller than the required error. Since we work with the sparse representation of the right-hand side and the sparse representation of the vector representing the solution, the method is adaptive. The computation of a sparse representation is simple and it insists in thresholding the smallest coefficients and working only with the largest coefficients. It is known that the coefficients in the wavelet basis are small in regions where the function is smooth and large in regions where the function has some singularity.

6 Numerical example

We solve the same example with real data as in [10], i.e. a basket put option with 60% Allianz stock ($\alpha_1 = 0.6$) and 40% Deutsche Bank stock ($\alpha_2 = 0.4$), the striking price is 40 Euro, and the option is maturing in 94 days. We assume that the current date is September 13, 2011. Then the two stocks trades at $S_1 = 59.79$ and $S_2 = 23.40$ Euro, respectively. The last year estimate of the Pearson linear correlation is $\rho = 0.88$ and market implied risk-free interest rate for a given horizon $r = 0.01557$ p.a. The corresponding volatilities are $\sigma_1 = 0.6392$ and $\sigma_2 = 0.9461$. We set $S_1^{max} = 130$, $S_2^{max} = 220$, $\theta = 1$ and the time step is one day, i.e. $\tau = 1/365$. We use quadratic spline wavelet bases from [5, 6, 13] and a cubic spline wavelet basis from [8]. The solution $\tilde{V}$ of the homogeneous problem and the function W satisfying boundary conditions are displayed in Figure 2.

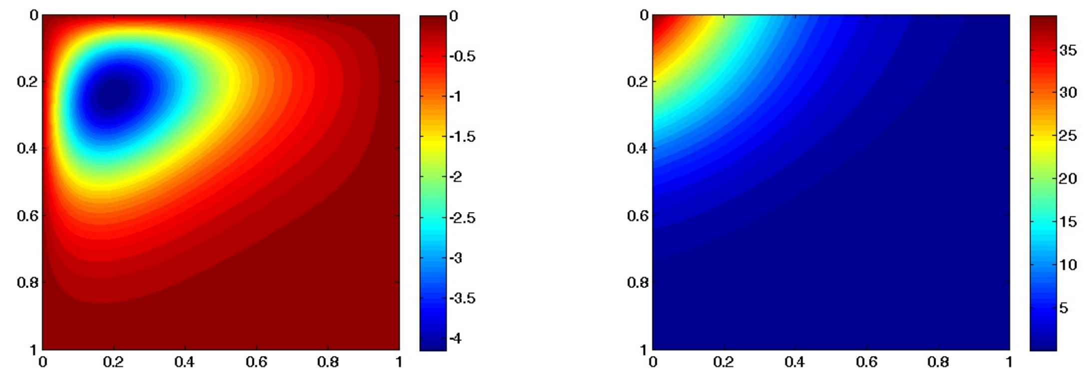


Figure 2 The solution $\tilde{V}$ of the homogenous problem (left) and the function W (right).

It can be seen that the gradient of the solution $\tilde{V}$ has largest values near the parts of the boundary Γ_1 and Γ_2. Therefore the largest wavelet coefficients correspond to wavelets with supports in regions near Γ_1 and Γ_2 and wavelet coefficients are small for wavelets on large levels that are not located in these regions. Thus many wavelet coefficients are thresholded and the representation of the solution is sparse. For cubic spline wavelets from [8] the number of parameters representing the solution $\tilde{V}$ in Figure 2 and thus also the price of the option V in Figure 3 is 1048. We have compared the number of iterations needed for solving the problem with a given residual for wavelets from [5, 6, 8, 13]. The least number of iterations has been needed for method with wavelets from [8]. Since wavelets from [5, 8] are piecewise cubic functions while wavelets from [6, 13] are piecewise quadratic, the convergence is faster for wavelets from [5, 8]. The resulting market value of the option for the above reference prices is 3.59 Euro.

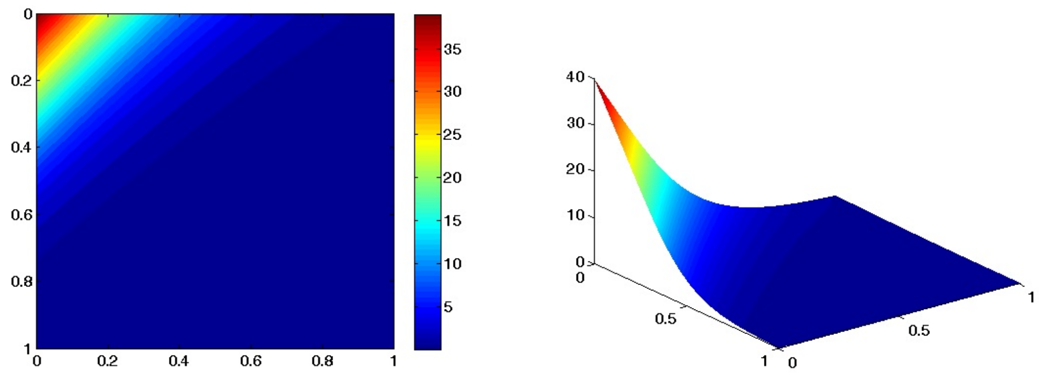


Figure 3 Contour plot (left) and 3D plot (right) of the price of the option for $t = 94$.

7 Conclusion

We have used an adaptive wavelet method for a numerical solution of the Black-Scholes equation for the pricing of a European option on a basket of two assets. Due to the compression property of wavelets the solution is represented by the small number of parameters. Since the used wavelets have been piecewise quadratic or cubic functions the method is high-order accurate. Our future aim is to develop the efficient solver for solving higher-dimensional equations for option pricing and also to use the adaptive wavelet method for option pricing using more accurate models such as Heston or Lévy model. These models also use the Black-Scholes operator (2) and thus they can be viewed as the generalization of the Black-Scholes equation. However, Lévy model uses also an integral operator that is non-local, see [4, 12]. Wavelet methods seem to be appropriate for solution of such equations, because they enable to represent integral operators by quasi-sparse matrices while the classical methods typically lead to full matrices.

Acknowledgements

This work was supported by grant No. GA16-09541S of the Czech Science Foundation.

References

[1] Achdou, Y., and Pironneau, O.: *Computational Methods for Option Pricing.* Society for Industrial and Applied Mathematics, Philadelphia, 2005.

[2] Black, F., and Scholes, M.: The pricing of options and corporate liabilities. *J. Polit. Econ.* **81** (1973) 637–659.

[3] Cohen, A., Dahmen, W., and DeVore, R.: Adaptive wavelet methods II - beyond the elliptic case. *Found. Math.* **2** (2002), 203–245.

[4] Cont, R., and Tankov, P.: *Financial modelling with jump processes.* Chapman & Hall/CRC, Boca Raton, FL, 2004.

[5] Černá, D., and Finěk, V: Construction of optimally conditioned cubic spline wavelets on the interval. *Adv. Comput. Math.* **34** (2011), 219–252.

[6] Černá, D., and Finěk, V.: Quadratic wavelets with short support on the interval. *AIP Conference Proceedings* **1497** (2012), 113–117.

[7] Černá, D., and Finěk, V.: Approximate multiplication in adaptive wavelet methods, *Cent. Eur. J. Math.* **11** (2013), 972–983.

[8] Černá, D., and Finěk, V.: Wavelet basis of cubic splines on the hypercube satisfying homogeneous boundary conditions. *Int. J. Wavelets Multiresolut. Inf. Process.* **3** (2015), 1550014 (21 pages).

[9] Hozman, J.: Analysis of the discontinuous Galerkin method applied to the European option pricing problem. *AIP Conference Proceedings* **1570** (2013), 227–234.

[10] Hozman, J., and Tichý, T.: A discontinuous Galerkin method for pricing of two-asset options. In: *33rd International Conference Mathematical Methods in Economics*, Cheb, ZČU Plzeň, (2015), 273–278.

[11] Koh, W. S., Sulaiman, J., and Mail, R.: Numerical solution for 2D European option pricing using quarter-sweep modified Gauss-Seidel method. *Journal of Mathematics and Statistics* **8** (2012), 129–135.

[12] Merton, R. C.: Option pricing when underlying stock returns are discontinuous. *J. Financ. Econ.* **3** (1976), 125–144.

[13] Primbs, M.: New stable biorthogonal spline-wavelets on the interval. *Result. Math.* **57** (2010), 121–162.

[14] Rometsch, M.: *A wavelet tour of option pricing.* PhD thesis, Universität Ulm, 2010.

Income and price elasticity of demand for transport services in rail passenger transport in the Slovak republic

Lenka Černá[1] – Vladislav Zitrický[2] – Ján Ponický[3]

Abstract. The elasticity is sensitivity of rate to change one variable to another variable. It`s the division of percentage changes two variables. These changes are expressed in percentage and the elasticity doesn`t matter on the units of variables.
The paper is focused on the analysis of income and price elasticity for services in rail passenger transport. The analysis is focused on the influence of selected factors for demand in transport services. Increase or decrease of transport volume in rail passenger transport depends on those factors. Price elasticity of demand for transport services is calculated as ratio of the changes of transport volume and the change of ticket price in the specified period.
Income elasticity for demand of transport services in rail passenger transport expresses the percentage change of demand to percentage change of income. By the income elasticity we will analyze the dependence of transport performance in rail passenger transport from population income in Slovakia in the specified period. The calculation takes into account the total income in selected types of households and the revenue per one household member and number of passengers in the specified period.

Keywords: rail passenger transport, elasticity, demand

JEL Classification: C44
AMS Classification: 90B20

Introduction

Passengers and their preferences on the transport market determine the demand for services to satisfy their transport needs. Sellers (in this case passenger operators) determine the range of services in order to maximize their profits. The success of service exchange is influenced by the price level on the transport market. Demand is the amount of services for which passengers are willing to pay at a given price. Important characteristic of demand is its price elasticity. Price elasticity of demand is an economic definition and it expresses the sensitivity of demanded quantity of certain goods to its price.

In the Slovak Republic the legislative framework in the financial, economic and social area has changed in the rail passenger transport in the last period. As Gašaprík said [3] and [4], an example of the support in rail passenger transport is the introduction of the free of charge transportation for selected groups of passengers. Free of charge transportation is applied on trains which are operated based on the Contract for Transport Services in the Public Interest.

The paper is focused on the change of price and income elasticity by the offered services in rail passenger transport. The secondary aim of the paper is recognition of transport customers (passengers), mathematical expression of demand elasticity and factor analysis that influenced demand for transport services. The change of price and income can cause various changes in transport volume in the rail passenger transport. From the fact mentioned above we can conclude that demand of transport services can be flexible – it has its elasticity. The result of transport demand elasticity is increase or decrease in number of passengers in the rail transport.

1 University of Žilina, Faculty of operation and economics of transport and communications, Department of railway transport, Univerzitná 1, 010 26 Žilina, lenka.cerna@fpedas.uniza.sk
2 University of Žilina, Faculty of operation and economics of transport and communications, Department of railway transport, Univerzitná 1, 010 26 Žilina, vladislav.zitricky@fpedas.uniza.
3 University of Žilina, Faculty of operation and economics of transport and communications, Department of railway transport, Univerzitná 1, 010 26 Žilina, jan.ponicky@fpedas.uniza.

The applied methodology of the paper is divided into two parts: preparatory phase and realization phase. First phase determines research problem and its objective. Preparatory phase consists of the analysis of demand elasticity in the rail passenger transport, pricing in rail passenger transport and definition of the price and income elasticity. Second phase includes data collecting, processing and its interpretation. Next step of realization phase includes mathematical expression of coefficients of price and income elasticity. End of the paper includes recommendations resulting from realization phase. Model of price formation based on the demand curve could help the rail companies to become more profitable.

Price and income elasticity is calculated based on different data and periods, because there are more rail passenger carriers operating on the rail passenger market: Železničná spúoločnosť Slovensko, a.s. (ZSSK), RegioJet, a.s. and Leo Expres a.s. Price elasticity is calculated only for state rail carrier ZSSK and only this company changed fares during the reported period and provided services during the all period. Also ZSSK published their transport performance per year 2015 [8]. Income elasticity is calculated for all passenger carriers; however, data for year 2015 will be available until end of 2016. For this fact income elasticity is calculated only for period 2011 – 2014.

1 Pricing in rail passenger transport in the Slovak Republic

Pricing policy in the Slovak Republic in rail passenger transport is regulated by Transport Authority. A price in national passenger transport is classified as a social category, i.e. application of state social policy by regulating prices. For example: special price for students, pensioners and etc.. The price determination in transport performance is one the most sensitive parts of transport policy. A state must take into account free market principles (reaction to demand and supply) and the need of competition (among rail operators and also other modes of transport).

Based on the [1], price in rail passenger transport is determined by the tariff. To determine a tariff in rail passenger transport economically justified costs and a reasonable profit must be taken into account. These two items must ensure development of rail transport and eliminate the dependence of rail on the state budget. To create competitive pricing and competition on the market, it is necessary to harmonize the conditions among transport modes. The harmonization means reimbursement of the transport infrastructure costs, price regulation and reimbursement of the performance in public interests.

Price is basic factor, which influences customer (passenger) decision about the use of specific transport mode. Nevertheless, the price in passenger transport might not be primary decision factor for customers. A customer can decide based on other factors, such as quality environmental impact. At present the price of the rail passenger transport in the Slovak Republic reaches its maximal tolerable level.

The last correction of prices in rail passenger transport was in 17. 11. 2014, when the government of the Slovak Republic decided about free of charge transport for selected category of passengers. The free of charge transport was introduced as a social measure to support the students, pensioners and commuting passengers (they travel at a reduced fare not for free). This decision was made also to support the environmentally friendly mode of transport [1].

2 Elasticity of demand for transport services in railway passenger transport.

According to the Gašparik [4], passenger's decision about the quantity of use of public transport services depends on the price of service. This fact represents demand for services.

Demand, based on the Gnap and Konečný [5][6], in transport represents the quantity of services which passenger are willing to purchase on the transport market at a certain price. Demand of the only one passenger, or demand for the only one service, represents individual demand. Demand of all passengers for the only one service represents a partial demand. Aggregate demand is the sum of demand on all partial markets. Graphical representation of demand is called the demand curve. Demand curve shows the relationship between the price of a certain commodity and the amount of it that consumers are willing and able to purchase at the given price. Demand curve has a decreasing course. It follows the law of a downward demand. That means if the prices of services are growing, the demand is decreasing and in reverse if the prices of services are decreasing, the demand is growing.

In [5] [6], demand for transport is determined by several factors and it is possible to express it as a function of several variables (factors):

$$Qi = f\,(Pi, Pj, ..., Pn, I, T, ...)$$

The significant factors which determine the demand for services in rail passenger transport are fare in rail transport (Pi), fare in other modes of transport, or other operators (Pj…Pn), income of population - pension of the population (I), habits of consumers - passengers preferences (T), number of households - market range (D), quality of transport services (Q), etc..

2.1 Expression of the coefficient of the price and income elasticity of demand

Based on the [6] and [7], the coefficient of price elasticity - $E_{Q,P}$ is calculated as a ratio of percentage change in quantity of the demand and percentage change in price (Formula 1). Income elasticity is calculated as the coefficient of income elasticity of demand $E_{Q,M}$, which is the ratio of change in quantity of the demand and change in the consumer's income. The coefficient is expressed in formula 2.

$$E_{Q,P} = \left| \frac{Q_2 - Q_1}{P_2 - P_1} \cdot \frac{P_1 + P_2}{Q_1 + Q_2} \right| \tag{1}$$

$$E_{Q,I} = \left| \frac{Q_2 - Q_1}{I_2 - I_1} \cdot \frac{I_1 + I_2}{Q_1 + Q_2} \right| \tag{2}$$

The change of purchased goods or service Q due to changes in price of the goods or service is reflected in the total volume of spending on the goods or service (Px.Q). Because the price and quantity move in the opposite direction – growing prices result in a decrease in quantity and vice versa. Income elasticity of demand is the elasticity where the examined factor is consumer's income. Various kinds of elasticity are shown in Table 1 [6].

Price elasticity of demand		**Income elasticity of demand**	
Value	**Kind of elasticity**	**Value**	**Kind of elasticity**
$E_{Q,P} = 0$	Perfectly inelastic demand	$E_{Q,I} < 0$	A negative income elasticity of demand is associated with inferior goods
$E_{Q,P} < 1$	Less elastic demand		
$E_{Q,P} = 1$	Unit-elastic demand	$E_{Q,I} > 0$	A positive income elasticity of demand is associated with normal goods
$E_{Q,P} > 1$	Elastic demand		
$E_{Q,P} = \infty$	Perfectly elastic demand		

Table 1 Kinds of price and income elasticity

The price elasticity of demand for transport services in rail passenger transport is calculated as the ratio of the changes in quantity (expressed as transport performance) and the change of price. Income elasticity is expressed as a percentage change of demand for transport services depending on a percentage change in income.

2.2 Methods for measuring the elasticity of demand functions

The demand function has a decreasing course therefore the elasticity of this function has a negative value. To avoid the absolute value when calculating the percentage change of the function minus mark is used to define elasticity of demand function.

Definition of demand by the Molnárová [7] is: When q=D (p) is a demand function, where p>0 is the price of services on the market and q>0 is a demand for the service. If there is D' (p) for p ∈ (0, ∞). Number -(D'(p0)/D (p0)). p0, which reflects the percentage by which demand D(p) of the service is reduced at p0 price increase by 1%, is called elasticity of demand function at point p0.

$$E_D = \eta(D(p_0)) = -\frac{D'(p_0)}{D(p_0)} \tag{3}$$

The elasticity of demand function could help to determine the appropriate price for service on the market. The demand function based on elasticity is characterized in Chapter 2.1.

To calculate price and income elasticity of demand it was necessary to use several methods depending on the character of the individual parts of solution. In the text below the applied methods and their intended use within the elasticity of demand of transport services in rail passenger transport are described:

- *Collecting of information method* – performance in rail passenger transport (passenger-kilometres, number of passenger), fare in the last four years (regular fare for 2nd class and special price for students) and income in selected household categories. Fares are from the official tariff of Slovak national rail passenger operator – Železničná spoločnosť Slovensko, a.s. (ZSSK), that have 99% share on the rail passenger market.
- *Information processing method* – statistical processing performance, prices and incomes data - expressed by the price and income coefficient of elasticity of demand. The coefficient of price elasticity of demand is calculated based on the formula 1 (P- price, Q – quantity) and the coefficient of income elasticity of demand is calculated based on the formula 2 (Q – amount of demand, I – income) [6] [7].
- *Simplification for discrete variables* – percentage change of quantity can be expressed by Formula 4 (Q_1 is the original quantity and Q_2 is quantity after the price change). The percentage change of price is expressed similarly – Formula 5 [6].

$$\%Q = \frac{Q_2 - Q_1}{\frac{1}{2}\left(Q_1 + Q_2\right)} \cdot 100 \tag{4}$$

$$\%P = \frac{P_2 - P_1}{\frac{1}{2}\left(P_1 + P_2\right)} \cdot 100 \tag{5}$$

The final coefficient of price elasticity of demand is calculated by dividing the percentage change of demand quantity and the percentage change of price. The final income elasticity of demand is calculated based on Formula 2.

3 Analysis price and income elasticity for transport services in railway passenger transport in the Slovak Republic

The analysis of price and income elasticity of demand for transport services in rail passenger transport is influenced by the change of prices for transport (regular and special price). The sensitivity of performance change (number of passenger) is influenced by the change of price for transport in rail transport - free of charge transport, which effective from 17. 11. 2014. Statistics data are from [9].

3.1 Analysis of price elasticity

The analysis of price elasticity of demand for transport was calculated as dependence of demand for transport services in rail passenger transport (number of passenger) to height of fares offered by ZSSK. Analysis was performed for the period 2011 -2015. To calculate the elasticity of demand two types of fares were used: regular price in 2nd class and special price for students in 2nd class. Table 2 and 3 shows the preview of transport performance (for national transport), calculation of average fares (regular and special) and the final coefficient of price elasticity of demand. The coefficient of price elasticity of demand was calculated as the ratio of the percentage change of demand quantity and the percentage change of price. The calculation of an average fare was conducted based on the tariff of ZSSK. All prices in the tariff depend on kilometrical distance, the tariff distances are from 1 km up to 510 km.

Year	**2011**	**2012**	**2013**	**2014**	**2015**
Number of transported passengers (mil. pass.)	45.959	43.445	44.287	47.286	53.7
Average price (regular price) – (*€/km)*	0.044193	0.044193	0.044193	0.044193	0.044193
Price elasticity of demand $E_{Q,P}$	-	0	0	0	0

Table 2 Number of transported passenger by ZSSK, average prices of regular prices and coefficient of price elasticity – regular price

Based on the calculated coefficients of price elastic of demand (basic fare) we can see that the demand is perfectly inelastic. The price has not changed compared to the transport performance during the monitored period. The fare is determined by Transport Authority of the Slovak Republic and during the monitored period they did not change the fares (Chapter 1).

Year	2011	2012	2013	2014	2015
Number of transported passengers (mil. pass.)	45.959	43.445	44.287	47.286	53.7
Average price (student) – (*€/km)*	0.022097	0.022097	0.022097	0.019901	0
Price elasticity of demand $E_{Q,P}$	-	0	0	0.000216	-0.09607

Table 3 Number of transported passenger by ZSSK, average price of special prices and coefficient of price elasticity – student

The coefficient of price elasticity of demand for student's price was perfectly inelastic until the year 2013. The fare and number of transported passenger show minimal differences and values are almost the same in the period 2011 – 2013. In 2014 we can see little change compared with the previous year and 2015 when the average fare per kilometre is 0€. This change is caused by the introduction of free of charge travel for students in the Slovak Republic. The number of transported passengers had a rising tendency since the beginning 2015 (approximately 18 % growth). More accurate results in the changed number of transported passengers by rail transport will be determined after the longer period (not just one year).

The preference of travellers has changed after the introduction of free of charge fare for transport for selected passenger and currently we can see the growth of travellers in the rail passenger transport.

3.2 Analysis of income elasticity

Income elasticity of demand for transport services in rail passenger transport shows the dependence of number of transported passengers (1000 pass.) in national transport on the disposable income. For the purposes of analysis of income elasticity of demand the disposable income in the Slovak Republic is divided into two types of households. The first is one-member household and the second is a household is with two dependent children (two adults with two children). The calculation of income elasticity of demand is in the table 4 and 5.

Year	2011	2012	2013	2014
Number of transported passengers (mil. pass.)	47.531	44.698	46.064	49.272
Disposable income - one member household (€)	499	505	513	523
Income elasticity of demand $E_{Q,I}$	0	-6.09857	1.745349	3.6342233

Table 4 Total number of passengers in the Slovak Republic and disposable income - one member of household

Based on the calculated coefficients of income elasticity of demand we can see that the demand for one member household in 2012 was slightly elastic. Disposable income increased but the number of passenger was decreasing. The reason could be increasing passenger demand for quality of transport services in the relation to prices for transport. On the other hand, in 2013 the demand was elastic and the significant change came in year 2014 when the demand was even more elastic – disposable income and number of passengers was increasing. As the increase of disposable income was not so high we can say that the rapid increase in number of passengers was caused by the free of charge fare for selected passengers introduced in November 2014.

Year	2011	2012	2013	2014
Number of transported passengers (mil. pass.)	47.531	44.698	46.064	49.272
Disposable income - household with two dependent children (€)	553	610	584	594
Income elasticity of demand $E_{Q,I}$	0	-0.74362	-0.62988	4.13235

Table 5 Total number of passengers in the Slovak Republic and disposable income - household with two dependent children

Based on the calculated coefficients of income elasticity of demand (household with two depended children) we can see smaller change of elasticity compared to the income elasticity of demand in one-member household. The demand 2012 was slightly elastic – income of household increased, but number of transported passengers was decreasing. The increase of disposable income does not influence the demand for transport services in rail passenger transport. In 2013 coefficient of income elasticity of demand was also just slightly elastic, but the change in income and number of passengers was different – income of household decreased and number of passengers in the Slovak Republic increased. The decrease in disposable income did not have any influence on the performance in rail passenger transport in that year. In 2014 the coefficient of income elasticity of demand was positive, i.e. demand for transport services was elastic – income and performance of rail transport were increasing. This growth was also influenced by free of charge fare for selected passengers introduced November 2014.

Conclusion

Free of charge fare for selected groups of passengers in rail transport changed the demand for transport services. Also it was necessary to take steps to maintain quality standards for paying passengers. Demand for transport services is perfectly inelastic; quantity demanded does not change with the change of price that also did not change. However, this fact does not apply to the selected category of passenger since November 2014. After that date the demand for transport services is perfectly elastic in the selected category of passengers – change of price influenced the number of transported passengers. Based on the current statistics the free of charge transport influences the major statistical indicators: number of passengers, sale of seat reservations, 42% share of free of charge passengers in 2015, decrease in revenues and increase of average transported distance by approx. 4.3 km.

Comparison of the performance of rail passenger transport in the Slovak Republic and disposable income in one-member household is – 0.23967 and we can say that the demand for transport services in 2011-2014 shows low elasticity. On the other hand, in four-member household income elasticity of demand for transport services is 0.919618. The change in income impacted the percentage change of performance in rail passenger transport.

Acknowledgements

The paper is supported by the VEGA Agency by the Project 1/0095/16 "Assessment of the quality of connections on the transport network as a tool to enhance the competitiveness of public passenger transport system" that is solved at Faculty of Operations and Economics of Transport and Communication, University of Žilina.

References

[1] Černá, L., Daniš, J., and Ponický, J.: Legislative changes in the rail public passenger transport in Slovakia In: *Railway transport and logistics, Journal of railway transport, logistic and management*, ISSN 1336-7943

[2] Dolinayová, A., Černá, L., and Daniš, J.: The present state of the price regulation of public railway transport in Slovakia, In: *Procurements in public transport: opportunities and pitfalls: proceeding of Telč 2015*, Brno: Masaryk Univerzity, 2015, ISBN 978-80-210-8003-4

[3] Gašparík, J., Abramivič, B., and Halás, M.: New graphical approach to railway infrastructure capacity analysis In: *Promet - Traffic& Transportation: scientific journal on traffic and transportation research,* Vol. 27, no. 4 (2015), p. 283-290., ISSN 0353-5320

[4] Gašparík, J., Stopka, O., and Pečený, L.: Quality evaluation in regional passenger rail transport In: *Naše more = Our sea: znanstveno-stručni časopis za more i pomorstvo*, Vol. 62, Iss. 3 (2015), p. 114-118., ISSN 0469-6255

[5] Gnap, J., and Sedláková, I.: Cenová elasticita dopytu v cestnej nákladnej doprave, In: *Doprava a spoje, electronically journal of Faculty of operation and economics of transport and communications*, Volume 2, Number 1, p. 17-25, 2006 ISSN 1336-7676

[6] Konečný, V.: Príjmová elasticity dopytu po soosbnej doprave v SR, In: *Doprava a spoje, electronically journal of Faculty of operation and economics of transport and communications*, Volume 3, Number 1, p. 30-37, 2007, ISSN 1336-7676.

[7] Molnárová, M.: *Matematika 1 a jej využitie v ekonómii*, 1st ed., Košice, TU, 2012. ISBN 9788055311685

[8] State trains transported more than 57 million passenger last year, In: *SME journal*, http:/ /ekonomika.sme.sk /c/20112857/statne-vlaky-vlani-prepravili-vyse-57-milionov-cestujucich.html#ixzz4CXSg7DNj, online 08.03.2016

[9] Yearbook of transport, Posts and telecommunications in 2015, Statistic office of Slovak Republic, https://www7.statistics.sk/PortalTraffic/fileServlet?Dokument=5959eb33-277c-4933-af2c-d78890b9fca0, online 01.04.2014

Finite-sample behavior of GLFP-based estimators for EIV regression models: a case with restricted parameter space

Michal Černý[1]

Abstract. We discuss the finite-sample properties of a recent consistent estimator for structural Errors-In-Variables (EIV) linear regression models with uniformly bounded error distributions. The estimator can be formulated as a family of generalized linear-fractional programs (GLFP). Here we investigate the special case when the parameter space is restricted to a known orthant of $\mathbb{R}^p$, where p is the number of regression parameters. In this case, the estimator reduces to a single GLFP and is thus computable in polynomial time. We perform a simulation study with independent uniformly distributed errors. We find out that the asymptotic convergence of the estimator to the true value and the reduction of its variance (which hold asymptotically by theory) can be empirically observed even for small datasets, but the speed of convergence heavily depends on p.

Keywords: Errors-in-Variables regression, bounded errors, generalized linear-fractional programming.

JEL classification: C46,C61
AMS classification: 62J05,62H12

1 Introduction

In Errors-in-Variables (EIV) regression we assume the regression relationship

$$y = X\theta + \varepsilon,$$

where the design matrix X is unobservable; we can only observe its contaminated form

$$Z = X + \Xi,$$

where Ξ is a matrix of error terms in the (observations) of regressors. So, observable data are (Z, y) and the task is to estimate the vector of regression parameters θ.

In our context, the matrix X is assumed to be stochastic (and the EIV model is called *structural EIV model*). The stochastic terms involved in the model are (X, Ξ, ε) and a "good" estimation method for the vector of regression parameters θ depends on particular assumptions on (X, Ξ, ε). Traditional estimation methods, such as Total Least Squares, are discussed e.g. in [3]; see also [4].

In the entire text, n stands for the number of observations and p stands for the number of regression parameters.

2 A recent estimation method for uniformly bounded error distributions

In this paper we discuss our recent estimation method [5, 6] (with J. Antoch and M. Hladík) based on Generalized Linear-Fractional Programming. The method yields a consistent estimator for the case of uniformly bounded errors (Ξ, ε). Recall that such types of errors are discussed in literature in various frameworks, see e.g. [1, 2].

[1]University of Economics, Prague, Department of Econometrics, Winston Churchill Square 4, 13067 Prague, Czech Republic, cernym@vse.cz

Before we state the assumptions rigorously (but they are somehow difficult-to-read in this form), we can summarize them informally: we assume that all errors in (Ξ, ε) are bounded by a uniform constant γ (which is unknown and is to be estimated) and we assume that if $n \to \infty$, then we can find (with probability tending to one) a line i_0 of the matrix (Ξ, ε) where the errors are very close to the bounds $\pm\gamma$, for every prescribed $(p+1)$-tuple of signs $\pm$. For example, if $p = 2$ and we have a sufficiently high number of observations, we assume that with a high probability we can find a line i_0 where the three errors $\Xi_{i_0,1}, \Xi_{i_0,2}, \varepsilon_{i_0}$ are close to $+\gamma, +\gamma, +\gamma$, respectively; we can also find another line i_0 where the three errors $\Xi_{i_0,1}, \Xi_{i_0,2}, \varepsilon_{i_0}$ are close to $-\gamma, +\gamma, +\gamma$; and so on, such a line can be found with a high probability for each possible choice out of the eight possibilities $(\pm\gamma, \pm\gamma, \pm\gamma)$.

Now we will be more formal. Let us assume:

(a) There exists (an unknown) constant $\gamma \geq 0$, called *error radius*, such that

  (a1) $|\varepsilon_i| \leq \gamma$ a.s., $i = 1, \ldots, n$,

  (a2) $|\Xi_{ij}| \leq \gamma$ a.s., $i = 1, \ldots, n$, $j = 1, \ldots, p$.

(b) Let $\|\cdot\|$ be a fixed vector norm. Assume that

$$\forall \alpha > 0 \; \exists c > 0 \; \forall u \in \mathbb{R}^p \text{ s.t.} \|u\| = 1 : \lim_{n \to \infty} \Pr[A_n(\alpha, c, u)] = 1,$$

where A_n is the following event: $\exists i_0 \in \{1, \ldots, n\}$ such that

  (b1) $|x_{i_0}^{\mathrm{T}} u| \geq c$,

  (b2) $-\operatorname{sgn}(x_{i_0}^{\mathrm{T}} u) \cdot \varepsilon_{i_0} \geq \gamma - \alpha$,

  (b3) $\operatorname{sgn}(x_{i_0}^{\mathrm{T}} u) \cdot \operatorname{sgn}(\theta_j + u_j) \cdot \Xi_{i_0 j} \geq \gamma - \alpha$, $j = 1, \ldots, p$,

where x_i^{T} is the ith row of X and

$$\operatorname{sgn}(\xi) = \begin{cases} 1, & \text{if } \xi \geq 0, \\ -1, & \text{if } \xi < 0. \end{cases}$$

Theorem 1 ([5, 6])**.** *Let $z_1^{\mathrm{T}}, \ldots, z_n^{\mathrm{T}}$ be the rows of Z. Given a sign vector*

$$s \in \{\pm 1\}^p, \tag{1}$$

consider the generalized linear-fractional program

$$\widehat{\gamma}_s^n = \min_{\theta \in \mathbb{R}^p} \left\{ \max_{\substack{i \in \{1,\ldots,n\} \\ k \in \{0,1\}}} \frac{(-1)^{1-k} z_i^{\mathrm{T}} \theta + (-1)^k y_i}{e^{\mathrm{T}} D_s \theta + 1} \;\middle|\; D_s \theta \geq 0 \right\}, \tag{2}$$

where $e = (1, \ldots, 1)^{\mathrm{T}}$ and $D_s = \operatorname{diag}(s)$. Let $\widehat{\theta}_s^n$ be the argmin of (2). Furthermore, let

$$\widehat{\gamma}^n = \min_{s \in \{\pm 1\}^p} \widehat{\gamma}_s^n \tag{3}$$

and let s^* be the argmin of (3). Now

$$\widehat{\gamma}^n \xrightarrow{P} \gamma \quad \text{and} \quad \widehat{\theta}^n := \widehat{\theta}_{s^*}^n \xrightarrow{P} \theta \quad \text{as} \quad n \to \infty.$$

□

Theorem 1 tells us that $(\widehat{\gamma}^n, \widehat{\theta}^n)$ is a consistent estimator of (γ, θ) and that the estimator can be computed by solving 2^p generalized linear-fractional programs (2), one GLFP per one choice of signs in (1). (Recall that generalized linear-fractional programming is solvable in polynomial-time by interior-point methods, see [7]).

3 An efficiently solvable case

Now we turn our attention to the case when θ is a priori known to be in a particular orthant of $\mathbb{R}^p$. Or, in other words, we assume that the *signs* of regression parameters are known in advance. Or we can also say that the parameter space Θ is restricted to a particular orthant

$$\Theta = \{\xi : D_{s_0}\xi \geq 0\} \ni \theta, \tag{4}$$

where $s_0 \in \{\pm 1\}^p$ is a *known* sign vector. Such cases are frequent in practice: many practical regression models have the property that we are able to say in advance whether a given regression parameter should be positive or negative. Now we get a "simpler" version of Theorem 1:

Corollary 2. If the parameter space has the form (4), consider the generalized linear-fractional program

$$\widehat{\gamma}^n = \min_{\theta \in \mathbb{R}^p} \left\{ \max_{\substack{i \in \{1,\ldots,n\} \\ k \in \{0,1\}}} \frac{(-1)^{1-k} z_i^{\mathrm{T}}\theta + (-1)^k y_i}{e^{\mathrm{T}} D_{s_0}\theta + 1} \;\middle|\; D_{s_0}\theta \geq 0 \right\}. \tag{5}$$

Let $\widehat{\theta}^n$ be the argmin of (5). Then

$$\widehat{\gamma}^n \xrightarrow{P} \gamma \text{ and } \widehat{\theta}^n \xrightarrow{P} \theta \text{ as } n \to \infty.$$

$\square$

Corollary 3. The estimates $(\widehat{\gamma}^n, \widehat{\theta}^n)$ can be computed from (Z, y) in polynomial time. $\square$

4 A simulation study

Currently we do not have a theorem on the speed of convergence of $(\widehat{\gamma}^n, \widehat{\theta}^n)$ to the true values (γ, θ). The speed certainly depends on the properties of $(X, \Xi, \varepsilon, \gamma, \theta)$. Observe that Assumptions (a) and (b) are very general; by the way, they admit many possible distributions and dependence structures (indeed, they require neither independence, nor zero means, nor identical distributions).

Here we restrict ourselves to the following case:

- rows of X are unit vectors;
- all errors in (Ξ, ε) are independently sampled from the uniform distribution on $(-\gamma, \gamma)$ with $\gamma = 1$;
- $\theta = e$;
- the number of simulated observations is $n \in \{60, 120, 180, 240, 300, 360\}$;
- the number of parameters is $p \in \{2, 3, 4\}$;
- for each (n, p) we perform 50 simulations.

In this setup we measure empirically the speed of convergence of $\widehat{\gamma}^n$ to the true value $\gamma = 1$ and the speed of convergence of $\widehat{\theta}^n$ in terms of the error $\|\widehat{\theta}^n - \theta\|_2$, for various sample sizes n and numbers p of parameters. We also measure the empirical standard error of $\widehat{\gamma}^n$ and $\|\widehat{\theta}^n\|_2$.

5 Results and conclusions

Results are summarized in Figure 1 ($p = 2$), Figure 2 ($p = 3$) and Figure 3 ($p = 4$).

We can observe that in all cases, the value of $\widehat{\gamma}^n$ indeed converges to the true value γ, but the speed of convergence is the slower the higher is p; moreover, the standard error of $\widehat{\gamma}^n$ also grows with p. The speed of convergence of $\widehat{\theta}^n$ to θ is also low; we can see that the mean error for $n = 400$ is only slightly lower than for $n = 100$ in all cases $p = 2, 3, 4$. Thus we can conclude that the asymptotic result of Theorem 1 (and Corollary 2) works for a quite high number of observations.

Notwithstanding, the convergence properties assured by Theorem 1 (and Corollary 2) are apparent even for small-sample datasets.

Finally, Fig. 4 depicts the distributions of $\widehat{\gamma}^n$ and $\|\widehat{\theta}_n - \theta\|_2$ for $p = 2$ and $n = 100$.

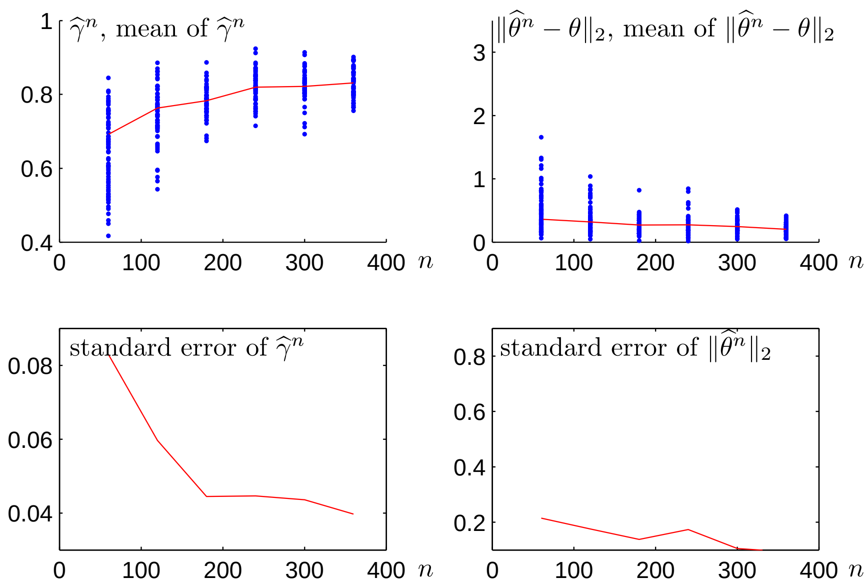


Figure 1 Simulation results for $p = 2$. True values: $\gamma = 1$, $\theta = (1, 1)^{\mathrm{T}}$.

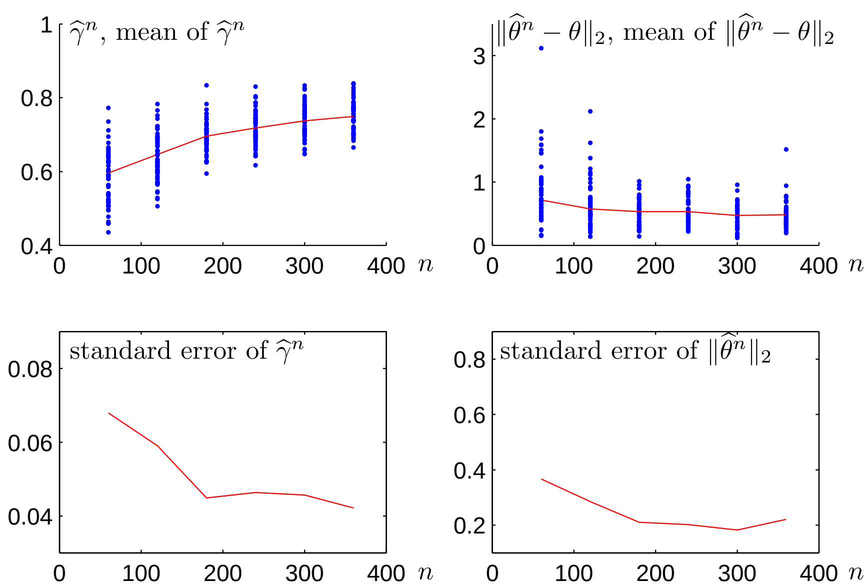


Figure 2 Simulation results for $p = 3$. True values: $\gamma = 1$, $\theta = (1, 1, 1)^{\mathrm{T}}$.

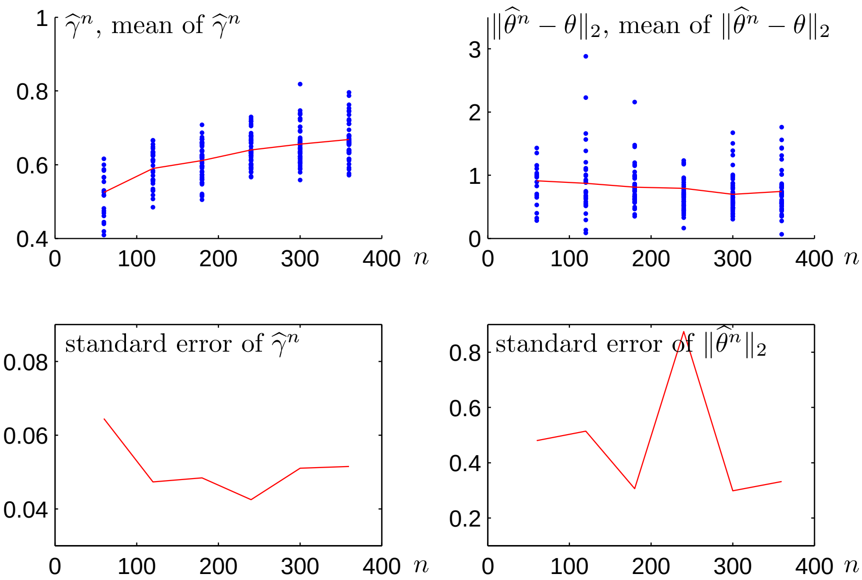


Figure 3 Simulation results for $p = 4$. True values: $\gamma = 1$, $\theta = (1, 1, 1, 1)^{\mathrm{T}}$. The high standard error for $n = 240$ is caused by presence of an outlier (not visible in upper right figure).

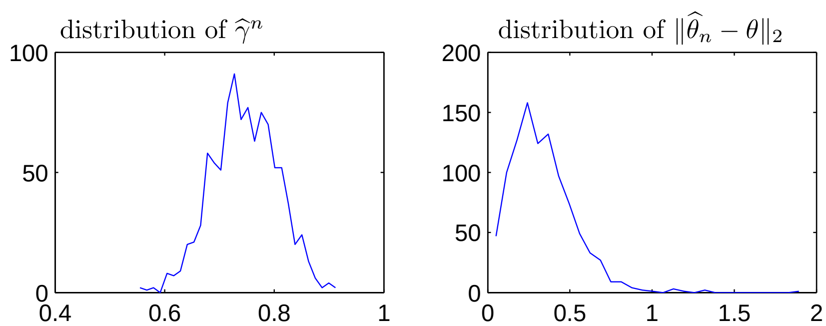


Figure 4 Histograms of simulated values $\widehat{\gamma}^n$ (left chart) and $\|\widehat{\theta}_n - \theta\|_2$ (right chart) for $p = 2$ and $n = 100$ with 1000 simulations.

Acknowledgment. The work was supported by the University of Economics, Prague under grant F4/54/2015 of the Internal Grant Agency (IGA).

References

[1] Bouda, M. and Formánek, T. Housing sector-specific DSGE model with applications to Czech and Slovak economies. *Ekonomický časopis* **62** (8), 805–822, 2014.

[2] Černý, M. Binary segmentation and Bonferroni-type bounds. *Kybernetika* **47** (1), 38–49, 2011.

[3] Cheng, C. I. and Van Ness, J. W. *Regression with Measurement Error.* Oxford University Press, New York, 1999.

[4] Kukush, A., Markovsky, I. and Van Huffel, S. *On consistent estimators in linear and bilinear multivariate Errors-In-Variables models.* In: Van Huffel, S. and Lemmerling, P. (eds.). *Total Least Squares and Errors-In-Variables Modeling: Analysis, Algorithms and Applications.* SIAM, Philadelphia, 155–164, 2002.

[5] Hladík, M. and Černý M. Total Least Squares and Chebyshev norm. *Procedia Computer Science* **51**, 1791–1800, 2015.

[6] Hladík, M., Černý, M. and Antoch, J. Linear regression with bounded errors in data: Total "Least Squares" with Chebyshev norm. Submitted, 2016. Preprint: `http://nb.vse.cz/~cernym/tls.pdf`.

[7] Nesterov, Y. E. and Nemirovski, A. S. An interior-point method for generalized linear-fractional programming. *Mathematical Programming* **69** (1B), 177–204, 1995.

Capacited Vehicle Routing Problem with Time Restriction Using Minimal Number of Vehicles

Zuzana Čičková[1], Ivan Brezina[2], Juraj Pekár[3]

Abstract: Classical capacited vehicle routing problem with time windows enables finding optimal set of routes of vehicle (vehicles) in order to serve given set of customers within a given time period. This paper is focused on a generalized problem that allows determining the minimal number of vehicles taking into account time limit when a vehicle is used (consider maximal working hours of a driver) and secondly also goal of minimizing total time of the distribution (based on lexicographic optimization). The route of vehicles are determined in such a way that one vehicle can pass through the center more than once but provided an additional service time at the center. Although the models dealing with multiple vehicles rides are known, proposed model uses only the binary variables with two indices instead of the commonly used three-indexed variables. The model is implemented in system GAMS and illustrative example is given.

Keywords: Vehicle Routing Problem, Time Restrictions, Minimal Number of Vehicles, Lexicographic Optimization

JEL Classification: C02, C61
AMS Classification: 90C11, 90B06

1 Introduction

The classical vehicle routing problem (VRP), also known as the capacitated VRP (CVRP), designs optimal set of routes aimed to serve a set of customers with a certain demand, where each vehicle travels exactly one route, each vehicle has the same characteristics and there is only one origin called depot ([1], [2], [4], [6], [8]). It is assuming the known shortest cost (time or distance) between origin and each customer's location, as well as between each pairs of customer's location. The goal is to find the optimal shortest route (starting and ending at the origin) for a vehicle (vehicles) so that each customer demand is met by exactly one vehicle (all the demands are met in full). The capacity of vehicle (or fleet of vehicles) is (are) known (if more than one vehicle are used, the same capacity of all of them is supposed) and that capacity must not be exceeded.

The VRP is an important combinatorial optimisation problem. Toth and Vigo have reported that the use of computerised methods in distribution processes often results in savings ranging from 5% to 20% in transportation costs and in [9] describe several case studies where the application of VRP algorithms has led to substantial cost savings.

Many of modification CVRP are known (e.g. [3], [5]), for all we list these: inventory routing, combinations of scheduling and routing, multi-echelon routing, multi-dimensional loading problem and routing with cross-docking, heterogeneous fleet VRP, also known as the mixed fleet VRP, VRP with time windows, VRP with pickup and delivery, VRP with backhauls, multi depot VRP, periodic VRP. Further on we mentioned two problems, which are relevant to model presented below: distance-constrained capacitated vehicle routing problem (DCVRP), where capacity restriction is replaced by a maximum length or by a time constraint and vehicle routing problem with multiple routes (VRPM), which consists in determining the routing of a fleet of vehicles where each vehicle can perform multiple routes at a specific time horizon. Using of those problems is relevant in applications where the duration of each route is limited; it can found when perishable goods are transported.

The above modification of the classical CVRP are usually aimed to find a minimum cost of routes, but do not reflect the problem how to set minimal number of vehicles to be able realizing delivery. In this article the authors present a way how to modify the above problems to identify the minimum number of vehicles which must be

[1] *Department of Operations Research and Econometrics, Faculty of Economic Informatics, University of Economics in Bratislava, Dolnozemská cesta 1, 852 35 Bratislava*, e-mail: zuzana.cickova@euba.sk
[2] *Department of Operations Research and Econometrics, Faculty of Economic Informatics, University of Economics in Bratislava, Dolnozemská cesta 1, 852 35 Bratislava*, e-mail: ivan.brezina@euba.sk.
[3] *Department of Operations Research and Econometrics, Faculty of Economic Informatics, University of Economics in Bratislava, Dolnozemská cesta 1, 852 35 Bratislava*, e-mail:, juraj.pekar@euba.sk.

available for distribution. Presented problem is based on lexicographic optimization, where the first priority is to determine the minimal number of vehicles and the second priority is aimed at minimizing of total time of the service. It considers not only vehicle capacity limit, but also vehicle usage time restriction (e.g. limiting the working hours of drivers). However, if the next customer service is not possible because of exceeding vehicle capacity limit but time restriction still allows its further use, the vehicle can be reloaded. Thus the time of loading is added to the total time of vehicle route. Proposed model enables using only the binary variables with two indices, instead of the commonly used three-indexed variables.

2 Vehicle routing problem and open vehicle routing problem

Various routing problems can be described by mathematical models using following notation: Let $N = \{1, 2, ...n\}$ be the set of served nodes (customers) and let $N_0 = N \cup \{0\}$ be a set of nodes that represents the customers also with the origin (depot). A shortest time distance d_{ij} is associated with pairs $i, j \in N_0$, $i \neq j$. One way how to mathematically describe routing problems is using binary programming formulations. The models involve binary variables x_{ij} $(i, j \in N_0, i \neq j)$ that enable to model if the node i precedes node j in a route of the vehicle $x_{ij} = 1$ and $x_{ij} = 0$ otherwise. Certain demand q_i, $i \in N$, which has to be met from the initial node ($i = 0$), is associated with each customer. The distribution is performed using a vehicles with a certain capacity (g). The goal is to identify such routes of vehicles where the total traveled distance (or time) is as low as possible with respect to the following restrictions: the origin represents initial node and also the final node of every route, from this node the demands q_i, $i \in N$ of all the other nodes are met (in full), each node (except origin) is visited exactly once and total demand on route must not exceed the capacity of the vehicle (g). The model implicitly assumes that $q_i \leq g$ for all $i \in N$, i.e. the demand of each customer does not exceed the capacity of the vehicle. Further on, the variables $u_i, i \in N$ that based on well-known Miller-Tucker- Zemlin's formulation, e.g. Miller et al. ([7]) are employed. Those variables represent cumulative demand of customers on one particular route.

Further on suppose following: let $K = \{1, 2, ...k\}$ be the set of vehicle, where k represent maximal number of vehicles (equal to maximal value of shuttle routes (n), or it may be subjected based on additional information). Suppose the service can be performed using vehicles on the understanding that one vehicle can also made more routes (starting at ending in the origin), but consider vehicle using time limit designated as P. But if the vehicle returns to the origin due to violation of capacity limit, and it is able to serve the nodes on the next route than the service time at the center (s) is added to the total time of vehicle route. The goal is to determine such minimal number of vehicles as possible with respect to minimizing of total time of service.

Besides the above mentioned variables let us using following:

- variables $t_i, i \in N$ are also based on Miller-Tucker- Zemlin's formulation, but they will represent the total cumulative time to corresponding i-th customer (including),
- variables z_{il}, $i \in N, l \in K$ represent the total time of such particular route of l-th vehicle, in which the i-th node served as the last,
- variables v_{il}, $i \in N, l \in K$ are binary variables that represent if the particular route, which ends with i-th nod, is served by l-th vehicle,
- variables c_l, $l \in K$ represent if the l-th vehicle is in use or not.

Now recapitulate the model parameters more clearly:

n – number of customers (served nodes),
$N = \{1, 2, ...n\}$ – set of customers (served nodes),
$N_0 = N \cup \{0\}$ – set of customers and the origin,
k – maximal number of vehicles,
$K = \{1, 2, ...k\}$ – set representing vehicles,
d_{ij}, $i, j \in N_0$, $i \neq j$ – shortest time moving from node i to node j,
q_i, $i \in N$ – demand of i-th customer,
g – capacity of vehicles,
s – service time at the centre,
P – time limit of vehicle usage,
M – big positive number.

The model describing aforementioned situation deals with those variables:

$$x_{ij} \in \{0,1\}, \; i, j \in N_0, i \neq j \tag{1}$$

$$v_{il},\ i \in N, l \in K \tag{2}$$
$$u_i \geq 0,\ i \in N,\ u_0 = 0 \tag{3}$$
$$t_i \geq 0,\ i \in N,\ t_0 = 0 \tag{4}$$
$$z_{il},\ i \in N, l \in K \tag{5}$$
$$0 \leq c_l \leq 1,\ l \in K$$

where objectives can be written as follows:

$$lex \min \left\{ \sum_{l \in K} c_l, \sum_{i \in N} \sum_{l \in K} z_{il} + s \sum_{l \in K} \left(\sum_{i \in N} v_{il} - 1 \right) \right\} \tag{6}$$

subject to:

$$\sum_{i \in N_0} x_{ij} = 1, \quad j \in N, \quad i \neq j \tag{7}$$

$$\sum_{j \in N_0} x_{ij} = 1, \quad i \in N, \quad i \neq j \tag{8}$$

$$u_i - u_j + q_j \leq g(1 - x_{ij}), \quad i \in N_0, \quad j \in N, \quad i \neq j \tag{9}$$

$$q_i \leq u_i \leq g, i \in N \tag{10}$$
$$t_i - t_j + d_{ij} \leq M(1 - x_{ij}), \quad i \in N_0, \quad j \in N, \quad i \neq j \tag{11}$$
$$t_i - z_{il} + d_{i0} \leq M(1 - v_{il}), \quad i \in N, l \in K \tag{12}$$
$$x_{i0} = \sum_{l \in K} v_{il}, \quad i \in N \tag{13}$$
$$v_{il} \leq c_l, \quad i \in N, l \in K \tag{14}$$
$$\sum_{i \in N} z_{il} + s \cdot \left(\sum_{i \in N} v_{il} - 1 \right) \leq P, \quad l \in K \tag{15}$$

Objective function (6) determines both of objectives: firstly the minimization of number of vehicles and secondly minimization of total time of service. Equations (7) and (8) ensure that each customer (except the origin) is visited exactly ones. Equations (9) and (11) are anti-cyclical conditions that prevent the formation of such sub-cycles which do not contain an initial node ($i = 0$). Equations (9) also ensure calculating of current load of vehicles in its route to i-th customer (including) and equations (10) ensure respecting maximal capacity of vehicle. The cumulative time of vehicle particular route is calculated by equations (11). Equations (12) enable calculating total time of such particular route of vehicle, in which the i-th node served as the last if it is served by l-th vehicle. Equations (13) ensure assignment the particular route ending with i-th node to exactly one vehicle. Equations (15) enable respecting time limit of vehicle usage (also if the vehicle pass through origin several time).

4 Illustrative Example

Consider scheduling in network consisting of origin from where 8 nodes (customers) need to be served. Values of input parameters were set as follows:

n =8, $N = \{1, 2, ...8\}$, $N_0 = N \cup \{0\}$ - number and sets of nodes (customers and also origin),

k =8, $K = \{1, 2, ...k\}$ - maximal number of vehicles (shuttle routes) and set representing vehicles,

$q_i = 10,\ i \in N$ - vector of customers' demands,

g =30 – capacity of vehicles,

s = 10 – service time at the center,

P = 80 – time limit of vehicle usage,

$\mathbf{D} = \{d_{ij}\},\ i, j \in N_0, i \neq j$ – matrix of shortest time moving from node i to node j

$$\mathbf{D} = \begin{bmatrix} 0 & 10 & 4 & 8 & 8 & 3 & 8 & 17 & 16 \\ 10 & 0 & 6 & 9 & 10 & 14 & 19 & 27 & 26 \\ 4 & 6 & 0 & 12 & 12 & 8 & 13 & 21 & 20 \\ 8 & 9 & 12 & 0 & 7 & 11 & 16 & 25 & 24 \\ 8 & 10 & 12 & 7 & 0 & 11 & 16 & 25 & 24 \\ 3 & 14 & 8 & 11 & 11 & 0 & 11 & 14 & 13 \\ 8 & 19 & 13 & 16 & 16 & 11 & 0 & 15 & 8 \\ 17 & 27 & 21 & 25 & 25 & 14 & 15 & 0 & 7 \\ 16 & 26 & 20 & 24 & 24 & 13 & 8 & 7 & 0 \end{bmatrix}$$

Model (1) – (15) was implemented in software GAMS (solver Cplex 12.2.0.0) on PC with Intel ® Core ™ i7-3770 CPU with a frequency of 3.40 GHz and 8 GB of RAM under MS Windows 8. The solution was realized according to both criteria. In first run minimal number of vehicles required to distribution was obtained (two). The second run minimize the total time of transportation reflecting computed number of vehicles from the first run.

Computed values of variables are those:

variables (1): $x_{01} = x_{02} = x_{06} = x_{13} = x_{25} = x_{34} = x_{40} = x_{50} = x_{68} = x_{70} = x_{87} = 1$, remaining variables were set to 0,

variables (2): $v_{41} = v_{51} = v_{72} = 1$, remaining variables were set to 0,

variables (3): $u_0 = 10, u_1 = 10, u_2 = 10, u_3 = 20, u_4 = 30, u_5 = 20, u_6 = 10, u_7 = 30, u_8 = 20$,

variables (4): $t_0 = 0, t_1 = 10, t_2 = 4, t_3 = 19, t_4 = 29, t_5 = 12, t_6 = 8, t_7 = 27, t_8 = 16$,

variables (5): $z_{41} = 34, z_{51} = 15, z_{72} = 40$, remaining variables were set to 0,

variables (6): $c_1 = c_2 = 1$, all the other were set to 0.

It is clear that two vehicles are used (but three routes are realized). The results are given in Table 1.

Route	Sequence of nodes	Number of distributed units	Total time of particular route
Route 1	0-1-3-4-0	30	34
Route 2	0-2-5-0	20	15
Route 3	0-6-8-7-0	30	40

Table 1 Summarization of solution. *Source:* Own compilation.

This distribution requires 99 time units. Total duration is calculated as the sum of the individual distribution routes (89 time units) to which service time at the origin is added (when using first vehicle which passed through center one time).

Conclusion

This paper considers the modification of capacited vehicle routing problem (CVRP). Modifications of the classical CVRP are usually aimed to find a minimal cost of routes. The present model considers two objectives based on their priority (minimizing of number of vehicles at first, minimizing total time of distribution at second) taking into account time limit when the vehicle is used. When the next customer service is not possible because its service would exceed vehicle capacity but its time restriction still allows vehicle further usage, the vehicle can be reloaded. In such case the time of loading is added to the total time of vehicle route. The mathematical formulation was provided on the base of mixed integer programming (MIP). Software implementation was realized in GAMS and also illustrative example is given.

Acknowledgements

This paper is supported by the Grant Agency of Slovak Republic – VEGA, grant no. 1/0245/15 „Transportation planning focused on greenhouse gases emission reduction“.

References

[1] Čičková, Z., Brezina, I.: An evolutionary approach for solving vehicle routing problem. In: *Quantitative methods in economics multiple criteria decision making XIV*. IURA EDITION, Bratislava, 2008, 40–44.

[2] Čičková, Z., Brezina, I., Pekár, J.: *Solving the Real-life Vehice Routing Problem with Time Windows Using Self Organizing Migrating Algorithm*. Ekonomický časopis **61**(5/2013), 497–513.

[3] Čičková, Z., Brezina, I., Pekár, J.: *Open vehicle routing problem.* In: *Mathematical Methods in Economics 2014*. Olomouc: Faculty of Science, Palacký University, 2014, 124 – 127.

[4] Desrochers, M., Desrosiers, J., Solomon, M.: A New Algorithm for the Vehicle Routing Problem with Time Windows. *Operations Research*. **40** (1992), 342-354.

[5] Eksioglu, B., Vural, A. V., Reisman, A.: *The vehicle routing problem: A taxonomic review.* In: Computers & Industrial Engineering **57**(4/2009), 1472–1483.

[6] Fábry, J., Kořenář, V., Kobzareva, M.: Column Generation Method for the Vehicle Routing Problem – The Case Study. In: *Mathematical Methods in Economics 2011 [CD-ROM]*. Praha, PROFESSIONAL PUBLISHING, (2011), 140–144.

[7] Miller, C.E., Tucker, A.W., Zemlin, R.A.: *Integer programming Formulation of Traveling Salesman Problems*. Journal of the ACM, **7**(4/1960), 326-329.

[8] Palúch, S., Peško, Š.: *Kvantitatívne metódy v logistike*. EDIS vydavateľstvo ŽU, Žilina, 2006.

[9] Toth, P., Vigo, D.: *The Vehicle Routing Problem* (Monographs on Discrete Mathematics and Applica-tions) SIAM. Philadelphia, PA, 2002.

Long-Run Growth in the Czech Republic

Ondřej Čížek[1]

Abstract. The goal of the paper is to analyze the long-run growth in the Czech Republic and to shed some light on relevant questions regarding the influence of the current economic crisis on the long-run growth. The relevant question is the extent to which output has recovered from the current crisis as well as to quantify the amount of the permanent loss in output. These issues will be approached using a bivariate unobserved components model of output and unemployment which enables decomposition of variables into transitory and permanent component. The significant finding is that a transitory component of output has already recovered after the initial shock in the beginning of the crisis. However, the opposite is true for the permanent component. The paper presents evidence of a persistently decreased growth of the trend component of output due to which there is a significant permanent loss in the level of output. The paper estimates this loss to equal approximately to 20 % of the quarterly real GDP.

Keywords: unobserved components model, economic crisis, recovery, transitory and permanent effects.

JEL Classification: E32
AMS Classification: 91G70

1 Introduction

This paper estimates the long-run growth in the Czech Republic by applying bivariate unobserved components model of output and unemployment. There is currently highly discussed issue of long-term effects of the current global depression on the long-run growth which will be to certain extent discussed in this paper as well. Cerra, Saxena [4] discuss this topic from a perspective of regime switching models for the case of the Asian crisis of 1997. Ball [1] estimates the long-term effects of the global depression of 2008 on the level of output of OECD countries. The concept of potential output is applied in his paper. Ball uses OECD methodology when obtaining estimates of potential output, which is based on a production function approach and is described in Beffy et al. [3]. Evidence favoring the hypothesis of permanent effects of deep recessions on output is found by Barro [2] as well as Ball [1] and by many others which are cited in these papers. Similar results are found in the presented paper for the case of the Czech Republic.

The paper is organized as follows. The model is formulated in chapter 2. Data is described in section 3 and the subsequent chapter 4 deals with econometric estimation. Economic discussion is contained in the chapter 5. The final chapter 6 concludes.

2 Model

The model decomposing real GDP and unemployment rate into trend and a cycle component is presented in this chapter. The formulation is based on Clark's [5] unobserved components model which was summarized in a textbook treatment by Kim, Nelson [8]. The model equations are given as follows:

$$y_t = n_t + x_t , \tag{1}$$

$$n_t = g_{t-1} + n_{t-1} + v_t ,\ v_t \sim i.i.d.N\left(0,\sigma_v^2\right), \tag{2}$$

$$g_t = g_{t-1} + w_t ,\ w_t \sim i.i.d.N\left(0,\sigma_w^2\right), \tag{3}$$

$$x_t = \phi_1 \cdot x_{t-1} + \phi_2 \cdot x_{t-2} + e_t ,\ e_t \sim i.i.d.N\left(0,\sigma_e^2\right), \tag{4}$$

where y_t is the logarithm of real GDP,

[1] University of Economics in Prague, Department of Econometrics, W. Churchill Sq. 4, 130 67, Prague 3, Czech Republic, cizeko@vse.cz.

n_t is a stochastic trend component,

x_t represents a stationary cyclical component,

v_t, w_t, e_t are independent white noise processes.

The autoregressive process of order two was chosen in the equation (4). This is the most common assumption used in empirical literature when modelling cyclical variables as an autoregressive process of the second order is a parsimonious way to model cyclical dynamics.

This standard univariate model is extended into a bivariate model of real GDP and unemployment. The unemployment rate is decomposed into trend and a cycle as well as follows:

$$U_t = L_t + C_t, \tag{5}$$

$$L_t = L_{t-1} + \varepsilon_t,\ \varepsilon_t \sim i.i.d.N\left(0, \sigma_\varepsilon^2\right), \tag{6}$$

$$C_t = \alpha_0 \cdot x_t + \alpha_1 \cdot x_{t-1} + \alpha_2 \cdot x_{t-2} + \eta_t,\ \eta_t \sim i.i.d.N\left(0, \sigma_\eta^2\right), \tag{7}$$

where L_t is a trend component of unemployment rate,

C_t is a stationary component of unemployment rate,

ε_t, η_t are independent white noise processes.

The cyclical component C_t is assumed to be a function of current and past transitory components of real output which represents a version of Okun's law. The number of lags used in the equation (7) was chosen rather arbitrarily. This choice, however, is quite common in the empirical literature (Kim, Nelson [8]).

3 Data

Quarterly data for the Czech Republic from 1996 Q1 to 2015 Q4 were used in the application part of this article. All such data is available at the database of the Eurostat.

Unemployment rate u_t was obtained from the series „ Unemployment rate by sex and age - quarterly average, % [une_rt_q] ". The relevant age structure was chosen to be "from 24 to 74 years". This time series was already seasonally adjusted by Eurostat. The series was originally in percents, but it was then divided by 100 for practical purposes.

The name of the time series for the real GDP in the Eurostat database is "GDP and main components (output, expenditure and income) [namq_10_gdp]". This time series was seasonally and calendar adjusted by Eurostat. The series is measured in chain linked volumes (2010) in millions euro. This time series was transformed by logarithms in order to obtain empirical counterpart to the variable y_t.

4 Econometric estimation

The model is estimated by the method of maximum likelihood. The transition and measurement equation of the state space representation can be written as

$$\begin{bmatrix} y_t \\ U_t \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & \alpha_0 & \alpha_1 & \alpha_2 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} n_t \\ x_t \\ x_{t-1} \\ x_{t-2} \\ g_t \\ L_t \end{bmatrix} + \begin{bmatrix} 0 \\ \eta_t \end{bmatrix}, \tag{8}$$

$$\begin{bmatrix} n_t \\ x_t \\ x_{t-1} \\ x_{t-2} \\ g_t \\ L_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & \phi_1 & \phi_2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} n_{t-1} \\ x_{t-1} \\ x_{t-2} \\ x_{t-3} \\ g_{t-1} \\ L_{t-1} \end{bmatrix} + \begin{bmatrix} v_t \\ e_t \\ 0 \\ 0 \\ w_t \\ \varepsilon_t \end{bmatrix}. \tag{9}$$

The Kalman filter algorithm was applied to this state space representation in order to calculate the likelihood function which was maximized numerically using standard numerical optimization procedures implemented in Matlab. Algorithms implemented in Matlab by Čížek [6] were used for this purpose.

Estimation results are summarized in the following table 1.

	σ_v	σ_e	σ_w	ϕ_1	ϕ_2	α_0	α_1	α_2	σ_ε	σ_η
Estimate	0.0026	0.0055	0.0022	1.5396	-0.6345	-0.1170	-0.1649	-0.1117	0.0015	0.0002
Standard error	0.0016	0.0010	0.0006	0.0672	0.0661	0.0425	0.0568	0.0364	0.0005	0.0011

Table 1 Econometric estimates of the model (1996Q1-2015Q4)

Relatively high value of the parameter $\sigma_e = 0.0055$ indicates that significant portion of the quarter-to-quarter innovations in real GDP are cyclical. Similar results were obtained for the U.S. economy (Clark [5], Kim, Nelson [8]). Nonetheless, the standard error $\sigma_w = 0.0022$ representing the variability of the GDP (long-run) growth is approximately 10 times higher than that reported by Clark or Kim and Nelson.

High volatility of the Czech long-run economic growth in the studied period can be caused by the process of economic transformation. The Czech economy had been opening to the rest of the world. Many international trade barriers had been removed by the entrance to the European Union. Lots of economic reforms had been realized. By neoclassical growth terminology, these transformations could influence the Czech economy steady state and so they could increase the growth volatility. For these reasons, long-run growth volatility of the transition economy is higher than the volatility of a stable economy which rests in a steady state. This is one possible explanation for why the U.S. economic long-run growth is less volatile than the long-run growth in the Czech Republic.

One might also argue that higher volatility of the long-run growth might be caused by the current economic crisis as the above mentioned studies of the U.S. economy did not analyze the crisis period. Detailed analysis of these questions are beyond the scope of this paper. Nevertheless, my suggestion is that the current crisis is not the cause for a higher long-run growth volatility but is the cause for decreased values of the long-run growth in the crisis period. Some arguments will be given later in this paper.

Negative values of the parameters α_i, $i = 0,1,2$ are in line with a priory assumption. Statistical significance of these parameters confirms strong negative relationship between output and unemployment referred to as Okun's law.

5 Economic discussion

The following figure 1 plots the log of real GDP together with its trend and a cyclical component. The graph on the right-hand side illustrates that the cycle component (output gap) decreased dramatically in 2008 and 2009. Nonetheless, the output gap has recovered after this initial shock and has been improving since then. The output gap has been even positive since 2014 Q1. This finding is in contrast with earlier business-cycle studies of univariate trend-cycle model of real GDP in the U.S. economy. For example, Watson [11], Kim, Nelson [8] and Clark [5] attribute most of the variation of U.S. output to the cyclical component.

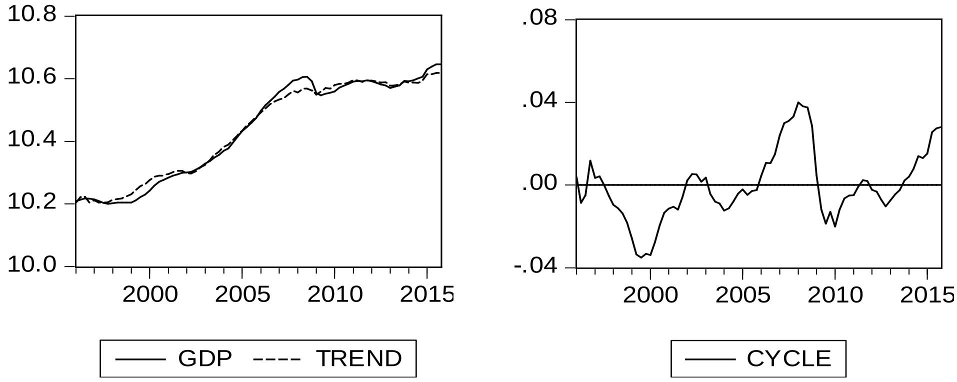


Figure 1 Log of real GDP y_t and its trend n_t and a cycle x_t component

Results found in this paper are similar to that found by Perron, Wada [10] who emphasized the importance of changes in the slope of the trend. Indeed, the graph in the left suggests that the current economic crisis is characterized by the change in the trend. This suggests that the huge output loss induced by the crisis is permanent and not only transitory. While the trend was upward-sloping from 1996 to 2008, it is practically constant from 2009 to 2015. This is confirmed in the figure 2 which shows that the quarter-to-quarter growth of the GDP trend component g_t has been very near to the value of zero since 2009. Specifically, the mean of the variable g_t in the pre-crisis time period from 1996 to 2008 is 0.0081 while in the post-crisis period from 2009 to 2015 the mean is 0.0018. These results are in line with other empirical studies analyzing the impact of the current global economic crisis (Barro [2], Ball [1]).

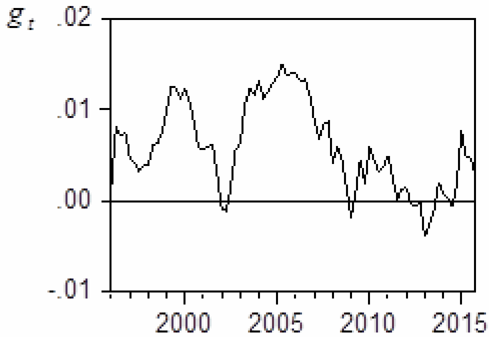


Figure 2 Long-run growth of the GDP trend component g_t

These findings suggest an adverse permanent influence of the crisis on the long-run economic growth. Nonetheless, this is only a suggestion. The rigorous evaluation of the influence of the crisis on the long-run growth is left for future research. Such a research would apply the growth theory according to which the growth rate depends on its determinants. Changing these determinants leads to a change in the long-run growth rate. The current economic crisis can be considered to be just one of many determinants of the long-run growth rate. Nonetheless, it is probably the case that the current economic crisis is indeed the most important factor which caused the decreased values of the long-run growth g_t after 2008. For this reason, it will be assumed for simplicity that the current crisis is the only factor which caused the lowered values of the long-run growth after 2008 which enables us to perform some illustrative calculations.

The following figure 3 distinguishes between the situations where the drop in the long-run growth is temporary and permanent. The graph in the left depicts the situation in which the decline in growth g_t is only temporary and lasts from 2009 to 2015. The graph in the right-hand side of the figure 3 shows the situation where the decline in growth g_t is permanent. The important thing to note is that even only a *temporal* decline in the growth g_t has a *permanent* long-run negative effect on the *level* of output. Therefore, the *level* of output in the Czech Republic will be permanently decreased even if the growth of the GDP trend component g_t would return to the pre-crisis values from now on. There are indeed certain signs that this might be the case as figure 2 shows that the growth of the GDP trend component g_t has been rising since 2013.

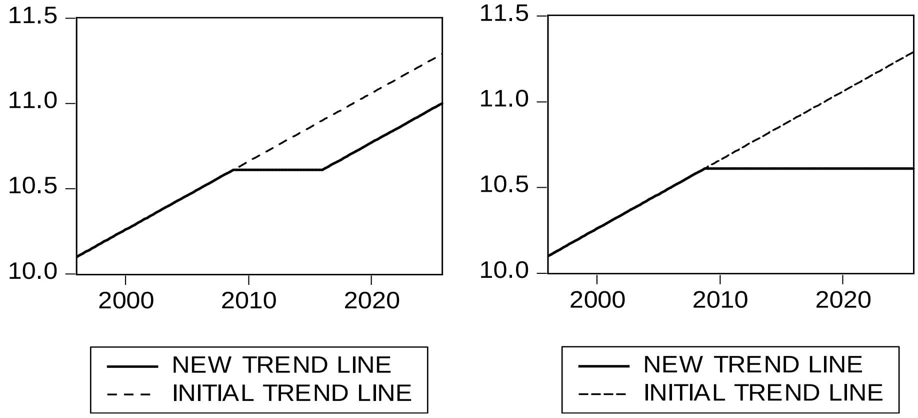


Figure 3 The distinction between the effect of the crisis on a level and a growth of (log) output

Let's perform some illustrative calculations in order to quantify the amount of permanent loss in the level of output due to the current economic crisis. As was already mentioned, the mean of the variable g_t in the pre-crisis time period was $g^1 = 0.0081$ and in the post-crisis period from 2009 Q1 to 2015 Q4 it was $g^2 = 0.0018$. The trend value of the (log) real output in 2008 Q4 was $n_{2008Q4} = 10.5627$. By neglecting the random error v_t, the equation (2) is modified as follows:

$$n_t = g^2 + n_{t-1},\ t = 2009\text{ Q1},...,2015\text{ Q4} \tag{10}$$

$$n_{t+k} = (k+1)\cdot g^2 + n_{t-1},\ t = 2009\text{ Q1},\ \text{k=27}. \tag{11}$$

The value of the trend in time period $t+k = 2015\text{ Q4}$ calculated according to (11) is $n_{2015Q4} = 10.6131$. Replacing g^2 by g^1 in the equation (11) would lead to the value of the trend in $t+k = 2015\text{ Q4}$ given by $n_{2015Q4} = 10.7895$. The estimated amount of the permanent loss for the quarter 2015 Q4 is therefore equal to $\exp(10.7895) - \exp(10.6131) = 7845$ millions euro. The real GDP in 2015 Q4 in the Czech Republic was 42023 millions euro. The calculated loss for 2015 Q4 thus represents approximately 20 % of the quarterly value of the real GDP. This loss is *permanent* in the sense that even if the growth of the time trend returned from $g^2 = 0.0018$ to the pre-crisis value of $g^1 = 0.0081$ for all the future time periods 2016 Q1, 2016 Q2,.... than the 20 % loss of the quarterly value of the real GDP would be experienced for all the future quarters 2016 Q1, 2016 Q2,.....

The following figure 4 documents the decomposition of the unemployment rate into the cyclical and trend component.

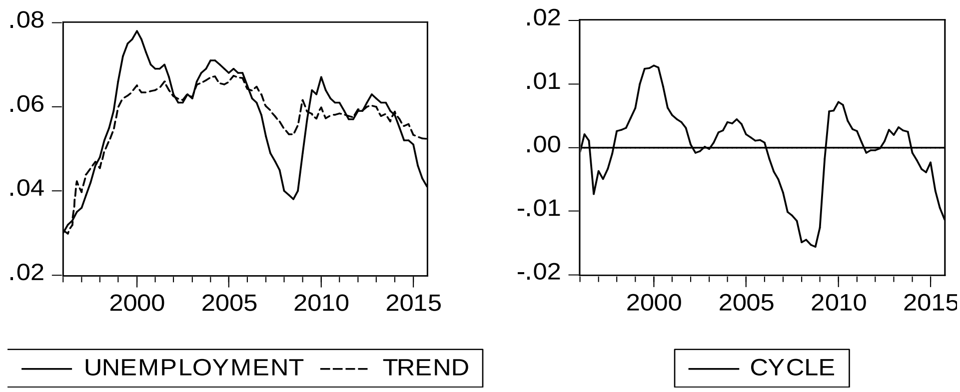


Figure 4 Unemployment rate U_t and its trend L_t and a cycle C_t component

The figure 4 shows that the cyclical component increased dramatically in the beginning of the crisis in 2009. Since then, however, the cyclical component of the unemployment rate has been decreasing steadily. The trend component has been practically constant since the beginning of the crisis which suggests that possible hysteresis effects haven't played important role yet. This supports the view that the effect of the current crisis on output might be better characterized by the left-hand side graph of the figure 3 rather than the graph depicted on the right side of the figure 3.

6 Conclusion

This paper finds that the long-run growth in the Czech Republic has been highly volatile and that it has decreased dramatically since 2008. The current economic crisis is viewed as an important factor for these lowered values. Nonetheless, detailed analysis along the lines of the growth theory is left for future research to confirm this hypothesis. The fall in the trend to its pre-crisis values has been almost as large as the fall in actual output. Consequently, the Czech Republic has definitely experienced severe long-term loss in output. The calculated permanent loss represents approximately 20 % of the quarterly value of the real GDP.

The methodology applied in this paper is standard and commonly used in the literature. Note however that measures of the output gap are always associated with a considerable level of uncertainty especially during the times of the current economic crisis. There is always considerable model uncertainty. Certain robustness check is made by studying other empirical studies dealing with the current economic crisis and its effects on output. Similar results of a huge long-term damage to output were found by Ball [1] as well as by many others cited in Ball's influential paper.

The model could also be expanded in many ways. Morley et al. [9] relax the presumption usually assumed in the literature on unobserved components model that there is no correlation between the shocks to the trend and the cycle. The authors find that relaxing this restriction makes the results even more strongly suggesting that there is a big long-term damage to output from recessions. Possible parameter instability due to the economic crisis could be taken into account by applying regime-switching methodology as by Cerra, Saxena [4]. Formánek, Hušek [7] investigate the behavior of GDP and unemployment in the Czech Republic taking into account spatial dependencies of its neighbors. Taking spatial dependencies explicitly into account could be an interesting modification of the presented model.

Acknowledgements

Financial support of VŠE IGA IG403036 is gratefully acknowledged by the author. Paper was processed with contribution of long term support of scientific work on Faculty of Informatics and Statistics, University of Economics, Prague (IP 400040).

References

[1] Ball, L. M.: Long-Term Damage from the Great Recession in OECD Countries. *NBER Working Paper 20185*, 2014.
[2] Barro, R. J.: Economic Growth in East Asia Before and After the Financial Crisis. *NBER Working Paper No. W8330*, Cambridge, Massachusetts: National Bureau of Economic Research, 2001.
[3] Beffy, P. O., et al.: New OECD Methods for Supply-Side and Medium-Term Assessments: A Capital Services Approach. *OECD Economic Department Working Paper 482*, 2006.
[4] Cerra, V., Saxena, S. C.: Did Output Recover from the Asian Crisis? *IMF Staff Paper* **52**(1), 2005, International Monetary Fund.
[5] Clark, P. K.: The Cyclical Component of U.S. Economic Activity. *The Quarterly Journal of Economics*, **102** (1987), 797-814.
[6] Čížek, O.: Macroeconometric Model of Monetary Policy (in Czech). Prague 2013. Dissertation thesis (Ph.D.). University of Economics, Prague, Faculty of Informatics and Statistics, Department of Econometrics, 2013-06-27.
[7] Formánek, T., Hušek, R.: The Czech Republic and its Neighbors: Analysis of Spatial Macroeconomic dynamics. In: *Mathematical Methods in Economics 2015* (MME 2015). Cheb, 09.09.2015 – 11.09.2015. Plzeň: University of West Bohemia, 2015, 190-195. ISBN 978-80-261-0539-8.
[8] Kim, C. J., Nelson, C. R.: *State Space Models with Regime Switching*. MIT Press, Cambridge, 1999.
[9] Morley, J. C, Nelson, C. R, Zivot, E.: Why Are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP So Different? *Review of Economics and Statistics*, **85**, 2003, 235-243.
[10] Perron, P., Wada, T.: Let's Take a Break: Trends and Cycles in US real GDP. *Journal of Monetary Economics*, **56**(6), 2009, 749-765.
[11] Watson, M.: Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*, **18**(1), 1986, 49-75.

Fractional Brownian Bridge as a Tool for Short Time Series Analysis

Martin Dlask[1]

Abstract. Traditional fractional stochastic processes represent suitable models for fractal analysis of long time series. However, due to their asymptotic behaviour, the estimation of Hurst exponent is often biased when the sample is too short. The novel approach is based on the construction of fractional Brownian bridge and thanks to its statistical properties and artificial extension to infinite length, it can be used for short time series investigation and resulting estimate was proven not to be burdened by bias. At first, the input signal is split into short stationary segments and the optimal interval length can be obtained via multiple statistical testing. Subsequently, the estimation of the Hurst exponent and its standard deviation is performed on the interval level. The methodology is applied to the stock market indices and based on the Hurst exponent variability in time, the decision about its predictability can be made. As a referential technique, the revisited zero-crossing method is presented and its performance is discussed in the context of obtained results.

Keywords: fractional Gaussian noise, fractional Brownian bridge, short time series, Hurst exponent, stock market indices

JEL classification: E44
AMS classification: 60G22, 62M10

1 Introduction

The dependence of time-series can be measured with Hurst exponent that determines its predictability and often carries more important information than autocorrelation. There are plenty of methods which are suitable for long time series analysis and provide unbiased estimation of Hurst exponent. However, the investigated sample has to be long enough to fulfil the prerequisites of the asymptotic methods.

The aim of the paper is to present a new method that can estimate Hurst exponent from short time series. Subsequently, the technique is used for analysis of stock market indices. Normally, the unbiasedness of this parameter can be guaranteed only when the investigated sample contains a lot of elements. However, the new methodology employs discrete signal that can be extended to infinite length using properties of fractional processes.

In fact, the majority of estimation methods are based on these fractional processes, that are continuous. Therefore the quality of estimate is strongly influenced by the amount of input data. Generally, the Hurst exponent estimate is accurate and its standard estimate decreases when the investigated time series has thousands of elements. The R/S method [6] can be considered as a very simple method and it is widely used till today in areas, where excessive precision is not required such as geography [1] or traffic flow modelling [14]. Utilizing statistical properties of fractional processes, Whittle estimator [13] is believed to be one of the most accurate approaches in the frequency domain, and with respect to their Holder continuity, the Quadratic variations method [7] can provide reasonable results.

The traditional methods can be used for financial time series modelling such as exchange rates [11], commodity prices [2] or stock market indices [12]. Recently, there were some attempts to refine the conventional methods using the self-affine property [3], spectrum of a signal [8] or self-similarity [5]. In this paper, the revisited zero-crossing method [4] is used as a referential method due to its robustness and signal segmentation principle.

[1]Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Trojanova 13, Prague, Czech Republic, dlaskma1@fjfi.cvut.cz

The novel approach based on fractional Brownian bridges described later in this work is especially designed to provide unbiased estimate of fractional parameter when dealing with short time series containing only few tens of elements. In contrast to traditional methods, it is useful for local Hurst exponent determination and analysis of its variability in time.

2 Traditional Fractional Processes

Fractional Brownian motion (fBm) and fractional Gaussian noise (fGn) belong to the family of fundamental fractional processes and are often used for time series analysis. In this section, the well-known statistical properties of fBm and fGn are summarized.

Fractional Brownian motion $B_H(t)$ [9] is a continuous Gaussian process defined for $t \in [0; +\infty)$, $H \in (0; 1)$ and $\sigma > 0$. The process starts at zero and has zero expected value for all positive times t. The autocovariance structure of fBm obeys for all $t, s > 0$

$$\mathrm{E}(B_H(t)B_H(s)) = \frac{\sigma^2}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right). \tag{1}$$

Parameter H is called Hurst exponent and influences the Hausdorff dimension of fBm graph that equals $D_H = 2 - H$. The parameter σ is often normalized to be unit and this special case the process is called standardized fractional Brownian motion. For $H = 1/2$, the fBm is standard Brownian motion.

Fractional Gaussian noise $G_H(t)$ is defined for all $t > 0$ as

$$G_H(t) = B_H(t+1) - B_H(t). \tag{2}$$

The Hausdorff dimension of fGn is, however, independent of H and equals $D_H = 2$. The process is still Gaussian, zero mean and if it is constructed on the basis of standardized fBm, it has also unit variance and the autocorrelation function can be expressed as

$$\mathrm{E}(B_H(t)B_H(t+k)) = \frac{1}{2}\left(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right). \tag{3}$$

3 Novel Method: Fractional Brownian Bridge

The novel methodology presents new stochastic process that we denote as fractional Brownian bridge (fBB) and uses its autocorrelation for Hurst exponent estimation. The first step is to discretize the domain, where the process is defined. Therefore we start with fGn sequence $G_H(k)$ of size N that is sampling of continuous standardized fGn for $k = 1, \ldots, N$. By means of cumulative sum and adding zero to the beginning of the sequence, we can obtain a sample $B_H(k)$ with $N + 1$ elements of standardized fBm for $k = 0, \ldots, N$. The fractional Brownian bridge $M_H(k)$ is subsequently defined as

$$M_H(k) = B_H(k) - B_H(0) - \frac{k}{N}(B_H(N) - B_H(0)). \tag{4}$$

In fractal analysis of time series, the fractional processes are often converted to fractional noises by means of differencing to simplify their covariance structure together with its spectral properties keeping the desired dependence on Hurst exponent. In this paper we follow this procedure and the final suggested process $X_H(k)$ is presented as the differentiation of fBB. We define the differenced fractional Brownian bridge (dfBB) $X_H(k)$ as

$$X_H(k) = M_H(k+1) - M_H(k) \tag{5}$$

for $k = 0, \ldots, N - 1$. This new process employs fBm sample in its definition and its construction is essential for several reasons. The first principle is based on the preservation of fractal character, therefore the suggested process needs still to have fractional properties. Additionally, the process is defined only in fixed points of finite interval, and therefore it is more suitable for short time analysis. Since the process both starts and ends with zero, the next feature of proposed process is its extension to infinite length.

Using properties of traditional fractional processes, one can deduce that dfBB has zero expected value and its variance can be expressed as

$$\gamma_0(H) = 1 - N^{2H-2} \tag{6}$$

and the autocovariance with lag m equals

$$\gamma_m(H) = N^{2H-2} + \frac{1}{2}\left(|m+1|^{2H} - 2|m|^{2H} + |m-1|^{2H}\right) + \frac{|m|^{2H} - |N-m|^{2H} - |N|^{2H}}{N(N-m)}, \tag{7}$$

which can be rewritten as a sum of correlation function of fGn and correction resulting from the discretization of the signal and construction of fractional bridge. The autocorrelation of dfBB can be subsequently expressed as

$$\rho_m(H) = \frac{\gamma_m(H)}{\gamma_0(H)}. \tag{8}$$

The procedure of Hurst exponent estimation involves the transformation of input fGn sample using equations 4 and 5 into dfBB sample $x_0, x_1, \dots, x_{N-1}$. The m-th autocovariance coefficient can be expressed for $m = 0, \dots, N-1$ as

$$\widehat{r}_m = \sum_{k=0}^{N-m-1} x_k x_{k+m} \tag{9}$$

and the estimation of m-th autocorrelation coefficient is obtained via

$$\widehat{\rho}_m = \frac{r_m}{r_0}. \tag{10}$$

Assuming model in the following form

$$\widehat{\rho}_m = \rho_m(H) + e_m \tag{11}$$

for $m = 1, \dots, N-1$ and Gaussian random variables e_m one can obtain the expected value of H together with its standard deviation using maximum likelihood method.

4 Referential Method: Zero-crossing Technique

As a referential technique, the revisited zero-crossing method [4] is used later in this paper. Consider fGn sample of length N that intersects the horizontal axis Z times. Than it is well-known, that the point estimate of H equals

$$H \approx 1 + \log_2 \cos \frac{\pi Z}{2N}. \tag{12}$$

The ratio Z denotes the estimate of the true probability of zero-crossing p in the signal. The improved method employs the Bayesian rule and expresses the posterior probability density function for p as

$$f(p) = \frac{p^Z (1-p)^{N-Z}}{\mathrm{B}(Z+1, N-Z+1)} \tag{13}$$

where Z still indicates the total amount of zero-crossings in the signal and B denotes the Beta function. The key idea of the revisited method is signal segmentation and averaging the $f(p)$ function over all intervals. Considering the division of the input signal into L non-overlapping segments of fixed length M, one can quantify the $f(p)$ function value in each segment assuming that the number of zero-crossings in each interval equals Z_k for $k = 1, \dots, L$. Therefore, the aggregate averaged probability of zero-crossing equals

$$f_{\mathrm{L}}(p) = \frac{1}{L} \sum_{k=1}^{L} \frac{p^{Z_k}(1-p)^{N-Z_k}}{\mathrm{B}(Z_k+1, N-Z_k+1)}. \tag{14}$$

The only remaining task is the determination of the parameter L that indicates the total number of segments. This is achieved thanks to the unimodality principle. The optimal segmentation L^* is defined as the smallest integer greater than two, for which the $f_{\mathrm{L}}(p)$ function has one unique peak. Only in the case, when the $f_L(p)$ is unimodal for $L = 2$ already, the L^* parameter is defined to be unit. The expected value of Hurst exponent can be subsequently expressed as

$$\mathrm{E}(H) = 1 + \int_0^1 f_{L^*}(p) \log_2 \cos \frac{p\pi}{2} \mathrm{d}p \tag{15}$$

together with its variance estimation

$$\sigma^2 = \int_0^1 \left(1 + \log_2 \cos \frac{p\pi}{2}\right)^2 f_{L^*}(p) \mathrm{d}p - (\mathrm{E}(H))^2. \tag{16}$$

5 Application to Stock Market Indices

The fBB estimation method can be used for analysis of stock market indices. Daily data of eight indices were investigated in the period between 1st April 2014 to 1st March 2016 (500 trading days). In the analysis, there are four European indices (FTSE100, SMI, AEX, DAX), two Asian (HSI,NIKKEI) and two North American (SP500, NASDAQ). The original time series is transformed via logarithmic differences and considered as fGn sample with unknown Hurst exponent.

5.1 Stationarity Testing

The key property of fGn process is stationarity. However, none of the samples is stationary in the whole range, and therefore it is necessary to divide the input time series into S non-overlapping intervals with constant length $L(S)$ and test the stationarity on the interval level. For each $S \geq 2$, we performed $S-1$ two-sample F-tests with null hypothesis

$$H_0 : s_i^2 = s_{i+1}^2 \tag{17}$$

where s_i^2 denotes the variance of time series values in i-th segment for $i = 1, \dots, S-1$. The suitable segmentation is the smallest integer S^*, for which we cannot refuse any of the S^*-1 null hypotheses on the 5% significance level. Table 1 shows the suitable segmentations and the respective interval lengths for each stock market index.

index	FTSE100	SP500	NASDAQ	SMI	AEX	NIKKEI	HSI	DAX
S^*	13	14	25	3	12	21	13	7
$L(S^*)$	38	35	20	166	41	23	38	71

Table 1 Maximal length of stationary segments of stock market indices.

5.2 Hurst Exponent Time Variability

The NASDAQ stock market index needed finest segmentation into intervals with 20 elements. Therefore, to keep the possibility of comparing markets between each other, we apply the fBB estimation procedure to all stock market indices with fixed segment size $N = 20$. Figure 1 shows the expected values of Hurst exponent together with the standard deviations illustrated by rectangles for FTSE100, DAX, NASDAQ and SMI.

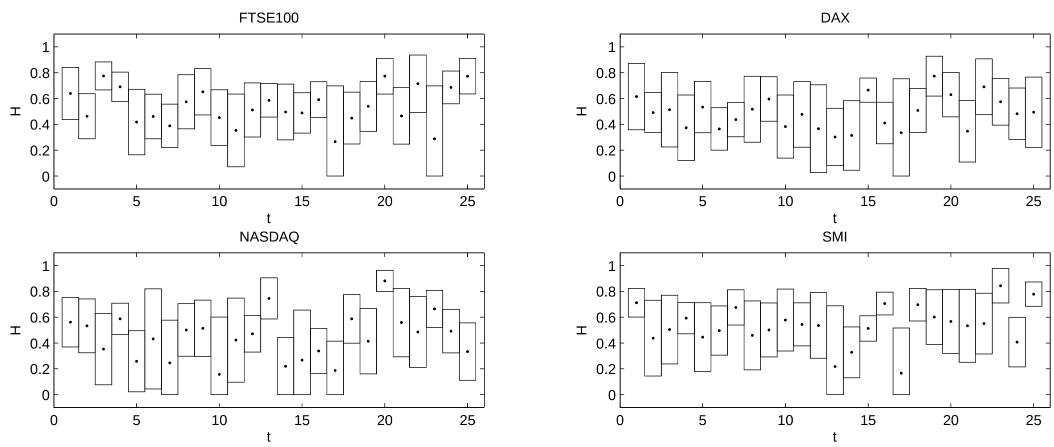


Figure 1 Hurst exponent time variability.

Based on figure 1 it is possible to notice, that the standard deviation is dependent on the value of estimated Hurst exponent. When the segment is evaluated as long-range dependent, the standard deviation is smaller in comparison to the case, when the stock market index fluctuates and shows strong negative correlation with H values close to zero. Table 2 provides results of stock market segment analysis. The ρ coefficient denotes the correlation between Hurst exponent expected value and its standard deviation sd, the r value refers to the percentage of cases, when the Hurst exponent was higher than 0.5, and the $\overline{H}$ denotes average for each stock market.

index	ρ	r	$\overline{H}$	rank
FTSE100	-0.7475	0.52	0.5358	2
SP500	-0.8081	0.60	0.5073	4
NASDAQ	-0.6554	0.40	0.4484	-
SMI	-0.7716	0.68	0.5397	1
AEX	-0.8473	0.48	0.4869	-
NIKKEI	-0.8549	0.64	0.5108	3
HSI	-0.8123	0.40	0.4942	-
DAX	-0.4659	0.44	0.4878	-

Table 2 Short segment analysis of stock market indices.

All analysed samples showed strong negative correlation between H and sd. Therefore, when the stock market exhibits predictable behaviour, the dfBB model provides very confident Hurst exponent estimation which could be subsequently used for prediction. The SMI stock market obtained largest percentage rate r of predictable segments together with highest average of H values. In terms of the analysis of short time series, the SMI stock market together with FTSE100 and NIKKEI are considered as predictable and therefore recommended for investments.

5.3 Comparison to Zero-crossing Method

The fBB tool can be used also for long time series analysis. The experimental autocorrelation function is averaged via all segments and the aggregate Hurst exponent estimation can be obtained. As a referential method, the revisited zero-crossing method is used that utilizes different property of fractional process to perform the estimate. Table 3 presents the H estimation together with its standard deviation and rank column sorts the stock market indices from the highest Hurst exponent in descending order.

index	fBB method			zero-crossing method		
	H	std	rank	H	std	rank
FTSE100	0.5487	0.0306	2	0.5607	0.0487	3
SP500	0.5226	0.0354	5	0.5432	0.0435	5
NASDAQ	0.4731	0.0487	8	0.5001	0.0428	8
SMI	0.5648	0.0364	1	0.6201	0.0321	1
AEX	0.5370	0.0288	3	0.5730	0.0350	2
NIKKEI	0.5232	0.0377	4	0.5523	0.0361	4
HSI	0.5172	0.0383	6	0.5323	0.0368	6
DAX	0.4894	0.0641	7	0.5067	0.0508	7

Table 3 Comparison of fBB and zero-crossing method.

Based on the fact, that the standard deviations are smaller than in the case of segment analysis, the Hurst exponent is more accurate and evaluates the dependence of the time series in the long term. Estimated H values and standard deviation results provided by zero-crossing method are similar to the estimates that were performed on the basis of dfBB. Therefore the dfBB technique has suitable asymptotic properties and provides unbiased Hurst exponent estimation in the case of both short and long time series analysis. From the long time period perspective, the SMI, AEX and FTSE stock market can be recommended for investments based on the results of both novel and referential method.

6 Conclusion

Numerical experiments showed that the stock market indices are not stationary in the whole range, however, they can be analysed in smaller stationary segments. The new methodology was applied to short intervals containing 20 values. Based on the dfBB model, unbiased estimate of Hurst exponent was obtained even from such a small sample and the SMI and FTSE100 stock markets were identified as the most predictable. In case of the aggregate estimate from the whole sample containing 500 values, experiments showed that the increasing number of data can substantially lower the standard deviation and the estimates were still unbiased as checked by the referential zero-crossing principle. Based on these results we conclude, that the SMI and FTSE100 stock market exhibited slowest changes and therefore they are the suitable for both short and long term investments.

Acknowledgements

The paper was created with the support of CTU in Prague, Grant SGS14/208/OHK4/3T/14.

References

[1] Bogusz, J., Klos., A, Figurski, M. and Kujawa, M.: Investigation of long-range dependencies in the stochastic part of daily GPS solutions. *Survey Review* **48**, 347 (2016), 140–147.

[2] Bolgorian, M. and Gharli, Z.: A Multifractal Detrended Fluctuation Analysis of Gold Price Fluctuations. *Acta Physica Polonica B* **42**, 1 (2011), 159–167.

[3] Campanharo, A.: Hurst exponent estimation of self-affine time series using quantile graphs. *Physica A - Statistical Mechanics and its Applications* **444**, 1 (2016), 43–48.

[4] Dlask, M., Kukal, J., Tran, Q.V.: Revisited Zero-Crossing Method for Hurst Exponent Estimation in Time Series. In: *33rd International Conference on Mathematical Methods in Economics* (D. Martincik, J. Ircingova, P. Janecek, eds.), University of West Bohemia, Plzen, 2015, 115 – 120.

[5] Fernandez-Martinez M. and Sanchez-Granero M.A.: An accurate algorithm to calculate the Hurst exponent of self-similar processes. *Physics Letters A* **378**, 32 (2014), 2355–2362.

[6] Hurst, H. E.: Methods of Using Long-Term Storage in Reservoirs. *Proceedings of the Institution of Civil Engineers* **5**, 5(1956), 519–543.

[7] Istas, J. and Lang, G.: Quadratic variations and estimation of the local Holder index of a Gaussian process. *Annales de l'Institut Henri Poincare B Probability and Statistics* **33**, 4 (1997), 407–436.

[8] Kristoufek, L.: Spectrum-based estimators of the bivariate Hurst exponent. *Physical Review E* **90**, 6 (2014), 1–6.

[9] Mandelbrot B. and Van Ness, J.V.: Fractional Brownian Motions, Fractional Noises and Applications. *SIAM Review* **10**, 4 (1968),422–437.

[10] Mansuy, R. and Yor M.: *Aspects of Brownian motion*. Springer, Heidelberg, 2008.

[11] Stosic D., Ludemir T., De Oliviera W. and Stosic T.: Foreign exchange rate entropy evolution during financial crises. *Physica A: Statistical Mechanics and its Applications* **449**, 1 (2016), 233–239.

[12] Tzouras, S., Anagnostopoulous C. and Mccoy E.: Financial time series modelling using the Hurst exponent. *Physica A: Statistical Mechanics and its Applications*, **425**, 3 (2015), 50–68.

[13] Whittle, P.: Estimation and information in stationary time series. *Arkiv for Matematik* **2**, 5 (1953), 423–434.

[14] Xue, Y., Jia, L., Teng, W. and Lu, W.: Long-range correlations in vehicular traffic flow studied in the framework of Kerners three-phase theory based on rescaled range analysis. *Communications in Nonlinear Science and Numerical Simulation* **22**, 1 (2015), 285–296.

[15] Yang, X.: On the large deviation principle of generalized Brownian bridges. *Journal of Mathematical Analysis and Applications* **430**, 2 (2015), 845–856.

Earnings effects of job and educational mismatch in the Czech graduate labour market

Zuzana Dlouhá[1]

Abstract. The paper provides the estimates of the impact of job and educational mismatches on earnings of graduates from the University of Economics, Prague. The dataset was obtained from the cross-sections REFLEX survey that gathered information about the labour market status of the graduates within the period 2008-2012. The earnings equations were estimated controlling the source of sample selection bias when using Heckman maximum likelihood procedure with the main Mincerian earnings equation, while the selection equation is a probit estimate of the probability to be employed rather than unemployed. We tested the hypothesis that educational and job mismatched workers earn less due to their lower competencies and skills in relative terms. We estimated earnings penalties associated to job and educational mismatches controlling for sample selection. The earnings penalty associated to educational / job mismatch is equal to 32.2 % and –8.2 %, respectively. This finding supports the hypothesis there is positive selection into employment of the most skilled among workers whose individual characteristics are less on demand on the labour market.

Keywords: Job and educational mismatch, Heckit model, REFLEX survey.

JEL Classification: I23, J24
AMS Classification: 62P20

1 Introduction

Higher education is the main provider of highly skilled human capital due to the close ties with economic growth and development, and it is also one of the sources of competitive advantages. In addition, students who graduated from university or higher composed are more likely to be employed with the higher salaries to receive or at least to expect. Preparing for a good paying job or career through tertiary education has been the original motivation of most students in receiving higher education. When such motivation is not satisfied sufficiently, the occurrence or feeling of educational and job mismatch can appear [2]. The educational mismatch together with job (skills) mismatch of workers lead to lower job satisfaction and wage differentials than in case of properly matched workers [6], but educational requirements for a certain type of job can rise over time, individuals can be overqualified due to low ability for that level of qualification or they chose to work at less stressful work, etc. [3]. Factors of dissatisfaction with chosen study programme of graduates from the University of Economics, Prague were investigated and variable horizontally mismatched at first job were confirmed as statistically significant in [1]. The problem of educational mismatch was also studied by [3] using data from European Social Survey ESS5 collected in the years 2010 and 2011 in selected EU countries.

Educational mismatch represents a mismatch between qualification necessary for a particular job and qualification actually acquired by an individual working on this position [4] or it can be defined as unfulfilled expectations of the educated concerning their career attainments as described above. When type or level of skills is different from that required to adequately perform the job we refer to job mismatch. We use self-assessed measure of educational and job mismatch in this study as this measurement is always up-to-date and corresponds with requirements in the individual firm. Its disadvantage is subjective bias: respondents may overstate job requirements, inflate their status, or reproduce actual hiring standards. We tested the hypothesis that salaries of educational and job mismatched workers are lower that non-mismatched due to their lower competencies and skills in relative terms.

The Czech tertiary education system has experienced a number of deep, dynamic changes and extensive development since 1993 similar to other developed countries such as Slovakia, Poland or Hungary. From a strictly uniform highly centralized and ideologically bound system under the communist regime it has been changed into the much more diversified and decentralized system with full academic freedom and self-governing bodies. We present trends in educational attainment of the population older than 15 years between 1993 and 2014 in the Czech Republic on the Figure 1 with focus on tertiary education level. The ratio of individuals with elementary

[1] University of Economics, Prague, Faculty of Informatics and Statistics, Department of Econometrics, W. Churchill Sq. 1938/4, 130 67 Prague, Czech Republic, e-mail: zuzana.dlouha@vse.cz.

or no education is decreasing, the same slightly lowering trend is visible for graduates from secondary schools with apprenticeship certificate, from 38.9 % to 34.3 %. Tertiary education (provided by universities and other higher education institutions, defined also as the level of education following secondary schooling) has progressed rapidly in the last 20 years (from 7.8 % to 17.3 %) and relatively more women than men follow tertiary level programmes.

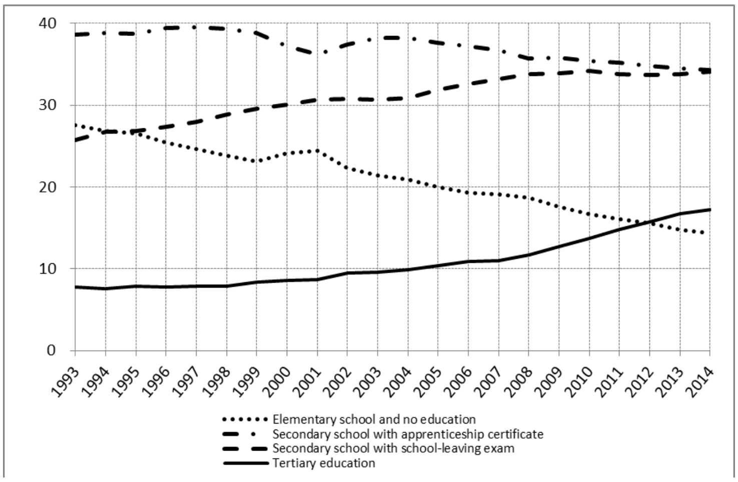


Figure 1 Educational attainment of the population older than 15 years old (in percent), source: www.czso.cz

The rest of the paper is structured as follows. The next section describes the methodology and data sources together with explanation of the variables included in the Heckit model. Section three explores factors of earnings equations among university graduates. In section four we offer concluding remarks and propose next steps.

2 Methodology and data

We briefly introduce methodology used for estimating earnings penalty of educational and job mismatch within the following section. Next we describe the analysed dataset and provide basic descriptive statistics of the variables incorporated into models.

2.1 Heckit model

In Heckman selection model (therefore Heckit model) we use in the first stage for prediction of the probability that someone is included or selected (therefore selection equation) in the sample. This data truncation occurs because sample selection is determined by the people's decision, not the surveyor's decision. Bias caused by this type of truncation is called the sample selection bias. We consider a class of binary response (self-selected into the sample / self-selected out of the sample) models of the form [7]:

$$\mathrm{P}(y = 1|\mathbf{x}) = G(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k) = G(\beta_0 + \mathbf{x}\boldsymbol{\beta}), \quad (1)$$

where G is a function taking on values strictly between 0 and 1: $0 < G(z) < 1$, for all real numbers z. For the estimation of response probabilities we use probit model, hence G is the standard normal cumulative distribution function expressed as integral:

$$G(\mathbf{x}\boldsymbol{\beta}) = \Phi\,(\mathbf{x}\boldsymbol{\beta}) = \int_{-\infty}^{\mathbf{x}\boldsymbol{\beta}} \phi(v)dv, \quad (2)$$

where

$$\phi(v) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{v^2}{2}\right) \quad (3)$$

is the standard normal density. In the second stage we the control for the probability of being selected when estimating extended Mincerian earnings response equation, augmented of the educational and job mismatches term in form:

$$lnEARN = \beta_0 + \mathbf{x}\boldsymbol{\gamma} + \delta M + \epsilon, \quad (4)$$

where *lnEARN* is the natural logarithm of the monthly earnings of individuals, **x** is a vector of variables assumed to affect earnings, *M* is a dummy variable taking a value of 1 if the individual is mismatched by education or by a job. As we use survey data, the standard OLS estimation in the second step would not result in consistent and unbiased estimates and we decided to apply maximum likelihood estimation method, see detailed in [5].

2.2 Data and variables

The dataset was gathered from the international cross-sectional Flexible Professional in the Knowledge Society (REFLEX) survey that was held by the Education Policy Centre, Charles University, Prague, in 2013. The project is a large-scale international project that had been carried out in 15 European countries and Japan and it was financed as a Specific Targeted Research Project (STREP) of the European Union's Sixth Framework Programme among higher education graduates. Altogether 21 public, and 15 private colleges and universities participated on this survey in the Czech Republic. We were provided only with the sub-sample of graduates from the University of Economics, Prague, which consists of 1,704 respondents who obtained their tertiary degree during the time period 2008–2012. Respondents were asked to exclude jobs they left within 4 months after graduation. Data was weighted by the proportion of faculties, type of study, gender, economical status and year of graduating. Due to lack of answers in some variables we use final dataset of 1,211 graduates for model estimation.

The natural logarithm of gross monthly earnings in CZK taking the mid-point of interval data is dependent variable in response function (4), while in selection function we consider as dependent variable *Employment* taking one when employed and zero otherwise. We included following demographic characteristics as independent variables: binary variable *Female*, *Age* is measured in years, binary variable *Children* coded as 1 when respondents has at least one child. Parental highest earned education is distinguishing between mother's and father's education, coded as follows: 1 – elementary school, 2 – secondary education with apprenticeship certificate, 3 – secondary education with school-leaving exam, 4 – tertiary level of education. Continuous variables related to *Work experience*, measured in months, differentiate between two options of the graduate's job experience in study-related and non-study-related activities. *Left first employment* is binary variable that distinguishes those respondents who still work in their first employment after five years of graduation from those who already left their jobs. We also included *Number of jobs* the graduate had during the labour market career until the data collection. Dummy variable *Type of contract* in the current occupation takes the value of 1 if the job is permanent in contrast to fixed-term employment or self-employment. Independent variable *Educational mismatch* presents mismatch between level of education and current graduate's job. Variable *Job mismatch* is based on the response to a question asking respondents to rate on a 1 (not at all) to 5 (very high extent) to which their skills and knowledge were utilized in their job with a response 1 or 2 deemed consistent with mismatching. Variable *Nationality* differentiates between Slovak nationality of the graduates and others and dummy variable *Prague* is equal to one for those living in Prague and zero otherwise.

In Table 1 we present means, standard deviations and number of observations of all variables included in the models. We can observe that 51.4% of the graduates considered themselves as job mismatched and 21.2% as skills mismatched. Average age of respondents is almost 28 years. Mother's highest earned education is a little bit higher than father's one, comparing 3.366 and 3.355. We can conclude that work experience of graduates in study related field is in average 6.6 months longer that in non-study-related field. Most of graduates have permanent type of contract (almost 77%) and work in Prague (78.8%).

Variable	**No. obs.**	**Mean**	**Std. dev.**
Monthly wage	1,316	35,856.38	20,545.39
Employment	1,704	0.967	0.180
Educational mismatch	1,381	0.212	0.396
Job mismatch	1,381	0.514	0.499
Female	1,704	0.620	0.485
Age	1,704	27.632	3.517
Children	1,704	0.107	0.309
Father's education	1,698	3.355	0.913
Mother's education	1,698	3.366	0.808
Work experience (study related, in months)	1,704	27.780	243.823
Work experience (not study related, in months)	1,704	21.096	242.530
Left first employment (= 1)	1,482	0.613	0.487
Number of jobs	1,482	1.573	0.869
Type of contract (permanent = 1)	1,332	0.767	0.422
Nationality (SK = 1)	1,698	0.060	0.237
Prague	1,381	0.787	0.410

Table 1 Descriptive statistics

3 Results

The main results of earnings equations with the Heckman correction in Model 1 and Model 2 specifications are presented below in Table 2. We used econometric software EViews 9 for all the estimations and further calculations and testing.

The main response equation is a typical Mincerian earnings equation with dependent variable log of gross monthly earnings (4), while the selection equation is a probit estimate of the probability to be employed rather than unemployed (1).

We apply the rule that variables in the response and selection equation should be the same, except for instrumental binary variable *Children*, assuming that having a child or children affects the probability to participate on labour market, but not wages. Estimated coefficient of this variable has the correct expected sign that having a child increases the reservation wage and therefore decreases the probability of being employed, but this instrument is not statistically significant. Being female leads to lower salary, according to economic theory, the higher age increases salary (statistically significant). Graduates working in Prague and with permanent type of contract have significantly higher earnings. Estimated earnings penalties associated to educational and job mismatch is 32.2 % and 8.2 %, respectively, in absolute terms, both coefficients are statistically significant. In the selection equation variable *Father's education* decreases the probability of being employed, more significant positive impact on employment has variable *Mother's education* together with variable study related working experience of the graduate.

Response equation (dependent variable – log of monthly earnings)		
Predictors	**Model 1**	**Model2**
Constant	8.717***	8.572***
Educational mismatch	–0.322***	
Job mismatch		0.082***
Female	–0.186***	–0.195***
Age	0.050***	0.052***
Father's education	0.043**	0.039**
Mother's education	0.010	0.020
Work experience (study related, in months)	<0.000	<0.000
Work experience (not study related, in months)	<0.000	<0.000
Left first employment (= 1)	–0.138***	–0.150***
Number of jobs	–0.075**	–0.084***
Type of contract (permanent = 1)	0.330***	0.342***
Nationality (SK = 1)	0.002	–0.017
Prague	0.187***	0.203***
Selection equation (dependent variable – Employment)		
Predictors	**Model 1**	**Model2**
Constant	0.364	0.304
Female	0.042	0.037
Age	0.031	0.036
Father's education	–0.197**	–0.228**
Mother's education	0.279***	0.285***
Work experience (study related, in months)	0.011**	0.012***
Work experience (not study related, in months)	0.003	0.004
Left first employment (= 1)	0.235	0.223
Nationality (SK = 1)	0.363	0.349
Children	–0.122	–0.145
Interaction term ρ	0.477***	0.484***
Interaction term σ	–0.856***	–0.843***
Log likelihood	–880.081	–899.910
Akaike inform. criterion	1.495	1.528
Wald test	1115.612	1046.431

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2 Earnings equations with correction for sample selection

Interaction term ρ presents the correlation coefficient of the error terms from the selection (1) and the response (4) function and according to results from Table 2 we can reject the null hypothesis at a statistically significant level and can conclude that ρ is not equal to zero. This suggests that applying Heckit model to the data is appropriate. The value of σ is the estimated variance of the response function (4). We also report results of the Wald test (χ^2) of all coefficients in the regression model (except constant) being zero. With $p < 0.0001$ we conclude that the covariates used in the model may be appropriate, and at least one of the covariates has an effect that is not equal to zero.

4 Conclusion

This paper has attempted to estimate the main factors of earnings of graduates from the University of Economics, Prague, within five years after their graduation. We control for the possible sample selection bias from measuring only among the graduates that were employed at the time of collection of data by the Heckit econometric specification of the earnings response function. The earnings penalty associated to educational / job mismatch is equal to 32.2 % and –8.2 %, respectively. This finding supports the hypothesis there is positive selection into employment of the most skilled among workers whose individual characteristics are less on demand on the labour market. Further research will exploit the panel dimension of the data when incorporating more universities and previous years using data from REFLEX survey to test in a different context the role of the ability bias and measurement errors.

Acknowledgements

The research is supported by the Internal Grant Agency of the University of Economics, Prague, project no. F4/62/2015 and was prepared with the institutional support for the long-term development of science and research at the Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] Dlouhá, Z., and Dlouhý, M.: Factors of dissatisfaction with chosen study programme. In: *Mathematical Methods in Economics 2014*, (2014), 151–155.

[2] Lin, X.: The Expansion of Higher Education and Overeducation in Taiwan: Evidence from 1997 to 2007. *International Journal of Information and Education Technology*, **5** (12), 2015.

[3] Maršíková, K., and Urbánek, V.: A comparison of educational mismatches across Europe. E+M Ekonomie a management, **18** (4), (2015), 24–38.

[4] McGuiness, S., and Sloan, P.J.: Labour market mismatch among UK graduates: An analysis using REFLEX data. *Economics of Education Review* **30** (1), (2011), 130–145.

[5] Nawata, K.: Estimation of sample selection bias models by the maximum likelihood estimator and Heckman's two-step estimator. *Economics Letters* **45** (1), (1994), 33–40.

[6] Robst, J.: Education and job match: The relatedness of college major and work. *Economics of Education review* **26,** (2007), 397–407.

[7] Wooldridge, J. M.: *Introductory Econometrics: A Modern Approach.* 5th edition, Cengage Learning, Mason, 2012.

Multicriteria Voting Game and its Application

Martin Dlouhý[1], Michaela Tichá[2]

Abstract. In the multicriteria voting game, we assume the set of political parties and the set of political programmes with multiple public policy dimensions. The coalitional programme is formulated as the weighted average of individual political programmes. The objectives of each political party are to minimize the maximal distance between the coalitional programme and its own programme; to maximize its own share of power in the winning coalition; and to maximize stability of the winning coalition, which is measured as the maximal distance between the coalitional and individual programmes of all political parties in the coalition. The model of multicriteria voting game is applied to the Chamber of Deputies of the Parliament of the Czech Republic. In total, 24 of 200 deputies responded and were able to describe fully or partially programmes of political parties represented in the Chamber of Deputies in five public policy dimensions: health care, public finance and tax policy, social policy, foreign policy and EU, and labour market. The data were used to identify the optimal winning coalition (ČSSD, ANO 2011, KDU-ČSL), which is the same as the real governing coalition in the Czech Republic.

Keywords: game theory, voting game, cooperative game, Chamber of Deputies.

JEL Classification: C71
AMS Classification: 91A12

1 Introduction

The theory of games can be defined as the study of mathematical models of conflict and cooperative decision making situations with more rational participants that are called players (see, for example [1, 2, 3]). These mathematical models include, for example, the game in normal form, game in extensive form, repeated game, and coalitional game. A game has a quite general meaning, including very diverse decision making situations such as auctions, competition between firms, military conflicts, playing chess, coalition formation in the parliament, conflicts among biological species.

A voting game is a special case of cooperative game applied to political or other decision making bodies. For example, Dlouhý and Fiala [2] studied the coalition formation in the Prague City Assembly 2006 - 2014; Turnovec, Mercik, and Mazurkiewicz [6] studied the case of the European Parliament, which has a dual structure, because its members represent their own countries and at the same time they are clustered in European political parties.

The objectives of this paper are: (a) to formulate a multicriteria voting game as a model that is able to predict the result of coalition formation; (b) to apply the multicriteria voting game to a real example. In the multicriteria voting game, we define the set of political parties and the set of political programmes characterized by multiple dimensions of public policies. The coalitional programme is formulated as the weighted average of individual political programmes.

We assume that the objectives of an individual political party are: (1) to minimize the maximal distance between the coalitional programme and its own political programme in all dimensions of public policy; (2) to maximize its own share of power in the coalition; and (3) to maximize stability of the coalition, which is measured as the maximal distance between the coalitional and individual political programmes for all political parties in the coalition.

The model of multicriteria voting game was applied to the Chamber of Deputies of the Parliament of the Czech Republic.

2 Model Formulation

Let **N** = {1, 2,…, n} is the set of political parties, a_0 is the total number of deputies, and a_i is the number of deputies of political party i. Ideology (political programme) of political parties is traditionally modelled geometrically. The simple example is the one-dimensional left-right model. More sophisticated models place political parties in multidimensional ideological space [4]. In our model, we assume that political parties can be characterized by the set

[1] University of Economics Prague, Dpt. of Econometrics, 4 W. Churchill Sq., Prague 3, e-mail: dlouhy@vse.cz.
[2] University of Economics Prague, Dpt. of Econometrics, 4 W. Churchill Sq., Prague 3, e-mail: xticm11@vse.cz.

of values in m public policy dimensions that can quantified within [0, 1] interval. The vector of values for each political party in m public policy dimensions is denoted as

$$k^i = (k_1^i, k_2^i, \dots, k_m^i), \tag{1}$$

where k_j^i is the value for political party i in j-th public policy dimension. The coalitional political programme is denoted as

$$K = (K_1, K_2, \dots, K_m), \tag{2}$$

where K_j is the value in j-th public policy dimension. If political parties form the winning coalition, the coalitional programme is a compromise of individual political programmes. Let us suppose that the political programme of winning coalition S is calculated as the weighted average of individual programmes where weights are determined by the relative power of individual political parties:

$$K_j = \sum_{i \in S} \frac{a_i}{\sum_{i \in S} a_i} \; k_j^i. \tag{3}$$

By knowing the coalitional programme K we are now able to evaluate the expected benefits of the political party from being a member of the winning coalition. Let us assume that each political party has three objectives: to implement its political programme, maximize its power in the coalition, and be a member of politically stable coalition. Quantitatively, we formulate these objectives in the following way:

(a) Each political party i in the winning coalition S has interest in minimizing the maximal difference between its individual political programme and the coalition programme:

$$f_1(i, S) = \max_{j \in \{1,2,\dots,m\}} |k_j^i - K_j|. \tag{4}$$

(b) Each political party i in the winning coalition S maximizes its power in the coalition. The power of political party is measured by the relative number of its deputies:

$$f_2(i, S) = \frac{a_i}{\sum_{i \in S} a_i}. \tag{5}$$

(c) Each political party is interested in the stability of the winning coalition as a whole. This interest is expressed as minimizing the maximal distance between all individual political programmes and the coalitional programme:

$$f_3(S) = \max_{i \in S} \left(\max_{j \in \{1,2,\dots,m\}} |k_j^i - K_j| \right). \tag{6}$$

In this multiple criteria model, each political party has interest in minimizing $f_1(i, S)$, maximizing $f_2(i, S)$, and minimizing $f_3(S)$. All these objective functions are defined on the interval [0, 1]. The preferences of the political party among these three objectives are expressed in the form of non-negative weights:

$$\sum_{j=1}^{3} v_j^i = 1 \;\text{ and }\; v_j^i \geq 0 \;\; \forall i \in \{1,2,\dots,n\}. \tag{7}$$

The utility (payoff) of each political party from being a member of the winning coalition is obtained by maximizing the global objective function:

$$f(i, S) = \max_{S \in \Omega} \left(v_1^i \big(1 - f_1(i, S)\big) + v_2^i f_2(i, S) + v_3^i (1 - f_3(S)) \right), \tag{8}$$

where Ω is the set of all winning coalitions. The utility $f(i, S)$ is non-transferable between political parties.

In the next step, the principle of group stability can be used to reduce the set of winning coalitions. The principle states that the utility $f(i, S)$ for each member of the winning coalition must be greater than the utility that they could obtain from any sub-coalition, which is also the winning coalition. However, this procedure does not guarantee a unique solution and other criteria has to be used to determine the optimal winning coalition.

In this model, we assume that the winning coalition S maximizes the average utility of political parties included in the coalition:

$$U(S) = \frac{\sum_{i \in S} f(i, S)}{|S|}. \tag{9}$$

The winning coalition with maximum utility will be identified as the optimal winning coalition. However, there are more alternatives how to decide which winning coalition is the optimal one. For example, we can assume the maximization of weighted average utility:

$$U(S) = \sum_{i \in S} f_2(\mathrm{i}, \mathrm{S})\, f(i, S). \tag{10}$$

It should also be noted that the optimal winning coalition does not have to be the minimal winning coalition in this model.

3 Application

The model of multicriteria voting game was applied to the Chamber of Deputies of the Parliament of the Czech Republic, which is the lower house of the bicameral Parliament of the Czech Republic. The Chamber of Deputies has 200 members that serve a four-year terms. The Czech government is primarily responsible to the Chamber of Deputies. The current deputies representing seven political parties were elected in 2013 (Table 1).

Political party	**Number of Deputies**
ČSSD	50
ANO 2011	47
KSČM	33
TOP 09 + STAN	26
ODS	16
KDU-ČSL	14
ÚSVIT	14

Table 1 Chamber of Deputies, 2013.

The data on political programmes of political parties were obtained by the questionnaire survey in 2015. All 200 deputies were contacted via e-mail with a kind request to fulfil a short questionnaire on ideological positions of political parties represented in the Chamber of Deputies [5]. In total, 24 deputies responded and were able to describe fully or partially programmes of political parties represented in the Chamber of Deputies in five public policy dimensions: health care, public finance and tax policy, social policy, foreign policy and EU, and labour market. In the questionnaire, each public policy dimension was defined on interval [0, 10], instead of interval [0, 1], because we assumed that it would be easier for respondents to work with integer values rather than with decimal numbers.

In the definition of each public policy dimensions, 0 represents extremely left-wing oriented political programme, and 10 represents extremely right-wing oriented political programme. The simplified description of five public policy dimensions in the questionnaire that was sent to respondents follows:

(1) health care:
0 – all health services are publicly financed, no above-standard health services;
10 – basic health services are publicly financed, private above-standard services are available;

(2) public finance and tax policy:
0 – high progressive taxes, high taxes for rich people, high degree of solidarity;
10 – low taxes, flat tax, mainly indirect taxes;

(3) social policy:
0 – high social benefits for unemployed, disabled, families with children, high pensions;
10 – low social benefits, lower state pensions, support of individual savings;

(4) foreign policy and EU:

0 – protection of internal market, the lower level of EU integration and the possibility of exit, no to Eurozone;

10 – the higher level of EU integration, yes to Eurozone;

(5) labour market

0 – strong rights of employees and labour unions, higher taxes and higher control of enterprises;

10 – strong support of small and medium-sized enterprises, tax reliefs, diminishing role of labour unions.

Using the data obtained from the questionnaire survey, we were able to calculate the values in each public policy dimension for each political party as the weighted average of answers. Because we assume that deputies know better the programme of their own political party than the programmes of other political parties, the weight of the values for their own party was three times higher.

The ideological positions of political parties are summarized in Table 2. The results show what can be expected: KSČM (Communist Party of Bohemia and Moravia) is the extreme left-wing party, ČSSD (Czech Social Democratic Party) is the pro-European left-wing party, TOP 09+STAN and ODS (Civic Democratic Party) are the right-wing political parties that differ in particular in foreign policy and their relation to EU integration, KDU-ČSL (Christian Democratic Union) and ANO 2011 are parties in the political middle, and ÚSVIT is the nationalist political party. The respondents differed most in their answers when characterizing the political programme of ÚSVIT, noting that the political programme of this political party was unclear for them.

Political Party/Public Policy	**ČSSD**	**ANO**	**KSČM**	**TOP 09 STAN**	**ODS**	**KDU ČSL**	**ÚSVIT**
Health care	1.9	5.3	0.7	7.9	8.4	4.7	5.0
Public finance and tax policy	2.7	5.7	0.9	8.2	8.8	4.8	5.0
Social policy	2.8	5.4	1.3	8.1	8.3	4.0	2.0
Foreign policy and EU	7.1	5.0	1.6	7.9	3.2	6.8	0.0
Labour market	2.9	5.4	1.5	8.5	8.2	5.4	7.0

Table 2 Political Programmes in Five Public Policy Dimensions.

In the next step, we constructed all possible winning coalitions in the Chamber of Deputies. There is 64 winning coalitions that have 101 or more deputies. By using the principle of group stability, we can reduce the set of winning coalitions. For example, the grand coalition (ČSSD, ANO 2011, KSČM, TOP 09+STAN, ODS, KDU-ČSL, ÚSVIT) with utilities (0.560, 0.617, 0.465, 0.450, 0.412, 0.531, 0.372) is not stable, because there is a possibility to form the winning sub-coalition (ČSSD, ANO 2011, KDU-ČSL) with utilities (0.711, 0.713, 0.647). Alternatively, we can work with the original set of all winning coalitions, and then check if the optimal winning coalition is stable. Such procedure is mathematically less demanding.

Because we do not have any information about the weights v_j^i, we set them in this experiment to be (0.5, 0.3, 0.2) for all political parties. These weights suppose that implementation of the political programme is the most important objective of the political party. On the other hand, the stability of coalition is the least important objective. For each winning coalition, we calculated the vector of coalitional programme K, for each political party and winning coalition we calculated values of $f_1(i, S)$, $f_2(i, S)$, $f_3(i, S)$ and $f(i, S)$. Finally, we calculated, for each winning coalition, the average utility $U(S)$.

The results of our calculations for top-five winning coalitions are shown in Table 3. According to assumptions of the model, the optimal winning coalition is (ČSSD, ANO 2011, KDU-ČSL), which maximizes the value of $U(S)$. This coalition is coincidentally the real governing coalition in the Czech Republic. The sensitivity analysis shows that the optimal winning coalition (ČSSD, ANO 2011, KDU-ČSL) is not dependent on the value of weights. It can be easily shown that the fourth coalition (ČSSD, ANO 2011, TOP 09+STAN, KDU-ČSL) is not stable because the winning sub-coalition (ČSSD, ANO 2011, KDU-ČSL) will be preferred by its members.

Winning Coalition	**ČSSD**	**ANO**	**KSČM**	**TOP09 STAN**	**ODS**	**KDU ČSL**	**ÚSVIT**	$\boldsymbol{U(S)}$
ČSSD, ANO 2011, KDU-ČSL	0.711	0.713				0.647		0.690
ANO 2011, TOP 09+STAN, ODS, KDU-ČSL		0.727		0.616	0.572	0.575		0.622
ČSSD, ANO 2011, TOP 09+STAN	0.625	0.672		0.521				0.606
ČSSD, ANO 2011, TOP 09+STAN, KDU-ČSL	0.612	0.659		0.516		0.622		0.602
ČSSD, ANO 2011, KSČM	0.639	0.618	0.539					0.599

Table 3 Utility of Winning Coalition $U(S)$

4 Conclusion

In the paper, we have formulated the multicriteria voting game that is able to model the process of coalition formation. The model of multicriteria voting game was applied to the Chamber of Deputies of the Parliament of the Czech Republic. The data on political programmes of political parties represented in the Chamber of Deputies were obtained by the questionnaire survey in 2015. The model identified the winning coalition (ČSSD, ANO 2011, KDU-ČSL) as the optimal winning coalition. Such coalition is the same as the real governing coalition in the Czech Republic. This shows that even the simplified voting model based on the game theory can serve for the prediction of governing coalition.

There is one methodological question to be answered if we would obtain the same optimal winning coalition in case that the questionnaire would be fulfilled before the existence of the winning coalition. We may consider the hypothesis that the members of coalitional political parties may view differences in political programmes of coalition parties less important and the members of opposition political parties can view differences with coalitional political programme as more important.

Acknowledgements

The research was supported by the Czech Science Foundation, project no. 16-01821S, and by the Internal Grant Agency of the University of Economics in Prague, the project no. 54/2015.

References

[1] Binmore, K.: *Fun and Games: A Text on Game Theory*. D. C. Heath and Company, Lexington, 1992.
[2] Dlouhý, M., and Fiala, P.: *Teorie ekonomických a politických her*. Oeconomica, Prague, 2015.
[3] Myerson, R. B.: *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge 1991.
[4] Straffin, P. D.: Chapter 32 - Power and Stability in Politics. In: *Handbook of Game Theory with Economic Applications - Volume 2* (Aumann, R., and Hart, S., eds.). North-Holland, Amsterdam, 1994.
[5] Tichá, M.: *Vícekriteriální hry*. PhD thesis, University of Economics, Prague, 2016.
[6] Turnovec, F., Mercik, J. W., and Mazurkiewicz, M.: National and ideological influence in the European Parliament. *Control and Cybernetics* **37** (2008), 3:585–606.

On queueing model of freight train classification process

Michal Dorda[1], Dušan Teichmann[2], Vojtěch Graf[3]

Abstract. The paper deals with a finite single-server queueing system with a server subject to non-preemptive failures. The queueing model was created to model classification processes in marshalling yards. Customers are inbound freight trains entering the marshalling yard (primary shunting) and non-preemptive failures correspond to classification of trains of wagons entering the yard from industrial sidings (secondary shunting). It is assumed that the customers come to the system in the Poisson stream, service times are Erlang distributed. The capacity of the system is given by the number of arrival tracks in the marshalling yards. All the requests for secondary shunting are considered to be non-preemptive failures of the system because primary shunting cannot be carried out during secondary shunting. Times between non-preemptive failures are exponentially distributed and times to repair are Erlang distributed. The queueing system is modelled as a quasi-birth death process for which its state transition diagram and linear equation system are presented. By solving the equation system in Matlab stationary probabilities are calculated and on their basis some performance measures can be computed.

Keywords: $M/E_n/1/m$, queue, freight train classification, non-preemptive failures

JEL Classification: C44
AMS Classification: 60K25

1 Introduction

Marshalling yards play an important role in freight railway transport because two important processes are carried out in them: classification of inbound freight trains entering the marshalling yard and forming outbound freight trains leaving the marshalling yard. A survey of operations related to the problem modelled in the article is provided in [2]. In the article we pay attention to modelling the classification process.

Each inbound freight train that has entered the marshalling yard has to be classified because the individual wagons that form the train usually differ in their destination. Classification is done by gravity – each inbound freight train is pulled by a shunting locomotive from reception sidings towards a hump. By treatment of gravity the individual wagons or group of wagons start to move separately and are sent to selected tracks of sorting sidings. Apart from classification of the inbound freight trains, trains of wagons coming from industrial sidings have to be also classified. Such trains of wagons are also pulled over the hump.

Let us call classification of the inbound freight trains primary shunting and classification of the trains of wagons coming from the industrial sidings secondary shunting. We consider primary shunting to be one of the primary functions of marshalling yards. That means primary shunting takes precedence over secondary shunting. Unfortunately, both primary and secondary shunting use the same infrastructure and, therefore, cannot be done at the same time. That is the reason why we decided to model the classification process as a queueing system subject to non-preemptive failures that are represented by requests for secondary shunting.

Modelling interruptions of the service process in queueing systems is quite frequent. A survey of queueing models subject to different types of service interruptions is given in [7]. A typical reason of the service interruption is a server breakdown. In the past, many articles were devoted to modelling server breakdowns. We can mention some of them. In article [5] multiple types of server breakdowns are assumed. Server breakdowns emptying the queueing system are called disasters or disastrous breakdowns – see [8]. In article [6] working

[1] VŠB – Technical University of Ostrava / Faculty of Mechanical Engineering, Institute of Transport, 17. listopadu 15, Ostrava – Poruba, phone: +420 597 325 754, e-mail: michal.dorda@vsb.cz.
[2] VŠB – Technical University of Ostrava / Faculty of Mechanical Engineering, Institute of Transport, 17. listopadu 15, Ostrava – Poruba, phone: +420 597 324 575, e-mail: dusan.teichmann@vsb.cz.
[3] VŠB – Technical University of Ostrava / Faculty of Mechanical Engineering, Institute of Transport, 17. listopadu 15, Ostrava – Poruba, phone: +420 597 325 754, e-mail: vojtech.graf@vsb.cz.

breakdowns that slow down the service process are introduced. Article [1] examines the effect of postponing service interruptions that are caused by server breakdowns.

The presented article continues in articles [3] and [4]. The referenced articles are devoted to a mathematical model of an M/E_n/1/m queueing system subject to non-preemptive failures. The model was applied to model classification processes in marshalling yards. The presented article further develops the original model; the difference lies in the fact that the improved model works on the assumption of Erlang distributed times to repair whilst the original model works with exponentially distributed times to repair. Such generalization brings a broader field of applicability.

2 Notation and mathematical model

Let us assume a finite queueing system consisting of a server and a queue with a finite capacity which is equal to m-1, where $m \geq 2$. That means the total capacity of the system is m and corresponds to the number of arrival tracks of the marshalling yard. Let us consider that the arrival process of customers is the Poisson process which is defined with the arrival rate λ. The customers are represented by inbound freight trains that have entered the marshalling yard and have been prepared for their classification.

The classification process of each inbound freight train (primary shunting) corresponds to the service provided in the modelled queueing system and takes a random time; we assume that the times can be modelled with the Erlang distribution of probability defined by the shape parameter $n \geq 2$ and the scale parameter $n\mu$. The Erlang distributed times are decomposed into n mutually independent exponentially distributed phases, each of the phases is defined by the parameter $n\mu$. That means the mean service time is equal to $\frac{n}{n\mu} = \frac{1}{\mu}$.

Trains of wagons that have entered the marshalling yard from industrial sidings leading to the marshalling yard have to be also classified in the marshalling yard (secondary shunting). As mentioned in the introduction, when secondary shunting is being carried out, primary shunting is interrupted and, therefore, secondary shunting is considered to be a failure of the server (all the elements of the infrastructure needed for primary shunting are occupied by secondary shunting). Moreover, it is more than obvious, that secondary shunting does not interrupt primary shunting immediately when a new request for secondary shunting is made. Such request for secondary shunting has to wait until primary shunting of the inbound freight train which is just being classified is finished and then secondary shunting can be initiated. Therefore we call such failures of the server non-preemptive. Let the arrival process of the non-preemptive failures be also the Poisson process but with the parameter η. If a failure already is in the system, then the parameter η is equal to zero. Times to repair are assumed to be Erlang distributed with the shape parameter s and with the scale parameter $s\zeta$. That means the times to repair can be modelled in the same manner as the service times. Please remind that the mean time to repair is equal to $\frac{s}{s\zeta} = \frac{1}{\zeta}$.

A summary of all the random variables used in the model is given in Table 1.

Random variables	**Meaning in practice**	**Probability distributions**	**Parameters**
Costumer inter-arrival times	Primary shunting	Exponentially distributed	$\lambda > 0$
Customer service times		Erlang distributed	$n \geq 2$, $n\mu > 0$
Times between failures	Secondary shunting	Exponentially distributed	$\eta > 0$
Times to repair		Erlang distributed	$s \geq 2$, $s\zeta > 0$

Table 1 The summary of all the random variables used in the model

To describe possible states of the modelled queueing system, let us define three discrete random variables denoted as K, P, F. The variable K describes the number of the customers (the inbound freight trains) found in the system, the variable can takes values from 0 to m. The variable P expresses how many phases of the Erlang distributed service time or time to repair have already been finished. The variable can take values from 0 to n-1 for the customer service and from 0 to s-1 for the repair of the non-preemptive failure. The third variable F can take its values from the set {0,1,2}. The meaning of the individual values of F is:

- If F is 0, then no failure is in the system.
- If F is 1, then the non-preemptive failure of the server is waiting for repair because we have to finish the service of the customer being in the service.
- If F is 2, then the non-preemptive failure of the server is under repair.

Using the discrete random variables K, P and F all the possible states of the system can be defined by triplets (*k*,*p*,*f*). The state space of the system can be expressed as:

$$\{(0,0,0)\}\cup\{(k,p,f):k\in\{1,\dots,m\},p\in\{0,\dots,n-1\},f\in\{0,1\}\}\cup\{(k,p,f):k\in\{0,\dots,m-1\},r\in\{0,\dots,s-1\},f=2\}.$$

Figure 1 presents a state transition diagram, the vertices represent the states of the system and the oriented edges indicate the possible transitions with the corresponding rate.

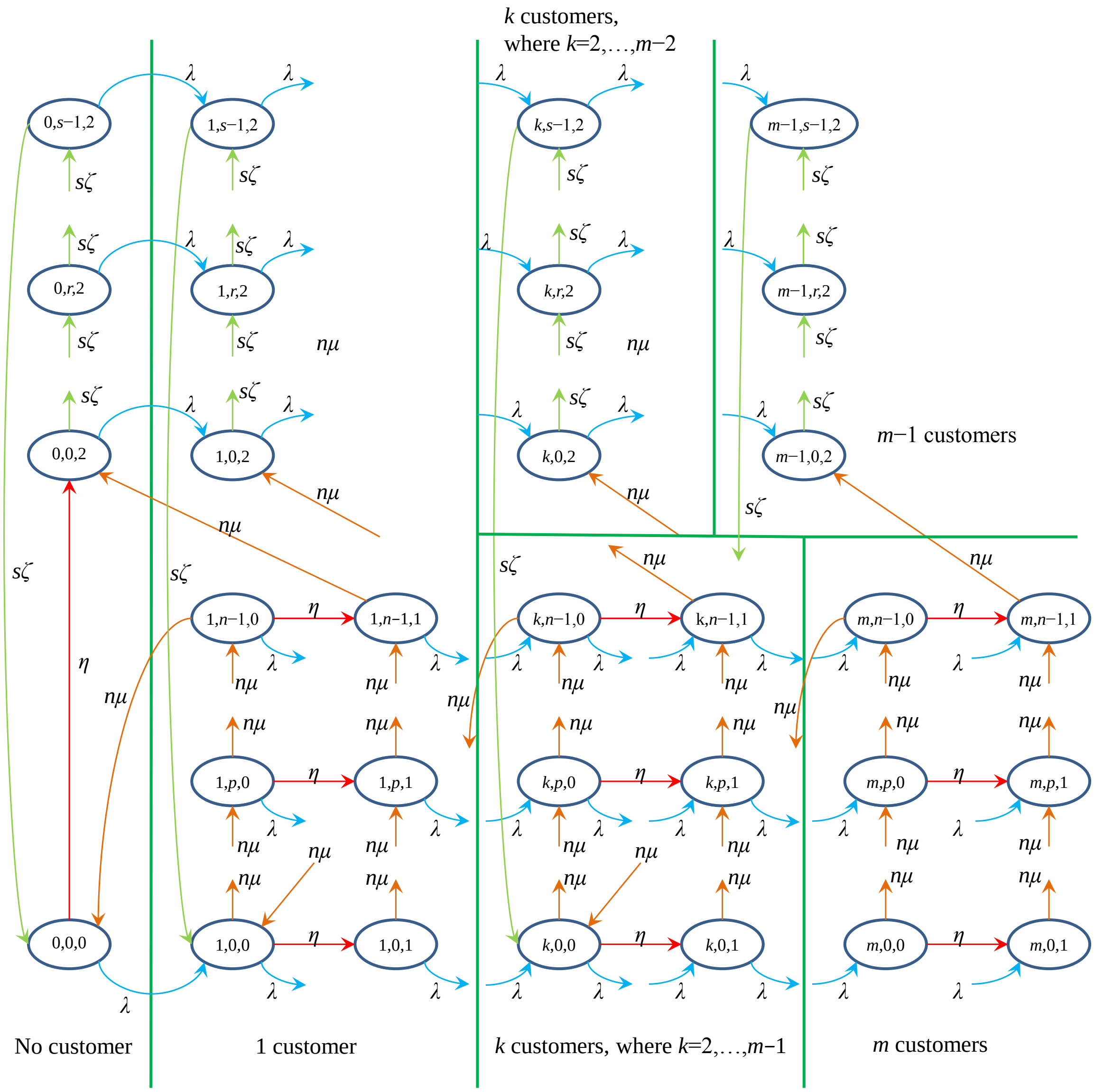


Figure 1 The state transition diagram of the system

On the basis of the state transition diagram we can write liner equations for the individual state probabilities. The linear equation system consists of the following equations:

$$(\lambda+\eta)\cdot P_{(0,0,0)}=n\mu\cdot P_{(1,n-1,0)}+s\zeta\cdot P_{(0,s-1,2)}\,, \tag{1}$$

$$(\lambda+n\mu+\eta)\cdot P_{(k,0,0)}=\lambda\cdot P_{(k-1,0,0)}+n\mu\cdot P_{(k+1,n-1,0)}+s\zeta\cdot P_{(k,s-1,2)}\ \text{ for }\ k=1,\dots,m-1, \tag{2}$$

$$(n\mu+\eta)\cdot P_{(m,0,0)}=\lambda\cdot P_{(m-1,0,0)}\,, \tag{3}$$

$$(\lambda+n\mu)\cdot P_{(1,0,1)}=\eta\cdot P_{(1,0,0)}\,, \tag{4}$$

$$(\lambda + n\mu)\cdot P_{(k,0,1)} = \eta \cdot P_{(k,0,0)} + \lambda \cdot P_{(k-1,0,1)} \text{ for } k = 2,\dots,m-1, \tag{5}$$

$$n\mu \cdot P_{(m,0,1)} = \eta \cdot P_{(m,0,0)} + \lambda \cdot P_{(m-1,0,1)}, \tag{6}$$

$$(\lambda + n\mu + \eta)\cdot P_{(1,p,0)} = n\mu \cdot P_{(1,p-1,0)} \text{ for } p = 1,\dots,n-1, \tag{7}$$

$$(\lambda + n\mu + \eta)\cdot P_{(k,p,0)} = \lambda \cdot P_{(k-1,p,0)} + n\mu \cdot P_{(k,p-1,0)} \text{ for } k = 2,\dots,m-1 \text{ and } p = 1,\dots,n-1, \tag{8}$$

$$(n\mu + \eta)\cdot P_{(m,p,0)} = \lambda \cdot P_{(m-1,p,0)} + n\mu \cdot P_{(m,p-1,0)} \text{ for } p = 1,\dots,n-1, \tag{9}$$

$$(\lambda + n\mu)\cdot P_{(1,p,1)} = \eta \cdot P_{(1,p,0)} + n\mu \cdot P_{(1,p-1,1)} \text{ for } p = 1,\dots,n-1, \tag{10}$$

$$(\lambda + n\mu)\cdot P_{(k,p,1)} = \eta \cdot P_{(k,p,0)} + \lambda \cdot P_{(k-1,p,1)} + n\mu \cdot P_{(k,p-1,1)} \text{ for } k = 2,\dots,m-1 \text{ and } p = 1,\dots,n-1, \tag{11}$$

$$n\mu \cdot P_{(m,p,1)} = \eta \cdot P_{(m,p,0)} + \lambda \cdot P_{(m-1,p,1)} + n\mu \cdot P_{(m,p-1,1)} \text{ for } p = 1,\dots,n-1, \tag{12}$$

$$(\lambda + s\zeta)\cdot P_{(0,0,2)} = \eta \cdot P_{(0,0,0)} + n\mu \cdot P_{(1,n-1,1)}, \tag{13}$$

$$(\lambda + s\zeta)\cdot P_{(k,0,2)} = n\mu \cdot P_{(k+1,n-1,1)} + \lambda \cdot P_{(k-1,0,2)} \text{ for } k = 1,\dots,m-2, \tag{14}$$

$$s\zeta \cdot P_{(m-1,0,2)} = n\mu \cdot P_{(m,n-1,1)} + \lambda \cdot P_{(m-2,0,2)} \tag{15}$$

$$(\lambda + s\zeta)\cdot P_{(0,r,2)} = s\zeta \cdot P_{(0,r-1,2)} \text{ for } r = 1,\dots,s-1, \tag{16}$$

$$(\lambda + s\zeta)\cdot P_{(k,r,2)} = \lambda \cdot P_{(k-1,r,2)} + s\zeta \cdot P_{(k,r-1,2)} \text{ for } k = 1,\dots,m-2 \text{ and } r = 1,\dots,s-1, \tag{17}$$

$$s\zeta \cdot P_{(m-1,r,2)} = \lambda \cdot P_{(m-2,r,2)} + s\zeta \cdot P_{(m-1,r-1,2)} \text{ for } r = 1,\dots,s-1. \tag{18}$$

An equation – for example equation (1) – of linear equation system (1) – (18) is redundant and, therefore, has to be omitted and replaced by normalization condition (19) in order to get a unique solution of the system. The normalization condition is given by the following formula:

$$P_{(0,0,0)} + \sum_{k=1}^{m}\sum_{p=0}^{n-1}\sum_{f=0}^{1} P_{(k,p,f)} + \sum_{k=0}^{m-1}\sum_{r=0}^{s-1} P_{k,r,2} = 1. \tag{19}$$

The resulting equation system consists of $2\cdot m\cdot n + m\cdot s + 1$ linear equations formed by equations (2) up to (19) with $2\cdot m\cdot n + m\cdot s + 1$ unknown stationary probabilities. To solve the linear equation system in Matlab it is necessary to employ an alternative one-dimensional state description in the following form:

- The states (k,p,f) for k=1,...,m, p=0,...,n-1, and f=0,1 are denoted $f\cdot m\cdot n + p\cdot m + k$,
- The states $(k,r,2)$ for k=0,...m-1 and r=0,...,n-1 are denoted $2\cdot m\cdot n + r\cdot m + k + 1$.
- The state (0,0,0) is labelled as $2\cdot m\cdot n + s\cdot m + 1$.

After numerical solving the linear equation system we get the steady-state probabilities, on the basis of them some performance measures can be computed. Let us focus on the following performance measures. The mean number of the customers in the service *ES* (utilization of the hump by primary shunting) is equal to:

$$ES = \sum_{k=1}^{m}\sum_{p=0}^{n-1}\sum_{f=0}^{1} P_{k,p,f}\ . \tag{20}$$

The mean number of the customers waiting in the queue *EL* (the mean number of the inbound freight trains waiting in the system in order to be classified) is given by expression:

$$EL = \sum_{k=2}^{m}(k-1)\cdot\sum_{p=0}^{n-1}\sum_{f=0}^{1} P_{k,p,f} + \sum_{k=1}^{m-1} k\cdot\sum_{r=0}^{s-1} P_{k,r,2}\ . \tag{21}$$

For the mean number of the customers found in the system *EK* (the mean number of the inbound freight trains found in the system) it has to hold that:

$$EK = ES + EL = \sum_{k=1}^{m} k\cdot\sum_{p=0}^{n-1}\sum_{f=0}^{1} P_{k,p,f} + \sum_{k=1}^{m-1} k\cdot\sum_{r=0}^{s-1} P_{k,r,2}\ . \tag{22}$$

And finally, for the mean number of the broken servers *EF* (utilization of the hump by secondary shunting) it holds:

$$EF = \sum_{k=0}^{m-1}\sum_{r=0}^{s-1} P_{k,r,2}\ . \tag{23}$$

3 Results of numerical experiments

To demonstrate solvability of the mathematical model several numerical experiments were carried out. The values of the individual parameters used for the experiments are given in Table 1. The results of the experiments are presented in Figures 2 and 3. The figures show the dependences of the individual performance measures *ES*, *EL*, *EK* and *EF* on the parameters η and $s\zeta$.

Random variables	Applied values of the parameters
Costumer inter-arrival times	$\lambda = 0.01520\ \text{min}^{-1}$
Customer service times	$n = 10$, $n\mu = 0.63622\ \text{min}^{-1}$
Times between failures	$\eta = 0.01 – 0.05\ \text{min}^{-1}$ (the step $0.004\ \text{min}^{-1}$)
Times to repair	$s = 5$, $s\zeta = 0.05 – 0.25\ \text{min}^{-1}$ (the step $0.02\ \text{min}^{-1}$)

Table 2 The parameters for the numerical experiments

The numerical experiments confirmed the following dependences:

- The values of *ES* decrease with the increasing value of η (that is because the non-preemptive failures are more frequent) and increase with the increasing value of $s\zeta$ (that is because the increasing value of $s\zeta$ makes the times to repair shorter for the constant value of s).
- The values of *EL*, *EK* and *EF* increase with the increasing value of η and with the decreasing value of $s\zeta$.

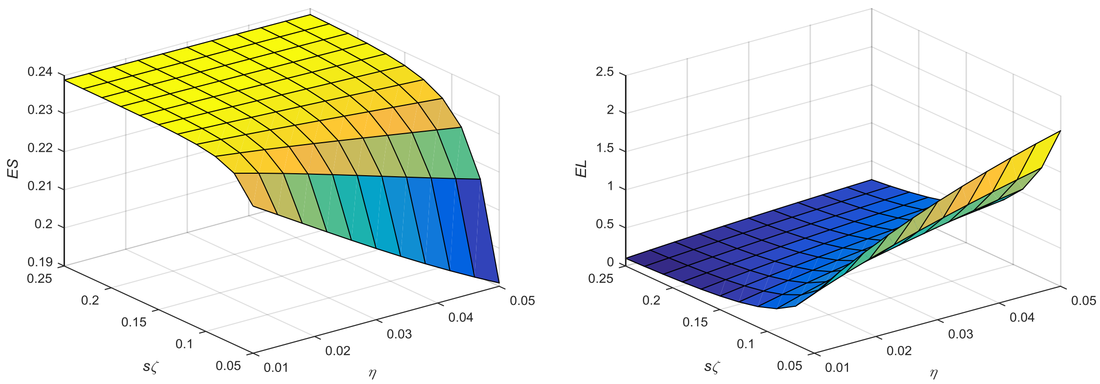


Figure 2 The dependence of *ES* and *EL* on η and $s\zeta$

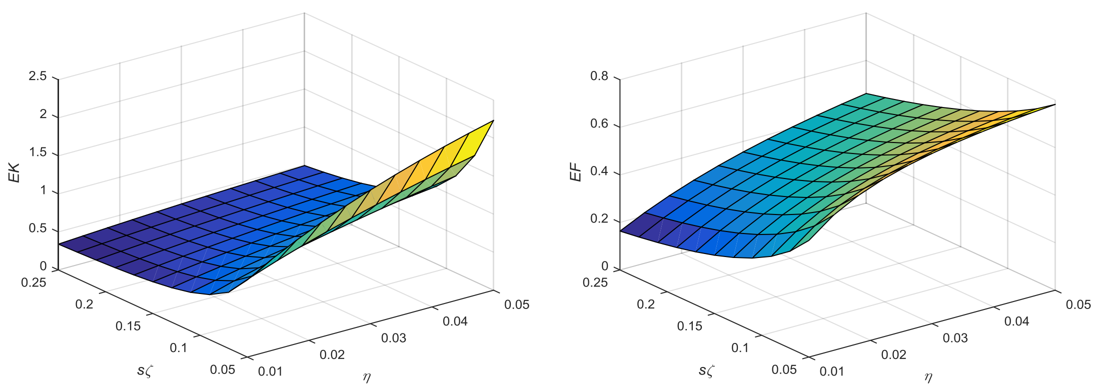


Figure 3 The dependence of *EK* and *EF* on η and $s\zeta$

We can state that the dependencies obtained by the numerical experiments are consistent with our logical expectations.

4 Conclusions

In the article the single-server queueing model subject to the non-preemptive failures is presented. The model was developed in order to model the classification processes in the marshalling yards. For the marshalling yards it is typical that two different input streams enter the marshalling yards – the inbound freight trains (requests for primary shunting) represent the customers and trains of wagons coming from industrial sidings (requests for secondary shunting) are modelled with the non-preemptive failures that interrupt primary shunting.

As regards our future research, other performance measures (for example the mean waiting time) could be taken into account. The results of the mathematical model can be verified by simulation. And finally, the inter-arrival times and the times between failures can be also modelled with the Erlang distribution of probability in order to create a more general mathematical model.

References

[1] Blanc, H. J.: M/G/1 queues with postponed interruptions. *ISRN Probability and Statistics*, vol. 2012, Article ID 653167, 12 pages, 2012. doi:10.5402/2012/653167.

[2] Boysen, N., Fliedner, M., Jaehn, F., and Pesch, E.: Shunting yard operations: Theoretical aspects and applications. *European Journal of Operational Research* **220.1** (2012), 1-14.

[3] Dorda, M., and Teichmann, D.: On finite single-server queue subject to non-preemptive breakdowns. In: *Proceedings of the 31st International Conference Mathematical Methods in Economics 2013, Part I.* (Vojáčková, H. ed.). College of Polytechnics Jihlava, Jihlava, 2013, 141-146.

[4] Dorda, M., and Teichmann, D. Modelling of freight trains classification using queueing system subject to breakdowns. *Mathematical Problems in Engineering*, vol. 2013, Article ID 307652, 11 pages, 2013. doi:10.1155/2013/307652.

[5] Jain, M., and Jain, A.: Working vacations queueing model with multiple types of server breakdowns. *Applied Mathematical Modelling, 34*(1) (2010), 1-13.

[6] Kalidass, K., and Kasturi, R.: A queue with working breakdowns. *Computers & Industrial Engineering, 63*(4) (2012), 779-783.

[7] Krishnamoorthy, A., Pramod, P. K., and Chakravarthy, S. R.: Queues with interruptions: a survey. *Top* 22.1 (2014), 290-320.

[8] Sudhesh, R.: Transient analysis of a queue with system disasters and customer impatience. *Queueing systems, 66*(1) (2010), 95-105.

Data Extension for Stock Market Analysis

Marek Dvořák[1] , Jaromír Kukal[2] , Petr Fiala[3]

Abstract. Time series of stock prices were subjects of multivarietal statistical analysis. After basic data description obtained by logarithmic differences, several methods of data augmentations were implemented. Several dozen of new variables were computed from sliding window of the differentiated time series using simple statistic tools like first and second moments as well as more complicated statistic, like auto-regression coefficients and residual analysis, followed by an optional quadratic transformation that was further used for data extension. Regularized logistic regression helped in prediction of Buy-Sell Index (BSI) from real stock market data. Various measures of prediction quality were discussed. The regularization gain affected both the number of descriptors and prediction accuracy. The prediction system was optimized on a group of stock series and then cross-validated on another group of stocks.

Keywords: Logistic regression, stock market, data augmentation, regularized estimation, cross validation.

JEL classification: C44
AMS classification: 90C15

1 Introduction

The role of predicting stock market behavior may seem a bit ambitious at first and we certainly are not the first one to tackle this theme. One of the most obvious models that could be used is the ARIMA model, which is autoregressive integrated moving average, to predict future time series values. An extension to that might be a simple seasonal adjustments, which could decrease some variability in predictions. More robust way for market value prediction can be found in vector autoregression models (VAR). These methods use multidimensional input data to estimate lagged interdependency on each other. Our model does not take multi-dimensionality into consideration, and instead focuses on prediction based on a single time series. This technique does not attempt to estimate the value of future market values directly though. We use a simple single dimensional time series source in addition to data extension and augmentation techniques, so that we could then estimate what we call a BSI (Buy/Sell Index). BSI is a simple indicator that tells us if and when we should invest into a stock, so that we can sell it in the future with a profit.

2 Data augmentation and preparation

The data consists of several time series imported using R's *quantmod* package [6]. One of the uses of this package is a Yahoo Finance API to import daily stock market data. As we mentioned in the introduction the model is not multidimensional. Instead those multiple series are merely used to obtain more data for training, validation and testing. At first, simple logarithmization and differentiation of time series is used. X_t is defined as the price of the stock

$$x_t := \log X_t \tag{1}$$

$$\Delta x_t = x_t - x_{t-1} \tag{2}$$

[1]University of Economics, Prague, W. Churchill Sq. 1938/4, Prague, Czech Republic, marek.dvorak@vse.cz
[2]Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Břehová 7, Prague, Czech Republic, kukal@fjfi.cvut.cz
[3]University of Economics, Prague, W. Churchill Sq. 1938/4, Prague, Czech Republic, pfiala@vse.cz

The preparation begins with definition of two windows and their sizes, w and k. w is the length of a window from which we extend the predictor variables and which we call a *lookbehind*. k denotes the length of a window which we call a *lookahead*. It is the window which will be used to extract dependent variable - that is the BSI index. n is the number of observations.

2.1 Preprocessing explanatory variables

For each lookbehind window of fixed length, we compute following statistics: mean (3), median (4), mid (5), logarithm of standard deviation (6), logarithm of mean absolute deviation (7), logarithm of median absolute deviation (8). The index of this new series is defined as s.

$$x_s^{mean} = \frac{\sum_{t=s}^{s+w-1} \Delta x_t}{w} \text{ for each } s = 1 \dots n-w-k-2 \tag{3}$$
$$x_s^{median} = \text{median}_{t=s}^{s+w-1} \Delta x_t \text{ for each } s = 1 \dots n-w-k-2 \tag{4}$$
$$x_s^{mid} = \frac{\max_{t=s}^{s+w-1} \Delta x_t + \min_{t=s}^{s+w-1} \Delta x_t}{2} \text{ for each } s = 1 \dots n-w-k-2 \tag{5}$$
$$x_s^{sd} = \ln\left(\sqrt{\frac{1}{w}\sum_{t=s}^{s+w-1} (\Delta x_t - x_s^{mean})^2}\right) \text{ for each } s = 1 \dots n-w-k-2 \tag{6}$$
$$x_s^{mad} = \ln\left(\frac{1}{w}\sum_{t=s}^{s+w-1} |\Delta x_t - x_s^{mean}|\right) \text{ for each } s = 1 \dots n-w-k-2 \tag{7}$$
$$x_s^{mad^*} = \ln\left(\text{median}_{t=s}^{s+w-1} |\Delta x_t - x_s^{mean}|\right) \text{ for each } s = 1 \dots n-w-k-2 \tag{8}$$

After that another set of variables were calculated. For each lookbehind window s, three autoregression models were estimated - AR(1), AR(2) and AR(3), with c_s and φ_s being regressors and $\epsilon_{s,t}$ being the error term – white noise.

$$\Delta x_t = c_s^1 + \varphi_s^{1,1} \Delta x_{t-1} + \epsilon_{s,t}^1 \tag{9}$$
$$\Delta x_t = c_s^2 + \varphi_s^{2,1} \Delta x_{t-1} + \varphi_s^{2,2} \Delta x_{t-2} + \epsilon_{s,t}^2 \tag{10}$$
$$\Delta x_t = c_s^3 + \varphi_s^{3,1} \Delta x_{t-1} + \varphi_s^{3,2} \Delta x_{t-2} + \varphi_s^{3,3} \Delta x_{t-3} + \epsilon_{s,t}^3 \tag{11}$$

These estimated AR parameters were used as another explanatory variables for the main model. One for each AR model and regressor. Six in total.

$$x_s^{ar(1,1)} = \varphi_s^{1,1} \tag{12}$$
$$x_s^{ar(2,1)} = \varphi_s^{2,1} \tag{13}$$
$$x_s^{ar(2,2)} = \varphi_s^{2,2} \tag{14}$$
$$x_s^{ar(3,1)} = \varphi_s^{3,1} \tag{15}$$
$$x_s^{ar(3,2)} = \varphi_s^{3,2} \tag{16}$$
$$x_s^{ar(3,3)} = \varphi_s^{3,3} \tag{17}$$

Finally last set of nine augmented predictors are calculated for each window using residuals $e_{s,t}^i$ from the previous autoregressive models. These are logarithms of standard deviation of residuals (18), logarithm of mean absolute difference of residuals (19) and logarithm of median absolute difference of residuals (20).

$$x_s^{ar(i)sd} = \ln\left(\sqrt{\frac{1}{w}\sum_{t=s}^{s+w-1}\left(e_{s,t}^1\right)^2}\right) \text{ for each } s = 1\dots n-w-k-2 \tag{18}$$

$$x_s^{ar(i)mad} = \ln\left(\frac{1}{w}\sum_{t=s}^{s+w-1} e_{s,t}^2\right) \text{ for each } s = 1\dots n-w-k-2 \tag{19}$$

$$x_s^{ar(i)mad} = \ln\left(\text{median}_{t=s}^{s+w-1} e_{s,t}^3\right) \text{ for each } s = 1\dots n-w-k-2 \tag{20}$$

Together that makes 21 variables, each of the length $n-w-k-2$, which makes up the predictor vector.

2.2 Preprocessing dependent variable

Dependent variable is what we call a BSI index and is estimated with (22). Worthiness of such investment is determined by looking ahead up to a fixed amount of time (lookahead window) and calculating if there is a time in which selling a stock would yield a profit. Our model takes into account transaction prices (q for buy and p for sell) and a threshold of determination Θ which determines minimum amount of yield. The buying strategy would then consist of monitoring when the model estimates 1 as a value of BSI index and buying the stock at that time. Then selling the stock as soon as the inequality (21) is satisfied. The threshold Θ is defined apriori as another parameter of the model, alongside w and k. The X_i is defined form the section 2 as the price of the stock.

$$\left(\frac{X_t\,(1+q)}{X_{t+i}\,(1-p)}\right)^{\frac{365}{i}} \geq \Theta \tag{21}$$

$$BSI_s = \begin{cases} 0 & \max_{j=1+w}^{k+w}\left(\frac{X_t(1+q)}{X_{t+j}(1-p)}\right)^{\frac{365}{j}} \leq \Theta \\ 1 & \max_{j=1+w}^{k+w}\left(\frac{X_t(1+q)}{X_{t+j}(1-p)}\right)^{\frac{365}{j}} \geq \Theta \end{cases} \text{ for each } s = 1\dots n-w-k-2 \tag{22}$$

2.3 Interaction

In section 2.1, we have shown how to augment dataset using 21 explanatory variables. This amount can further be augmented by considering an interaction between them. For each of 21 variables, we create another set of explanatory variables, which are a multiplication of each of the original 21 explanatory variables. That adds another $\frac{21^2-21}{2}$ explanatory variables bringing the total number of explanatory variables to 231.

3 Logistic regression and LASSO

Considering the nature of dependent variable, a simple binary classificator could be used. The choice was a logistic regression (as can be seen in [4])

A logistic regression can be estimated using a generalized linear model as such:

$$\text{logit}\,(BSI_s) = \ln\left(\frac{BSI_s}{1-BSI_s}\right) = \beta_0 + \beta_1 x_t^{mean} + \beta_2 x_t^{median} + \dots \tag{23}$$

where the time indices t and s refer to the equivalent variables x_t and BSI_s defined in (1) and (22). However with 231 explanatory variables might result to over fitting (depending on the number of observations). There are numerous ways to deal with this issue as can be seen in [2], [5], [7] or [8]. With our data, LASSO technique was used to restrict the number of explanatory variables. This means imposing a L^1-norm constraint on the vector of model coefficients as shown in (24).

$$\sum_{j=1}^{231} |\beta_j| \leq t \quad t > 0 \tag{24}$$

More specifically, the package *glmnet* [1] for the R software solves following problem (25) [3], where the λ is the strength of regularization penalty, $l\,(y, \eta)$ is the negative log-likelihood contribution for observation i, β are regressors, w is a weight matrix, and $\alpha = 1$ for L^1-norm LASSO regression.

$$\min_{\beta_0,\beta} \frac{1}{N} \sum_{i=1}^{N} w_i l\left(y_i, \beta_0 + \beta^T x_i\right) + \lambda \left[(1-\alpha)\, ||\beta||_2^2/2 + \alpha ||\beta||_1\right] \tag{25}$$

4 Cross-Validation and results

The estimation uses Cross-Validation as another method to deal with over fitting.

The model was fitted and cross-validated on data of several Nasdaq stock market series, namely VIAB, GOOG, FB and NFLX. The area under the curve (AUC) of ROC of the logistic binary classificator was used as a model quality measure. Model parameters were estimated on the sample of 1000 data points from those series. Number of cross-validation sets were 5. After that, out of sample test on the SBUX stock was predicted and its AUC was computed.

The regularization parameter λ was chosen as such that it maximizes the AUC of the validation sample. If the λ is too large, the number of dependent variable gets too low and the model looses its prediction capabilities. If the regularization parameter is too small, the number of dependent parameters is too large and the model is prone to over fitting, losing its generalization properties and also providing suboptimal results.

The choice of apriori parameters w, k and Θ of the out of sample AUC was tested. The base parameters $w = 14; k = 14; \Theta = 0.2$ were chosen. And then several variations along each parameter were changed. The results are in the table 1.

w	k	Θ	AUC
14	7	0.2	0.8375
14	14	0.2	0.757
14	21	0.2	0.7282
14	30	0.2	0.7185
14	60	0.2	0.571
5	14	0.2	0.7426
9	14	0.2	0.7747
17	14	0.2	0.7557
14	14	0.1	0.7726
14	14	0.15	0.7769
14	14	0.25	0.7808

Table 1 AUC dependency on the apriori parameters

Figure 1 shows example ROC curve of prediction using the parameters $w = 14; k = 14; \Theta = 0.2$ on SBUX out of sample dataset.

Figure 2 shows dependency of area under curve metric on the number of explanatory variables and the regularization parameter λ.

5 Conclusion

We have introduced a model useful for prediction of univariete time series. We have described dependent and explanatory variables as well as method for estimation. Since the dependent was a binary variable,

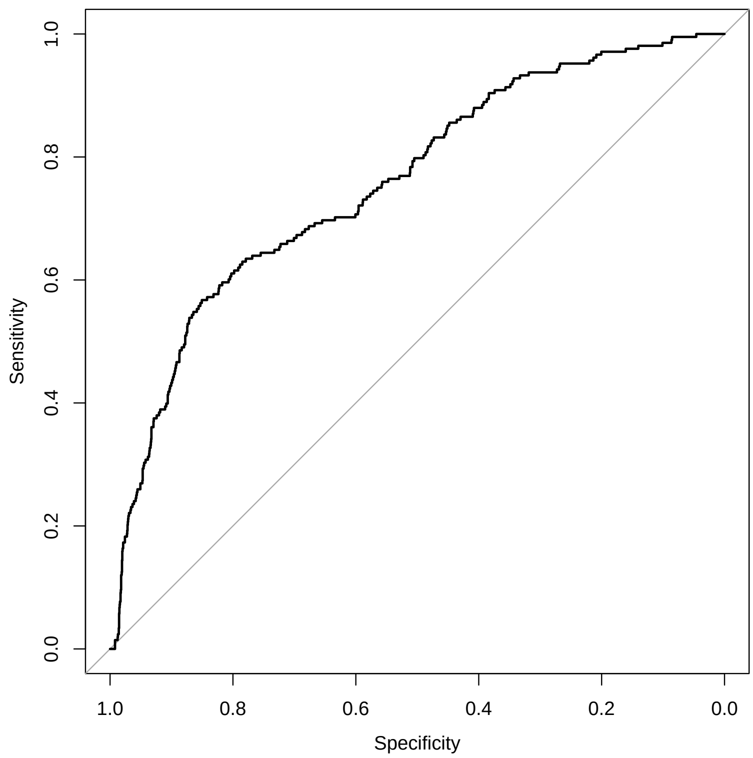


Figure 1 ROC curve of VIAB out of sample test dataset
Source: own R output

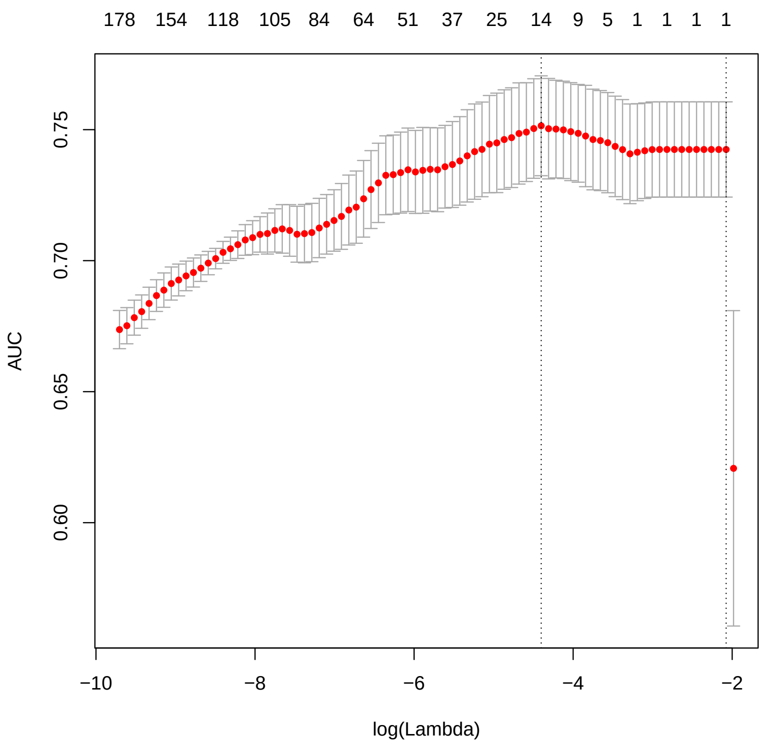


Figure 2 Dependency of AUC on λ
Source: own R output

area under the curve of receiver operating characteristic was used as a measure of quality. Further discussions may focus on the threshold of predicted variables that would yield best investment strategy together with some additional aposteriori simulations might be conducted.

Acknowledgements

Supported by the grant No. IGA F4/54/2015 of the Internal Grant Agency, Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] Friedman, J., Hastie, T., and Tibshirani, R.: package 'glmnet'. *R package version 2.0-5. URL http://cran. r-project. org/web/packages/glmnet* (2016).

[2] Gold, C., Holub, A., and Sollich, P.: Bayesian approach to feature selection and parameter tuning for support vector machine classifiers. *Neural Networks* **18** (2005), 693–701.

[3] Hastie, T., and Qian, J.: Glmnet vignette, 2014. URL `http://web.stanford.edu/~hastie/glmnet/glmnet_alpha.html`.

[4] Hosmer Jr, D. W., and Lemeshow, S.: *Applied logistic regression.* John Wiley & Sons, 2004.

[5] Majumdar, A., and Ward, R. K.: Classification via group sparsity promoting regularization. In: *2009 IEEE International Conference on Acoustics, Speech and Signal Processing.* IEEE, 861–864.

[6] Ryan, J. A.: quantmod: Quantitative financial modelling framework. *R package version 0.3-5. URL http://www. quantmod. com URL http://r-forge. r-project. org/projects/quantmod* (2016).

[7] Wheeler, G.: The lasso logistic regression model: Modifications to aid causality assessment for adverse events following immunization. *London School of Hygiene and Tropical Medicine* (2010).

[8] Zou, H., and Hastie, T.: Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **67** (2005), 301–320.

Multi-parametric simulation of a model of nonconventional public projects evaluated by cost-benefit analysis

Stanislava Dvořáková[1], Petr Jiříček [2]

Abstract. The paper will deal with simulation of cash-flows of investment character public projects subsidized from public funds. In the model of a project, the influence of changes in the amount of subsidy and changes of economic cash-flows generated by the project on the result of public investment social efficiency evaluation will be examined by the economic internal rate of return (ERR) method. The constructed deterministic model based on a rational fractional function will be used to examine the characteristics of projects when alternation of positive and negative cash-flows occurs in the course of an investment process. For modelling the project, we will use the EU methodology for assessing socio-economic evaluation of project utility by means of cost-benefit analysis issuing from Kaldor–Hicks social optimality test. Using the Maple program, the simulation of individual scenarios of nonconventional projects will be presented in 3D presentation and assessed by means of the value of the root function in relation to project parameters. The analysis of results by the simulation of public project scenarios will be interpreted in relation to economic conditions possible in real life.

Keywords: Non-conventional public projects, root function, subsidies from European funds, cost-benefit analysis.

JEL classification: C20, H43
AMS classification: 65H04, 68R10

1 Introduction

The aim of public projects is to raise the utility for the stakeholders and to increase social well-being. Unlike commercial projects, we consider not only net financial benefits, but we also define benefits and costs (losses) issuing from external impacts of the project on the society. The principles for evaluating public projects are provided by theoretical principles of welfare economics. The welfare function as the sum of discounted utilities is featured e.g. by Perman et al. [7]. These utilities are given by the difference between benefits and costs (losses) from a social change caused by implementing a public project. Public project efficiency evaluation uses in its mathematical basis commercial methods – the net present value (NPV) method and the internal rate of return method (IRR), the issue of negative cash flows in the area of commercial nonconventional projects being summarised by Osborne [6] or Magni [5]. The commercial net present value method (NPV) is after applying the theoretical principles of welfare economics available for evaluation of public projects of investment character and under the name of economic net present value ($ENPV$) it is generally used as the preferred method for evaluating projects from European funds. These are characterized by certain amount of subsidy to the supported subjects. Another preferred method for evaluating projects subsidized from the European funds is the so called method of economic internal rate of return (ERR), which is a modification of the IRR commercial method. This is based on finding a real root of the $ENPV$ function plotted from the parameters of the proposed project. The present methodology of the European Union issues namely from works of Florio and Vignetti [2, 3]. Public projects are supported from the EU funds in the situation when a project generates negative FRR and $FNPV$ values and so it is inefficient from purely commercial perspective. By adding positive cash-flows from positive externalities we then arrive in the phase of economic analysis from negative

[1]Department of Mathematics, College of Polytechnics Jihlava, 586 01 Jihlava, Czech Republic, stanislava.dvorakova@vspj.cz
[2]Department of Economic Studies, College of Polytechnics Jihlava, 586 01 Jihlava, Czech Republic, petr.jiricek@vspj.cz

values of IRR and NPV at positive values on the basis of ERR and $ENPV$. In the real economic situation, however, there may be scenarios when in a public project, usually proposed as a conventional one, negative cash-flows are generated even on the basis of economic cash-flows (e.g. due to unexpectedly high negative environmental externalities). Those negative cash-flows will cause a nonconventional nature of the project, i.e. there may not be exactly one positive real-valued root of the $ENPV$ function (i.e. the value of $ERR > 0$). Then it is not possible to decide conclusively about a project on the basis of the ERR method. The paper focuses on this very area of public projects where a partial subsidy form European funds is expected and it examines basic behaviour of the ENPV function roots by means of 3D simulation in fundamental model states of nonconventional public projects.

2 Aims and methods

The presented paper aims at simulating a nonconventional project in selected economically legitimate scenarios on project evaluation by the ERR method. To describe the behaviour of a public project under the conditions of subsidy from the European funds, a proposed deterministic model of the ENPV function will be used. To perform a complex analysis of the $ENPV$ function and its roots (ERR) especially in the positive real area in relation to parameter changes, 3D images will be used within numerical solving of the model by the Maple program. The simulated model scenarios will capture the changes of the $ENPV$ function with a view to the amount of subsidy (30%, 70% a 100%) related to the rise of negative cash-flows invoked by negative externalities in operational and liquidation phase of a public investment project.

2.1 Assumption

Conventional projects – they are characterized by only one change of their polarity in the sequence of the stream of cash-flows generated within the project (e.g. $[-, - - + + ++]$).

Non-conventional projects – they are characteristic by more than one change of project generated cash-flows in the sequence of the CF stream (e.g. $[-, - + + + +-]$).

Investment curve of the project – defined by the rationally fractional function

$$ENPV = \sum_{t=0}^{n} \frac{CF_t}{(1+k)^t} = CF_0 + \frac{CF_1}{1+k} + \frac{CF_2}{(1+k)^2} + \cdots + \frac{CF_n}{(1+k)^n} \ . \tag{1}$$

It expresses the relation between economic net present value $ENPV$ of the project (dependent variable) and the social discount rate k of a public project (independent variable). The investment curve will be analysed with a view of solving real roots.

The degree (n) of the $ENPV$ investment curve (1) – it represents the period of economic lifespan of a public project (in the model, n will equal 6 years).

The constant term (CF_0) – it represents capital expenditures of the project in the model. It will acquire non-positive values, which corresponds to a negative cash-flow reflecting project investment costs or as the case may be various (i.e. up to 100%) amount of subsidy from the donor.

Coefficients (CF_1 to CF_n) – in the model represent cash-flows caused by the investment and they can generally acquire both positive and negative values, while $CF_1, \ldots, CF_{n-1}$ represent the operational phase of the project and CF_n the liquidation phase of the project.

Years of project's lifetime (t) – the model is built on the initial prerequisite of the possibility to separate the investment phase (year 0), the operational phase (years 1 to $n-1$) and the liquidation phase (year n) of the project.

Economic internal rate of return of the project (ERR) – the root of the investment curve (1) of the project.

Stream of CFs – the sequence of cash-flows generated by the project in the form $CF = [CF_1, CF_2, CF_3, CF_4, CF_5, CF_6]$, in the model in the basic form $CF = [-6, 1, 1, 1, 1, 1, 1]$, which corresponds to realistically predictable flows of proposed public projects.

The function $CF_j = \varphi(ERR)$, $0 \le j \le n$ – the function of roots for the function of two variables of

$ENPV(ERR, CFj) = 0$, i.e.

$$CF_j = \varphi(ERR) = -\sum_{t=0}^{j-1} \frac{CF_t}{(1+ERR)^{t-j}} - \sum_{t=j+1}^{n} \frac{CF_t}{(1+ERR)^{t-j}} \ . \tag{2}$$

It is therefore a curve corresponding to the intersection of the surface $ENPV(k, CF_j)$ and the plane $ENPV = 0$.

2.2 Discussing mathematical functions

To describe the behaviour of a public project under the conditions of subsidy of investment costs from European funds in order to find out its efficiency by the cost-benefit analysis method we will use a rational fractional function. This function corresponds to the definition of the investment curve (1) as the sum of discounted utilities from a public project and roots ERR of this function can be used for assessing a project. When calculating the ERR it is, however, preferable to convert the rational fractional function $ENPV$ (1) by the substitution $x = \frac{1}{1+k}$ to a polynomial function

$$g(x) = CF_0 + CF_1x + \cdots + CF_nx^n \ . \tag{3}$$

By using suitable mathematical software (e.g. Maple or Matlab), we can calculate the roots of this polynomial and by reverse transform $k = \frac{1}{x} - 1$ the real roots back to the k variable. The only problem with the conversion is when $x = 0$, then $k = \infty$.

We do not consider imaginary roots yet because their importance is ambiguous in economic practice. This procedure is preferable because there are more algorithms for seeking the root of a polynomial and they are simpler than algorithms for seeking the root of a rational fractional function (Novotná and Trch [4]). To calculate the roots it is necessary to use numerical methods because, as well known, there are no formulas for analytical calculation of roots of polynomials of a degree higher than $n = 3$.

Descartes rule of signs: By means of this rule, it is possible to determine how many positive roots there are for a polynomial function. Assuming that there is a polynomial function $g(x)$, then the number of positive roots is equal to the number of variations of the sign between individual terms of the polynomial. This method finds even the solutions which are not from the set of real numbers, but from the set of complex numbers. The rule says that the number of positive real roots of a polynomial is equal to the maximum of the number of the sign variations.

Nevertheless, for economic interpretation we are not interested in the roots of the polynomial (3), but the roots of the rational fractional function (1), i.e. the investment curve of the project. Taking into account the substitution Descartes rule can be adjusted for the $ENPV$ rational fractional function: **The number of real roots of the function (1) which are larger than -1 or infinity is equal to the maximum of the number of variations of the sign of the stream of CFs.**

Next we are interested in the course of the investment curve (1). Due to the economic interpretation, we will focus only on the positive half-plane of the roots of the polynomial, i.e. $x \geq 0$, or rather for rational fractional function we will focus on $k > -1$.

An interesting point is the one where the curve (1) intersects the y axis. It can be easily verified that this point is equal to the sum of coefficients, i.e. $CF_0 + CF_1 + \cdots + CF_n$.

The investment curve $ENPV$ (1) has one vertical asymptote, namely in the value $k = -1$. The horizontal asymptote of the function (1) equals to the constant term, i.e. $\lim_{k\to\infty} ENPV = CF_0$.

3 Results and discussion

For the proposed model from the set of variations of public investment projects, we will consider only some selected economically legitimate scenarios in the paper, when a change of originally conventional project to a project of nonconventional nature can occur on the level of economic cash-flows on the basis of social evaluation of utility. This situation may occur in the liquidation phase of a project, when there is a negative cash-flow due to increased compensation of negative externalities upon completion of the project. a similar process of an originally conventionally proposed project change into an unconventional one can, for example, be caused by unexpected investment revenue due to preventing environmental damage in the course of project implementation. In such cases, cash-flows from positive externalities cannot balance such decrease and so negative cash-flow is generated on the ENPV level during the operational phase.

3.1 Modelling the course of the project investment curve

To present a more realistic idea we will display the course of the project investment curve $ENPV$ (1), i.e. of the relation of the utility of the project on the social discount rate. We will use 2D imaging and we will look for a relation of the curve course and the root position to the amount of subsidy from public resources, demonstrated as the change of CF_0 for the following cases:

- Nonconventional project with negative cash-flow in the operational phase $CF = [-, + + - + ++]$
- Nonconventional project with negative cash-flow in the liquidation phase $CF = [-, + + + + +-]$

We issue from the basic conventional model $CF = [-6, 1, 1, 1, 1, 1, 1]$, which was solved in Dvořáková and Jiříček [1]. According to the assumptions we only consider subsidies of 30%, 70% and 100%.

The course of the investment curve of the nonconventional project in the situation of cash-flow change in the operational phase (change CF_3 – Figure 1 on the left) does not differ in the right half of the curve from the course of a conventional project (Figure 1 in Dvořáková and Jiříček [1]). According to the Descartes rule of signs this project has three or one real root greater than -1.

The course of the $ENPV$ function is quite different in case of negative cash-flow in the liquidation project phase (Figure 1 on the right), when the curves "flip" at the vertical asymptote in -1 to $-\infty$. According to the Descartes rule of signs this project has two or no real roots greater than -1.

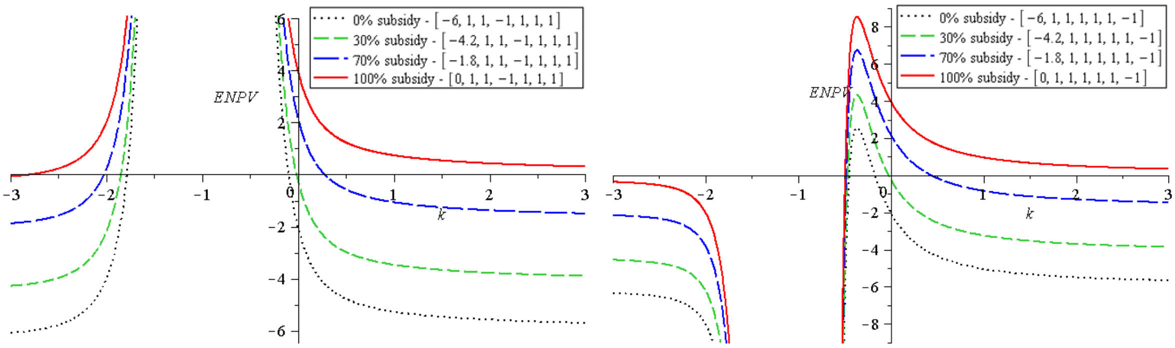


Figure 1 Courses of $ENPV$ investment curves by amount of subsidy of nonconventional project with negative CF

3.2 Multi-parametrical scenario analysis

To analyse the roots of $ENPV$, i.e. ERR, of a public project, a 2D model is not suitable and we are not able to distinguish the influence of individual variations of the CF stream and its parameters on the position and number of roots. For that reason we will analyse the ERR roots in 3D, where we will monitor the influence of multiple project parameters on the position and number of ERR for two above mentioned nonconventional projects. According to assumptions, we only consider the subsidy amount of 30%, 70% and 100% in the 3D model. Thus we are looking for the number and position of the roots of the $ENPV(k, CF_j)$ function for individual levels of subsidy in reaction to a change of a critical parameter of individual scenarios, i.e. to:

- a change of cash-flow in a selected year of the operational phase of the project, e.g. due to further investment costs to prevent environmental damage (ΔCF_j), in our case $j = 3$,
- a change of cash-flow in the liquidation phase of the project (ΔCF_6), e.g. due to compensation of negative externalities.

For a detailed analysis of a scenario for a classical conventional project, see Dvořáková and Jiříček [1].

The rationally fractional function $ENPV$ (1) is in 3D projection displayed in the real space from $k = -1$. From the derived display of the root function in 2D projection the position and number of roots of the $ENPV$ function in the real plane can be determined depending on the CF_j parameter (see (2)). For using the ERR method for public project evaluation, the situation when there is only one real root greater than 0 is essential.

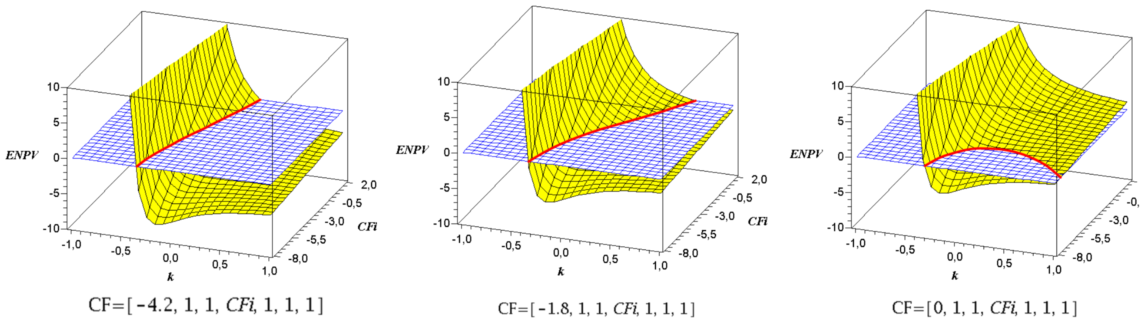


Figure 2 3D model of $ENPV$ function for nonconventional project with critical operational phase for various subsidy amounts of 30%, 70% and 100%

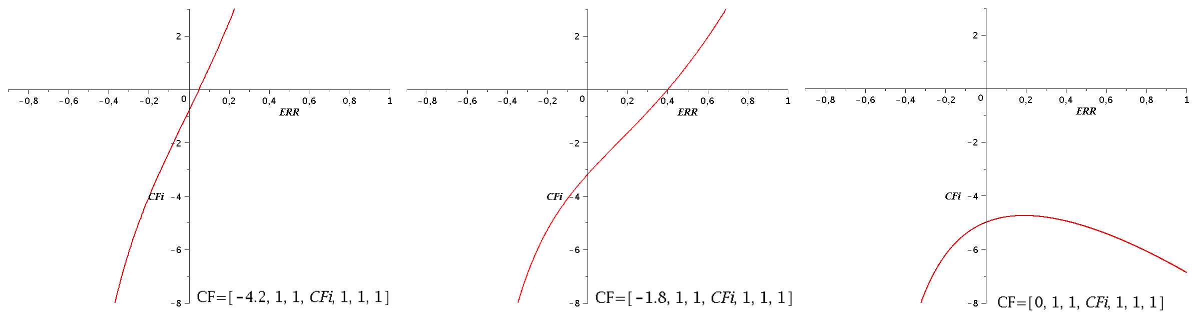


Figure 3 Root function $CF_3 = \varphi(ERR)$ for nonconventional project with critical operational phase for various subsidy amounts of 30%, 70% and 100%

In Figure 3 we can see the proof of the Descartes rule as in the CF stream for conventional model with a CF change in the operational part there are either three changes (for $CF_3 < 0$) or there is one change of the sign (for $CF_3 > 0$). The $ENPV$ function thus has either three or one real root $ERR > -1$ for a particular value of the CF_3 parameter. From Figure 2 and Figure 3 it follows that there is always only one root and we can define the area of the CF_3 parameter for the use of the ERR method, when there is only one real positive ERR root. For instance, for the 70% subsidy amount it is the interval of $CF_3 \in (-3.2, +\infty)$. On the other side, the $ENPV$ function has these two real roots when CF_3 is constant in case that the public project is fully subsidized, i.e. $CF_0 = 0$.

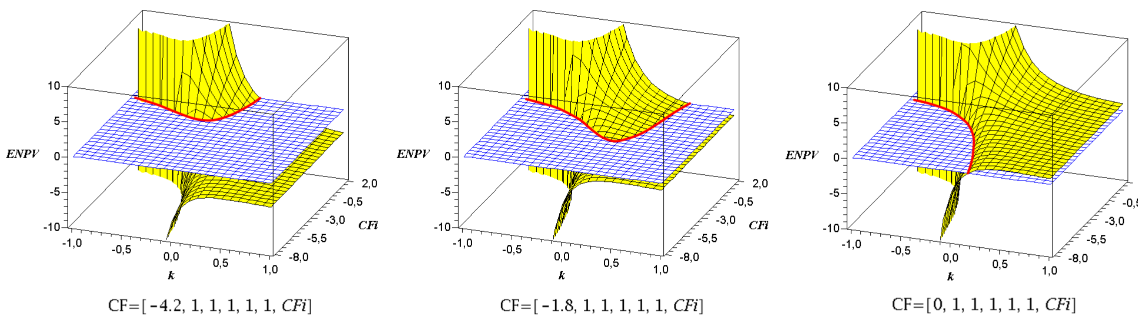


Figure 4 3D model of $ENPV$ function for nonconventional project with critical liquidation phase for various subsidy amounts of 30%, 70% and 100%

From the root function graphs in Figure 5 it is evident that for nonconventional project with negative CF in the liquidation phase the rationally fractional function $ENPV$ has two real roots $ERR > -1$ for a constant value of the $CF_6 < 0$ parameter, which corresponds to the Descartes rule (see above). There is an exception of projects with 100% subsidy (i.e. $CF_0 = 0$), where, as shown in Figure 5, for $CF_6 < 0$ the $ENPV$ function always has only one real root $ERR > -1$. From the perspective of using the ERR method, the $ENPV$ function has in this case one positive root for $CF_6 < 0$. An anomalous situation can be seen in case of 70% subsidy, where two positive real roots can be found for a certain limited interval (CF_6 around -3.5), we cannot therefore decide about the use of the ERR method.

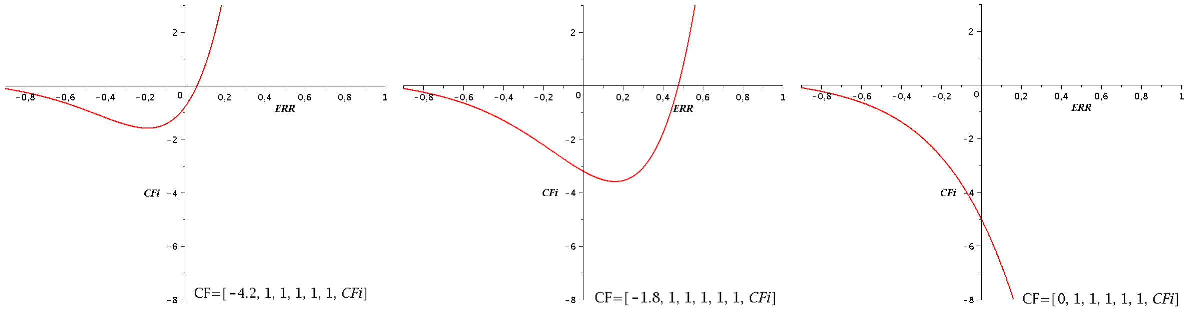


Figure 5 Root function $CF_6 = \varphi(ERR)$ for nonconventional project with critical liquidation phase for various subsidy amounts of 30%, 70% and 100%

4 Conclusion

The paper deals with nonconventional public projects during the rise of negative cash-flows due to crisis situations in operational and liquidation phases of a public project. A multi-parametric analysis of a project is performed based on the $ENPV$ function, modelling a public project defined by a stream of cash-flows in the fundamental version of $CF = [-6, 1, 1, 1, 1, 1, 1]$ in 3D projection, where the influence of a change of project parameters on the position and curving of the $ENPV$ plane can be seen. The function of the roots of $ENPV$ then in the derived 2D projection provides information on the position, number and value of ERR in individual model scenarios. This value can be taken directly from the graph of the root function $CF_j = \varphi(ERR)$, $0 \leq j \leq n$ dependent on the change of individual model parameters (i.e. cash-flows of a real project). From the results of the analysis, it is possible to verify that there is an area while changing the parameters (CF_j) which allows us to apply the ERR method even on nonconventional projects, i.e. the $ENPV$ function gives exactly one real root > 0. As the results of the analysis of functions of the $ENPV$ roots in the model show, the occurrence of multiple positive real roots would probably be caused by much bigger instability of projects, which is a significant conclusion for practical proposing and evaluating of public projects. The conditions of such instabilities occurring can be determined by further analysis on the grounds of the proposed model generally by means of the function $CF_j = \varphi(ERR)$. The model based on applying rationally fractional function to describe a public project thus allows us, by means of 3D projection and consequent analysis of the $CF_j = \varphi(ERR)$ function, to perform an analysis of different variants of public projects not described in the paper. These are for instance nonconventional projects with multiple changes of cash-flow polarity which are characterized by more than two real roots.

References

[1] Dvořáková, S., Jiříček, P.: *Modelling Financial Flows of Development Projects Subsidized from European Funds*. In: 31st International Conference on Mathematical Methods in Economics. Jihlava: VPJ. 2013. p. 119–128.

[2] Florio, M., Vignetti, S.: *Cost-benefit analysis of infrastructure projects in an enlarged European Union: Returns and Incentives*. Economic change and restructuring 38, 2006, 179–210.

[3] Florio, M., Vignetti, S.: *The use of ex-post Cost-Benefit Analysis to assess the long-term effects of Major Infrastructure Projects*. 2013, Available at http://www.csilmilano.com/docs/WP2013_02.pdf

[4] Novotná, J., Trch, M.: *Algebra a teoretická aritmetika. Polynomická algebra*. Univerzita Karlova, Praha 2000

[5] Magni, CA.: *Average Internal Rate of Return and investment decisions: a new perspective*. The Engineering Economist 55(2), 2011, 150–180

[6] Osborne, M.: *A resolution to the NPV-IRR debate?* The Quarterly Review of Economics and Finance 50(2), 2010, 234–239.

[7] Perman, R., Ma, Y., McGilvray, J.: *Natural Resource and Environmental Economics*. Longman Publishers, London, 1996.

On the Extremal Dependence between Stock and Commodity Markets

Krzysztof Echaust[1]

Summary: Many stock market investors do not pay enough attention to the current price of the various commodities such as oil, gold, silver, gas and copper, for example. However, these current prices can have a tremendous impact on the value of the main stock market indices. This paper uses copula approach to investigate the dependence between commodities and stock markets. More specifically, we focus on the tail dependence between both markets. We used two-dimensional Gumbel copula and censored log-likelihood optimization procedure with Generalized Pareto Distribution as a marginal distribution to obtain upper and lower tail dependence coefficients. Our results show that tail dependence between commodity and stock markets exists and safe-haven role of gold is evidenced. We find the asymmetric tail dependence with the larger upper than lower tail dependence. Empirical findings in this paper have potentially important implications for investors and other financial market participants.

Key words: extremal dependence, extreme-value copula, stock indices, commodity.

JEL Classification: C13, F30, G15
AMS Classification: 82C31

Introduction

Many stock market investors do not pay enough attention to the current price of the various commodities such as oil, gold and copper, for example. However, these current prices can have a tremendous impact on the value of the main stock market indices. In this paper we investigate the extremal dependence between commodities and stock markets. We use five main commodities like gold, silver, cooper, crude and gas which have the greatest impact on financial markets. Hence, it is important to understand the linkage between commodities and stock markets especially in situations when extreme risk appear. Extreme dependence, analyzed here, is a key factor to properly asses the extreme risk the investors bear. There is an extensive literature on commodity-stock market relationship, but the extreme dependence structure is not properly investigated [10]. We use a Generalized Pareto Distribution to model marginal distribution and Gumbel copula and tail dependence coefficients to handle the extreme dependence. The main objective of this paper is to study the extreme dependence between commodity prices and the world leading stock markets.

The remaining structure of the paper is as follows. In Section 1 the basics of copula and tail dependence are introduced. In Section 2 a class of copulas called extreme-value is described. Section 3 discusses Neyman-Pearson test for extremal dependence. Section 4 concerns the method of estimation of parametric copula. Section 5 illustrates the empirical results, and Section 6 presents the conclusions.

1 Two-dimensional copulas and tail dependence

Copula functions represent a methodology which has recently become the most significant new tool to handle in a flexible way the comovement between markets, risk factors and other relevant variables considered in finance [2]. The copula *C*, in the probability theory and statistics, is a multivariate probability distribution, which is used to describe the dependence between random variables apart from their marginal distribution. The extensive considerations about copula can be found in e.g. [2, 16]. In our work only its two-dimensional case called 2-copulas is considered. 2-copula joins marginal distributions of random variables X_1 i X_2 to form bivariate distribution function as follows:

$$F(x_1, x_2) = C\big(F_1(x_2), F_2(x_2)\big), \tag{1}$$

where:

F_j (j = 1, 2) – marginal distribution functions of X_1 i X_2,
F – two-dimensional distribution function of X_1 i X_2.

[1] Poznań University of Economics and Business, Department of Operations Research, al. Niepodległości 10, Poznań , k.echaust@ue.poznan.pl.

The Sklar's Theorem guarantees the existence and uniqueness of copula for continuous random variables. The copula is a flexible tool allowing to handle the large market movements occurring together. The concept of tail dependence coefficients (TDC), easy to obtain from copula, is the basic measure of dependence between extremal values of random variables. Loosely speaking, tail dependence describes the limiting proportion that one margin exceeds a certain threshold given that the other margin has already exceeded that threshold. Because the extreme returns can appear both in the left and in the right tail, there are two types of TDC – lower and upper ones. Formally, lower and upper tail dependence coefficients are defined as [4]:

$$\lambda_L = \lim_{v\to 0} \mathbb{P}(X_1 \le F_1^{-1}(v) | X_2 \le F_2^{-1}(v)) = \lim_{v\to 0} \frac{C(v,v)}{v}, \tag{2}$$

$$\lambda_U = \lim_{v\to 1} \mathbb{P}(X_1 > F_1^{-1}(v) | X_2 > F_2^{-1}(v)) = \lim_{v\to 1} \frac{1+2v+C(v,v)}{1-v}. \tag{3}$$

If TDC is equal to zero, $\lambda_L = 0$ or $\lambda_U = 0$, the variables X_1, X_2 are said to be asymptotically independent in lower or upper tail. In the case of $\lambda_L \in (0,1]$ or $\lambda_U \in (0,1]$ there exists the extremal dependence, the strongest the closer to one the value of TDC is.

2 Extreme-value copulas

There are many parametric types of copulas in statistical considerations. The most popular in finance applications are Archimedean, elliptical and extreme-value copulas. In this paper we focus on Gumbel copula which belongs to both Archimedean and extreme-value copula families. In this section we present only the two-dimensional case of extreme-value copula.

Let $X_i = (X_{i1}, X_{i2}), i = 1, \dots, n$, be a sample of independent and identically distributed (iid) random vectors with common distribution function F, margins F_1, F_2, and copula C_F. For convenience we assume that F is continuous. Consider the vector of componentwise maxima:

$$M_n = \max\{M_{n,1}, M_{n,2}\}, \text{ where } M_{n,j} = \max\{X_{1j}, \dots, X_{nj}\}, j = 1, 2. \tag{4}$$

Since the joint and marginal distribution functions of M_n are given by F_n and $F_1^n, \dots, F_d^n$ respectively, it follows that the copula, C_n, of M_n is given by

$$C_n(u_1, u_2) = C_F\left(u_1^{1/n}, u_2^{1/n}\right)^n. \tag{5}$$

Any extreme-value copula can be represented in terms of Pickands dependence function [18]. For a bivariate copula C, this representation takes the form:

$$C(u_1, u_2) = \exp\left(\ln(u_1 u_2) A\left(\frac{\ln u_2}{\ln(u_1 u_2)}\right)\right), u_1, u_2 \in (0,1), \tag{6}$$

where: $A: [0,1] \to \left[\frac{1}{2}, 1\right]$ the Pickands dependence function is convex and satisfies $A(0) = A(1) = 1$ and $\max(t, 1-t) \le A(t) \le 1$ for all $t \in [0,1]$. The upper bound $A(t) = 1$ corresponds to independence, whereas the lower bound corresponds to perfect dependence (comonotonicity). The tail dependence coefficient can have the following representation:

$$\lambda_U = 2\left(1 - A\left(\frac{1}{2}\right)\right). \tag{7}$$

The most popular copula among the extreme-value copulas are Gumbel, Galambos, Hüsler-Reiss and Student's t copulas described in i.a. [8, 9]. In this paper only the Gumbel copula is used in calculations. It is of the form:

$$C_\theta^{Gu}(u_1, u_2) = \exp\left(-\left((-\ln u_1)^\theta + (-\ln v_2)^\theta\right)^{1/\theta}\right), \theta \in [1, \infty). \tag{8}$$

For $\theta = 1$ the random variables are independent, whereas for $\theta \to \infty$ they are comonotonic. The Gumbel copula does not have the lower tail dependence and the upper tail dependence is equal to $\lambda_U = 2 - 2^{1/\theta}$.

3 Testing of asymptotic independence

Modeling the extreme dependence between two variables should be preceded first by testing whether any structure of extremal dependence exists. The popular procedure used in such situations follows from Falk and Michel [6]. Let (X_1, X_2) be a random vector which follows in its upper tail a bivariate extreme-value distribution with reverse exponential margins i.e.

$$H(x_1, x_2) = \exp\left((x_1 + x_2) A\left(\frac{x_1}{x_1 + x_2}\right)\right), \; x_1, x_2 \le 0, \tag{9}$$

where A is a Pickand's function,

$$H(x,0) = H(0,x) = \exp(x), x \le 0. \tag{10}$$

Conditional distribution function (df) of $X_1 + X_2$ given that $X_1 + X_2 > c$, converges to the df $F(t) = t^2$ for $t \in [0,1]$, as $c \uparrow 0$ if and only if X_1 and X_2 are tail independent. Otherwise, the limit is $F(t) = t$. Asymptotic independence was tested by the authors using four different statistics i.e. Neyman-Pearson, Fisher, Kolmogorov-Smirnov and Chi-square. Numerous simulations indicated that the Neyman–Pearson test has the smallest type II error rate.

Neyman-Pearson test. Suppose we have n independent copies of $(X_{1,1}, X_{2,1}), \dots, (X_{1,n}, X_{2,n})$ of random vector (X_1, X_2). Fix $c < 0$ and consider only those observations $X_{1,i}, X_{2,i}$ among the sample that satisfy $X_{1,i} + X_{2,i} > c$. Denote these by $C_1, \dots, C_m$ in the order of their outcome. One has to decide whether df of $W_i = C_i/c$ is equal to either the null hypothesis $F_{(0)}(t) = t^2$ or the alternative $F_{(1)}(t) = t$, $t \in [0,1]$. The test statistics is as follows:

$$\mathrm{NP} = -\sum_{i=1}^{m} \log(W_i) - m\log(2) \tag{11}$$

The p value of the test is, therefore:

$$p_{\mathrm{NP}} = \Phi\left(\frac{2\sum_{i=1}^{m}\log(W_i) + m}{\sqrt{m}}\right), \tag{12}$$

where $\Phi(\cdot)$ denotes the df of the standard normal distribution.

4 Estimation of extreme-value copula parameters

In order to estimate the parameters of parametric copula $\{C_{\boldsymbol{\theta}}: \boldsymbol{\theta} \in \Theta\}$, where Θ is a set of parameters, two main approaches can be distinguished: the one-step and the two-step maximum likelihood estimation. In our calculations we used the second one. Let us consider the copula model of the form

$$F(x_1, x_2;\, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}) = C(F_1(x_1; \boldsymbol{\theta}_1), F_2(x_2; \boldsymbol{\theta}_2); \boldsymbol{\theta}), \tag{13}$$

where F_1 and F_2 are marginal distribution functions with the vector of parameters $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ respectively, and C is the copula with parameters $\boldsymbol{\theta}$. Suppose there are not any restriction between marginal distribution parameters and copula parameters. In the first step the marginal distribution parameters $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$ are estimated using maximum likelihood method and in the second step the vector $\boldsymbol{\theta}$ is estimated by maximizing the likelihood which has the form:

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln c\left(F_1(x_{1i}; \widehat{\boldsymbol{\theta}}_1), F_2(x_{2i}; \widehat{\boldsymbol{\theta}}_2); \boldsymbol{\theta}\right), \tag{14}$$

where c is the copula density $c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}$. Hereby we obtain the consistent and asymptotically normal estimation of $\boldsymbol{\theta}$. The two-step estimator is asymptotically less efficient than one-step estimator, but this approach has the obvious advantage of reducing the dimensionality of the problem, which is particularly useful when one has to resort to numerical maximization [16].

In the empirical study we focus only on the extreme returns represented by exceedances of high threshold q, therefore the Generalized Pareto Distribution (GPD) is used to model a marginal distribution. The explanation of that choice implies from Balkema-deHann Theorem [1]. The unconditional marginal distribution for returns is the following:

$$F(x) = G_{\xi,\beta}(y)\left(1 - \frac{n - N_u}{n}\right) + \frac{n - N_u}{n}. \tag{15}$$

where:

$$G_{\xi,\beta}(y) = 1 - \left(1 + \xi\frac{y}{\beta}\right)^{-1/\xi}, \tag{16}$$

is the GPD with a two parameters $\xi \in \mathbb{R}$ and $\beta > 0$, satisfying $1 + \xi\frac{y}{\beta} > 0$, and $\beta > 0$, $y \ge 0$ for $\xi \ge 0$ and $0 \le y \le -\beta/\xi$ for $\xi < 0$; $y = x - q$ is the exceedance of a high threshold q; N_q is a number of exceedances and n is a sample size.

After the marginal distribution parameters are estimated the following transformation of the variables X_1 and X_2 is necessary [3]:

$$U_1 = -\log\left(1 - \frac{N_{q_1}}{n}\left(1 + \hat{\xi}_1\frac{X_1 - q_1}{\hat{\beta}_1}\right)^{-\frac{1}{\hat{\xi}_1}}\right)^{-1},\ X_1 > q_1, \tag{17}$$

$$U_2 = -\log\left(1 - \frac{N_{q_2}}{n}\left(1 + \hat{\xi}_2 \frac{X_2 - q_2}{\hat{\beta}_2}\right)^{-\frac{1}{\hat{\xi}_2}}\right)^{-1}, \ X_2 > q_2. \tag{18}$$

The vector (U_1, U_2) has the distribution function $\tilde{F}$ and margins that are approximately standard Fréchet distribution and satisfies:

$$\tilde{F}(u_1, u_2) = F(x_1, x_2) = C(u_1, u_2), \tag{19}$$

where C is extreme-value copula.

The inference for this model is complicated because a bivariate pair may exceed a specified threshold in just one of its components. There appears the problem of applicability of $\tilde{F}$ in the particular regions:
$R_{0,0} = (-\infty, q_1) \times (-\infty, q_2)$, $R_{1,0} = (q_1, \infty) \times (-\infty, q_2)$, $R_{0,1} = (-\infty, q_1) \times (q_2, \infty)$,
$R_{1,1} = (q_1, \infty) \times (q_2, \infty)$. Hence the censored likelihood of the form:

$$\ell(\theta) = \prod_{i=1}^{n} \psi(u_1, u_2; \boldsymbol{\theta}), \tag{20}$$

where:

$$\psi(u_1, u_2; \boldsymbol{\theta}) = \begin{cases} c(u_1, u_2), & \text{dla } (u_1, u_2) \in R_{1,1}, \\ \frac{\partial C}{\partial u_1}(u_1, q_2), & \text{dla } (u_1, u_2) \in R_{1,0}, \\ \frac{\partial C}{\partial u_2}(q_1, u_2), & \text{dla } (u_1, u_2) \in R_{0,1}, \\ C(q_1, q_2), & \text{dla } (u_1, u_2) \in R_{0,0}, \end{cases} \tag{21}$$

has to be used.

5 Empirical results

The goal of the work is to measure the extremal dependence between commodity and stock markets. We took into account following commodity: gold, silver, copper, crude and gas and a stock indices: DAX, CAC40, FTSE100, SP500, NIKKEI and B-shares. The sample period of January 04, 2011 to February 26, 2016 is used and calculation on the daily log-returns are conducted. Of course only those data which are common for each asset are taken into account. Therefore the total number of observations is 1121 for the whole sample.

In the first step of analysis the univariate analysis is conducted. The parameters of the marginal distribution are estimated. As a high threshold q the 95th percentile is taken for upper tail. In the case of lower tail the estimation runs exactly in the same way after taking into account minus returns. Table 1 presents the parameter estimates for GPD distribution. The key parameter of the GPD is a shape parameter ξ. When $\xi > 0$, the underlying distribution belongs to fat tail domain of attraction (Fréchet family), when $\xi = 0$ then the distribution belongs to the thin tail domain of attraction (Gumbel family) but when $\xi < 0$ the upper tail has finite right endpoint (Weibull family). The estimates contained in the Table 1 indicate that distributions of the commodities have fat left tails and Weibull or Gumbel type upper tails. Only for gas series we received a Fréchet type distribution. It proves that the extreme events represented by large profits and large losses differ significantly. It seems to be clear from intuitive point of view the left tail must be heavier than the right one. However there are studies indicating the presence of heavier left tails of distributions, e.g. [5, 13], but on the other hand, there are many papers which have failed to demonstrate such asymmetry between the thickness of the left and right tails, e.g. [7, 12, 14, 15]. Also the ambiguous conclusions imply from the stock markets analysis. The thickness of the tail depends on the index is considered and it is not possible to notice any universal property.

The main results of the paper concern an extremal dependence. The outcomes of the calculations are contained in the Table 2. In the two cases it was not possible to calculate the test statistics and p-value. The Neyman-Pearson test indicates the lack of the extremal dependence between gold and stock indices. In this case there is not any dependence and such statement is additionally supported by the zero correlation coefficient. The extremal independence is evidenced either in the lower tail of the copper and stock indices or in the gas and stock indices as well apart from B-shares index. This conclusion is a bit surprising taking into account the connection of cooper and Chinese economy. China is responsible for nearly half of global metals demand and its slowing economy is at the heart of the commodities sell-off. It turns out that only positive high returns of copper and indices are strongly extreme-dependent. Crude oil seems to be the most extreme-dependent commodity on stock markets but only when the high positive returns are taken into account. Positive shocks of oil prices can make investors perceive them as a symptom to increase on financial markets. Such perception is strengthened by permanent falling of oil prices since 2011 and stagnation of world economy. Silver in turn, exhibits a significant but not strong extremal dependence with stock markets, especially for positive movements.

	Left tail		Right tail	
	$\hat{\xi}$	$\hat{\beta}$	$\hat{\xi}$	$\hat{\beta}$
Gold	0,286 (0,169)	0,687 (0,146)	-0,285 (0,152)	0,962 (0,191)
Silver	0,245 (0,220)	1,171 (0,299)	-0,088 (0,151)	1,906 (0,383)
Copper	0,125 (0,153)	0,873 (0,177)	0,078 (0,168)	0,846 (0,181)
Crude	0,367 (0,176)	1,565 (0,338)	-0,205 (0,107)	1,419 (0,240)
Gas	0,196 (0,156)	1,050 (0,214)	0,158 (0,163)	1,099 (0,231)
DAX	-0,127 (0,167)	1,106 (0,236)	-0,020 (0,148)	0,892 (0,178)
CAC40	0,088 (0,197)	0,784 (0,187)	0,034 (0,181)	0,887 (0,200)
FTSE100	-0,069 (0,134)	0,802 (0,151)	-0,027 (0,134)	0,782 (0,148)
SP500	0,067 (0,132)	0,801 (0,150)	0,323 (0,208)	0,569 (0,139)
NIKKEI	0,227 (0,156)	0,839 (0,170)	0,157 (0,155)	0,763 (0,155)
B-shares	-0,169 (0,154)	2,853 (0,578)	0,217 (0,243)	1,244 (0,344)

Table 1 GPD estimates and its standard errors in parenthesis

	Left tail	Right tail	Correlation	λ_L	λ_U
Gold-DAX	-0,711 (0,376)	-1,352 (0,499)	-0,026	-	-
Gold-CAC40	0,142 (0,158)	-1,292 (0,533)	-0,024	-	-
Gold-FTSE100	0,220 (0,145)	-1,426 (0,624)	0,003	-	-
Gold-SP500	-0,539 (0,322)	-2,563 (0,797)	0,029	-	-
Gold-NIKKEI	-0,265 (0,115)	-0,587 (0,364)	-0,014	-	-
Gold-Bshares	-1,234 (0,604)	-	0,037	-	-
Silver-DAX	0,556 (0,097)	1,857 (0,020)	0,128	-	0,118
Silver-CAC40	1,841 (0,018)	2,669 (0,007)	0,134	0,113	0,115
Silver-FTSE100	2,426 (0,008)	4,044 (0,001)	0,175	0,118	0,106
Silver-SP500	1,459 (0,032)	4,293 (0,000)	0,183	0,145	0,098
Silver-NIKKEI	0,146 (0,159)	1,856 (0,014)	0,061	-	0,047
Silver-B-shares	-0,699 (0,392)	-1,031 (0,666)	0,090	-	-
Copper-DAX	-2,388 (0,436)	2,670 (0,008)	0,382	-	0,223
Copper-CAC40	-1,346 (0,292)	0,739 (0,056)	0,392	-	-
Copper-FTSE100	-0,150 (0,144)	1,971 (0,017)	0,416	-	0,243
Copper-SP500	0,330 (0,075)	2,691 (0,008)	0,470	-	0,274
Copper-NIKKEI	-	-0,471 (0,312)	0,093	-	-
Copper-Bshares	-1,034 (0,400)	-0,358 (0,251)	0,176	-	-
Crude-DAX	1,909 (0,019)	-0,606 (0,173)	0,282	0,172	-
Crude-CAC40	1,409 (0,034)	1,482 (0,029)	0,314	0,174	0,206
Crude-FTSE100	0,623 (0,071)	2,082 (0,015)	0,379	-	0,242
Crude-SP500	1,824 (0,021)	2,871 (0,006)	0,442	0,216	0,223
Crude-NIKKEI	-0,734 (0,352)	1,349 (0,035)	0,063	-	0,062
Crude-Bshares	1,056 (0,053)	0,091 (0,164)	0,108	-	-
Gas-DAX	-1,192 (0,401)	0,597 (0,088)	0,256	-	-
Gas-CAC40	-1,988 (0,522)	1,337 (0,038)	0,286	-	0,133
Gas-FTSE100	0,926 (0,056)	1,552 (0,029)	0,346	-	0,148
Gas-SP500	-0,549 (0,221)	1,017 (0,056)	0,382	-	-
Gas-NIKKEI	0,471 (0,108)	1,501 (0,028)	0,086	-	0,059
Gas-Bshares	1,710 (0,021)	-0,539 (0,322)	0,120	0,096	-

Table 2 Tail-dependence test, correlation coefficient and tail dependence coefficients

A surprising results concern the dependence between commodities and B-shares. Apart from negative returns on gas in each considered pair the Neyman-Pearson test did not reject the null hypothesis of extremal independence. However, such results must be treated with caution because of different parts of the world and their respective trading time zones.

Conclusions

Although the linkages between extreme returns of commodity markets and stock markets seem to be obvious, such argument cannot be acceptable from statistical point of view. Our findings show that the tail dependence between commodity and stock markets can be evidenced but it depends on the type of commodity and stock market are considered. Gold is entirely extreme independent on stock markets which emphasizes its safe-haven role. Surprisingly, we did not detect left tail dependence between either cooper and stock indices or natural gas and stock indices. But when the cooper prices are expected to remain high, investors may expect high increase of the stock prices. The strongest extremal dependence is evidenced between positive returns of crude and indices excluding DAX and B-shares. Positive and negative large changes of silver price occur together with stock market movements in most considered cases but the strength of the dependence is not high.

References

[1] Balkema, A.A. and de Haan, L.: Residual Life Time at Great Age. *Annals of Probability*, 2(5), 1974.

[2] Cherubini, U., Luciano, E. and Vecchiato W.: *Copula methods in finance*, Wiley, New York, 2006.

[3] Coles, S.G.: *An Introduction to Statistical Modeling of Extreme Values*, Springer, London, 2001.

[4] Coles, S.G., Heffernan, J. and Tawn, J.A.: Dependence Measures for Extreme Value Analyses. *Extremes*, 3, 1999, 5-38.

[5] Cotter, J.: Margin exceedances for European stock index futures using extreme value theory, *Journal of Banking & Finance*, 25(8), 2001,1475-1502.

[6] Falk, M. and Michel, M.,R.: Testing for tail independence in extreme value models. *Annals of the Institute of Statistical Mathematics*, 58, 2006, 261–290.

[7] Francq, C. and Zakoïan, J.M.: Estimating the Marginal Law of a Time Series with Applications to Heavy Tailed Distributions. *Journal of Business & Economic Statistics*, 31, 2013, 412-425.

[8] Gudendorf, G. and Segers, J.: Extreme-Value Copulas. In: Jaworski, P., Durante,F., Härdle, W.K., Rychlik, W. (ed.), *Copula Theory and Its Applications. Lecture Notes in Statistics*, 198, Springer, Berlin 2010.

[9] Haug, S., Klüppelberg, C. and Peng, L.: Statistical models and methods for dependence in insurance data. *Journal of the Korean Statistical Society*, 40(2), 2011, 125-139.

[10] He Y. and Zhao J.: *Extreme dependence between crude oil and the stock markets in China: A sector Investigation*, Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, Working Papers no. 361005, 2013.

[11] Jansen, D. and de Vries, C.: On the frequency of large stock market returns: putting booms and busts into perspective. *Review of Economics and Statistics*, 23, 1991, 18-24.

[12] Jondeau, E. & Rockinger, M., Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance*, 10, 2003, 559- 581.

[13] LeBaron, B. and Ritirupa, S.: *Extreme Value Theory and Fat Tails in Equity Markets*. International Business School, Brandeis University, 2005.

[14] Longin, F.M.: The asymptotic distribution of extreme stock market returns. *Journal of Business*, 69, 1996, 383-408.

[15] Longin, F.: Stock market crashes: Some quantitative results based on extreme value theory. *Derivatives Use, Trading and Regulation*, 7, 2001, 197-205.

[16] Malevergne, Y. and Sornette, D.: *Extreme Financial Risk,* Springer-Verlag, Berlin, 2006.

[17] Nelesen, B.R.: *An Introduction to Copulas*, Springer, New York, 2006.

[18] Pickands, J.: *Multivariate extreme value distributions*. With a discussion in Proceedings of the 43rd Session of the International Statistical Institute, Bulletin of the International Statistical Institute, 49, 1981, 859-878, 894-902.

Cointegration of Interdependencies Among Capital Markets of Chosen Visegrad Countries and Germany

Marcin Fałdziński[1], Adam P. Balcerzak[2], Tomáš Meluzín[3], Michał Bernard Pietrzak[4], Marek Zinecker[5]

Abstract. Identification of linkages among capital markets is crucial for forming policies that take into account risk associated with international financial markets interdependencies. Thus, the aim of the article is to analyse interdependencies among capital markets of Germany, Poland, Czech Republic and Hungary. The research hypothesis was given as follows: There is a similar course and changes in the interdependencies among capital markets of Germany and the markets of the mentioned countries of the Visegrad Group. In the research a DCC-GARCH model was applied. The model allowed to estimate conditional correlations that indicate strength of the interrelationship among the markets. Then, the cointegration analysis of the conditional correlations was conducted. The proposed econometric procedure allowed to verify the research hypothesis. It confirmed that the capital markets of Germany, Poland, Czech Republic and Hungary are characterised with similar long-term path. Additionally, the research showed that changes in the direction and strength of the interrelationships among the studied markets are determined by the German capital market in the long-term, which is a leader in the region.

Keywords: cointegration of interdependencies among capital markets, conditional correlation, DCC-GARCH model, conditional variance

JEL Classification: G15, C58
AMS Classification: 90C15

1 Introduction

Identification of linkages among capital markets and evaluation of their variability over time are crucial for forming guidelines, which can help to "manage" globalized economy [18, 14, 2, 19, 28]. The research in that field is important from the perspective of forming rules and regulations of financial markets that can support their positive influence on the development processes and socio-economic sustainability [3, 5, 31, 7, 8]. It is necessary to propose policies that take into account the risk associated with international financial markets interdependencies. This risk is especially important during the time of global uncertainty, where the instability of one foreign capital market can lead to recession in real sphere in different economies [12, 29, 13, 16, 1, 4] and influence negatively the crucial macro and micro economic spheres such as countries fiscal stability [24, 10,11], situation on labour markets [26, 35, 6, 9, 32, 33] or international competitiveness of national industries [34]. As a result, the purpose of the article is to describe interdependencies among capital markets of Germany, Poland, Czech Republic and Hungary. The analysis was conducted for the years 1997-2015. The research hypothesis was given as follows: There is a similar course and changes in the interdependencies among capital markets of Germany and the markets of the mentioned countries of the Visegrad Group.

In the first step of the research DCC-GARCH model was used in order to estimate conditional correlations that indicate strength of the interrelationship among the analysed markets. In the second stage, the cointegration analysis of the conditional correlations was conducted. The research is continuation of previous studies of the authors in the field [17, 36].

[1] Nicolaus Copernicus University/ Department of Econometrics and Statistics, ul. Gagarina 13a, Toruń, Poland. marf@umk.pl.
[2] Nicolaus Copernicus University, Department of Economics, ul. Gagarina 13a, Toruń, Poland, adam.balcerzak@umk.pl.
[3] Brno University of Technology, Kolejní 2906/4, 612 00 Brno, Czech Republic, meluzint@vutbr.cz.
[4] Nicolaus Copernicus University/Department of Econometrics and Statistics, ul. Gagarina 13a, Toruń, Poland, michal.pietrzak@umk.pl.
[5] Brno University of Technology, Kolejní 2906/4, 612 00 Brno, Czech Republic, zinecker@fbm.vutbr.cz.

2 Method of the Research

In the research parameters of DCC=GARCH model were estimated [15], which enabled to assess conditional variances and conditional correlations for the pairs of indices for the capital markets of Germany, Poland, Czech Republic and Hungary.

The next step was the cointegration analysis and proposition of VECM model (Vector Error Correction Model), which is a VAR model for cointegrated processes. Let's assume *p* rank VAR model with deterministic component d_i .

$$y_t = d_t + A_1 y_{t-1} + ... + A_p y_{t-p} + \varepsilon_t \tag{1}$$

where: y_t is *N*-dimensional stochastic process, A_i for $i = 1...p$ are matrixes of parameters $N \times N$ and ε_t is N-dimensional white noise.

VAR model in the form of differences of the process y_t can be given with equation 2.

$$\Delta y_t = \mu_t + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{2}$$

where: $\Pi = \sum_{i=1}^{p} A_i - I$ and $\Gamma_i = -\sum_{j=t+1}^{p} A_j$.

The rank "*r*" of a matrix Π determines the interpretation of the model given with equation 2. When the rank *r* equals *N*, then process y_t is not integrated [$I(0)$]. When rank r equals zero, then process y_t is integrated of order one [$I(1)$], but it is not cointegrated. There is a cointegration when rank r is between $0 < r < N$ and matrix Π can be given as $\alpha\beta'$. Matrix α is a matrix of adjustments matrix and β is a cointegrating vector.

In order to test cointegration the existence of Π matrix is determined and then the rank *r* of matrix Π is calculated. For this purpose Johansen [21, 22, 23] procedure is applied, where a maximum eigenvalue test is used. The procedure is an iterative process and it is applied until the rejection of the null hypothesis. In the first step of the procedure the null hypothesis that there is no cointegrating vector is adopted (lack of cointegration, H_0: r = 0), against the alternative that there is at least one such vector H_1: $(r > 1)$. In the case of the absence of evidence that enable to reject the null hypothesis, the following assumptions are made successively: $(H_0 : r \le 1, H_1 : r = 1), (H_0 : r \le 2, H_1 : r = 2)$, which in next steps results in is successive increasing of *r*. The max-eigenvalue test statistics is characterized with nonstandard distribution, where the critical values can be found in the work of Osterwald-Lenum [27].

3 Results of the research

In the research time series for four stock market indexes (BUX, PX 50, WIG and DAX) were used. For each of the indexes a daily rate of return were set for the period: 1 July 1997 to 30 September 2015, which gave a total number of 4592 observations[6]. In the first step of the research parameters of DCC-GARCH model were estimated with application of maximum likelihood method with the t-student conditional distribution. The results are presented in table 1. Additionally for all the indices a constant in the equation conditional mean were estimated, where all the parameters corresponding to the constant were statistically significant at the 5% significance level. For each of the indices the parameters corresponding to the conditional variances and conditional correlations in DCC-GARCH model were also statistically significant.

Parameter (stock index)	Estimate	p-value	Parameter (stock index)	Estimate	p-value
const (PX 50)	0.0493	0.0124	const(WIG)	0.0526	0.0013
ω_1 (PX 50)	0.0221	0.0053	ω_3 (WIG)	0.0209	0.0055

[6] The time series for the research were retrieved from: http://www.finance.yahoo.com (30.10.2015).

α_1 (PX 50)	0.0694	0.0000	α_3 (WIG)	0.0771	0.0000
β_1 (PX 50)	0.9250	0.0000	β_3 (WIG)	0.9145	0.0000
const (BUX)	0.0580	0.0027	const (DAX)	0.0799	0.0000
ω_2 (BUX)	0.0641	0.0003	ω_4 (DAX)	0.0259	0.0001
α_2 (BUX)	0.1058	0.0000	α_4 (DAX)	0.0947	0.0000
β_2 (BUX)	0.8732	0.0000	β_4 (DAX)	0.8960	0.0000
a	0.0084	0.0000	b	0.9891	0.0000
v	8.4787	0.0000	-	-	-

Table 1 The results of the estimation of the DCC-GARCH model

The estimation of parameters of DCC-GARCH model allowed to determine the values of conditional correlations for the next pairs of indices. The correlation values for a given pair of indices indicate the strength of the relationship between the two capital markets. Additionally, it gives information on changes of upward or downward trends of these interrelationships over time. In the next stage, based on the aim of the current paper the values of conditional correlations were used to analyze the cointegration.

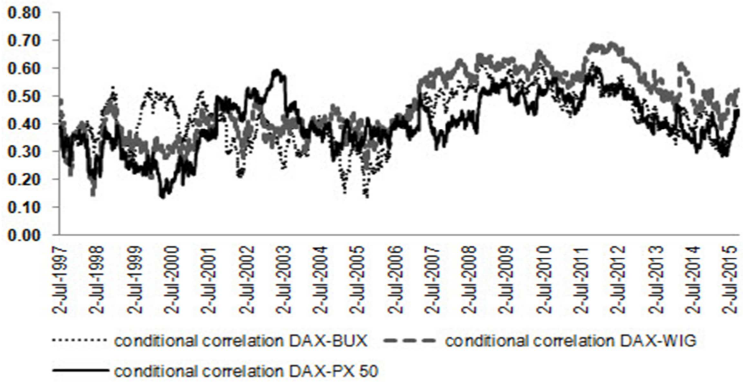


Figure 1 The conditional correlation between pairs of chosen markets

Figure 1 shows the conditional correlations between the DAX index and PX 50, WIG, BUX indices. The analysis of correlations shows that the relationships between the capital markets of Germany and the markets of the Czech Republic, Poland and Hungary are quite similar. It can be said that since 2004 the shocks on German capital market have been transferred in a similar way to Czech, Polish and Hungarian markets. This means that the stock valuation on every analyzed capital market is largely dependent on the situation on the other markets.

In the next step an analysis of cointegration of conditional correlations obtained after application of DCC_GARCH model was conducted. For this purpose a stationarity of time series of conditional correlations was tested with application of Phillips-Perron test [30]. The test results showed that all the time series of conditional correlations are integrated in the order one $I(1)$. In the next step a Johansen procedure was carried out, which allowed to test the number of cointegrating relations. Table 2 presents the test results for the number of cointegrating vectors [25]. The results of the max-eigenvalue test indicate the presence of one cointegrating vector.

Hypothesized number of cointegrating vectors	Eigenvalue	Statistic	p-value
None	0.005	26.519	0.007
At most 1	0.002	13.147	0.074
At most 2	0.001	4.880	0.027

Table 2 Johansen test results

After determination of the number of cointegrating vectors a long-term equation for the model VECM(2) was assessed. The empirical model given with equation 3 was obtained.

$$\rho_{DAX-BUX} = -0.25 + \underset{[4.606]}{1.18}\,\rho_{DAX-PX50} - \underset{[-6.561]}{1.38}\,\rho_{DAX-WIG} + \varepsilon_t \tag{3}$$

where t-statistics are given in square brackets.

The parameters of the long-term equation given with equation 3 are statistically significant. The equation indicates that the conditional correlations between indices DAX and BUX combine a positive relationship with respect to conditional correlation between DAX and PX50, and negative with respect to conditional correlation between DAX and WIG. In the long term, analyzed capital markets form one cointegrating relation with the dominant role of German capital market.

Next the parameters of short-term equations were estimated. Table 3 presents the estimates of parameters and t-statics in brackets for three equations.

Variables	Equations		
	$\Delta\rho_{DAX-BUX}$	$\Delta\rho_{DAX-PX50}$	$\Delta\rho_{DAX-WIG}$
Const	0.0000076 [0.06301]	0.0000105 [0.08551]	0.0000102 [0.08561]
ECM_{t-1}	-0.002538 [-2.00532]	-0.001877 [-1.46076]	0.004031 [3.21920]
$\Delta\rho_{DAX-BUX,t-1}$	0.004411 [0.27181]	0.027074 [1.64340]	0.020635 [1.28504]
$\Delta\rho_{DAX-BUX,t-2}$	-0.007255 [-0.44694]	-0.019378 [-1.17615]	0.010238 [0.63750]
$\Delta\rho_{DAX-PX50,t-1}$	0.002909 [0.19442]	-0.013303 [-0.87597]	-0.006694 [-0.45224]
$\Delta\rho_{DAX-PX50,t-2}$	-0.012671 [-0.84720]	0.003501 [0.23059]	0.003495 [0.23621]
$\Delta\rho_{DAX-WIG,t-1}$	0.028176 [1.70311]	-0.018993 [-1.13100]	0.020098 [1.22783]
$\Delta\rho_{DAX-WIG,t-2}$	0.016676 [1.00766]	0.013382 [0.79662]	-0.018205 [-1.11183]

Table 3 Vector Error Correction Estimates

Negative estimates of parameters for variable ECM_{t-1} in equations for $\Delta\rho_{DAX-PX50}$ and $\Delta\rho_{DAX-BUX}$ were obtained. The obtained sign of estimates means that the pairs of markets are characterized with similar long-term path. On the other hand, in the case of variable ECM_{t-1} in the equation for $\Delta\rho_{DAX-WIG}$, a positive estimate was obtained. It means that the conditional correlation for the pair of indices DAX-WIG is influenced by additional determinants than the one included in the given model VECM. It can indicate that the German capital market is not the only one that has significant influence Polish capital market in the context of short term deviations.

4 Conclusions

The article concentrates on the problem of interrelations among capital markets of chosen Visegrad countries and Germany. The identification of the international relationships among markets should be treated as an important scientific problem. Its adoption is crucial for determination of strategies that can be useful in counteraction of consequences of potential crises. The profound research in that field is a condition for proposing some tools that can be useful in risk management both at micro and macroeconomic level.

The conducted analysis enabled to verify the research hypothesis. It confirmed that the capital markets of Germany, Poland, Czech Republic and Hungary are characterised with similar long-term path. The research showed that changes in the direction and strength of the interrelationships among the capital markets are determined by the German market in the long-term, which can be considered as a leader in the region.

References

[1] Balcerzak, A. P.: Monetary Policy Under Conditions of NAIRU "Flattening". *Olsztyn Economic Journal* **4(1)** (2009), 95-105.

[2] Balcerzak, A. P.: Limitations of the National Economic Policy in the Face of Globalization Process. *Olsztyn Economic Journal* **4(2)** (2009), 279-290.

[3] Balcerzak, A. P.: Structure of Financial Systems and Development of Innovative Enterprises with High Grow Potential. *Research Papers of Wrocław University of Economics - Global Challenges and Policies of the European Union-Consequences for the " New Member States"* **59** (2009), 30-39.

[4] Balcerzak, A. P.: Przegląd i wstępna ocena teoretycznych stanowisk dotyczących źródeł globalnego kryzysu gospodarczego. In: *Gospodarka w warunkach kryzysu* (Antkiewicz, S. and Pronobis, M. (eds.). Wydawnictwo Naukowe CeDeWu.pl, Warszawa, 2009, 257-274.

[5] Balcerzak, A. P. (2009). *Effectiveness of the Institutional System Related to the Potential of the Knowledge Based Economy*. *Ekonomista* **6** (2009), 711-739.

[6] Balcerzak, A. P.: Multiple-criteria Evaluation of Quality of Human Capital in the European Union Countries. *Economics & Sociology* **9(2)** (2016), 11-26.

[7] Balcerzak, A. P., and Pietrzak, M. B.: Human Development and Quality of Institutions in Highly Developed Countries. In: *Financial Environment and Business Development. Proceedings of the 16th Eurasia Business and Economics Society* (Bilgin, M. H., Danis, H., Demir, E. and Can, U. eds.). Springer International Publishing, 2016.

[8] Balcerzak, A. P., and Pietrzak, M. P.: Application of TOPSIS Method for Analysis of Sustainable Development in European Union Countries. In: *The 10th International Days of Statistics and Economics. Conference Proceedings* (Loster, T. and Pavelka, T. (eds.). September 8-10, Prague, 2016.

[9] Balcerzak, A. P., and Pietrzak, M. B.: Quality of Human Capital in European Union in the Years 2004-2013. Application of Structural Equation Modeling. In: *Proceedings of the International Scientific Conference Quantitative Methods in Economics Multiple Criteria Decision Making XVIII*. Letra Interactive. Vratna, 7-12.

[10] Balcerzak, A. P., Pietrzak, M. B., and Rogalska, E.: Fiscal Contractions in Eurozone in the years 1995-2012: Can non-Keynesian effects be helpful in future deleverage process? In: *Business Challenges in the Changing Economic Landscape - Vol. 1. Proceedings of the 14th Eurasia Business and Economics Society* (Bilgin, M. H., Danis, H., Demir, E. and Can, U., eds.). Springer International Publishing, 2016, 483-496.

[11] Balcerzak, A. P., and Rogalska, E.: Non-Keynesian Effects of Fiscal Consolidations in Central Europe in the Years 2000-2013. In: *Entrepreneurship, Business and Economics - Vol. 2. Proceedings of the 15th Eurasia Business and Economics Society* (Bilgin, M. H., and Danis, H., eds.). Springer International Publishing, 2016, 271-282.

[12] Baur, D.: Testing for Contagion – Mean and Volatility Contagion. *Journal of Multinational Financial Management* **13** (2003), 405-422.

[13] Billio, M., and Caporin, M.: Market Linkages, Variance Spillover and Correlation Stability: Empirical Evidences of Financial Contagion. *Computational Statistics & Data Analysis* **54(11)** 2010, 2443–2458.

[14] Corsetti, G., Pericoli, M., and Sbracia, M.: *Some Contagion, Some Interdependence: More Pitfalls in Testing for Contagion. Journal of International Money and Finance* **24** (2005), 1177-1199.

[15] Engle, R. F.: Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics* **20** (2002), 339-350.

[16] Engle, R. F.: *Anticipating Correlations A New Paradigm for Risk Management*. Princeton University Press, Princeton and Oxford, 2009.

[17] Fałdziński, M., and Pietrzak, M. B.: The Multivariate DCC-GARCH model with interdependence among markets in conditional variances' equations. *Przeglad Statystyczny* **62(1)** (2015), 397-413.

[18] Forbes, K., and Rigobon, R.: No Contagion, Only Interdependence: Measuring Stock Market Comovements. *The Journal of Finance* **57(5)** (2002), 2223-2261.

[19] Heryán, T. and Ziegelbauer, J.: Volatility of Yields of Government Bonds Among GIIPS Countries During the Sovereign Debt Crisis in the Euro Area. *Equilibrium. Quarterly Journal of Economics and Economic Policy* **11(1)** (2016), 61-75, DOI: http://dx.doi.org/10.12775/ EQUIL.2016.003.

[20] http://www.finance.yahoo.com (30.10.2015).

[21] Johansen, S.: Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* **12(2–3)**, (1988), 231-254.

[22] Johansen, S.: Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* **59** (1991), 1551–1580.

[23] Johansen, S.: *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford, 1995.

[24] Mackiewicz-Łyziak, J.: Fiscal Sustainability in CEE Countries – the Case of the Czech Republic, Hungary and Poland. *Equilibrium. Quarterly Journal of Economics and Economic Policy* **10(2)** (2015), 53-7.

[25] MacKinnon, J. G., Haug, A. A., and Michelis, L.: Numerical Distribution Functions of Likelihood Ratio Tests for Cointegration. *Journal of Applied Econometrics* **14** (1999), 563-577.

[26] Müller-Frączek, I., and Pietrzak, M. B.: Przestrzenna analiza stopy bezrobocia w Polsce w latach 2004-2008. In: *Hradecké ekonomické dny 2011. Ekonomický rozvoj a management regionu* (Jedlicka, P. ed.) Gaudeamus, Hradec Králové, 2011, 205-209.

[27] Osterwald-Lenum, M.: A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics. *Oxford Bulletin of Economics and Statistics* **54** (1992), 461-472.

[28] Osińska, M., Fałdziński M., and Zdanowicz T.: Econometric Analysis of the Risk Transfer in Capital Markets. The Case of China. *Argumenta Oeconomica* **2(29)** (2012), 139-164.

[29] Pericoli, M., and Sbracia, M.: A Primer on Financial Contagion. *Journal of Economic Surveys* **17(4)** (2003), 571-608.

[30] Phillips, P. and Perron, P.: Testing for a Unit Root in Time Series Regression. *Biometrica* **75(2)** (1988), 335-346.

[31] Pietrzak, M. B., and Balcerzak, A. P.: Assessment of Socio-Economic Sustainability in New European Union Members States in the years 2004-2012. In: *The 10th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings* (Papież, M., and Śmiech, S. (eds.). Foundation of the Cracow University of Economics, Cracow, 2016, 120-129.

[32] Pietrzak, M. B., and Balcerzak, A. P.: A Spatial SAR Model in Evaluating Influence of Entrepreneurship and Investments on Unemployment in Poland. In: *Proceedings of the International Scientific Conference Quantitative Methods in Economics Multiple Criteria Decision Making XVIII*. Letra Interactive, Vratna, 303-308..

[33] Pietrzak, M. B., and Balcerzak, A. P.: Quality of Human Capital and Total Factor Productivity in New EU Member States. In: *The 10th International Days of Statistics and Economics. Conference Proceedings* (Loster, T., and Pavelka, T., eds.). Prague, 2016.

[34] Pietrzak, M. B., and Łapińska, J.: Determinants European Union's trade – evidence from a panel estimation of the gravity model. *E & M Ekonomie a Management* **18(1)** (2015), 18-27.

[35] Pietrzak, M. B., Wilk, J., Kossowski, T. and Bivand, R.: The identification of spatial dependence in the analysis of regional economic development - join-count test application. In: *Proceedings of the 8th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena* (Papież, M., and Śmiech, S., eds.). Foundation of the Cracow University of Economics, Cracow, 2014, 135-144.

[36] Zinecker, M., Balcerzak, A. P., Fałdziński, M., Meluzín, T., and Pietrzak, M. B.: Application of DCC-GARCH Model for Analysis of Interrelations Among Capital Markets of Poland, Czech Republic and Germany. In: *Proceedings of the International Scientific Conference Quantitative Methods in Economics Multiple Criteria Decision Making XVIII*. Letra Interactive, Vratna, 416-421.

Microeconomic Optimization Equilibrium Models of Demand and Supply for Network Industries Markets

Michal Fendek[1], Eleonora Fendeková[2]

Abstract. For the network industries market equilibrium models a certain segregation of the market is characteristic, resulting in the network industries production usually not being substitutable. Therefore a satisfaction gained from their consumption can be uniquely quantified. In principle, it is such a representation of utility function where a network industry product is considered to be a good with its own and exactly formulated utility function and the other goods are considered as a consumption of one calculated aggregated good with standardized unit price.

In this paper we point out a specific interpretation of *consumer surplus* in case of network industries production demand analysis. For the complex characteristics of the network industries market we now discuss a formalized analysis of the producer's activities on this market. Note that also a cost function of the supplier or the producer in this model of the network industry producer is of somewhat atypical pragmatic interpretation.

For the optimization problems we will formulate the Kuhn-Tucker optimality conditions and we will study their interpretation options.

Keywords: Utility function, consumer preferences, consumers' and producers' surplus, social welfare, network industries, Kuhn – Tucker optimality conditions.

JEL Classification: D43, L11, L22
AMS Classification: 93C30

1 Introduction

Currently a significant attention in scholarly discussions on various levels is being paid to the subject of network industries. It is understandable as network industries in fact ensure the production and distribution of energy sources which play a key role in an effective operation of the developed economies. Blum, Müller and Weiske [2] and Pepall, Richards and Norman [7] show, that the discussions are usually focused on the question of a reasonable profit of the network industries companies and on the other hand on the question of prices which are determined by the reasonable and generally acceptable costs of their production.

Naturally, equilibrium on the network industries market, as well as on any market, is being created based on the level of demand and supply on said market. In this paper we will discuss the analysis of microeconomic optimization models of consumers and producers behavior on the network industries market, i.e. the analysis of demand and supply phenomena on this specific market, as well as the general questions of network industries sector effectivity.

2 Optimization Model of the Consumers for the Network Industries Market

For the network industries market equilibrium models a certain segregation of the market is characteristic, resulting in the network industries production usually not being substitutable. Fendek and Fendeková [4] show, that therefore a satisfaction gained from their consumption can be uniquely quantified. In principle, it is such a representation of utility function where a network industry product is considered to be a good with its own and exactly

[1] University of Economics Bratislava, Department of Operations Research and Econometrics, Dolnozemská cesta 1/b, 852 35 Bratislava, Slovakia, e-mail: fendek@dec.euba.sk.
[2] University of Economics Bratislava, Department of Business Economics, Dolnozemská cesta 1/b, 852 35 Bratislava, Slovakia, e-mail: nfendek@dec.euba.sk.
.

formulated utility function and the other goods are considered as a consumption of one calculated aggregated good with standardized unit price.

Suppose there is m consumers S_i for $i = 1, 2... m$ on the relevant network industry market. The network industry good or service of a homogenous character, let it be electricity distribution, is provided by n subjects, suppliers D_j for $j = 1, 2\dots n$. We examine market of the homogenous good where a consumption of the good in the volume of x_i is a consumption of the homogenous good of the i-th consumer S_i and a consumption of all the other goods in a market basket of this consumer is represented by calculated aggregate variable x_{0i}. If a utility of consuming the network industry product for a consumer S_i is given by $u_i(x_i)$, which represents utility in monetary units, and the price of the calculated good is standardized to "1" then total utility of consumer S_i in monetary units is represented by the function $v_i(x_i, x_{0i})$ as follows

$$v_i(x_i, x_{0i}) = u_i(x_i) + x_{0i}$$

where

$$v_i(x_i, x_{0i}): R^2 \to R$$
$$u_i(x_i): R \to R$$

This perception of the utility function allows us to interpret it as a total utility in "monetary units" of a consumer buying x_i units of the network industry product and concurrently buying x_{0i} units of aggregate other goods of consumer basket which are priced by standardized price of one monetary unit.

Behavior of the i-th consumer S_i for every i=1, 2, ..., m will be examined through optimization problem of total utility function maximization of the i-th consumer S_i with limited consumption expenditures with the limit w_i and a price of the network industry product p. This problem is for strictly positive variables x_i a x_{0i} represented:

$$v_i(x_i, x_{0i}) = u_i(x_i) + x_{0i} \to \max$$

subject to

$$px_i + x_{0i} = w_i$$
$$x_i, x_{0i} \geq 0$$

In Jarre and Stoer [5], we can see, that the above stated mathematical programming optimization problem is a maximization problem of bounded extrema. Let us modify this problem to standardized form i.e. minimization problem:

$$-v_i(x_i, x_{0i}) = -u_i(x_i) - x_{0i} \quad \to \min \tag{1}$$

subject to

$$px_i + x_{0i} = w_i \tag{2}$$
$$x_i, x_{0i} \geq 0 \tag{3}$$

For this problem we will formulate generalized Lagrange function. In Bazaraa and Shetty [1], we can see, that generalized Lagrange function does not explicitly include conditions of the variables being strictly positive, but they are included implicitly in Kuhn-Tucker optimality conditions The generalized Lagrange function of this mathematical programming problem is:

$$\begin{gathered} L_i(x_i, x_{0i}, \lambda_i) = -v_i(x_i, x_{0i}) + \lambda_i(px_i + x_{0i} - w_i) = \\ -u_i(x_i) - x_{0i} + \lambda_i(px_i + x_{0i} - w_i) \end{gathered} \tag{4}$$

We formulate the Kuhn-Tucker optimality conditions for the Lagrange function (4) for the i-th consumer S_i as follows

$$\begin{array}{lll} \dfrac{\partial L_i(x_i, x_{0i}, \lambda_i)}{\partial x_i} \geq 0 & \dfrac{\partial L_i(x_i, x_{0i}, \lambda_i)}{\partial x_{0i}} \geq 0 & \dfrac{\partial L_i(x_i, x_{0i}, \lambda_i)}{\partial \lambda_i} = 0 \\ x_i \dfrac{\partial L_i(x_i, x_{0i}, \lambda_i)}{\partial x_i} = 0 & x_{0i} \dfrac{\partial L_i(x_i, x_{0i}, \lambda_i)}{\partial x_{0i}} = 0 & \\ x_i \geq 0 & x_{0i} \geq 0 & \end{array} \tag{5}$$

And after other updates we finally get this form of the Kuhn-Tucker optimality conditions:

$$\begin{array}{llllll} \lambda_i p - mu_i(x_i) \geq 0 & \text{(a)} & \lambda_i - 1 \geq 0 & \text{(d)} & px_i + x_{0i} = w_i & \text{(g)} \\ x_i(\lambda_i p - mu_i(x_i)) = 0 & \text{(b)} & x_{0i}(\lambda_i - 1) = 0 & \text{(e)} & & \\ x_i \geq 0 & \text{(c)} & x_{0i} \geq 0 & \text{(f)} & & \end{array} \tag{6}$$

In other words if a consumer is willing to identify the optimal consumer strategy (x_i*, x_{0i}*), that means that the consumption of x_i* units of the network industry product with a price p and the consumption of x_{0i}* units of remaining aggregated goods with price 1 maximize his total utility v (x_i*, x_{0i}*), multiplier λi* must exist, for which the Kuhn-Tucker optimality conditions (6) hold, thus variables vector (x_i*, x_{0i}* , λ_i*) is a solution to (a), (b), ..., (g).

We now point out certain interesting economically interpretable consequences of the Kuhn-Tucker optimality conditions in the context of consumer behavior analysis on the network industries market:

1. Condition (g) implies that optimum market basket of the *i*-th consumer (x_i^*, x_{0i}^*) at a price *p* of the network industry product and the price of aggregated good being 1can be purchased for customers budget w_i.
2. Condition (e) implies that for positive optimum consumption of the aggregated good x_{0i}^* the optimum value of the Lagrange multiplier is 1, i.e. $\lambda_i^*=1$.
3. Condition (b) then implies that at positive consumption of the aggregated good x_{0i}^* and at positive consumption of the network industry product x_i^* which maximizes utility necessarily holds that in the point of maximum utility of the good its marginal utility equals its price, since Lagrange multiplier equals one $\lambda_i^*=1$ and holds
$$mu_i(x_i^*) = \left[\frac{du(x_i)}{dx_i}\right]_{x_i = x_i^*} = p \tag{7}$$
4. Consequence (3) also confirms another important theoretical postulate, namely that a consumer increases his consumption, of the network industry product in this case, until the marginal utility reaches the level of good's market price.

3 Optimization Model of the Producers for the Network Industries Market

For the complex characteristics of the network industries market we now discuss a formalized analysis of the producer's activities on this market. Note that also a cost function of the supplier or the producer in this model of the network industry producer somewhat atypical pragmatic interpretation. Pindyck and Rubinfeld [6] and Waldman and Jensen [8] show, that the essence of this conception is market perception already explained above within the consumer's behavior analysis, where all other goods and services but the network industry product are expressed as an aggregated good with a price equal to one.

Based on this conception the *j*-th supplier in order to produce y_j units of the network industry product at current technology uses an adequate number of aggregated inputs with a price equal to one. These costs are generally the sum of the fixed and variable costs. The production costs of the supplier D_j related to the production of y_j units are represented by the total costs function $n_j(y_j)$. We use the standard cost functions for the further analysis, namely average cost function

$$np_j(y_j) = \frac{n_j(y_j)}{y_j}, \quad y_j > 0$$

and marginal cost function

$$nm_j(y_j) = \frac{dn_j(y_j)}{dy_j}$$

while supposing these functions are convex, smooth and differentiable.

Examine now the behavior of the *j*-th company in terms of its natural tendency towards profit maximization. The behavior of *j*-th company D_j for every *j*=1, 2, ..., *n* is examined through optimization problem of profit function maximization of the *j*-th company D_j at a condition that a network industry product market price *p*, which is a parameter in this model, at least covers the average production costs $np_j(y_j)$. Fendek and Fendeková [3] show, that this optimization problem is for nonnegative variable y_j formulated as follows:

$$z_j(y_j) = py_j - n_j(y_j) \ \rightarrow \ \max$$

at constraints (8)

$$\frac{n_j(y_j)}{y_j} \leq p$$
$$y_j \geq 0$$

The abovementioned optimization problem is a problem of maximization at a bounded extrema. We modify this problem to a standard form, i.e. minimization problem:

$$-z_j(y_j) = -py_j + n_j(y_j) \ \rightarrow \ \min \tag{9}$$

at constraints

$$\frac{n_j(y_j)}{y_j} - p \leq 0 \tag{10}$$

$$y_j \geq 0 \tag{11}$$

For this mathematical programming problem (9, ... (11) we formulate generalized Lagrange function in the following form:

$$L_j(y_j, \lambda_j) = -py_j + n_j(y_j) + \lambda_j\left(\frac{n_j(y_j)}{y_j} - p\right) =$$
$$= -py_j + n_j(y_j) + \lambda_j\left(n_j(y_j)y_j^{-1} - p\right)$$

We formulate the Kuhn-Tucker optimality conditions for the Lagrange function for the j-th supplier D_i as follows

$$\begin{array}{ll} \dfrac{\partial L_j(y_j,\lambda_j)}{\partial y_j} \geq 0 & \dfrac{\partial L_j(y_j,\lambda_j)}{\partial \lambda_j} \leq 0 \\ y_j \dfrac{\partial L_j(y_j,\lambda_j)}{\partial y_j} = 0 & \lambda_j \dfrac{\partial L_j(y_j,\lambda_j)}{\partial \lambda_j} = 0 \\ y_j \geq 0 & \lambda_j \geq 0 \end{array} \tag{12}$$

We modify the Kuhn-Tucker optimality conditions in the form (12) after introducing the analytical form of the Lagrange function (19) and after other modifications as follows

$$\begin{array}{llll} -p + nm_j(y_j) + \dfrac{\lambda_j}{y_j}\left(nm_j(y_j) - np_j(y_j)\right) \geq 0 & (h) & np_j(y_j) - p \leq 0 & (k) \\ y_j\left(-p + nm_j(y_j) + \dfrac{\lambda_j}{y_j}\left(nm_j(y_j) - np_j(y_j)\right)\right) = 0 & (i) & \lambda_j\left(np_j(y_j) - p\right) = 0 & (l) \\ y_j \geq 0 & (j) & \lambda_j \geq 0 & (m) \end{array} \tag{13}$$

If a supplier offers y_j^* units of the network industry product on the market which at a market price p guarantee maximum profit then a Lagrange multiplier must exist λ_j^*, for which the Kuhn-Tucker optimality conditions (13) hold therefore the vector of variables $(y_j, \lambda_j)^*$ is a solution to the system of equations and inequations (h), (i), ..., (m). Examine now some interesting economically interpretable consequences of the Kuhn-Tucker optimality conditions in the context of supplier's behavior analysis on the network industries market:

1. First of all we reflect the possibilities of the Kuhn-Tucker optimality conditions analysis from the angle of the variable y_j, i.e. the volume of production. The profit of the company is the difference between its returns and costs while the Range of the cost function $n_j(y_j)$ are nonnegative numbers therefore the profit can be positive at any positive market price p only for the positive volume of the output $y_j^* > 0$. The condition (k) at the same time guarantees that with the optimum volume of the output y_j^* correspond the average costs $np_j(y_j^*)$ that do not exceed the market price p.
2. Assume that for the optimum volume of the supply y_j^* following holds $y_j^* > 0$ and $np_j(y_j^*) \leq p$, then for the optimum value of the Lagrange multiplier one of the two situations may occur:
   a. Assume that the optimum Lagrange multiplier is $\lambda_j^* = 0$, then the condition (l) holds and the validity of the condition (i) implies that for the optimum volume of the supply the marginal costs are equal

to the market price of the production $nm_j(y_j^*) = p$, while the condition (h) meets as an equality and the condition (k) may also meet as a strict inequality. We can now see that in case of zero Lagrange multiplier a company can reach positive maximum profit on the level of $py_j^* - n_j(y_j^*)$.

b. In case the Lagrange multiplier is positive $\lambda_j^* > 0$, the validity of the condition (l) implies that for the optimum supply volume the average costs are equal to the market price $np_j(y_j^*) = p$, while the condition (k) meets as an equality and the condition (i) is formulated as follows:

$$y_j\left(-p + nm_j(y_j) + \frac{\lambda_j}{y_j}\left(nm_j(y_j) - np_j(y_j)\right)\right) =$$

$$= y_j\left(-p + nm_j(y_j) + \frac{\lambda_j}{y_j}\left(nm_j(y_j) - p\right)\right) =$$

$$= y_j\left(\left(-p + nm_j(y_j)\right)\left(1 + \frac{\lambda_j}{y_j}\right)\right) = 0$$

Since

$y_j > 0$ and

$1 + \frac{\lambda_j}{y_j} > 0$

then the condition (i) implies that with the optimum output volume y_j^* correspond the marginal costs $nm_j(y_j^*)$ equal to the market price p, i.e. $nm_j(y_j^*) = p$, which together with the validity of the condition *(l)* implies the equality of the marginal and average costs of the company $nm_j(y_j^*) = np_j(y_j^*)$ for optimum supply volume y_j^*. In this case, however, the optimum profit is zero.

4 Conclusion

In this article we examined the optimality conditions for the partial models of the consumers and producers behavior on the network industry market as well as the optimality conditions for the model of the product effective allocation on this market. We showed how the findings resulting from the Kuhn-Tucker optimality conditions analysis formulated for the relevant problems of mathematical programming can be effectively used to interpret the factual relations, principles and strategic decisions in consumers and producers behavior on the network industry market.

Breaking of the equilibrium conditions can of course not be eliminated, the everyday economic life even expects some development and instability on every market and thus also the relative momentariness of the conditions validity, which in fact is not an unsolvable problem. It is however important to identify the situation competently and evaluate the possible reactions to the changed parameters of the system.

If the abovementioned situations wouldn't occur, certain consumers could increase their utility by an exchange. In that case a consumer with a higher marginal utility and thus a higher marginal readiness to pay for the aggregated good could gain a corresponding amount of the network industry product for an adequate volume of the aggregated good from a consumer with a lower marginal utility and therefore both consumers would better their market positions.

Similarly in the situation when the marginal costs of all the producers are not equal, it would be possible to reach the higher total supply on the market at the fixed total sector costs by a simple transfer of the production from a producer with the higher marginal costs to a producer with the lower marginal costs.

Therefore we showed that the use of the model approach and the optimization theory at the network industries market supply and demand equilibrium conditions analysis allows us to effectively study the equality conditions as well as the consequences of the market parameters change which results in reassessment of equilibrium attributes.

Acknowledgements

Supported by the grant No. 1/0697/15 *Optimization models and methods as tools of effective regulation in the modern theory of network industries organization* of the Grant Agency VEGA, Slovakia. (Scientific Grant Agency of the Ministry of Education, science, research and sport of the Slovak Republic and the Slovak Academy of Sciences)

References

[1] Bazaraa, M., and Shetty, C.M.: *Nonlinear Programming: Theory and Algorithms.* Wiley−Interscience, New York, 2006.
[2] Blum, U., Müller, S., and Weiske, A.: *Angewandte Industrieökonomik: Theorien − Modelle − Anwendungen*. Verlag Gabler, Berlin, 2006.
[3] Fendek, M., and Fendeková, E.: Kuhn –Tucker optimality conditions in equlibrium models of network industry markets. *Politická ekonomie: teorie, modelování, aplikace* **60**, 6, (2012), 801−821.
[4] Fendek, M., and Fendeková, E.: Models of regulation of network industries in Slovakia. *International journal of economics and business research* **1**, 4, (2009), 479−495.
[5] Jarre, F., and Stoer, J.: *Optimierung.* Springer Verlag, Berlin, 2012.
[6] Pindyck, R. S., and Rubinfeld, D. L: *Microeconomics.* Prentice Hall, New York, 2004.
[7] Pepall, L., Richards, D., and Norman, D.: *Industrial Organization: Contemporary Theory and Empirical Applications.* Wiley-Blackwell, New York, 2008.
[8] Waldman, D. E, and Jensen, E. J.: *Industrial Organization: Theory and Practice.* Addison Wesley, New York, 2006.

Models of Quantitative Analysis of the Competitive Environment in the Slovak Banking Sector

Eleonora Fendeková[1], Michal Fendek[2]

Abstract. The competitive environment is occurred in each field of life. A dynamism and flux are typical features, which is caused by efforts to achieve maximum competitiveness. The functioning of the market mechanism is subjected to the existence of good market conditions, for which to comply with conditions of a competition is necessary. The aim of this paper is to evaluate the competitive environment and analyze the concentration of the banking sector in Slovakia.

The idea that the market system in banking sector is functioning automatically and particularly effectively is misleading and dangerous. One of the most important faultless attributes, but from the perspective viewpoint of harmonically developing economic system, consists in protection of economic competition principles by the government.

Under the conditions of the economy of Slovakia the guarantee represents the Antimonopoly Office of Slovakia, which systematically takes into account the analysis of the competition state in the banking sector. The purpose of this paper is to present some results of quantitative analysis of the state and development of the banking sector in Slovakia during 2010 – 2014.

Keywords: Competitive environment, competition, concentration indices, banking sector, quantitative analysis.

JEL Classification: L43, L16, G18, G34
AMS Classification: 91B24

1 Introduction

Economically developed countries realize the necessity of reacting to this potential possibility of negative effects of the market environment in terms of protection of the competition and thus systematically create the institutional conditions and legislative guarantees for protecting the economic competition. They do not, however, settle for the mere statement of the nature of the current state of the competition. They exactly evaluate its level and report to the relevant institutions the risks of the negative trends in the development of various industries. The guarantor of the compliance with the competitive conditions in Slovakia is the Antimonopoly Office of the Slovak Republic,

The banking sector of the Slovak republic underwent the significant changes in the last decades. After finishing the process of the restructuration and privatization of the banks with huge asset volumes, the situation of the Slovak banking institutions was subsequently stabilized in 2003 - 2008. The development of the Slovak banking sector was significantly affected by an economic and financial crisis in 2009.

Considering a structure and a current state of the banking sector in Slovakia, we can say that the most frequent events on the market are mergers with the foreign banks which represent horizontal mergers, given that they were mergers of companies – financial institutions operating in one bank sector on transnational level. All the mergers are assessed and approved by the Antimonopoly Office of the Slovak Republic. The Antimonopoly Office is an independent central body of state administration of the Slovak Republic which was established, same as in other developed countries, for the protection and guarantee of competition. The Office intervenes in cases of cartels, abuse of a dominant position, vertical agreements; it controls mergers that meet the notification criteria and assesses actions of state and local administration authorities if they restrict competition.

[1] University of Economics Bratislava, Department of Business Economics, Dolnozemská cesta 1/b, 852 35 Bratislava, Slovakia, e-mail: nfendek@dec.euba.sk.

[2] University of Economics Bratislava, Department of Operations Research and Econometrics, Dolnozemská cesta 1/b, 852 35 Bratislava, Slovakia, e-mail: fendek@dec.euba.sk.

The basic prerequisite for the effective functioning of competition on the market is an existence of a legislative framework which defines the rules of behavior of various subjects in the area of competition practices. This legislative framework constitutes of laws – Act No. 136/2001 on Protection of Competition and on Amendments and Supplements to Act of the Slovak National Council No. 347/1990.

2 Methods of Measuring the Degree of Concentration

In literature we can find many methods which can be more or less successfully applied to evaluate the consequences of concentration in the conditions of imperfect competition. The nature of most methods is to quantify the indicators that somehow describe the position of an individual subject within the industry on the relevant market of a certain commodity respectively characterize a state of competition environment in said industry. Fendek and Fendeková [2], and Kufelova [5] show, that with a certain simplification we can say that all the methodological tools for measuring the concentration de facto quantify the share of a relevant characteristic (i.e. turnover) of a particular subject in a certain hierarchy structure (i.e. an enterprise within an industry) in a total value of these characteristics for all the analyzed subjects.

Depending on whether the indicators quantify the degree of concentration regarding all subjects or just their subset with a certain attribute, we can divide the indicators to two groups: indicators to measure absolute concentration (e.g. concentration ratio, Herfindahl index) and indicators to measure relative concentration (e.g. dispersion index, coefficient of variation).

In Fendek and Fendeková [3], we can see, that in terms of the concept of practical competition policy, the significance of the difference between the indicators of absolute and relative concentration lies in a fact that on the one hand while using the absolute indicators we accept the influence of the number of subjects, on the other hand while using the relative indicators this influence is eliminated and the influence of value dispersion in relation to the level of concentration is taken into account.

2.1 Market Share of the Producer within the Industry

O'Sullivan, Sheffrin and Perez [6] show, that basic concentration measure is a share of monitored indicator of a particular subject on a value of this indicator within an industry or a specifically defined group of subjects. The analytical expression of the indicator is as follows:

$$r_k = \frac{q_k}{Q} \tag{1}$$

where

n - number of subjects,

q_k – volume of the indicator of the k-th subject, $k=1,...,n$,

Q - volume of the indicator for the industry,

$$Q = \sum_{k=1}^{n} q_k$$

r_k – share of the k-th subject within the industry.

2.2 Concentration Ratio

Concentration ratio CR_ψ is a measure of so called *absolute concentration* in the industry and it is calculated for ψ "strongest" subjects in terms of measured indicator share of the homogenous production, in other words it represents the share of the first ψ subjects with the highest indicator on the value of this indicator for the whole industry

$$CR_\psi = \sum_{k=1}^{\psi} r_k \tag{2}$$

2.3 Herfindahl-Hirschman Index

Herfindahl-Hirschman index is probably the most common analytical measure to assess the level of concentration in an industry. It reflects the number of the subjects in the industry as well as their market share. It is based on the hypothesis that the significance of the subject within the industry is a function of the squares of its market share. In

Besanko and Braeutigan [1], we can see, that this philosophy of the Herfindahl index obviously increases the influence of the "strong" subjects and other way round, it eliminates the influence of the "small" subjects. The analytical expression of the Herfindahl index is following:

$$H = h(\mathbf{r}) = \sum_{k=1}^{n} r_k^2 \tag{3}$$

where $h(\mathbf{r})$ is a real function of n variables, h: $R^n \to R$. The usual classification of the concentration according to the Herfindahl index is following: unconcentrated industry – if Herfindahl index $H < 0.1$, moderately concentrated industry – if Herfindahl index ranges $H \in \langle 0.1, 0.18 \rangle$, concentrated industry – if Herfindahl index $H \geq 0.18$.

To avoid the one-sided preference of the strong subjects, which we tend to do when we formulate the conclusions on competition while using *concentration ratio*, it is useful to use the Herfindahl index as well as the methodological apparatus to measure the relative concentration. It should be kept in mind that the two measures of concentration with the same values can be viewed differently if the number of the subjects differs. This information can be somehow objectified by using the *measures for the relative concentration.*

3.4 Index of Dispersion

Index of dispersion represents how the concentration ratio of ψ subjects with the highest value of the measured indicator influences the uneven distribution of the indicator in particular subjects. The analytical form of the index of dispersion is following:

$$DR_\psi = \frac{CR_\psi - \frac{\psi}{n}}{CR_\psi} \tag{4}$$

While the index of dispersion DR_ψ as a measure for relative concentration represents certain equivalent to the measure of absolute concentration CR_ψ, the relative concentration measure equivalent for the Herfindahl index is a coefficient of variation.

3.5 Coefficient of Variation

Coefficient of variation represents a very effective tool to measure concentration while, and this is important, taking into account the specific number of subjects. In Shy [7], we can see, that compared to the Herfindahl index which reflects squares of the individual concentration levels r_k and thereby *does not reflect the influence of less dominant subjects*, the coefficient of variation represents the *rate of distribution of all subjects in terms of the measured indicator*, while the median or arithmetical average of the individual influence of the particular subjects is reciprocal of their number, that is *1/n*. Analytical expression of the coefficient of variation is following:

$$V^2 = n \times \sum_{j=1}^{n} \left(r_j - \frac{1}{n} \right)^2 \tag{5}$$

Abovementioned measures are now used to analyze the state and development of competition in the banking sector in the Slovak republic from the perspective of volumes of deposits, loans, assets and profit of the commercial banks.

3 State of Concentration in the Slovak Banking Sector

The state of concentration in the Slovak banking sector was more complexly assessed by the authors of this article within the project of the Scientific Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic and the Slovak Academy of Sciences "Optimization Models and Methods as Tools of the Effective Regulation in the Modern Theory of the Network Industries Organization", 2015-2017, grant no. 1/0697/15VEGA. In Fendek and Fendeková [4], we can see, that the concentration was assessed based on several relevant indicators such as:

- balance sheet,
- bank and client deposits and loans,
- equity,
- number of employees,
- profit before tax,
- profit after tax,

- number of cash-machines.

The limited scope of this article does not allow us to present the results of the complex analysis of the Slovak banking sector concentration for all the indicators, so following we will present the results of the analysis for one of the crucial indicators which is the amount of assets, respectively the balance sheet of the commercial banks in the Slovak republic in the period of 2010 – 2014. On Fig. 1 we present the market shares of the banks according to the balance sheet in the period of 2010 – 2014.

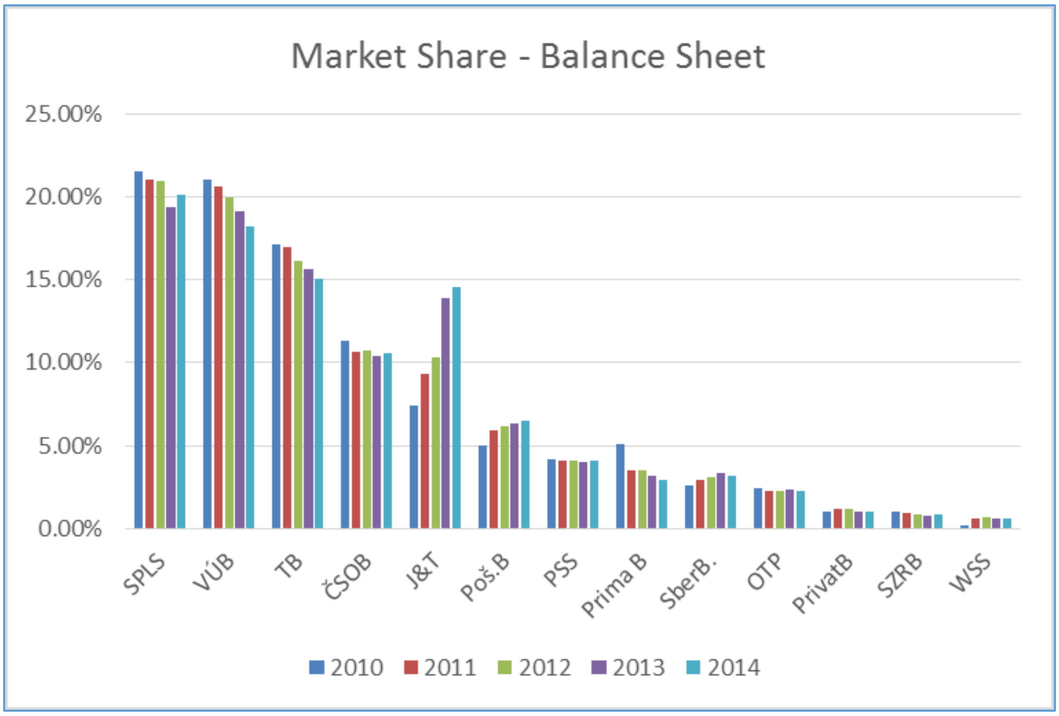


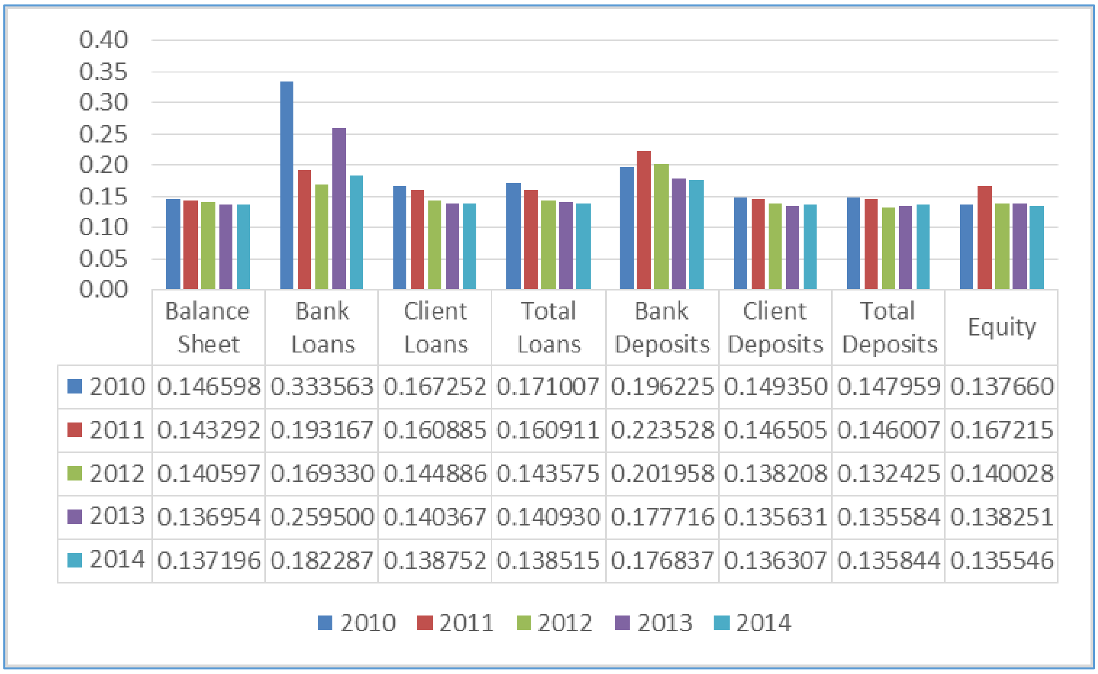
Fig. 1: Market share of the Slovak banks according to the balance sheet (2010 – 2014)

Based on the realized analysis we can state that despite the strong concentration in the banking sector, which was confirmed by several indicators, mostly deposits and loans, profit and number of branches and cash-machines, the clear monopoly position of neither of the subjects within the banking sector in Slovakia is indicated.

Based on the acquired values of the Herfindahl-Hirschman index the competition environment in the banking sector shows reasonable attribute of competition, which creates the satisfying conditions for the healthy economic competition, even despite of occurring slight concentration.

Taking into account the variety of examined indicators we can assess the concentration as low to high. The concentration according to the balance sheet and bank and client deposits and loans was low, ranging in the interval <0.13; 0.15>. However, according to profit and number of cash-machines the Herfindahl-Hirschman index exceeds 0.18 and the concentration is marked as high. In the monitored period, most of the Herfindahl-Hirschman index values show the decreasing tendency, at the utmost only minimal increase, which means there is an improving tendency in the competition within the banking sector.

The concentration of three and five of the most dominant banks also confirmed the high concentration in the Slovak banking sector. Based on the balance sheet, three of the most influential banks are Slovenská sporiteľňa (SLSP), Všeobecná úverová banka (VUB) and Tatra banka (TB). The values of concentration ratio CR_3 clearly show a strong competition as all values exceed 40 %. Among the five dominant banks are Slovenská sporiteľňa, Všeobecná úverová banka, Tatra banka, Československá obchodná banka (ČSOB) and J&T Banka. Concentration ratio CR_5 for all of the indicators was higher than 70 % which clearly indicates high concentration on the Slovak banking market.



	Balance Sheet	Bank Loans	Client Loans	Total Loans	Bank Deposits	Client Deposits	Total Deposits	Equity
2010	0.146598	0.333563	0.167252	0.171007	0.196225	0.149350	0.147959	0.137660
2011	0.143292	0.193167	0.160885	0.160911	0.223528	0.146505	0.146007	0.167215
2012	0.140597	0.169330	0.144886	0.143575	0.201958	0.138208	0.132425	0.140028
2013	0.136954	0.259500	0.140367	0.140930	0.177716	0.135631	0.135584	0.138251
2014	0.137196	0.182287	0.138752	0.138515	0.176837	0.136307	0.135844	0.135546

Fig. 2: Herfindahl – Hirschman index for chosen indicators of the banking sector in 2010 – 2014

. Based on the processed data and calculated concentration ratios was calculated the index of dispersion (DR_3 and DR_5), which describes the concentration from the point of view of relative degree of variability in the concentration. The advantage of the index of dispersion is that the results are not distorted by the external data. The index of dispersion favors the banks with high market share and characterizes the market in terms of even and uneven distribution. The index of dispersion is high enough, exceeds 60 % at almost all indicators, which shows that the distribution of banks on the market is uneven and there is a domination of some of the banks. In the monitored time period, the decreasing or stable tendency of the index of dispersion prevailed over increasing; therefore we can say that the competition on the Slovak banking market is relatively stable.

4 Conclusion

The structure of the subjects operating on the Slovak banking market develops dynamically. Over time and due to globalization the recently state owned banks are becoming subsidiaries of the transnational companies. These changes influence the development of the whole banking sector.

The phenomenon of bank mergers is becoming more and more frequent. In the past couple of years the structure of the market changed every year and several new banks began to operate on the market. A significant change was an establishment of the biggest foreign bank branch J&T (2005) and other foreign banks BKS Bank AG (2011), KDB Bank Europe, new internet banks mBank (2007), AXA (2010), Zuno (2011), merger of some Slovak and foreign banks (ČSOB – Istrobanka, Unicredit – HBC banka), establishment of the new banks and extinction or change in ownership of the new banks (Dexia - Prima banka (2011), Volksbank – Sberbank (2013)), Banco Mais, S.A - Banco Banif Maisa S.A., The Royal Bank of Scotland N.V. - The Royal Bank of Scotland plc., extinction of some foreign banks (Crédit Agricole Corporate and Investment Bank S.A , HSBC Bank plc., UNIBON), which caused the changes in the banking sector. Given that these changes mostly did not impact the most dominant banks within the Slovak banking sector, they had insignificant impact on the total competitiveness on the relevant market.

The main tools of the above presented analyzes of the concentration in the Slovak banking sector were the following measures: concentration ratio, Herfindahl-Hirschman index, index of dispersion and coefficient of variation, while a very important indicator with high informative value was the market share of the individual banks.

The measures were chosen to characterize the key activities of the subjects within the banking sector with regards to the goal of the analysis, which was the exact evaluation of the competition in the Slovak republic. Following indicators were used: balance sheet, bank and client deposits and loans, equity, profit before tax, profit after tax, interest income, net interest income, net fees and provisions income, money and money equivalents, number of branches, number of employees, and number of cash-machines.

We can conclude that the outcomes resulting from this analysis are an addition in the field of competition analysis in the banking sector. The results obtained based on the quantitative analysis of the concentration between 2010 and 2014 undoubtedly give very interesting information about the state and the trends of the competition environment in the Slovak banking sector in the past years.

Acknowledgements

The article was published with the support of the Scientific Grant Agency of the Ministry of Education, science, research and sport of the Slovak Republic and the Slovak Academy of Sciences, grant VEGA, No. 1/0697/15 *Optimization models and methods as tools of effective regulation in the modern theory of network industries organization*.

References

[1] Besanko, D. A., and Braeutigan, R. R.: *Microeconomics. An integrated Approach*. John Willey and Sons, Inc., New York, 2002.
[2] Fendek, M., and Fendeková, E.: Models of regulation of network industries in Slovakia. *International journal of economics and business research* **1**, 4, (2009), 479−495.
[3] Fendek, M., and Fendeková, E.: Modely cenovej regulácie sieťových odvetví. *Ekonomický časopis: časopis pre ekonomickú teóriu a hospodársku politiku, spoločensko-ekonomické prognózovanie* **58**, 10 (2010), 1039–1055.
[4] Fendek, M., and Fendeková, E.: Protection and formation of competitive environment in Antimonopoly Office of the Slovak Republic. *Global business & economics anthology* **2**, 1, (2010), 200−211.
[5] Kufelová, I.: Analýza cien a cenotvorba vo vybraných sieťových odvetviach. *Modely rovnováhy v podmienkach produktovej a cenovej diferenciácie na regulovaných trhoch sieťových odvetví*. Vydavateľstvo EKONÓM, Bratislava, 2015, 154−169.
[6] O'Sullivan, A., Sheffrin, S., and Perez, P.: *Microeconomics: Principles, Applications, and Tools.* Prentice Hall, New York, 2006.
[7] Shy, O.: *The Economics of Network Industries.* Cambridge university press, Cambridge, 2001.

Efficient project portfolio designing

Petr Fiala[1], Renata Majovska[2]

Abstract. In an accelerating economic world, projects become tools for promoting the objectives of the organization. Project portfolio is set of all projects that are implemented in the organization at a time. Projects in the project portfolio are interconnected by priorities, dependences, and organization's resources utilization. We propose a new approach for project portfolio designing based on a systemic combination of Data Envelopment Analysis (DEA) and De Novo optimization approach. Possible projects are characterized by sets of inputs and outputs. The DEA is an appropriate approach to select efficient projects. Inputs are resources for project realization. The organisation has its total resources in limited quantities. It is possible to buy the resources at given prices. A total available budget is a restriction on project portfolio. De Novo optimization is an approach for designing efficient systems by reshaping the feasible set. The proposed combination of DEA a De Novo approaches ensure solving of the efficient project portfolio designing problem. The proposed concept provides designing of optimal project portfolio by given budget. Possible extensions of the problem are formulated and discussed. These extensions include restricted weights, goal restrictions, and time dependent budget.

Keywords: Project portfolio, Data Envelopment Analysis, De Novo optimization.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Project management is the discipline of planning, organizing, securing and managing resources to bring about the successful completion of specific project goals. In an accelerating economic world, projects become tools for promoting the goals of the organization. There is a very extensive literature on the management of individual projects and project portfolios. We start from a publication [6] that describes very clearly project management as a managerial process. Projects are in accelerating world rhythm the right option of solving problems of lot of organizations. Nothing is permanent, everything is temporary, and that makes pressure on companies to finish new products or services faster, cheaper and definitely not to fail. Risk is a very important factor in project management. Most project organizations exist in a multi-project environment. This environment creates the problems of project interdependency and the need to share resources. Projects are the way for implementing the organization's strategy. Strategic alignment of projects is of major importance to effective use of organization resources. Selection criteria need to ensure each project is prioritized and contributes to strategic goals.

Ensuring alignment requires a selection process that is systematic, open, consistent, and balanced. All of the projects selected become part of a project portfolio that balances the total risk for the organization. Management of the project portfolio ensures that only the most valuable projects are approved and managed. Projects are considered as a tool for achieving the strategic goals of organizations. Continuous innovation, renewal and organizational learning are considered vital for survival. Intense global competition is forcing many organizations to look for new methods of management. According [3] the key to success in project portfolio management is to select the right projects at the right time. The project selection process is considered a major component of project portfolio management. This should be accompanied by periodically repeated inspections of project portfolio, which would identify projects that should be terminated. Effective portfolio management helps to achieve outperformance, making strategy real through organizational change. Strategic project portfolio management enables to present a framework for organization to complete significant strategic projects. Portfolio management is a process. This process must improve over time. Building feedback into every stage of the process is critical for the improvement.

To select a portfolio of projects are basically two approaches, one is based on standard methods used in practice, the second approach is based on searching and applying new sophisticated methods based on quantitative

[1] University of Economics, Prague, Department of Econometrics, W. Churchill Sq. 4, Prague 3, Czech Republic, pfiala@vse.cz.
[2] VŠB - Technical University of Ostrava, Department of Mathematical Methods in Economics, Sokolská třída 33, Ostrava, Czech Republic, renata.majovska@vsb.cz.

analysis. The paper focuses on the of project portfolio selection problem solved by applying sophisticated models.

The aim is to develop a general model, which would be completed for the specific needs of problems. This is not about managing individual projects, but their portfolios where relationships exist among projects. This paper aims to verify the ability to model and solve the problem of project portfolio using a combination of the Data Envelopment Analysis (DEA) model and De Novo optimization. Portfolio management is a process. This process must improve over time. Building feedback into every stage of the process is critical for the improvement. The organization must decide under risk whether to assign all available resources to present proposals or to reserve a portion of the funds unused for some time and wait for better alternatives that may occur later. We propose to complete our model by periodically repeated inspections of project portfolio.

The rest of the paper is organized as follows. In Section 2, the project portfolio problem is formulated. A basic Data Envelopment Analysis (DEA) model is summarized in Section 3. A new approach for project portfolio designing based on a systemic combination of DEA and De Novo optimization approach is proposed in Section 4. Section 5 presents possible extensions of the proposed approach. Conclusions are summarized in Section 6.

2 Project portfolio management

The portfolio management domain encompasses project management oversight at the organization level through the project level. Full insight of all components of the organization is crucial for aligning internal business resources with the requirements of the changing environment. Project portfolios are frequently managed by a project office that serves as a bridge between senior management and project managers and project teams. Project opportunities come in time and it is necessary to decide which will be accepted for creating a dynamic portfolio of projects and which will be rejected (Fig. 1).

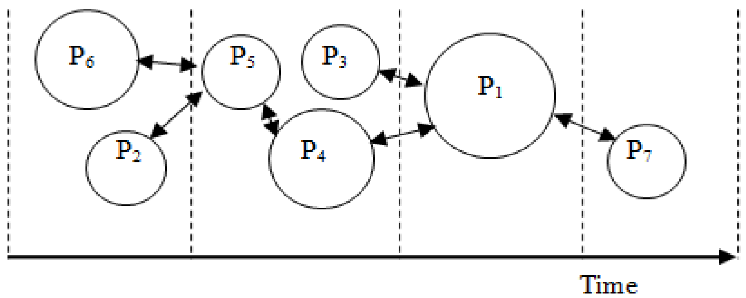


Figure 1 Dynamic flow of projects

Project portfolio is set all projects that are implemented in the organization at that time. The basic objectives of the project portfolio management include:

- optimize the results of the entire project portfolio and not individual projects,
- the selection of projects to start;
- interruption or discontinuation of projects;
- defining priorities for projects;
- coordinate internal and external sources;
- organization learning from each other project.

It is generally expected that the portfolio should be designed in such a way as to maximize the possibility of achieving the strategic goals of the company. This is consistent with the notion that portfolio selection problem is a multi-criteria decision making. The main goal of each project is to increase the value of the organization, so most managers prefer financial criteria for project evaluation. The most commonly used indicators include net present value, internal rate of return, payback period, rate of return.

In addition to these financial indicators, however, in selecting a portfolio of projects should be taken into account other characteristics, which include for example:
The probability of completing the project on time, within budget and within the proposed quality;

- Consistency between strategic and tactical plans;
- The balance between investment projects and maintenance projects;
- Efficient use of resources;
- Relations between projects;

- The scope of each project;
- Time-dependent consumption of resources on projects;
- Allocation of expenditure and resources for research and development;
- Allocation of marketing spending and resources.

Lot of professionals tried to find sophisticated way to improve techniques for project management in different ways. We propose a new approach for project portfolio designing based on a systemic combination of Data Envelopment Analysis (DEA) and De Novo optimization approach.

3 Basic DEA model

In [2], we can see that Data Envelopment Analysis (DEA) encompasses a variety of models and methods to evaluating performance. The essential characteristic of the DEA model is the reduction of the multiple input and multiple output using weights to that of a single ”virtual” input and a single ”virtual” output The method searches for the set of weights which maximize the efficiency of the project. The DEA may be characterized as method of objective weight assessment.

The first DEA model was developed by Charnes, Cooper and Rhodes [1]. Suppose there are n projects each consuming r inputs and producing s outputs and (r,n)- matrix X, (s,n)-matrix Y of observed input and output measures. The essential characteristic of the CCR ratio model is the reduction of the multiple input and multiple output to that of a single ”virtual” input and a single ”virtual” output. For a particular project the ratio of the single output to the single input provides a measure of efficiency that is a function of the weight multipliers (u,v). Instead of using an exogenously specified set of weights (u,v), the method searches for the set of weights which maximize the efficiency of the project P_0. The relative efficiency e_0 of the project P_0 is given as maximization of the ratio of single output to single input to the condition that the relative efficiency of every project is less than or equal to one. The formulation leads to a linear fractional programming problem.

$$e_0 = \frac{\sum_{i=1}^{s} u_i y_{i0}}{\sum_{j=1}^{r} v_j x_{j0}} \to \max$$

$$\frac{\sum_{i=1}^{s} u_i y_{ih}}{\sum_{j=1}^{r} v_j x_{jh}} \le 1 \quad h = 1,2,\ldots,n \tag{1}$$

$$u_i, v_j \ge \varepsilon, \quad i = 1,2,\ldots,s, \quad j = 1,2,\ldots,r$$

If it is possible to find a set of weights for which the efficiency ratio of the project P_0 is equal to one, the project P_0 will be regarded as efficient otherwise it will be regarded as inefficient. The set of efficient projects is designed by this way.

Solving of this nonlinear nonconvex problem directly is not an efficient approach. The following linear programming problem with new variable weights (μ,ν) that results from the Charnes - Cooper transformation gives optimal values that will also be optimal for the fractional programming problem.

$$e_0 = \sum_{i=1}^{s} \mu_i y_{i0} \to \max$$

$$\sum_{j=1}^{r} \nu_j x_{j0} = 1$$

$$\sum_{i=1}^{s} \mu_i y_{ih} - \sum_{j=1}^{r} \nu_h x_{jh} \le 0 \quad h = 1,2,\ldots,n$$

$$\mu_i, \nu_j \ge \varepsilon, \quad i = 1,2,\ldots,s \quad j = 1,2,\ldots,r \tag{2}$$

For some reasons it can be very useful to search the efficient frontier in the DEA model. The set of efficient projects is called the reference set. The set spanned by the reference set is called the efficient frontier. Searching the efficient frontier in the DEA model can be formulated as a multiobjective linear programming problem [4]. Different multiobjective linear programming methods can be used for solving of the problem.

The problem is defined as maximization of linear combination of outputs and minimization of linear combination of inputs. The combination vector is $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

$$Y\lambda \to \text{"max"}$$

$$X\lambda \to \text{"min"}$$

$$\lambda \geq 0 \tag{3}$$

A solution λ_0 is efficient if there does not exist another λ such that

$$Y\lambda \geq Y\lambda_0, \quad X\lambda \leq X\lambda_0, \quad (Y\lambda, X\lambda) \neq (Y\lambda_0, X\lambda_0).$$

4 De Novo optimization

Adding efficient projects in a portfolio will provide an inaccurate measure of the portfolio's true efficiency. The portfolios collective inputs and outputs must be compared against the set of all portfolios. The set of all portfolios is a set of combinations of single projects. Combinations of single projects to all portfolios are given by the vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, where $\lambda_i = 1$ (the project i is included in the portfolio) or $\lambda_i = 0$ (the project i is not included in the portfolio). The cumulative inputs $x = X\lambda$ and the cumulative outputs $y = Y\lambda$.

The model for efficient project portfolio designing is based on a combination of a DEA model and De Novo optimization. Traditional concepts of optimality focus on valuation of already given systems. New concept of designing optimal systems was proposed in [8]. As a methodology of optimal system design can be employed De Novo programming for reshaping feasible sets in linear systems. The approach is based on reformulation of problems by given prices of resources and the given budget.

The reformulation of the problem introduces new values: p is an r-vector of resource prices, q is an s-vector of output evaluations, and B is the given total available budget for the whole project portfolio.

Efficient project portfolio problem can be formulated as following linear programming problem:

$$q^T y \to \max$$

$$X\lambda = x$$

$$Y\lambda \geq y$$

$$p^T x \leq B$$

$$x \geq 0, \quad y \geq 0, \quad \lambda \in \{0,1\}^n \tag{4}$$

From the condition $q > 0$ and the formulation of the problem (4) the optimal solution of problem (4) must satisfy $Y\lambda = y$. The problem (4) can be reformulated into the problem (5):

$$q^T Y\lambda \to \max$$

$$X\lambda = x$$

$$p^T x \leq B$$

$$x \geq 0, \quad \lambda \in \{0,1\}^n \tag{5}$$

Inserting $X\lambda = x$ in the budget restriction becomes $p^T X\lambda \leq B$ and gets the following equivalent problem (6) in variables λ.

$$q^T Y\lambda \to \max$$

$$p^T X\lambda \le B$$

$$\lambda \in \{0,1\}^n \tag{6}$$

The model (6) only consists of one constraint with n binary variables $\lambda_i, i = 1,2,\ldots,n$.

Example 1.

An organization considers 5 potential projects $(P_1, P_2, \ldots, P_5)$ that are characterized by two inputs $(I1, I2)$ and two outputs $(O1, O2)$. The parameters of projects are given in Table 1.

	P_1	P_2	P_3	P_4	P_5
$I1_i$	6	3	8	9	5
$I2_i$	8	4	2	4	6
$O1_i$	9	7	6	10	8
$O2_i$	12	10	15	8	12
e_i	0.643	1	1	1	0.761

Table 1 Parameters of potential project

The efficiency ratios e_i of projects were computed using the model (2). The set of efficient projects consists of projects P_2, P_3, P_4 .

De Novo optimization is used. New values are introduced: the vector of resource prices $p = (5,6)$, the vector of output evaluations $q = (10,12)$, the given total available budget $B = 200$. The efficient project portfolio was computed using the model (6). The portfolio consists of projects P_1, P_2, P_3. The project portfolio is different from the set of efficient projects.

5 Extensions

The proposed model is based on a DEA model and De Novo approach. The basic model allows possible extensions. We can increase the flexibility of the model in several important ways.

According [7], the weights in the DEA can be restricted by the decision maker's judgements by the AHP. The comparison matrix $C = (c_{jk})$, where elements c_{jk} are judgements of w_j / w_k. It is known that the preference region W is structured by column vectors of the comparison matrix C. Any weight vector from W can be obtained as linear combination of column vectors

$$w = C\mu . \tag{7}$$

where μ is a nonnegative vector of coefficients $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$. If the matrix C is consistent, the consistency index $C.I. = 0$, the preference region is a line through origin. If the matrix C is inconsistent, the consistency index C.I. > 0, the preference region is a convex cone, the greater consistency index, the greater preference cone.

Each project portfolio is characterized by cumulative outputs y. Let us assume that $g = (g_1, g_2, \ldots, g_n)$, represents organizational goals to meet or exceed cumulative output y. The following constraint ensures that only portfolios whose cumulative outputs meet or exceed stated goals are feasible and can be evaluated:

$$Y\lambda \ge g , \tag{8}$$

where all or some outputs are compared with goals.

There are situations where it is not possible to evaluate outputs and combine the different outputs to one. Outputs should be considered separately. The problem (6) is reformulated to multiobjective linear programming problem with binary variables:

$$z = Y\lambda \to \text{"max"}$$

$$p^T X\lambda \le B$$

$$\lambda \in \{0,1\}^n \tag{9}$$

In [5], we can see that the problem (9) is possible to solve as multiobjective De Novo linear programming problem. Searching for a better portfolio of resources leads to a continuous reconfiguration and reshaping of systems boundaries. Technological innovations bring improvements to the desired objectives and the better utilization of available resources. Multiobjective optimization can be taken as a dynamic process. These changes can lead to beyond tradeoff-free solutions. Dynamization of the problem is very important but generally difficult.

It is easy to track changes of the budget B at the time by recalculating the appropriate model for the new budget levels. It is also possible to use linear parametrization of the budget depending on time $B(t)$. The resources for projects can be purchased from the budget in time. It is important to monitor the arrival of new projects, completion of old projects, and early termination of non-perspective projects in time. In these cases, the models will be recalculated with new parameters in time. Future work will focus on the elaboration of the dynamic approaches of project portfolio designing.

6 Conclusions

An approach for efficient project portfolio designing is proposed in the paper. The paper focuses on the of project portfolio selection problem solved by applying sophisticated models. The approach is based on a combination of a DEA model and De Novo optimization. Some extensions of the basic model are proposed and discussed. The problem can be formulated as a multiobjective linear programming problem and solved by De Novo approach for the specific problem. An important factor of the project portfolio designing problem is time. Some approaches for dynamization are proposed. Combination of the methods for searching an efficient project portfolio and methods for specific requirements gives a powerful instrument to capture managerial problems. The experiments show that this approach can be an appropriate instrument for analyzing project portfolio problems with dynamic changes and can produce interesting results in comparison with other approaches.

Acknowledgements

Supported by the grant No. P402/12/G097 (DYME - Dynamic Models in Economics) of the Czech Science Foundation and by Grant No. IGA F4/54/2015, Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] Charnes, A., Cooper, W.W., and Rhodes, E.: Measuring Efficiency of Projects. *European Journal of Operation Research* **1** (1978), 429-444.

[2] Cooper, W. W., Seiford, L. M., and Tone, K.: *Introduction to Data Envelopment Analysis and its Uses: With DEA-Solver Software and References*. Springer, New York, 2006.

[3] Enoch, C., N.: *Project Portfolio Management: A Model for Improved Decision Making*. Business Expert Press, New York, 2015.

[4] Fiala, P.: Data Envelopment Analysis by Multiobjective Linear Programming Methods. In: *Multiple Objective and Goal Programming* (Trzaskalik, T., and Michnik, J., eds.). Physica-Verlag, Heidelberg, 2002, 39–45.

[5] Fiala, P.: Multiobjective De Novo Linear Programming. *Acta Universitatis Palackianae Olomucensis*. *Facultas Rerum Naturalium. Mathematica* **50** (2011), 29–36.

[6] Larson, E. W., and Gray, C. F.: *Project Management: The Managerial Process*. Post and Telecom Press, China.

[7] Saaty, T. L.: *The Analytic Hierarchy Process*. RWS Publications, Pittsburgh, 1990.

[8] Zeleny, M.: Multiobjective Optimization, Systems Design and De Novo Programming. In: *Handbook of Multicriteria Analysis* (Zopounidis, C., and Pardalos, P. M., eds.). Springer, Berlin, 2010, 243–262.

On the stability of spatial econometric models: Application to the Czech Republic and its neighbors

Tomáš Formánek[1], Roman Hušek[2]

Abstract. Regional macroeconomic processes may not be properly analyzed without accounting for their spatial nature: regional distances, interactions between neighbors, spill-overs and interdependencies. Spatial econometrics is a unique tool for a broad range of quantitative analyses and evaluations with geo-coded (spatially defined) data. This contribution focuses on various factors of spatial model stability. Different spatial dependence matrices are used to evaluate the stability of the estimated unemployment dynamics in the Czech Republic, the Slovak Republic and all their EU-neighbors. The analysis is performed at the NUTS2 level. Spatial econometrics and regional competitiveness paradigms are combined into a robust model describing unemployment dynamics. Alternative spatial structures (i.e. neighbor definitions) are used for verification of stability of the estimated model and its properties. Spatial approach to econometric analysis provides important additional insight and robustness to a potentially broad range of spatially defined unemployment analyses that may be carried out using regional (spatial) data.

Keywords: spatial econometrics, model stability, regional dynamics.

JEL Classification: C23, C31, C52, E66
AMS Classification: 91B72

1 Introduction

Spatial econometric models explicitly address the presence of spatial effects (such as economic spill-overs) when analyzing the relationships between variables using regression models and other quantitative estimation methods. Spatial models play an ever more important role in regional macroeconomic and social analyses, real estate studies, ecological applications and in many other non-economic fields of research. For this type of analyses, data usually need to be geo-coded by means of the latitude/longitude geographic coordinates system, as distances and common borders are used to estimate spatial dependencies.

It may be argued that much of the spatial effects (spatial dependencies) are attributable to omitted variable factors. However, spatial autocorrelation may be conveniently interpreted as a proxy for many real and theoretically sound, yet practically unobservable spatial effects - many spatial interactions and their dynamic features are very difficult to explicitly define and properly structure in a way that would facilitate informative and harmonized quantification. Tasks such as consistently measuring cross-border work commuting preferences, accounting for administrative/qualification employment barriers between countries, quantifying the impact of language differences, aerial distances vs. transportation, etc. would inherently introduce many subjective decisions and – in practical terms – many disputable features to quantitative models. Hence, spatial models may provide a useful, interpretable and functional approach towards regional (macroeconomic) data analysis.

The variety of available approaches towards modelling functional forms of spatial dependencies and the underlying diverse spatial neighbor definition possibilities imply that researchers usually need to consider several different choices (spatial structure settings) in order to verify model stability and robustness. As far as spatial models are concerned, there usually isn't a single right solution and researches often look for the most useful or interpretable model setup. On the other hand, some authors argue in favor of model stability, inherent to many types of flexible spatial models. In this contribution, we discuss the potential for stability in estimated spatial models and we provide empirically-oriented arguments to the conclusions reached by LeSage and Pace in [7], who call out the instability in estimated spatial models as the "Biggest myth in spatial econometrics".

The remainder of this paper is structured as follows: Section two covers selected key topics of the spatial approach to econometrics and provides references to fundamental literature, section three discusses stationarity and stability in spatial models and section four provides an illustrative application to the stability topics outlined. Section five and the list of references conclude our paper.

[1] VŠE Praha, FIS, nam. W. Churchilla 4 Praha 3 Prague 130 67 Czech Republic, e-mail: formanek@vse.cz.
[2] VŠE Praha, FIS, nam. W. Churchilla 4 Praha 3 Prague 130 67 Czech Republic, e-mail: husek@vse.cz.

2 Spatial econometrics and the Spatial Durbin model

Before estimating spatial models, we should apply preliminary tests for spatial autocorrelation in the observed cross-sectional data. Many types of spatial autocorrelation test statistics are available. See [1], for the most common test statistics, such as *Moran's I* and *Geary's C*. Once significant spatial dependence in observed data is verified, spatial regression may be used to account for such situation. Again, various model specifications and estimation methods are available and detailed discussion is provided in [3]. **Spatial Durbin model** (SDM) is used when spatial interactions are present both in the dependent variable and in regressors. A general formula for the SDM and its reduced form may be written as

$$\boldsymbol{y}_t = \rho \boldsymbol{W}\boldsymbol{y}_t + \boldsymbol{X}_t\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_t\boldsymbol{\theta} + \alpha\boldsymbol{\iota} + \boldsymbol{u}_t \tag{1}$$

$$(\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y}_t = \boldsymbol{X}_t\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_t\boldsymbol{\theta} + \alpha\boldsymbol{\iota} + \boldsymbol{u}_t \tag{2}$$

where $\boldsymbol{y}_t$ is the vector of all y_{it} spatial unit observations at time t, $\boldsymbol{I}$ is the $n \times n$ identity matrix, $\boldsymbol{X}_t$ is a matrix of regressors (excluding the intercept elements $\alpha\iota$) and $\boldsymbol{W}_t$ is a spatial weights matrix. Maximum likelihood (ML) approach is used to estimate all model parameters: α, ρ, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$; α and ρ are scalars, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are $k \times 1$. As usual, $\boldsymbol{u}_t$ and its elements u_{it} describe the random part of the regression model.

Spatial weights matrix ($\boldsymbol{W}$) as in (1) or (2) is the corner stone of spatial econometrics and, perhaps surprisingly, its construction is the most ambiguous part of the otherwise well rooted methodology of spatial model specification and estimation. $\boldsymbol{W}$ is usually calculated in a two-step approach: First, a square spatial matrix $\boldsymbol{S}$ is used to define neighbors (spatially close observations) using a dummy variable technique, where each element of the symmetric spatial matrix equals 1 if the two spatial units are neighbors and 0 otherwise. Then, a spatial weights matrix $\boldsymbol{W}$ is constructed by row-standardizing $\boldsymbol{S}$, so that the row weights sum up to 1, while diagonal elements of $\boldsymbol{S}$ and $\boldsymbol{W}$ are set to zero by definition (units are not neighbors to themselves).

The first step (spatial matrix $\boldsymbol{S}$ construction) often requires extensive geographical (polygon-based) mapping datasets and specialized software. *Contiguity approach* is a theoretically simple yet computationally complex rule, defining two units as neighbors if they share a common border. A generalization of this approach is based on the premise that a second order neighbor (which may also be considered as a neighbor) is the neighbor of the first order (actual contiguity-based) neighbor – where the maximum allowed order of neighborhood (neighborhood lag) may be set arbitrarily. *Distance-based approach* usually constructs the spatial matrix by defining two units as neighbors if their distance does not exceed some ad-hoc predefined threshold. This is a relatively popular approach, yet it generates “islands” (units with zero neighbors), unless the defined threshold for distance between neighbors is greater than the maximum of “first nearest neighbor” distances. The distance-based approach is less convenient for analysis of regions with uneven geographical densities – i.e. for diverse sizes of units and distances between them. Distances are measured using *centroids*, conveniently chosen representative positions for each unit. Depending on model focus, data availability and researcher's individual preferences, centroids may be pure geographical center points, locations of main cities, population-based weighted positions, transportation network based, etc. Also, we may apply a *k-nearest neighbors* approach (*k*NN), where we denote a preset number of k nearest units as neighbors. This method conveniently solves for differences in areal densities (k neighbors are ensured for each unit), yet it usually leads to asymmetric $\boldsymbol{S}$ matrices with potentially flawed neighborhood interpretation (simple transformation algorithms for asymmetric $\boldsymbol{S}$ matrices are available).

The second step ($\boldsymbol{W}$ construction) usually consists of row-standardizing the binary 0/1 neighborhood indicators of the $\boldsymbol{S}$ matrix into matrix $\boldsymbol{W}$ so that all rows sum to unity. However, with increasing variance in units' neighbor-count (e.g. for distance-based neighbors with uneven geographical density), this widely adopted approach suffers from allocating excessive influence to links from units with few neighbors. To overcome this drawback, sometimes the non-zero elements in $\boldsymbol{S}$ matrix are “generalized” before the row-standardization. For example, distances to neighbors are used to reflect some prior information concerning the spatial dependency processes: often we assume that spatial influence is inversely proportional to distance (linear, quadratic or other functional forms of influence decay may be used). The efficiency of any $\boldsymbol{S}$ and $\boldsymbol{W}$ generalization crucially depends on the accuracy/validity of the prior information used.

Once $\boldsymbol{W}$ is properly introduced and defined, we may further elaborate on the description of the model as per equations (1) or (2): : $\boldsymbol{W}\boldsymbol{y}_t$ is the spatial lag of $\boldsymbol{y}_t$. For the i-th spatial unit, we may write: SpatialLag(y_{it}) = $\boldsymbol{w}_i^T\boldsymbol{y}_t$, where $\boldsymbol{w}_i^T$ is a row vector – the i-th row of $\boldsymbol{W}$ matrix is used for calculation. Similarly, $\boldsymbol{W}\boldsymbol{X}_t$ describes spatial interactions among the regressors. It has to be stressed out that the estimated parameters of the model (1) do not form a basis for a complete and proper description of model dynamics, especially if we want to focus on spillover effects: As we simulate a change in the j-th explanatory variable for spatial unit i (Δx_{ijt}), we expect the dependent variable in the i-th unit to change (direct effect) and at the same time - for $\rho \neq 0$ and/or $\theta_j \neq 0$ - effects on the dependent variables in neighboring units may be expected (indirect effects). The complexity of this type of dynamics for the SDM model may be demonstrated using partial derivatives. First, re rewrite the equation (2) into

$$\boldsymbol{y}_t = (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}(\boldsymbol{X}_t\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_t\boldsymbol{\theta}) + \boldsymbol{R}_t\,, \tag{3}$$

where $\boldsymbol{R}_t = (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\alpha\boldsymbol{\iota} + (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{u}_t$ and the spatial multiplier matrix may be decomposed as follows: $(\boldsymbol{I} - \rho\boldsymbol{W})^{-1} = \boldsymbol{I} + \rho\boldsymbol{W} + \rho^2\boldsymbol{W}^2 + \rho^3\boldsymbol{W}^3 + \ldots$ Hence, the partial derivatives for the expected values of $\boldsymbol{y}_t$ with respect to a chosen *j*-th explanatory variable $\boldsymbol{x}_{jt}$ in an *i*-th spatial unit $\{i = 1,2, \ldots \text{N}\}$ may be outlined as in (4), where time subscripts are omitted for readability:

$$\left[\frac{\partial E(y)}{\partial x_{1j}}, \frac{\partial E(y)}{\partial x_{2j}} \quad \cdots \quad \frac{\partial E(y)}{\partial x_{Nj}}\right] = (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\begin{bmatrix} \beta_j & w_{12}\theta_j & \ldots & w_{1N}\theta_j \\ w_{21}\theta_j & \beta_j & \ldots & w_{2N}\theta_j \\ \ldots & \ldots & \ldots & \ldots \\ w_{N1}\theta_j & w_{N2}\theta_j & \ldots & \beta_j \end{bmatrix}, \tag{4}$$

where the diagonal elements of the RHS in (4) represent direct effects and the off-diagonal elements represent the indirect effects (spillovers). Direct effects and spillovers differ across spatial units *i*, provided that $\rho \neq 0$. For presentation purposes, all the N×N partial derivative matrices (*k* matrices are produced - one for each of the x_j regressors) are summarized as follows: diagonal elements are averaged into a single direct effect indicator; similarly, row sums of the off-diagonal elements may be averaged into a summary spillover indicator. Testing for statistical significance of the direct and indirect effect is discussed in [3], along with many additional topics covering SDMs and other spatially-augmented models.

3 Stability and stationarity topics for SDMs

As discussed above, the spatial weights matrix $\boldsymbol{W}$ is a nonnegative matrix with zeros on the diagonal. This matrix is not estimated using the ML method – instead, it needs to be specified prior to the estimation of model (1) parameters. The variety of available matrix $\boldsymbol{S}$ definition approaches and $\boldsymbol{W}$ standardization methods imply that researchers usually need to consider alternative spatial structure settings in order to verify model stability and robustness of the results. On the other hand, in [7], the authors argue that SDMs and other flexible spatial models allow for accurate estimation of the spatial effects, even if *both* the spatial matrix $\boldsymbol{W}$ and the spatial regression model are misspecified. In this article, we focus on possible misspecification of $\boldsymbol{W}$ and its potential consequences. In [7], the authors postulate that for a given model - estimated using two similar (highly correlated) weight matrices $\boldsymbol{W}_a$ and $\boldsymbol{W}_b$ - it would be unlikely to reach materially different partial derivatives as in equation (4). Their argument in [7] is supported by an empirical (micro-level) housing-prices example (data from [5]), with 506 spatial units and for three alternative $\boldsymbol{W}$ matrices, generated using the *k*NN approach (for k = 5, 6 and 7). It is our belief that the conclusions presented in [7] are supported by an insufficient and relatively narrow class of spatial models (with a *k*NN-based spatial structure). Therefore, in the next section, we provide an empirical application demonstrating the conclusions presented in [7], using macroeconomic data at the NUTS2 level, with distance-based $\boldsymbol{W}$ matrices.

Also, regardless of the method chosen for spatial matrix definition and the $\boldsymbol{W}$ specification as described in the previous paragraph, certain conditions need to be observed for stationarity of the estimated model (1). Specifically, row and/or column sums of the binary neighbor-indicators in the spatial matrix $\boldsymbol{S}$ should be uniformly bounded (should not diverge to infinity) even if the number of spatial units (N) goes to infinity. This condition reflects the fact that correlation between two spatial units should converge to zero with increasing distance between units. As far as parameter space for ρ is concerned, most authors assume that ρ values are restricted to lie within the (-1, 1) interval. On the other hand, [6] argue that $\rho \in (r_{min}^{-1}; 1)$, where r_{min} is the minimum (most negative) purely real characteristic root of $\boldsymbol{W}$. Additional detailed discussion of the stability conditions and parameter spaces is provided in [3].

4 Application to the Czech Republic and its neighbors

In this section, we estimate a SDM of unemployment for the Czech Republic, the Slovak Republic, Poland, Hungary, Austria and Germany. Data from Eurostat are used (NUTS2 level, 82 spatial units). The unemployment model is derived along the concept of regional competitiveness (for detailed description of competitiveness and other labor market aspects, see [4] and [2]). Therefore, unemployment is modelled as a function of relative GDP dynamics, a convenient measure of technological advantage and we employ spatial clustering analysis as in [8] to describe specific behavior in high unemployment cluster(s). Specifically, variables for the estimation of model (1) are as follows: y_{it} is the general rate of unemployment (2012 data are used) in a given spatial unit *i* and the $\boldsymbol{X}_t$ matrix contains the following three variables: GDP_{it} is the region's GDP per capita expressed as a percentage of

EU average (we use 2011-2010 first differences to proxy region's macroeconomic competitiveness dynamics), $TechEmp_{it}$ describes the percentage of employees working in the "high-tech industry" (NACE r.2 code HTC) in a given region and time period and $H\_U\_cluster_{it}$ is a dummy variable based on the G^* (local G as in [8]) that discerns local clusters of high values of the variable being analyzed – as we searched for high unemployment clusters, we found a single cluster, containing the following NUTS2 regions: HU10, HU31, HU32, HU33, PL21, PL22, PL32, PL33, SK02, SK03 and SK04 (generally speaking this cluster consist of units close to or bordering with Ukraine).

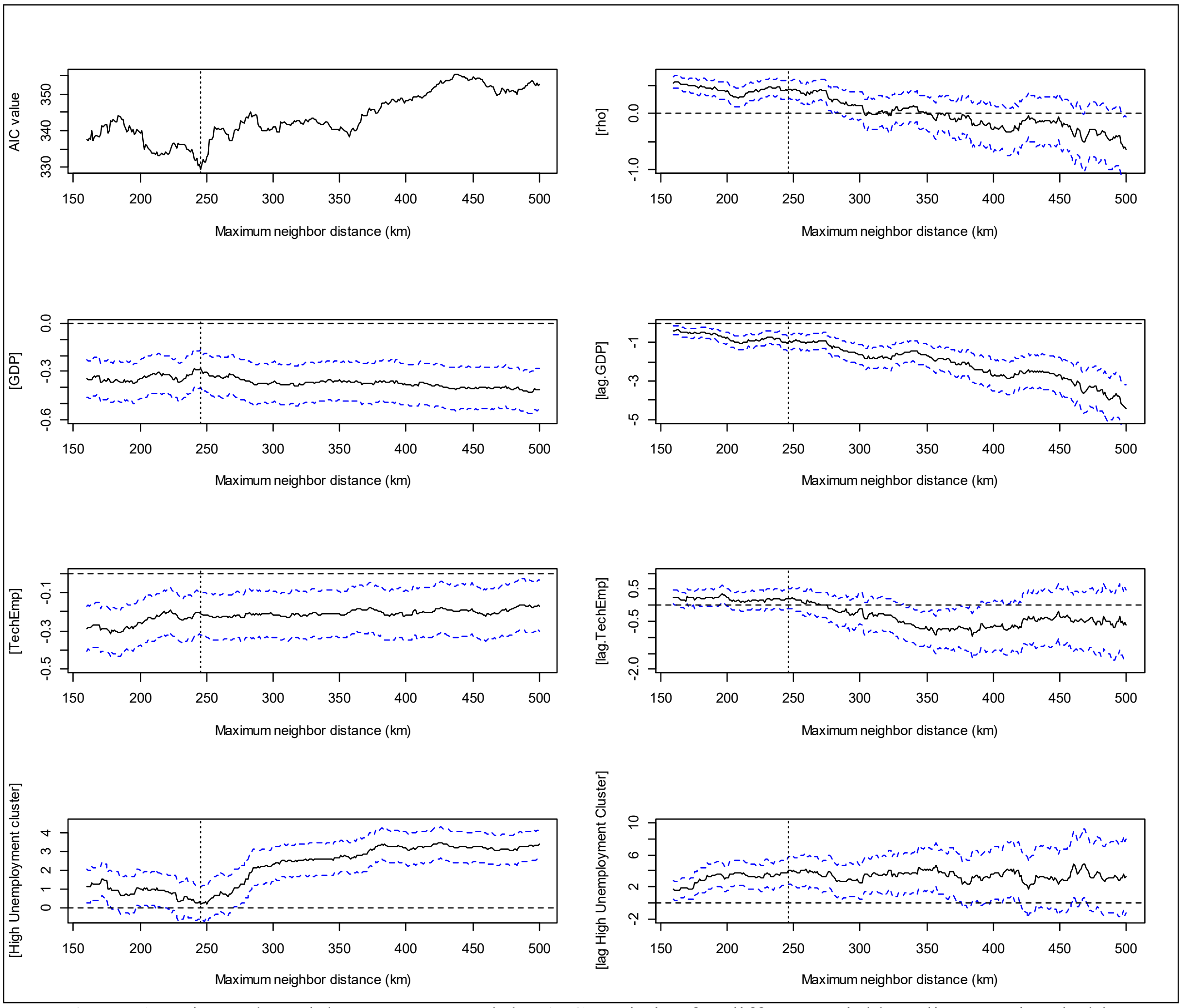


Figure 1 Estimated model parameters and the AIC statistics for different neighbor distance thresholds

Model (1) was estimated using R software. Different distance-based $\boldsymbol{W}$ matrices were used, with maximum neighbor distance thresholds varying from 160 km to 500 km. Thresholds lower than 160 km lead to the existence of at least one island unit (a region with zero neighbors), which breaks down the ML estimation. Beyond the 500 km threshold, unemployment-related spatial effects are not very realistic and the estimation does not lead to interpretable results. For $\boldsymbol{W}$ construction, 2-km distance threshold increments were used, hence a total of 171 versions of model (1) were estimated and used for comparison. The complex estimation output is summarized in figure 1, where the estimated model parameters are shown (intercept excluded) along with corresponding Akaike information criteria (AIC) for the model (1) estimated at each distance threshold. Among the models compared, $\boldsymbol{W}$ constructed using the maximum neighbor distance threshold of 246 km minimizes the AIC statistics, hence we choose this specification as our final model – the estimated coefficients along with their standard errors (heteroskedasticity consistent) and p-values may be observed from table 1.

From figure 1, we may see that the plotted AIC values have various local minima and upon considering such local minima, we may reach different coefficient estimates that can differ in their magnitudes and – more dramatically – in their significance levels. Nevertheless, the estimated coefficients are reasonably stable (economically speaking) across an interval of 220 - 270 km distance thresholds (beyond 270 km, ρ is not statistically significant). For proper interpretation of the estimated model (1), partial derivatives as per equation (4) – i.e. the direct and indirect effects – are provided in table 2 ($\boldsymbol{W}$ constructed using the 246 km threshold). While the standard errors and p-values in table 1 are analytical, errors and significance measures in table 2 are simulated (see [3] for detailed discussion).

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	5.8273	1.9332	3.014	0.0025
GDP	-0.2833	0.1169	-2.423	0.0153
TechEmp	-0.2094	0.1119	-1.871	0.0613
H_U_cluster	0.1625	0.9338	0.1741	0.8617
lag.GDP	-1.0025	0.3873	-2.5881	0.0096
lag.TechEmp	0.1913	0.3070	0.6234	0.5330
lag. H_U_cluster	4.1063	1.7046	2.4089	0.0159
ρ	0.4128	0.1618	2.5517	0.0107

Table 1 Estimated SDM (1) parameters for maximum neighbor distance threshold of 246 km

Impact measures	Mean	Std. Error	z value	Pr(>|z|)
Direct effects				
GDP	-0.3306	0.1153	-2.8684	0.0041
TechEmp	-0.2021	0.1169	-1.7284	0.0838
H_U_cluster	0.3673	0.8708	0.4218	0.6731
Indirect effects				
GDP	-1.923	0.6000	-3.2042	0.0013
TechEmp	0.198	0.6468	0.3060	0.7595
H_U_cluster	7.029	2.4064	2.9210	0.0034

Table 2 Estimated SDM (1) direct and indirect impacts for maximum neighbor distance threshold of 246 km

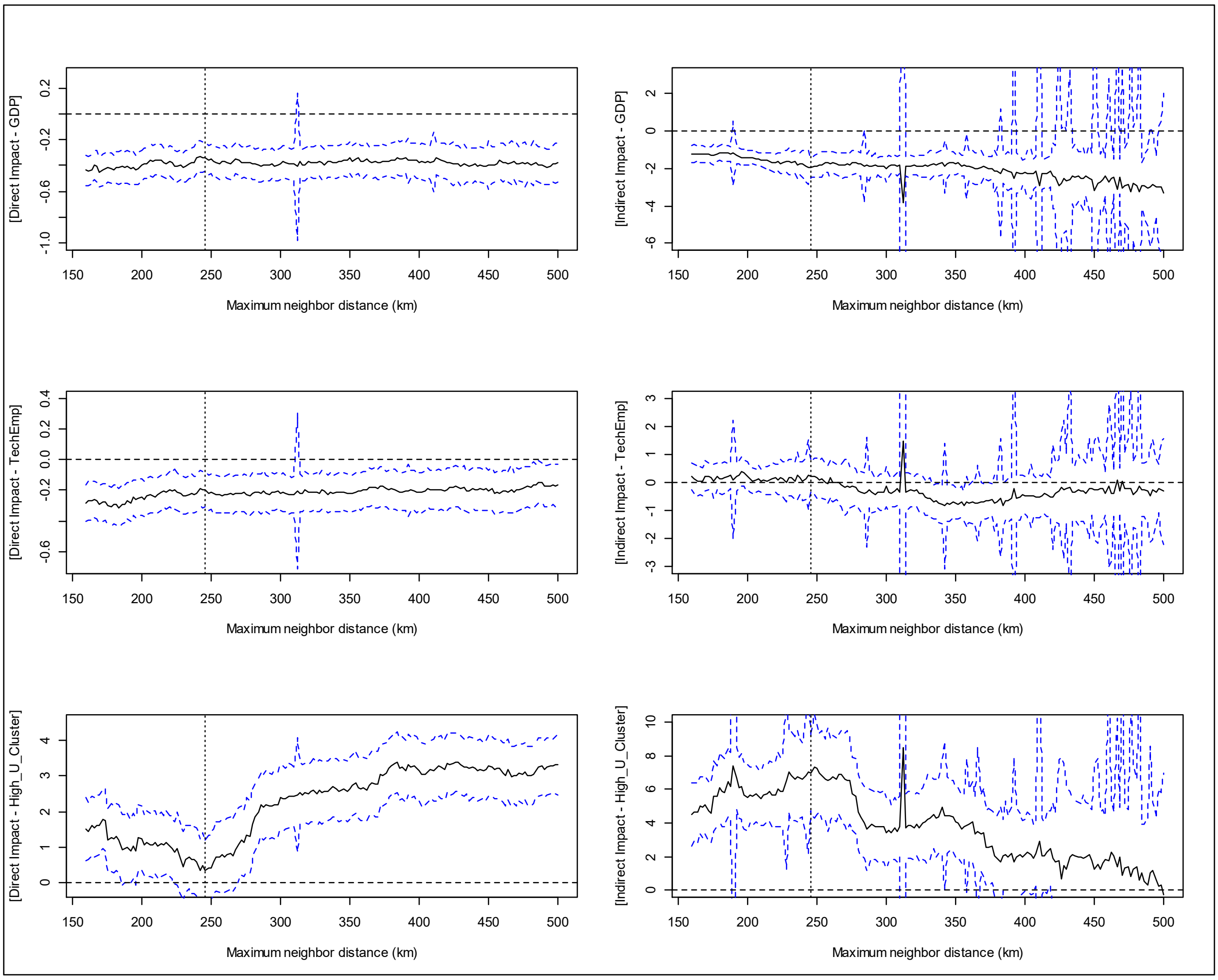


Figure 2 Estimated direct and indirect impacts for different neighbor distance thresholds

Figure 2 summarizes the partial derivatives - direct and indirect impacts – for models estimated along the 160 - 500 km maximum neighbor distance thresholds. Again, the estimated impacts vary to a relatively small extent across the 220 - 270 km interval of distance thresholds. However, geographical heterogeneity in spatial units (uneven geographical densities – sizes of units and distances between centroids vary substantially in our dataset) leads to a number of significant jumps in the estimated impact means (and simulated standard errors) between models estimated using "adjacent" $\boldsymbol{W}$ matrices, even as those are based on vey similar distance thresholds. It may be argued that the number of neighbors to each spatial unit is a non-linear function of the maximum distance neighbor threshold (this is in sharp contrast to kNN-based $\boldsymbol{W}$ matrices constructed using different k values). To summarize, we are able to identify a set of spatial structures ($\boldsymbol{W}$ matrices) where the estimation of model (1) leads to relatively stable (similar) coefficient estimates and partial derivatives as per equation (4).

Our conclusion is especially important given the heterogeneous geographical nature of the economies considered: NUTS2 regions are bound by the number of inhabitants (800,000 to 3 million) and there are prominent differences among geographical areas of Germany (densely populated and relatively small regions) on one hand and the NUTS2 units in Poland and Hungary on the other hand. At the same time, our results should not be interpreted as if stability analysis (based on different spatial structures) is a redundant task for spatial econometrics - it is quite the opposite. Stability verification as outlined in this section provides two important layers to the analysis: robustness assurance for the estimation results presented and a better description of spatial dependence in terms of statistical significance of the interactions over different distance-based spatial structures.

5 Conclusions

Spatial econometric models provide a useful estimation framework that allows for improved analyses of regional macroeconomic data. Spatial models have a unique ability to discern between geographical determination and the influence of relevant macroeconomic variables, many of which may be subject to or directly controlled by economic policy actions as undertaken by the central authorities at different levels. Our results support the conclusion of LeSage and Pace [7] in favor of spatial models leading to robust estimated spatial dependencies. For a relatively large interval of maximum neighbor distance thresholds, we find reasonably stable direct and indirect impacts. This study provides relevant implications towards the stability and robustness of spatial analyses for diverse macroeconomic indicators and for different types of spatial structures considered.

Acknowledgements

Supported by the grant No. IGA F4/73/2016, Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] Anselin, L., and Rey, S. J. (eds.): *Perspectives on Spatial Data Analysis*. Springer Verlag, Berlin, 2010.
[2] Čížek, O.: Nonlinear model of the Eurozone labor market. In: *Mathematical Methods in Economics 2015 (MME)*. Plzeň : University of West Bohemia, 2015, 109–114.
[3] Elhorst, J. P.: *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*. Springer Briefs in Regional Science, Berlin, 2014.
[4] Formánek, T. and Hušek, R.: The Czech Republic and its neighbors: Analysis of spatial macroeconomic dynamics. In: *Mathematical Methods in Economics 2015 (MME)*. Plzeň : University of West Bohemia, Cheb, 2015, 190–195.
[5] Harrison, D. and Rubinfeld, D. L.: *Hedonic prices and the demand for clean air*. J. Environ. Econ. Manag. **5** (1978), 81-102.
[6] LeSage, J. P. and Pace, R. K.: *Introduction to spatial econometrics.* CRC Press, Taylor & Francis Group, Boca Raton, 2009.
[7] LeSage, J. P., and Pace R. K.: *The biggest myth in spatial econometrics.* Econometrics **2**(4) (2014), 217-249.
[8] Ord, J. K. and Getis, A.: *Local spatial autocorrelation statistics: distributional issues and an application.* Geographical Analysis, **27** (1995), 286–306.

Application of selecting measures in data envelopment analysis for company performance rating

Jiří Franek[1]

Abstract. Data Envelopment Analysis (DEA) measures efficiency of a finite number of decision making units (DMUs). This approach is very sensitive to the choice of measures. Traditionally the rule of thumb has been used for initial decision about number of inputs and outputs. Recently a growth of interest in the DEA methodology have shown a need to create an approach or model which could identify the optimal number of inputs and outputs and which identifies the contribution of each variable to the measure of efficiency. The practice has been to select the variables by simply choosing the ones that make economic sense or use them according to preferred performance measures. Variable selection is further complicated because it is difficult to measure some attributes of the inputs and outputs that contribute to the efficiency of a DMU. Some approaches have been developed recently for BCC model. The goal of this contribution is to apply and compare aforementioned proposed methodology on a performance rating of selected nonfinancial companies using envelopment form of the selecting model and MCDM approach using objective weighting for a comparison.

Keywords: data envelopment analysis, multiple criteria decision making, selective measures, rating.

JEL Classification: C44, M20
AMS Classification: 90B50, 62J15

1 Introduction and literature review

Two approaches are commonly used to measure efficiency: the parametric approach, which relies on statistical techniques to estimate the parameters of a production function, and the non-parametric approach, which compares the observed inputs and outputs of each firm with that of the most performing firms in the information set. The parametric approach has been subject to persistent criticism, centred on two points; the assumption that the production function has the same functional form for all the firms, and the fact that econometric estimation of efficiency can produce biased and inconsistent parameter estimates (since an econometric measure of efficiency reflects the average performance and not the best performance). Data Envelopment Analysis (DEA) is now the most popular method used to measure efficiency. DEA is a non-parametric method, which does not assume any specific production function. Instead, it uses linear programming to identify points on a convex hull defined by the inputs and outputs of the most efficient firms (any productive unit, like a firm, is called a Decision Making Unit (DMU) in the literature of DEA).

The goal of this contribution is to apply and compare proposed DEA selecting model methodology on a performance rating of selected nonfinancial companies using envelopment form of the selecting model and MCDM approach using objective weighting for a comparison. This paper is focused on measuring efficiency when the number of firms is small and or when the number of explanatory variables needed to compute the measure of efficiency is too large to allow for the statistical approach. Hence, a selective model is used to determine appropriate number of inputs and outputs. This approach is mirrored by and MCDM approach using entropy and CRITIC method to estimate criteria weights and TOPSIS method for ranking of firms. First, a definition of selected methods is presented then a selection model [7] is described. Finally, the approach is applied to the assess leading automotive and telecommunication firms in a period of 4 years to analyse changes in applied models under different conditions.

Modelling efficiency with non-parametric tools was first introduced as an extension of activity analysis. The CCR model Charnes, Cooper, and Rhodes, [2] formally introduced the linear programming to measure technical efficiency with the assumption of constant returns to scale. In the CCR model, DMUs adjust either their use of inputs or their outputs to reach the production frontier. The BCC model [Banker, Charnes, and Cooper, [1] removed the assumption of constant returns to scale, and proposed the additive DEA model, where both inputs and outputs can be adjusted simultaneously [3]. All models use the distance to one of the facets of the production or cost frontier to generate an efficiency index. The DEA approach also has also earned its own set of criticisms, focused

[1] VŠB – Technical University of Ostrava, Faculty of Economics, Department of Business Administration, Sokolská třída 33, 702 00 Ostrava, Czech Republic, jiri.franek@vsb.cz.

on two points: (i) measures of technical efficiency are very sensitive to the omission of variables from the production function and (i) efficiency scores are biased if the degree of freedom is not uniform for all DMUs. One of the other problems was to deal with the ratio of the number of measures and DMUs. The problem of selection of appropriate number of outputs and inputs is frequently based on so called rule of thumb. A rough rule of thumb expresses the relation between the number of DMUs and the number of performance measures (see Cooper et al. [3]):

$$n \geq \max\left\{3(m+s), m \times s\right\}. \tag{1}$$

Exhaustive review of studies and used numbers of measures and DMUs can be found in Toloo et al. [8]. However, in this paper the focus is on application and validation of a lately proposed selective model based on BCC model of efficiency on empirical data.

2 Methodology

In this part the basic methodology and research approach will be described and discussed. It will be focused on selected models of DEA and MCDM that will be applied on the problem.

2.1 The BCC model

Suppose that, we have n DMUs (Decision Making Units), where each DMU utilizing m factors of production (inputs) $x_1, x_2, \dots, x_m$ (all positive) and $y_1, y_2, \dots, y_m$ (all positive) products (outputs) to produce the BCC model as introduced in Banker et al. (1984) for evaluation of efficiency of a specific DMU$_o$ is presented as follows

$$\begin{aligned}
&\max z = \sum_{r=1}^{s} u_r y_{ro} - u_o, \\
&\text{subject to} \\
&\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - u_o \leq 0, \quad j = 1, \dots, n, \\
&\sum_{m=1}^{m} v_i y_{i0} = 1, \\
&v_i \geq \varepsilon,\ u_r \geq \varepsilon,\ u_o \text{ free in sign.}
\end{aligned} \tag{2}$$

In above formulation Y is an $s \times n$ sample output matrix and X is $m \times n$ sample input matrix, u, v are multiplier imputed output and input, respectively.

2.2 The proposed selective model

In practice, it is hardly possible to increase the number of DMUs and consequently we resort to formulate some models which optimally decrease the number of performance measures. In some cases, the decision maker believes that some performance measures are more important than others and can use MCDM approach to decide which measures are important or not. There are two main approaches for decision maker: subjective or objective. Hence, to compare differentiating power of the DEA based selective model, the objective MCDM approach was followed using entropy [4][6] and CRITIC methods [5]. Multiplier form of selecting model proposed by Toloo & Tichy [7] uses extension of the multiplier form of the BCC model and multiplier form of selecting model to deal with selective measures:

$$\begin{aligned}
&\max z_3 = \sum_{r=1}^{s} u_r y_{ro} + u_0 \\
&\text{s.t.} \\
&\sum_{i=1}^{m} v_i x_{io} = 1 \\
&\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} \leqslant 0 \qquad j = 1, \dots, n \\
&\sum_{r \in s_2} b_r^y + \sum_{i \in m_2} b_i^x \leqslant \min\left\{\left[\frac{n}{3}\right], 2\sqrt{n}\right\} - (|m_1| + |s_1|) \\
&\varepsilon b_i^x \leqslant v_i \leqslant M b_i^x \qquad i \in m_2 \\
&\varepsilon b_r^y \leqslant u_r \leqslant M b_r^y \qquad r \in s_2 \\
&b_i^x, b_r^y \in \{0, 1\} \qquad i \in m_2, r \in s_2 \\
&v_i, u_r \geqslant \varepsilon \qquad i \in m_1, r \in s_1
\end{aligned} \tag{3}$$

In the next section a real data set of leading automotive and telecommunication firms over a period of four years is utilized in order to show the applicability of the proposed multiplier and MCDM approaches.

3 Application of DEA selecting model and MCDM approach

This application uses data from 6 leading companies in the automotive and telecommunication industry. The automotive industry is represented by leading companies: TOYOTA MOTOR CORP., VOLKSWAGEN AG, HONDA MOTOR CO LTD, DAIMLER AG, NISSAN MOTOR CO LTD, BAYERISCHE MOTOREN WERKE AG. The telecommunications industry is represented by another group of leading providers: AT&T INC, VODAFONE GROUP PLC, VERIZON COMMUNICATIONS INC, NIPPON TELEGRAPH & TELEPHONE, TELEFONICA SA, DEUTSCHE TELEKOM.

Following table includes performance measures used for ranking and assessment see table 1.

Measure	Character
Sales	Input
Assets	Input
R&D expenditures	Input
Stockholders' equity	Input
Employees	Input
Rank	Output
Profits	Output
Market Value	Output

Table 1 Performance measures

Given the number of DMUs and performance measures by rule of thumb we can expect to get more than one efficient DMU using BCC model for calculation efficiency.

3.1 MCMD approach to selecting performance measures

Firstly, the MCDM approach is used to investigate whether there are some measures that are estimated to have greater importance based on entropy and CRITIC approaches. Following Figure 1 and 2 show results of objective weighting of investigated measures and their changes throughout the 4-year period.

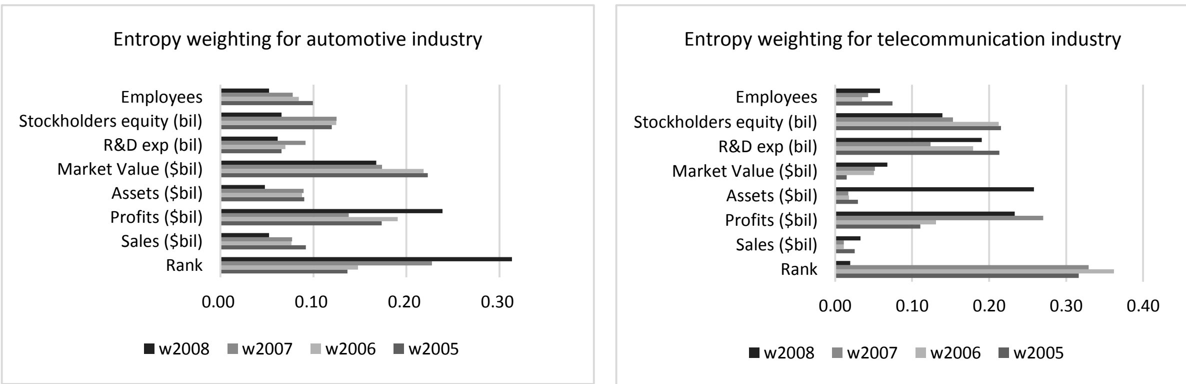


Figure 1 Entropy based weighting for automotive and telecommunication industry

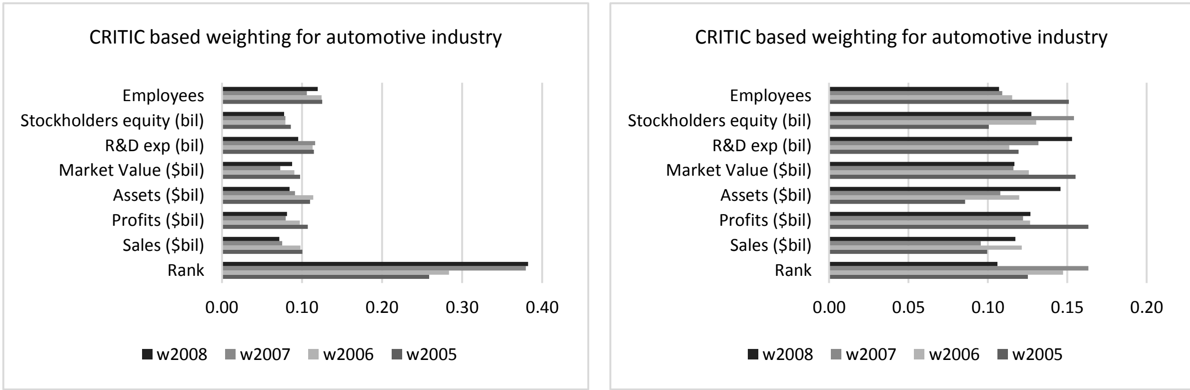


Figure 2 CRITIC based weighting for automotive and telecommunication industry

Looking at the objective weights calculated using selected MCDM methods, we can see that entropy has a more differentiating effect since it uses information based approach rather that correlation as the CRITIC method which less differentiate the weights except for the automotive industry where the “rank” measure has by far the highest weighting scores. Based on the findings it is better to used entropy based weights, so that we can rank the measures as follows (see table 2):

Automotive		Telecommunication	
measure	**geomean value**	**measure**	**geomean value**
Rank (O)	0.195	Stockholders equity (I)	0.177
Market value (O)	0.194	Profits (O)	0.174
Profits (O)	0.182	R&D exp (I)	0.174
Stockholders’ equity (I)	0.105	Rank (O)	0.165
Employees (I)	0.076	Employees (I)	0.051
Assets (I)	0.076	Market Value (O)	0.040
Sales (I)	0.073	Assets (I)	0.039
R&D Expenditures (I)	0.071	Sales (I)	0.018

Table 2 Ranking of performance measures based on entropy

The difference between both industries suggests that using geometrical mean of calculated weights over a 4-year period the output measures have more priority in automotive industry that in telecommunications. However, rank and profits have high priority. The lowest priorities are assigned to sales and assets (see table 3).

Automotive		Telecommunication	
measure	**geomean value**	**measure**	**geomean value**
Rank (O)	0.321	Profits (O)	0.134
Employees (I)	0.119	Rank (O)	0.134
R&D exp (I)	0.110	R&D exp (I)	0.129
Assets (I)	0.099	Market Value (O)	0.128
Profits (O)	0.091	Stockholders equity (I)	0.127
Market Value (O)	0.087	Employees (I)	0.119
Sales (I)	0.085	Assets (I)	0.113
Stockholders equity (I)	0.080	Sales (I)	0.108

Table 3 Ranking of performance measures based on CRITIC

Looking at the table 3 results suggest that in automotive rank is by far the most important measure followed by closely bunched other measures. In telecommunication industry the situation is less clear, when all measures weights are separated by sole percentages. This results also confirms that automotive companies using those performance measures are less correlated then companies in telecommunication industry, hence the financial conditions and capital structure is more variable.

In case of MCDM approaches it can be suggested that in an industry where performance measures are more correlated the entropy approach is more appropriate for selection of criteria.

3.2 DEA approach to selecting measures

Using proposed model (3) we can estimate what measures (input and output) have more influence on efficiency of particular DMUs. The model was tested on the same dataset. Results are summarized in the table 4.

Automotive	**2008**	**2007**	**2006**	**2005**	**Telecommunications**	**2008**	**2007**	**2006**	**2005**
(I)Sales	5	4	3	3	(I)Sales	4	3	2	3
(I)Assets	2	3	2	1	(I)Assets	3	2	4	3
(I)R&D exp	2	1	1	1	(I)R&D exp	2	0	2	1
(I)Stockholders equity	2	3	4	4	(I)Stockholders equity	2	5	3	3
(I)Employees	1	1	2	3	(I)Employees	1	2	1	2
(O)Rank	5	5	5	4	(O)Rank	5	5	4	4
(O)Profits	5	5	6	6	(O)Profits	5	2	3	4
(O)Market Value	2	2	1	1	(O)Market Value	2	5	4	4

Table 4 Selection of measures using DEA selecting model

Regarding the results, we can again see a clear distinction between both industries. The higher the number (1 to 5) the more the measure has been selected by the model as appropriate. However, this model also clearly shows that from the efficiency perspective R&D and Employees are not as important for these leading companies. In automotive the Sales and Stockholder Equity seems likely input and Rank and Profits as outputs. In the telecommunication industry the result is not easy to read. Closer look at results (considering the maximum and minimum values) suggests selection of Stockholders Equity, Rank and Market value.

3.3 Comparison of results and discussion

When comparing results from both approaches the decision maker has to have a clear understanding of both methodological and theoretical principles of particular methods. The MCDM approaches derive weights using different assumptions to DEA. However, investigated methods can be seen as substitute or alternative approaches to selection of appropriate measures of efficiency.

In MCDM as well as DEA approaches several measures are deemed to be more important than others in automotive inputs: Stockholders Equity, Sales; and outputs: Rank, Profits. In the telecommunication industry we can distinguish Stockholders Equity as the most important input and Profits and Rank as the most important outputs.

Thus we can run the BCC DEA model to see if the results show fewer efficient DMUs. We can test both industries separately.

In most cases we tend to get better differentiation between efficient non-efficient DMUs. In general, in such a case when the number of DMUs is very small we would need to use only two or three measures as maximum. Also there are differenced in particular years and industries. This suggests to use MCDM or DEA approach case by case since generalization seems impossible.

4 Conclusion

Thus we can conclude that applied selecting method is useful tool for determination of input and output measures when using DEA efficiency assessment. The MCDM approach is useful when the number of DMUs is smaller thus these methods are appropriate and easy to uses in such cases when ranking is the goal of the problem. However, for measuring efficiency it is necessary in such cases to use selecting model based on DEA which can in particular case help to identify more appropriate number of inputs and outputs. Investigation of changes in 4-year period suggests that generalization of measures of efficiency (inputs and outputs) may not be correct.

Further research could be focused on changes in time when using selecting models on same lists of measures but with different data. Testing under real conditions may reveal further insights in to efficiency measures.

Acknowledgements

Supported by the grant No. GA16–17810S „Selective measures in DEA: Theory and Applications“ of the Czech Science Foundation. The project is a result of European Social Fund project CZ.1.07/2.3.00/20.0296 and Student Grant Competition project No. SP2016/123.

References

[1] Banker, R. D., Charnes, A., and Cooper, W. W.: Models for estimation of technical and scale inefficiencies in data envelopment analysis. *Management Science* **30** (1984), 1078–1092.
[2] Charnes, A., Cooper, W. W., and Rhodes, E.: Measuring the efficiency of decision making units, *European Journal of Operations Research* **2** (1978), 429–444.
[3] Cooper, W. W., Seiford, L. M., and Tone, K.: *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. 2nd ed. Springer, Berlin, 2007.
[4] Deng, H., Yeh, C. H., and Willis, R. J.: Inter-company comparison using modified TOPSIS with objective weights. *Computers and Operations Research* **27** (2000), 963–973.
[5] Diakoulaki, D., Mavrotas, G., and Papayannakis, L.: Determining Objective Weights in Multiple Criteria Problems: The Critic Method. *Computers Operations Research* **22** (1995), 763–770.
[6] Shannon, C. E.: A Mathematical Theory of Communication. *Bell Systems Technology Journal* **27** (1948), 379–423.
[7] Toloo, M., and Tichý, T.: Two alternative approaches for selecting performance measures in data envelopment analysis. *Measurement*, **65** (2015), 29–40.
[8] Toloo, M., Barat, M., and Masoumzadeh, A.: Selective measures in data envelopment analysis. *Annals of Operations Research* **226** (2015), 623–642.

Granger Causality in Macroeconomics Panel Data

Lukáš Frýd [1]

Abstract. The shortage of macroeconomics time series poses a crucial problem which can be avoided by extending the time series including cross-sectional data to panel data. The first panel data models were built for microeconomic applying large number of cross-section units and short time series. With random sampling being an essential prerequisite for microeconomic panel data, the observations are independent. In macroeconomic data, we face a problem with non-random sampling and thus the possibility of dependent variables. This problem can be solved with the factor and spatial models.

In this paper, we focus on the factor models and answer to the question how energy supply influences GDP in the long term period. To test stationarity, co-integration and Granger causality of GDP, energy supply and capital formation, we follow methodologies by Pesaran and Ederhardt. It was found out that energy and capital have positive long-term effect on GDP and vice versa the GDP and capital have positive long-term effect on energy consumption.

Keywords: Cross-sectional dependence, panel Granger causality, energy nexus, panel cointegration

JEL classification: C44
AMS classification: 90C15

1 Introduction

Relationship between GDP and energy [2] was first studied by Kraft and Kraft [4]. The main problem of their work related to spurious regression [1]. In the early 90th researchers overcame the problem of spurious regression by testing stationarity and co-integration in time series. The interesting papers in this topic were elaborated by Adjaye [1], Soytas and Sari [13]. They found long-term relationship between GDP and energy consumption and proposed energy nexus as shown in Table 1. The nexus describes causal relationship between GDP and energy.

We can see four possible causality relationships. According to a conservative hypothesis, the energy shortage has a negative impact on economic growth. On the contrary, a growth hypothesis assumes that economic growth has impact on energy but not vice versa.

Hypothesis	Description
Neutral hypothesis	Non relationship E and GDP
Conservative hypothesis	E $\rightarrow$ GDP
Growth hypothesis	E $\leftarrow$ GDP
Bivariate hypothesis	E $\leftrightarrow$ GDP

Table 1 Eco-Energy relationship

The problem with these studies lies within the length of time series. The time series are relatively short which have a negative impact on power of the tests. Lee shows that extending time series by cross-section data increases the power of test [15].

[1]University of Economics in Prague,Department of Econometrics, Winston Churchill Square 4, 13067 Prague, Czech Republic, xfryl00@vse.cz

[2]we measure energy in oil equivalent

Examples of studies working with panel data are [7], [5], [6]. However, these studies are struggling with substantially different results probably arising from the existence of cross-sectional dependence. The violation of cross-sectional independence may lead to the inconsistency of fixed effect models.

We analyze long-term relationship between GDP and energy for upper-middle-income countries using models robust enough for cross-sectional dependence. This methodology allows for testing unit roots and co-integration in cross-sectional dependent data. In the last part, we set up the energy nexus via Granger causality test.

2 Model and data set

In our paper, we follow production function proposed by Stern [14]. The production function is expressed as:

$$GDP_{it} = \gamma_i + \alpha E_{it} + \beta CAP_{it} + \epsilon_{it} \tag{1}$$

where GDP is the natural logarithm of GDP per capita in constant price of 2005 in USD, E is energy use per capita in the natural logarithm and measure as oil equivalent in kg per capita, CAP is the natural logarithm of gross capital formation in constant price of 2005 in USD per capita. We approximate capital per capita with gross capital formation via Lee [5]. The index i represents cross-section units and index t represents time dimension of data. The data is intended for upper-middle-income economies ($4,126 to $12,735) and the time period is 1991-2014. [1]

Estimating equation (1) with standard panel techniques, we get inconsistent estimates in the case of cross-sectional dependence. From this reason, we follow Ederhardt and Teal [2] who used "unobserved common factor model" expressed as:

$$y_{it} = \beta_i^{'} x_{it} + \epsilon_{it} \tag{2}$$

where $i = 1, ..., N,\ t = 1, ..., T$

y_{it} is vector of GDP

x_{it} is matrix of independent variables $Energy$ and CAP

Ederhardt and Teal supposed that cross-sectional dependence is the consequence of common shocks. From this perspective, the cross-sectional dependence has a multi-factored structure:

$$\begin{aligned} \epsilon_{it} &= \alpha_i + \gamma_i^{'} f_t + u_{it} \\ x_{it} &= \pi_i + \rho_i f_t + v_{it} \end{aligned} \tag{3}$$

where f_t is vector of latent variables $m \times 1$, where m is number of factors. α_i and π_i are fixed effects, $u_{it} \sim iid(0, \sigma^2_{uit})$ and $v_{it} \sim iid(0, \sigma^2_{vit})$ are random variables for a specific country. A possible solution to obtain consistent estimate is the use of Common correlated effect model by Pesaran [10].

3 Methodology

3.1 Cross-sectional dependence

The existence of cross-sectional dependence leads to the inconsistent estimator and misleading first generation[2] unit root test. From this reason, we have to test the existence of cross-sectional dependence. We use Pesaran CD test in which the test statistic is expressed as:

$$CD = \sqrt{\frac{2}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \tag{4}$$

[1] http://data.worldbank.org/data-catalog/world-development-indicators

[2] the first generation does not assume cross-sectional dependence

where $\hat{\rho}_{ij}$ is estimated residuals correlations. The residuals origin in the equation (1) estimated by the fixed effect model. CD statistic has $N(0,1)$ distribution and H_0 represents cross-sectional independence and H_1 represents cross-sectional dependence.

3.2 Unit-root test

For unit-root test, we use augmented Dickey-Fuller test called Common Correlated Augmented Dickey-Fuller (CADF). This test is the part of the second generation panel unit root test. The test assumes the process:

$$\Delta y_{it} = \alpha_i + \beta_i y_{it-1} + c_i t + \sum_{j=1}^{p} d_{ij} \Delta y_{it-j} + \sum_{j=0}^{p} g_{ij} \Delta \bar{y}_{t-j} + \epsilon_{it} \tag{5}$$

where $\bar{y}_{t-1} = 1/N \sum_{i=1}^{N} y_{it-1}$. $H_0 : \beta_i = 0, i = 1, ..., N$ versus $H_1 : \beta_i < 0$ and t is time trend

Pesaran proposed test statistic:

$$CIPS = \frac{1}{N} \sum_{i=1}^{N} \tilde{t}_i \tag{6}$$

where $\tilde{t}_i$ is value of t-statistic from regression (5) for subject i. H_0 represents non-stationarity and H_1 represents stationarity. The critical values are tabulated in Pesaran [10].

3.3 Cointegration test

We test existence of co-integration via stationarity of residuals. For this purpose, we estimate equation (1) by Common Correlated Effect estimator (CCE) by Pesaran [11]. Pesaran has proposed two CCE estimators CCEMG (mean group) and CCEP (pool). Both estimators are robust for the existence of cross-sectional dependence. In our paper, we use CCEMG estimator because it allows heterogeneity in slope coefficients.

The main idea of CCEMG consists in the approximation of latent variables f_t by arithmetic mean:

$$\bar{y}_t = \bar{\alpha} + \bar{\beta}' \bar{x}_t + \bar{\gamma}' \bar{f}_t \tag{7}$$

which can be expressed as

$$\bar{f}_t = \bar{\gamma}^{-1} (\bar{y}_t - \bar{\alpha} + \bar{\beta}' \bar{x}_t) \tag{8}$$

where *bar* represents cross-sectional averages.

If $\bar{\gamma} \neq 0$ is true then the unobserved common factors are captured by cross-sectional means of $\mathbf{y}$ and $\mathbf{x}$ since $\bar{f}_t \xrightarrow{p} f_t$. [10].

$$GDP_{it} = \gamma_i + \alpha_i E_{it} + \beta_i CAP_{it} + \gamma_i \bar{E}_t + \delta_i C\bar{A}P_t + \theta_i G\bar{D}P_t + \epsilon_{it} \tag{9}$$

where $G\bar{D}P_t = 1/N \sum_{i=1}^{N} GDP_{it}$, $\bar{E}_t = 1/N \sum_{i=1}^{N} E_{it}$ and $C\bar{A}P_t = 1/N \sum_{i=1}^{N} CAP_{it}$ for all t. This method is robust for the structural changes and autocorrelation by Pesaran and Tosetti [12].

In the final step, we test the residuals from equation (9) for existence of unit root. We use the test from **section 3.2**. If we reject the null hypothesis that the residuals are stationary which points out the long run relationship among GDP, energy and capital.

3.4 Granger causality

As the main hypothesis is about causal relationship between GDP and energy, we use panel version of Vector Error Correction Model proposed by Pedroni [8] Panel Error-Corection Model (PECM). We test the significance of parameters from equation (11) and (12). The third equation is not important for our analysis. The parameters θ reveal short-run relationship among dependent variables and lags of GDP, energy and CAP in differences. For long-run relationships, we test parameters λ_1 and λ_2. Significant λ_1 and λ_2 with negative sign indicate the long-run relationship.

The equation (11) and (12) will be estimated with Pooled Mean Group estimator by Pesaran [8]. The optimal lags structure is determined by information criteria.

$$\begin{aligned}\Delta GDP_{it} =&\theta_{1i} + \sum_{k=1}^{p} \theta_{11ik} \Delta GDP_{it-k} + \sum_{k=1}^{p} \theta_{12ik} \Delta E_{it-k} \\ &+ \sum_{k=1}^{p} \theta_{13ik} \Delta CAP_{it-k} + \lambda_{1i} \epsilon_{it-1} + v_{1it}\end{aligned} \tag{10}$$

$$\begin{aligned}\Delta E_{it} =&\theta_{2i} + \sum_{k=1}^{p} \theta_{21ik} \Delta GDP_{it-k} + \sum_{k=1}^{p} \theta_{22ik} \Delta E_{it-k} + \\ &+ \sum_{k=1}^{p} \theta_{23ik} \Delta CAP_{it-k} + \lambda_{2i} \epsilon_{it-1} + v_{2it}\end{aligned} \tag{11}$$

4 Empirical part

Cross-sectional dependence

In the Table 2, we can see the result from Pesaran CD test for cross-sectional dependence. We can accept the alternative hypothesis about cross-sectional dependence. The GDP variable means that if we regress GDP on E and CAP by fixed effect estimator, the residuals will be highly correlated on average.

Variable	CD-test	p-value	corr
GDP	93.27	0.000	0.804
E	77.10	0.000	0.664
CAP	71.45	0.000	0.616

Table 2 Cross-Section dependence

Stationarity

As we reject hypothesis about non-existence of cross-sectional dependence, the second generation panel unit root test has to be used. The results of Pesaran Unit root test are shown in Table 3. We assume process from equation (5) with time trend t.

We cannot accept the alternative hypothesis about stationarity for all variables. In the next step, we test the first differences of variables. In this case we can reject the null hypothesis.[2] The final result is that GDP, Energy and CAP are I(1) process.

Cointegration

We have found that the variables are integrated of order 1. The next step is to test the existence of long term relationship. We estimate equation (1) by CCEMG model. The results are shown in Table 4. The

[2]test in not part of the paper

Variable	lags	Zt-bar	p-value
GDP	0	4.562	1.000
GDP	1	3.150	0.999
Energy	0	-0.694	0.244
Energy	1	1.139	0.873
CAP	0	0.901	0.816
CAP	1	-0.898	0.185

Table 3 Unit root test

crucial row is called Residuals. We can see that residuals are integrated of order 0[1] and we accept the hypothesis about existence of co-integration relationship.

The next interesting result is in row CD. In the same manner we use residuals from regression to test existence of cross-sectional dependence. The number 0.059 is value of Pesaran CD statistic from **section 3.1** and the number in brackets is p-value. CCEMG estimator thus provides cross-sectional independent residuals as opposed to the fixed effect estimator. The coefficients α and β from equation (1) are significant.

Variable	Coef.	Std. Error	z	p-value
Energy	0.397	0.164	2.41	0.016
CAP	0.706	0.133	5.29	0.000
Residuals	I(0)			
CD	0.059 (0.000)			

Table 4 CCEMG estimation

Granger causality

In the last part, we test hypothesis from Table 1. The estimated equation (11) and (12) by PEC model are displayed in the Table 5. [3]. In the brackets are standard errors and the optimal lags structure is one.

For the short-run period, we can see in the first row that energy does not have significant impact on GDP. The estimated coefficient is 0.026 and standard error is 0.065. Capital has significant impact on GDP.

In the second row, we can see that GDP does not have significant impact on energy but the capital has significant impact on energy. The coefficient value is 0.574 and standard error is 0.231. It can be interpreted as follows: one percentage point change in capital formation will lead to the 0.574 percentage point change in energy use. This value is relatively high comparied to results obtained from developed countries.

For long run relationship, we have to look at ECT coefficient. In both cases the sign is negative and significant. We can accept hypothesis that energy and capital have long run impact on GDP and that capital and GDP have significant long run relationship with energy.

	ΔGDP	$\Delta Energy$	ΔCAP	ECT
ΔGDP	-	.026 (0.065)	0.890 (0.168)	-0.043 (0.016)
$\Delta Energy$	0.134(0.200)	-	0.574 (0.231)	-0.156 (0.040)

Table 5 Granger causality

[1] we used CADF test

[3] equation with CAP as a dependent variable is not displayed

5 Conclusion

We analyzed Granger causality between GDP and energy for upper-middle-income economies from 1991-2014. The data were analyzed on cross-sectional dependence, unit root test and co-integration. Because of the existence cross-sectional dependence we used second generation unit root test and Common correlated effect estimator by Pesaran. We identified co-integrating relationship among variables. This conclusion is consistent with empiric results. We applied Granger-causality on panel error correction model and accepted the bivariate hypothesis for upper-midle income countries. The bivariate evidence is not typical form of causality in the literature. Narayan and Popp [7] argue that the data for developed countries supports the conservative hypothesis and the data for lower income countries supports growth hypothesis. One possible explanation of our results lies in the methodology. The econometric techniques in this paper do not distinguish cross-section units and give the same weight to all of them. The possible solution consists in the spatial econometrics for modelling relationship among the countries Formánek [3].

Acknowledgements

I gratefully acknowledge the support of this project by the research grant VŠE IGA F4/73/2016, Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] Asafu-Adjaye, J.: The relationship between energy consumption, energy prices and economic growth: time series evidence from Asian developing countries. *Energy Economics* **22** (2000), 615–625.

[2] Ederhardt, M., and Teal, F.: Common Factor Approach to Spatial Heterogeneity in Agricultural Productivity Analysis. MPRA Paper 15810, University Library of Munich, Germany (2009)

[3] Formánek, T., and Hušek, R.: The Czech Republic and its neighbors: Analysis of spatial macroeconomic dynamics. *Mathematical Methods in Economics 2015, MME 2015, Conference Proceedings*, Cheb, 2015, 190–195.

[4] Kraft, J., and Kraft, A.: On the relationship between energy and GNP. *Journal of Energy and Development* **3** (1978), 401–403.

[5] Lee, C. C.: Energy consumption and GDP in developing countries: a cointegrated panel analysis. *Energy Economics* **27** (2005), 415–427.

[6] Mohammadi, H., and Parvaresh, S.: Energy consumption and output: Evidence from a panel of 14 oil-exporting countries. *Energy Economics* **41** (2014), 41–46.

[7] Narayan, P. K., and Paresh, K., and Popp, S.: The energy consumption-real GDP nexus revisited: Empirical evidence from 93 countries. *Economic Modelling* **29** (2012), 303–308.

[8] Pedroni, P.: Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics* **61** (1999), 653–670.

[9] Pedroni, P.: Panel cointegration; asymptotic and finite sample properties of pooled time series tests with an application to the purchasing power parity hypothesis. *Econometric Theory* **20** (2004), 597–625.

[10] Pesaran, M. H.: General Diagnostic Tests for Cross Section Dependence in Panels. *Cambridge Working Papers in Economics 0435* Faculty of Economics, University of Cambridge (2004)

[11] Pesaran, M. H.: A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics* **22** (2007), 265–312.

[12] Pesaran, M. H., and Tosetti, E.: Large panels with common factors and spatial correlation. *Journal of Econometrics* **161** (2011), 182–202.

[13] Sari, R., and Soytas, U.: The growth of income and energy consumption in six developing countries. *Energy Policy* **35** (2007), 889–898.

[14] Stern, D. I.: Energy and economic growth in the U.S.A. *Energy Economics* **15** (1993), 137–150.

Incorrectly Posed Optimization Problems under Extremally Linear Equation Constraints

Mahmoud Gad [1], Josef Jablonský [2], Karel Zimmermann [3]

Abstract.

In this paper we propose an approach for solving (max, min)-separable linear equation systems. The concept of attainable set for (max, min)-separable linear equation systems will be introduced. Properties of the attainable sets will be studied in detail. The (max, min)-separable linear equation systems, in which the function of unknown variable occur only on one side, will be consider. In this case, attainable set means that "the set of all real vectors on the right hand side of linear separable equation, which make the separable linear equation system solvable". Optimization problem consisting in finding the nearest point of an attainable set to a given fixed point will be considered. An algorithm for solving the optimization problem will be proposed. Motivational example from the area of operations research, which shows possible applications of the optimization problem solved in the paper, will be given. Two numerical examples illustrating the proposed algorithm are included. Hints for further research will be briefly discussed in the conclusions.

Keywords: Attainable Sets, Incorrectly Posed Problems, (max,min)-Separable Equations.

JEL classification: C44, C14, C30
AMS classification: 90C47, 90C31

1 Introduction

Problems on algebraic structures, in which pairs of operations $(\max, +)$ or $(\max, \min)$ replace addition and multiplication of the classical linear algebra have appeared in the literature approximately since the sixties of the last century (see e.g. [1]). In this paper we will study so called attainable sets of such systems, i.e. the sets of right-hand sides, for which there exists a solution of the given system. Let us note that problems, the original formulation of which has no solution were called sometimes in the literature incorrectly posed problems (see e.g. [3]). Such problems are neither linear nor convex in usual algebraic sense. Such problems for $(\max, +)$-linear equation system were considered using a different approach in (see e.g. [2], [7]). Problems for $(\max, \min)$-linear equation and inequality system were considered using a different approach in (see e.g. [4], [5], [6]). Our purpose in this paper is to present an approach to incorrectly posed $(\max, \min)$-separable equation systems.

Let us introduce the following notations: $I = \{1, 2, \ldots, m\}, J = \{1, 2, \ldots, n\}$. Let A be a matrix with finite elements $a_{ij} \in R = (-\infty, +\infty), \forall i \in I, j \in J$, let $\alpha \wedge \beta \equiv \min(\alpha, \beta)$ for any $\alpha, \beta \in R$. Vector $A \otimes x \in R^m$ for $x = (x_1, \ldots, x_n)^T \in R^n$ will be defined as follows:

$$(A \otimes x)_i \equiv \max_{j \in J}(a_{ij} \wedge x_j) \forall i \in I.$$

The system of $(\max, \min)$-separable equations with right-hand side $b \in R^m$ is an equation system of the form

$$A \otimes x = b.$$

[1] Sohag University, Faculty of Science, Sohag 82524, Egypt; University of Economics, Faculty of Informatics and Statistics, Department of Econometrics, W. Churchill Sq. 1938/4, 130 67 Prague 3, Czech Republic, Mahmoud_attya_or@yahoo.com.
[2] University of Economics, Faculty of Informatics and Statistics, Department of Econometrics, W. Churchill Sq. 1938/4, 130 67 Prague 3, Czech Republic, jablon@vse.cz.
[3] Charles University, Faculty of Mathematics and Physics, Malostranské nám. 2/25, 118 00 Praha 1, Czech Republic, Zimmermann@seznam.cz.

The set of all solutions of the system will be denoted $M(b)$, (i.e. $M(b) = \{x \in R^n \ ; \ A \otimes x = b\}$.)

Definition 1. Set $R(A) \equiv \{b \in R^m \exists \ x \in R^n$ such that $A \otimes x = b\}$ is called attainable set of matrix A.

In what follows we will solve the following optimization problem:

$$\left|\left|b - \hat{b}\right|\right| = \max_{i \in I} \left|b_i - \hat{b}_i\right| \longrightarrow \min \quad \text{subject to} \qquad b \in R(A) \tag{1}$$

The optimal solution of problem (1) will be denoted b^{opt}. Let us note that if $\hat{b} \in R(A)$, it is evidently $b^{opt} = \hat{b}$. Therefore we will assume in what follows that $\hat{b} \notin R(A)$. Before investigating properties of attainable sets and analysis of problem solution, we will bring an example, which shows one possible application, which leads to solving the system given above.

Example 1. Let us assume that m places $i \in I^{(1)} \equiv \{1, \ 2, \ \ldots, \ m\}$ are connected with n places $j \in J \equiv \{1, \ 2, \ \ldots, \ n\}$ by roads with given capacities. The capacity of the road connecting place i with place j is equal to $a_{ij} \in R$. We have to extend for all $i \in I, \ j \in J$ the road between i and j by a road connecting j with a terminal place T and choose an appropriate capacity x_j for this road. If a capacity x_j is chosen, then the capacity of the road from i to T via j is equal to $a_{ij} \wedge x_j = \min(a_{ij}, x_j)$. We require that the connection between places i and T is for at least one j equal to a given number $b_i \in R$ and the chosen capacity x_j lies in a given finite interval i.e. $x_j \in [\underline{x}_j, \overline{x}_j]$, where $\underline{x}_j, \overline{x}_j \in R$ are given finite numbers. Therefore feasible vectors of capacities $x = (x_1, \ x_2, \ \ldots, \ x_n)$ (i.e. the vectors, the components of which are capacities x_j having the required properties) must satisfy an equation system of the form $A \otimes x = b$.

2 Properties of Attainable Sets and Analysis of the Problem

In this section we will study in more detail some properties of attainable sets and analyze the problem (1). The properties are formulated as the following three lemmas.

Lemma 2.1. Set $R(A)$ has the maximum element, i.e. an element $b^{\max} \in R(A)$ such that $b \leq b^{\max} \ \forall \ b \in R(A)$.

Proof. Let $\alpha_i = \max_{j \in J} a_{ij} \ \ \forall i \in I$. Let $x \in R^n$ be arbitrarily chosen. Then $a_{ij} \wedge x_j \leq a_{ij}$ for all $i \in J, \ j \in J$. Therefore for any $i \in I$ we obtain that

$$\max_{j \in J}(a_{ij} \wedge x_j) = \max_{j \in J} a_{ij} = \alpha_i.$$

Therefore if we set $b_i^{\max} = \alpha_i \ \ \forall \ \ i \in I$, then $b^{\max} \in R(A)$, since e.g. if $\hat{x} \in R^n$ and $\hat{x}_j \geq \max_{i \in I} \alpha_i$ we have $\max_{j \in J}(a_{ij} \wedge \hat{x}_j) = \alpha_i = b_i^{\max}$. For an arbitrary $b \in R(A)$ there exists $x \in R^n$ such that $b = A \otimes x \leq A \otimes \hat{x} = b^{\max}$, so that $b^{\max}$ is the maximum element of $R(A)$, this ends the proof.

Lemma 2.2. Let $b \in R^m$, $I_j^{>} = \{i \in I \ ; \ a_{ij} > b_i\} \ \ \forall \ \ j \in J$. Let $M(b) = \{x \in R^n \ ; \ A \otimes x = b\}$ be nonempty. Let vector $x(b) \in R^n$ be defined as follows:

$$x_j(b) = \min_{i \in I_j^{>}} b_i \ \ \forall \ \ j \in J \text{ if } I_j^{>} \neq \emptyset.$$

We set the minimum equal to infinity if $I_j^{>} = \emptyset$. Then $x(b)$ is the maximum element of set $M(b)$.

Proof. Let us note that if $x \in M(b)$, then it must be $a_{ij} \wedge x_j \leq b_i$ for all $i \in I, \ j \in J$. Therefore it must be $x \leq x(b) \ \ \forall x \in M(b)$ so that $x(b)$ is the upper bound for elements of $M(b)$. It remains to prove that if set $M(b)$ is nonempty it must be $x(b) \in M(b)$. Let us set

$$S_j(x_j) \equiv \{k \in I \ ; \ a_{kj} \wedge x_j = b_k\} \ \ \forall j \in J.$$

If $I_j^{>} \neq \emptyset$, then

$$S_j(x_j(b)) = \{k \in I \ ; \ x_j(b) = b_k = \min_{i \in I_j^{>}}(b_i)\}.$$

If $I_j^{>} = \emptyset$, then $x_j(b) = \infty$ and $S_j(x_j(b)) = \{k \in I \ ; \ a_{kj} = b_k\}$. We will show further that

$$x(b) \in M(b) \iff \bigcup_{j \in J} S_j(x_j(b)) = I.$$

Really if $\bigcup_{j \in J} S_j(x_j(b)) = I$ and $p \in I$ is arbitrary, then there exists index $j(p) \in J$ such that $p \in S_{j(p)}(x_{j(p)}(b))$ and therefore $a_{pj} \wedge x_j(b) \leq b_p$ for all $j \in J$ and $a_{pj(p)} \wedge x_{j(p)}(b) = b_p$ so that $\max_{j \in J}(a_{pj} \wedge x_j(b)) = b_p$. Since p was arbitrary, we obtain that $x(b) \in M(b)$. To prove the opposite implication let us assume that $\bigcup_{j \in J} S_j(x_j(b)) \neq I$ so that there exists index $i_0 \in I$ such that $i_0 \notin \bigcup_{j \in J} S_j(x_j(b))$ and therefore $a_{i_0 j} \wedge x_j(b) \neq b_{i_0} \quad \forall j \in J$ and therefore $\max_{j \in J}(a_{i_0 j} \wedge x_j(b)) \neq b_{i_0}$ and thus $x(b) \notin M(b)$.
Let us note that if $x_j \leq x_j(b)$ for any $j \in J$, then $S_j(x_j) \subseteq S_j(x_j(b))$. Therefore if $\bigcup_{j \in J} S_j(x_j(b)) \subset I$, then for any $x \leq x(b)$ we have

$$\bigcup_{j \in J} S_j(x_j) \subseteq \bigcup_{j \in J} S_j(x_j(b)) \subset I$$

and thus $M(b) = \emptyset$, since all elements of $M(b)$ must satisfy the inequality $x \leq x(b)$. It follows that

$$M(b) \neq \emptyset \iff x(b) \in M(b).$$

In other words if $M(b) \neq \emptyset$, then $x(b) \in M(b)$ and $x \leq x(b)$ for all $x \in M(b)$, so that $x(b)$ in the maximum element of $M(b)$, what was to be proved.

Lemma 2.3. Let $b^{\max}$ be the maximum element of $R(A)$, $\hat{b} \in R^m$ such that $\hat{b}_p \geq b_p^{\max}$ for some $p \in I$, b an arbitrary element of $R(A)$. Then $\left| b_p - \hat{b}_p \right| \geq \left| b_p^{\max} - \hat{b}_p \right|$.

3 Algorithm - A Parametric Version for Problem Solution

In what follows we will replace problem (1) with the following parametric optimization problems:

$$\text{Minimize} \quad t \quad \text{subject to} \quad \left\| b - \hat{b} \right\| \leq t, \quad b \in R(A) \tag{2}$$

$$\text{Minimize} \quad t \quad \text{subject to} \quad \hat{b}_i - t \leq \max_{j \in J}(a_{ij} \wedge x_j) \leq \hat{b}_i + t, \ \ \forall \ \ i \in I. \tag{3}$$

Let $M(t)$ denote the set of feasible solutions of problem (2). We have then

$$M(t) = \{x \ ; \ \hat{b}_i - t \leq x_j \leq \hat{b}_i + t, \ \forall \ i \in I\}.$$

And let us set for all $i \in I, \ j \in J$.

$$T_{ij}(t) \equiv \{x_j \mid \hat{b}_i - t \leq a_{ij} \wedge x_j \leq \hat{b}_i + t\}.$$

Note that $\max_{j \in J}(a_{ij} \wedge x_j) \leq \hat{b}_i + t, \ \ \forall \ \ i \in I,$ implies that for each fixed $j \in J$ it is $a_{ij} \wedge x_j \leq \hat{b}_i + t, \ \ \forall \ \ i \in I,$ so that for each fixed $j \in J$ and t it must be

$$x_j \leq x_j(\hat{b} + t) \equiv min_{i \in I_j(t)}(\hat{b}_i + t),$$

where $I_j(t) \equiv \{i \in I \mid a_{ij} > \hat{b}_i + t\}$, and we set the minimum equal to infinity if $I_j^{>}(t) = \emptyset$. Let us note that if $a_{ij} > \hat{b}_i + t$ i.e. $t < a_{ij} - \hat{b}_i$, then $\hat{b}_i + t$ is the upper bound for $x_j \in T_{ij}(t)$ and if $t \geq a_{ij} - \hat{b}_i$, then $a_{ij} \leq \hat{b}_i + t$ so that also $a_{ij} \wedge x_j \leq \hat{b}_i + t$ and $\hat{b}_i + t$ is no more an upper bound for x_j, i.e. the upper bound for x_j is higher. Let us note further that if $T_{ij}(t) \neq \emptyset$, then it must be fulfilled two inequalities

$$a_{ij} \geq \hat{b}_i - t \ \text{ and } \ \hat{b}_i - t \leq x_j(\hat{b}_i + t)$$

If either $a_{ij} < \hat{b}_i - t$ or $\hat{b}_i - t > x_j(\hat{b}_i + t)$, then $T_{ij}(t)$ is empty.
We will find minimum value of t, for which the inequalities $a_{ij} \geq \hat{b}_i - t$ and $\hat{b}_i - t \leq x_j(\hat{b}_i + t)$ hold. The minimum value of t, for which $a_{ij} \geq \hat{b}_i - t$ holds is evidently $\tau_{ij}^{(1)} \equiv \hat{b}_i - a_{ij}$. To define the minimal value of t, for which $\hat{b}_i - t \leq x_j(\hat{b}_i + t)$ holds, we will investigate $x_j(\hat{b}_i + t)$ as a function of t. We have for any fixed $j \in J$ and $t \geq 0$:

$$x_j(\hat{b} + t) = \min_{i \in I_j(t)}(\hat{b}_i + t) = \hat{b}_{k(j,t)} + t,$$

where we set $x_j(\hat{b}+t)=\infty$, if $I_j(t)=\emptyset$. Note that $x_j(\hat{b}+t)=\infty$, for all $t\geq \max_{i\in I}(a_{ij}-\hat{b}_i)$. We will consider therefore only values $t\leq \max_{i\in I}(a_{ij}-\hat{b}_i)$. Let us set
$I_j^{(1)}\equiv\left\{k \mid \max_{k\in I}(a_{ij}-\hat{b}_i)=a_{kj}-\hat{b}_k=\alpha_j^{(1)}\right\}$, $I_j^{(2)}\equiv\left\{k \mid \max_{k\in I\setminus I_j^{(1)}}(a_{ij}-\hat{b}_i)=a_{kj}-\hat{b}_k=\alpha_j^{(2)}\right\}$,
$I_j^{(3)}\equiv\left\{k \mid \max_{k\in I\setminus(I_j^{(1)}\cup I_j^{(1)})}(a_{ij}-\hat{b}_i)=a_{kj}-\hat{b}_k=\alpha_j^{(3)}\right\}$, and so on ....
$I_j^{(p)}\equiv\left\{k \mid \max_{k\in I\setminus\bigcup_{h=1}^{p-1}I_j^{(h)}}(a_{ij}-\hat{b}_i)=a_{kj}-\hat{b}_k=\alpha_j^{(p)}\right\}$, where $I\setminus\bigcup_{h=1}^{p-1}I_j^{(h)}\neq\emptyset$, and $\bigcup_{h=1}^{p}I_j^{(h)}=I$.
Values $\alpha_j^{(p)},\ h=1,\cdots,p$ are therefore different values, which occur in the set $a_{ij}-\hat{b}_i,\ i\in I$ and holds $\alpha_j^{(1)}>\alpha_j^{(2)}>\cdots>\alpha_j^{(p)},\ 1\leq p\leq m$.
Having determined values $\alpha_j^{(1)},\ \alpha_j^{(2)},\cdots,\ \alpha_j^{(p)}$, we can find explicitly $I_j(t)$ in dependence of t : $I_j(t)=\emptyset$ if $t\geq\alpha_j^{(1)}$, $I_j(t)=I_j^{(1)}$ if $\alpha_j^{(2)}\leq t<\alpha_j^{(1)}$, $I_j(t)=I_j^{(1)}\cup I_j^{(2)}$ if $\alpha_j^{(3)}\leq t<\alpha_j^{(2)}$, and so on .... $I_j(t)=\bigcup_{h=1}^{p-1}I_j^{(h)}$ if $\alpha_j^{(p)}\leq t<\alpha_j^{(p-1)}$ and $I_j(t)=\bigcup_{h=1}^{p}I_j^{(h)}=I$ if $t<\alpha_j^{(p)}$.
Now we can find the explicit form of $x_j(\hat{b}+t)$ as a function of t : $x_j(\hat{b}+t)=\infty$ where $t\geq\alpha_j^{(1)}$, $x_j(\hat{b}+t)=min_{i\in I_j(t)}\hat{b}_i+t=\hat{b}_{k(j,t)}+t$ where $\alpha_j^{(2)}\leq t<\alpha_j^{(1)}$, and $k(j,t)\in I_j^{(1)}$ $x_j(\hat{b}+t)=min_{i\in I_j(t)}\hat{b}_i+t=\hat{b}_{k(j,t)}+t$ where $\alpha_j^{(3)}\leq t<\alpha_j^{(2)}$, and $k(j,t)\in I_j^{(1)}\cup I_j^{(2)}$ and so on .... $x_j(\hat{b}+t)=min_{i\in I_j(t)}\hat{b}_i+t=\hat{b}_{k(j,t)}+t$ where $\alpha_j^{(p)}\leq t<\alpha_j^{(p-1)}$, and $k(j,t)\in\bigcup_{h=1}^{p-1}I_j^{(h)}$,
$x_j(\hat{b}+t)=min_{i\in I_j(t)}\hat{b}_i+t=\hat{b}_{k(j,t)}+t$ where $0\leq t<\alpha_j^{(p)}$, and $k(j,t)\in\bigcup_{h=1}^{p}I_j^{(h)}=I$
The following numerical example enlightens the definition of $I_j^{(k)},\ k=1,...,p$.

Example 2. Let $m=5$, $(a_{1j},a_{2j},a_{3j},a_{4j},a_{5j})^T=(5,8,8,16,20)^T$, $\hat{b}=(3,5,5,12,14)$ so that $(a_{1j}-\hat{b}_1,a_{2j}-\hat{b}_2,a_{3j}-\hat{b}_3,a_{4j}-\hat{b}_4,a_{5j}-\hat{b}_5)^T=(2,3,3,4,6)^T$, and we obtain $p=4$ and $I_j^{(1)}=\{5\}$, with $\alpha_j^{(1)}=6$, $I_j^{(2)}=\{4\}$, with $\alpha_j^{(2)}=4$ $I_j^{(3)}=\{2,3\}$, with $\alpha_j^{(3)}=3$ $I_j^{(4)}=\{1\}$, with $\alpha_j^{(4)}=2$.
Let us find $I_j(t)$ and $x_j(\hat{b}+t)$ for the numerical data of this example, we obtain

$$I_j(t)=\emptyset,\quad x_j(\hat{b}+t)=\infty\qquad\text{if}\qquad t\geq 6,$$

$$I_j(t)=\{5\},\quad x_j(\hat{b}+t)=14+t\qquad\text{if}\qquad 4\leq t<6,$$

$$I_j(t)=\{4,5\},\quad x_j(\hat{b}+t)=12+t\qquad\text{if}\qquad 3\leq t<4,$$

$$I_j(t)=\{2,3,4,5\},\quad x_j(\hat{b}+t)=5+t\qquad\text{if}\qquad 2\leq t<3,$$

$$I_j(t)=\{1,2,3,4,5\}=I,\quad x_j(\hat{b}+t)=3+t\qquad\text{if}\qquad 0\leq t<2.$$

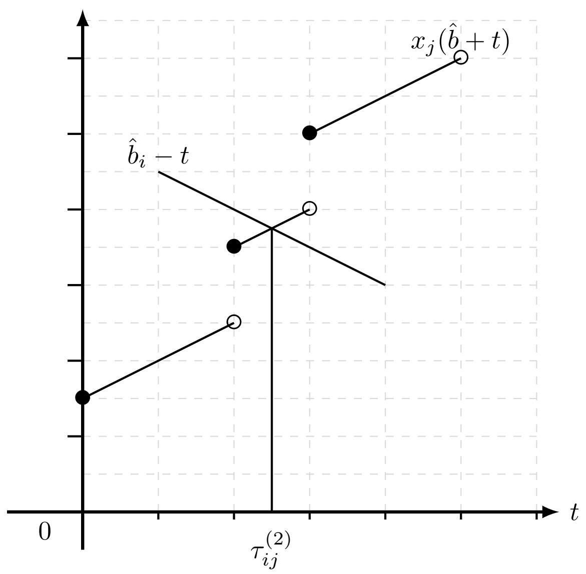


Figure 1 a: Graph of $x_j(\hat{b}+t)$ and $\hat{b}_i-t$

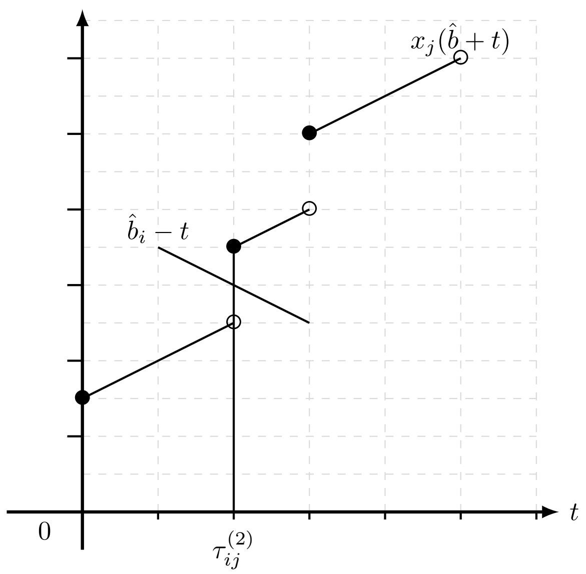


Figure 1 b: Graph of $x_j(\hat{b}+t)$ and $\hat{b}_i-t$

It follows that $x_j(\hat{b}+t)$ is for each $j \in J$ a strictly increasing, partially continuous function of t with at most m discontinuity points, in which it is continuous from above. The explicit expression of $x_j(\hat{b}+t)$ makes possible to find $\tau_{ij}^{(2)}$ such that $\hat{b}_i - \tau_{ij}^{(2)} \leq x_j(\hat{b}+\tau_{ij}^{(2)})$ and $\hat{b}_i - t > x_j(\hat{b}+t)$ if $t < \tau_{ij}^{(2)}$. Figures 1 a and 1 b show the following two possibilities:
Possibility (1) as in Figure 1 a: in this case $\hat{b}_i - \tau_{ij}^{(2)} = x_j(\hat{b}+\tau_{ij}^{(2)})$.
Possibility (2) as in Figure 1 b: in this case $\hat{b}_i - \tau_{ij}^{(2)} < x_j(\hat{b}+\tau_{ij}^{(2)})$ and $\hat{b}_i - t > x_j(\hat{b}+t)$ if $t < \tau_{ij}^{(2)}$.
New Let us set

$$\tau_{ij} \equiv \max(\tau_{ij}^{(1)}, \tau_{ij}^{(2)}),$$

Since we obtained that

$$T_{ij}(t) \neq \emptyset \text{ if and only if } t \geq \tau_{ij}.$$

In other words τ_{ij} is the optimal solution of the minimization problem

$$\text{Minimize} \quad t \quad \text{subject to} \quad T_{ij}(t) \neq \emptyset.$$

Note that it follows from Lemma 2.4 in [6] that for any fixed t ,

$$M(t) \neq \emptyset \quad \text{if and only if} \quad \forall\, i \in I\ \exists\, j(i) \in J \quad \text{such that} \quad T_{ij(i)}(t) \neq \emptyset \tag{4}$$

Which leads us to provide the next lemma.

Lemma 3.1. Let us set $I_j^{>} = \{i \in I\ ;\ a_{ij} > \hat{b}_i + t\}$, $T_{ij}(t) = \{x_j\ ;\ \hat{b}_i - t \leq a_{ij} \wedge x_j \leq \min_{k \in I_j^{>}} \hat{b}_k + t = \hat{b}_{k(j)} + t\}$ for any $i \in I, j \in J$. Then

$$M(t) \neq \emptyset \iff \forall i \in I\ \exists\ j(i) \in J, \quad \text{such that} \quad T_{ij(i)}(t) \neq \emptyset \ \ \& \ \ x \leq x(\hat{b}+t),$$

where we set $\hat{b}+t = (\hat{b}_1+t, \cdots, \hat{b}_m+t)$.

Proof. Let t be arbitrary and fixed. Since according to Lemma 2.2 $x(\hat{b}+t)$ is the maximum element of set $M(t)$, then $M(t) \neq \emptyset$ if and only if $x_j \leq x_j(\hat{b}+t)\ \forall\ x_j \in T_{ij}(t)$ or in other words the upper bound of $T_{ij}(t)$ must not be violated if x is in $M(t)$. Let us assume know that $x \in M(t)$ and at the same time there exists index $k \in I$ such that $T_{kj}(t) = \emptyset\ \ \forall\ j \in J$. Since $x \in M(t)$, it must be $x_j \leq x_j(\hat{b}+t)$ for all $j \in J$ and therefore if $T_{kj}(t)\ \forall\ j \in J$ is empty, we have $a_{kj} \wedge x_j < \hat{b}_k - t\ \ \forall\ j \in J$ and therefore $\max_{j \in J}(a_{kj} \wedge x_j) < \hat{b}_k - t$ and $x \notin M(t)$, which is a cotradiction. To prove the oppsite assertion, we assume that for each $i \in I$, there exists at least one index $j(i) \in J$ such that $T_{ij(i)}(t) \neq \emptyset$ and $x \leq x(\hat{b}+t)$. We will prove that $M(t) \neq \emptyset$. In this case it is e.g. $\max_{j \in J}(a_{ij} \wedge x_j(\hat{b}+t)) \geq \hat{b}_i - t$. Since $x(\hat{b}+t)$ evidently satisfies the upper bound condition $x_j \leq x(\hat{b}+t)$, we obtain that $x(\hat{b}+t) \in M(t)$ and thus $M(t) \neq \emptyset$, which completes the proof.

As a consequeance of (4) and Lemma 3.1 we obtain:

$$M(t) \neq \emptyset \quad \text{if and only if} \quad t \geq \tau \equiv \max_{i \in I} \min_{j \in J} \tau_{ij}.$$

Therefore the necessary and sufficient condition of Lemma 3.1 will be satisfied for $t \geq \max_{i \in I} \min_{j \in J} \tau_{ij}$. Therefore the optimal solution t^{opt} of problem (2) is

$$t^{opt} = \max_{i \in I} \min_{j \in J} \tau_{ij}.$$

We will illustrate the theoretical result by a small numerical example.

Example 3. Let $m = n = 3$, $\hat{b} = (0, 1, 1)^T$,

$$A = \begin{pmatrix} 3 & 1 & 5 \\ 4 & 4 & 6 \\ 7 & 7 & 3 \end{pmatrix}$$

In this case $b^{\max} = (5, 6, 7)^T$ and
$\min_{j \in J} \tau_{1j} = \min(0, 0, 0) = 0$,

$\min_{j\in J} \tau_{2j} = \min(1/2, 1/2, 1/2) = 1/2$,
$\min_{j\in J} \tau_{3j} = \min(1/2, 1/2, 1/2) = 1/2$,
so that $t^{opt} = \max_{i\in I} \min_{j\in J} \tau_{ij} = \max(0, 1/2, 1/2) = 1/2$ and the optimal solution of PROBLEM I is: $b^{opt} = A \otimes (0, 1/2, 1/2)^T = (1/2, 1/2, 1/2)^T$. Note that since for $\tilde{x} = (1/2, 1/2, 1/2)^T$ we have $A \otimes \tilde{x} = b^{opt}$, we obtain that $b^{opt} \in R(A)$. The optimal value of the objective function of problem (1) is $\left\| b^{opt} - \hat{b} \right\| = \max(1/2, 1/2, 1/2) = 1/2$.

4 Conclusions

In this paper the idea that arises in connection with practical applications, which are described by (max, min)-separable linear equation systems, is introduced. The problem what to do if the given (max, min)-separable linear equation systems has no feasible solution is analysed. We have to modify the original system (i.e. to modify its input coefficients) in such a way that the new problem has a solution. In this situation it is natural trying to modify the problems in such a way that the original goals of the given system (e.g. bounds on costs or arrival times) will be violated as little as possible. We introduced a technique through it we can modify the values on the right hand side of the systems until the system is solvable, and it will be violated as little as possible. In the future work we will introduce and propose an algorithm for solving problem (1) by using Threshold Version Technique. Moreover, we will try to apply the technique introduced here for real life problems and introduce a new technique that allows modify the values of the left hand side coefficients in order the system has a feasible solution. Another possibility to make the research closer to practical requirements would be considering stochastic or interval input coefficients of the $(\max, \min)-$linear systems.

Acknowledgements

The research is supported by the Cultural Affairs and Missions Sector - Egyptian Ministry of Higher Education, Sohag University in Sohag - Egypt, Internal Grant Agency of the Faculty of Informatics and Statistics, University of Economics, Prague project No. F4/62/2015, and Grant No. #14-02424S of the Grant Agency Czech Republic.

References

[1] Butkovič, P.: *Max-linear Systems: Theory and Algorithms.* Springer Monographs in Mathematics & Springer-Verlag, London - Dodrecht - Heudelberg - New York, 2010.

[2] Cuninghame-Green, R.A.: *Minimax Algebra, Lecture Notes in Economics and Mathematical Systems.* **166**, Springer-Verlag, Berlin, 1979.

[3] Eremin, I. I., Mazurov, V. D., and Astaf́ev, N. N.: Linear Inequalities in Mathematical Programming and Pattern Recognition, *Ukr. Math. J.* **40**, 3 (1988), 243–251. Translated from *Ukr. Mat. Zh.* **40**, 3 (1988), 288–297.

[4] Gad, M.: Optimization Problems under One-Sided $(\max, \min)$-Linear Equality Constraints, *WDS'12 Proceedings of Contributed Papers Part I*, (2012), 13–19. 21^{st} *Annual Student Conference, Week of Doctoral Students Charles University*, Prague, May 29 - June 1, 2012.

[5] Gad, M.: Optimization Problems under Two-Sided $(\max, \min)$-Linear Inequalities Constraints, *Academic Coordination Centre "ACC" JOURNAL* **18**, 4 (2012), 84–92. *International Conference Presentation of Mathematics '12 "Conference ICPM'12", Liberec*, June 21-22, 2012.

[6] Gavalec, M. , Gad, M. and Zimmermann, K.: Optimization Problems under $(\max, \min)$-Linear Equations and / or Inequality Constraints, *Journal of Mathematical Sciences* **193**, 5 (2013), 645–658. Translated from Russian Journal *Fundamentalnaya i Prikladnaya Matematika(Fundamental and Applied Mathematics)* **17**, 6 (2012), 3–21.

[7] Zimmermann, K. and Gad, M.: Optimization Problems under One-Sided $(\max, +)$-Linear Constraints, *International Conference Presentation of Mathematics '11 "ICPM'11, Liberec"*, October 20 - 21, 2011, 159–165.

One Type of Activity Synchronization Problems

Martin Gavalec[1], Karel Zimmermann[2]

Abstract. In this paper we will consider one type of equation systems, in each equation of which (max,plus)-linear and/or (min,plus)-linear functions occur. Such functions are expressed as the maximum (minimum) of a finite number of simple linear functions of one variable. Each equation of the system has a (max,plus)-linear function on one side and/or (min,plus)-linear function on the other side. So called one-sided systems with (max,plus)- or (min,plus)-linear functions on one side only and a constant on the other side were studied in the literature approximately since the sixties of the last century (see e.g. publications by N. Vorobjov or R.A.Cunninghame-Green). The systems we are going to study in this paper have (max,plus)- linear functions on one side and (min,plus)-linear function on the other side. We will call such systems two-sided. We will present method for finding a solution of such systems in case the systems are solvable. Such equation systems can be applied for synchronizing release and completion times of activities with given processing times. A motivating example will be presented. Results are illustrated by a small numerical example.

Keywords: Two-sided (max,plus)/(min,plus)-linear equation systems, Activity synchronisation.

JEL classification: C44, C41
AMS classification: 90B35, 90B20

1 Introduction

In this paper we will consider one type of equation systems, in each equation of which $(\max, +)$-linear and/or $(\min, +)$-linear functions occur. Such functions are expressed as the maximum (minimum) of a finite number of simple linear functions of one variable, more exactly, if $x = (x_1, \ldots, x_n) \in R^n$ and $\alpha_j \in R,\ j = 1, \ldots, n$, then function $f(x) = \max_{1 \leq j \leq n}(\alpha_j + x_j)$ is called $(\max, +)$-linear and similarly $(\min, +)$-linear function is defined. Each equation of the system has a $(\max, +)$-linear function on one side and/or $(\min, +)$-linear function on the other side. So called one-sided systems with $(\max, +)$- or $(\min, +)$-linear functions on one side only and a constant on the other side were studied in the literature approximately since the sixties of the last century (see e.g. [3], [5]). A relatively up to date state of art of the theory of such systems can be found in [2]. The systems we are going to study in this paper have $(\max, +)$- linear functions on one side and $(\min, +)$- or $(\max, +)$-linear function on the other side, we will call them two-sided. We will present method for finding a solution of such systems in case the systems are solvable. Such equation systems can be applied for synchronizing release and completion times of activities with given processing times. In the next section, we present a motivating example of such application.

2 Motivation

The following example shows a motivation for the research presented in this contribution.

Example 1. We assume that passengers should be transported from places $P_j,\ j \in J \equiv \{1, \ldots, n\}$ to destinations $D_k,\ k \in J$ via transit points $T_i,\ i \in I \equiv \{1, \ldots, m\}$. The passengers must change for

[1]University of Hradec Kralove/Faculty of Informatics and Management, Deaprtment of Information Technologies, addressRokitanova 62, Hradec Kralove,Czech Republic, e-mail: Martin.Gavalec@uhk.cz
[2]Charles University Prague/Faculty of Mathematics and Physics, Department Applied Mathematics, Malostranske namesti 25, 118 00 Praha 1, Czech Republic, e-mail: Karel.Zimmermann@mff.cuni.cz

another means of transport at the transit points. The times necessary to cover the distances between P_j and T_i are equal to $a_{ij} > 0$ and the times necessary for transfer between T_i and D_k are equal to $c_{ik} > 0$. The components of $x = (x_1, \ldots, x_n)$ denote departure times at $P_j,\ j \in J$ and the components of $y = (y_1, \ldots, y_n)$ denote the arrival times to destinations $D_k,\ k \in J$. We assume that the passengers can continue their journey from T_i to D_k only after all passengers who must change at T_i had arrived to T_i so that nobody will miss the connection at any transit point. The departure times z_k from T_i to D_k are under our assumption equal to $z_k = y_k - c_{ik}$. Feasible departure times from a fixed transit point T_i to a fixed D_k must satisfy the inequality

$$\max_{j \in J}(x_j + a_{ij}) \le z_k \ \forall k \in J$$

so that

$$\max_{j \in J}(x_j + a_{ij}) \le \min_{k \in J} z_k = \min_{k \in J}(y_k - c_{ik}).$$

It arises the question whether there exists a pair $x,\ y \in R^n$ satisfying the equation system

$$\max_{j \in J}(x_j + a_{ij}) = \min_{k \in J}(y_k - c_{ik}),\ i \in I.$$

If $x,\ y$ satisfy the equation system, no unnecessary delays between arrivals and departures in the transit points will occur.

Note that we can assume that the number of components is on both sides of the inequalities and equations is equal, since the missing variables can be added with coefficients $-\infty$ on the left and ∞ on the right hand side. In this paper we will investigate a special solution of the equation system mentioned above, in which $y = x + \alpha$ for some$\alpha \in R^n$. If the transportation of passengers is periodically repeated, such solution synchronizes the arrivals and departures in such a way that the repeating arrival and departure time vectors are equal to $x,\ x + \alpha,\ x + 2\alpha,\ \ldots$ etc.

3 Problem Formulation, Preliminary Results

Let us introduce the following notations:

R is the set of real numbers, $I = \{1,\ \ldots\ m\}$, $J = \{1,\ \ldots\ n\}$, $A,\ B$ are matrices with elements $a_{ij}, b_{ij} \in R\ \forall i \in I,\ j \in J$

$$(A\ o\ x)_i = \max_{j \in J}(a_{ij} + x_j),\ (B\ o'\ y)_i = \min_{j \in J}(b_{ij} + y_j),\ i \in I,$$

$x^T = (x_1,\ \ldots,\ x_n)$, $A\ o\ x = ((A\ o\ x)_1,\ \ldots,\ (A\ o\ x)_n)^T$, $B\ o'\ y = ((B\ o'\ y)_1,\ \ \ldots,\ (B\ o'\ y)_n))^T)$ (superscript T denotes transposition).

We will study equation system $A\ o\ x = B\ o'\ y$, which was applied in the motivating example in the preceding section. Especially we will be intrerested in so called steady state solutions, in which $y = x + \alpha$ for some $\alpha \in R$. Let us note that system

$$A\ o\ x = B\ o'\ (x + \alpha)$$

is equivalent with system

$$A\ o\ x = \tilde{B}\ o'\ x,$$

where $\tilde{B} = B + \alpha$, $(B + \alpha)_{ik} = b_{ik} + \alpha$. Therefore we will first consider equation system, in which $y = x$, i.e. system

$$\max_{j \in J}(a_{ij} + x_j) = \min_{k \in J}(b_{ik} + x_k),\ i \in I \tag{1}$$

or using the matrix-vector notation The set of all solutions x of system (3) will be denoted $M(A, B)$. Our aim is to investigate properties of the set $M(A, B)$ and find one of its elements if $M(A, B) \neq \emptyset$. For this purpose, we will use some preliminary definitions and results, the proofs of which can be found in the literature (see [2], [3], [4]).

Let

$$M_1(A, b) = \{x \in R^n \mid A\ o\ x = b\},\ M_2(B, b) = (y \in R^n \mid B\ o'\ y = b),$$

$$x_j(A,b) = \min_{k \in I}(b_k - a_{kj}),\ \hat{y}_j(B,b) = \max_{k \in I}(b_k - b_{kj}),\ j \in J,$$

$x(A,b)$ can be expressed using the matrix-vector notation as follows:

$$x(A,b) = -A^T\ o'\ b \texttt{ or } x(A,b)^T = b^T\ o'\ -A. \tag{2}$$

$$A\ o\ x = B\ o'\ x \tag{3}$$

Proposition 1. *([2], [3], [4]) It holds:*

(a) $A\ o\ x(A,b) \le b$.

(b) $M_1(A,b) \neq \emptyset$ if and only if $x(A,b) \in M_1(A,b)$.

(c) Let $M_1(A,b) \neq \emptyset$ and $x \in M_1(A,b)$. Then $x \le x(A,b)$, i.e. $x(A,b)$ is the maximum element of $M_1(A,b)$.

Definition 1. Let $z \in R^n,\ \lambda \in R,\ H = ||h_{jk}||,\ j,\ k \in J$ be such that $H\ o'\ z = \lambda\ o'\ z$. Then λ is called $(\min,+)$- eigenvalue and z $(\min,+)$- eigenvector of H corresponding to λ.

Proposition 2. *([2], [3]) Let H be a real $(n \times n)$-matrix, $\lambda(H)$ be a $(\min,+)$-eigenvalue of H, $H^p = H\ o'\ H\ \ldots,\ o'\ H$ ($p-$times), let $h_{ik}^{(p)}$ be elements of H^p. Then*

$$\lambda(H) = \min_{p \in J} \min_{i \in J}(h_{ii}^{(p)}/p),$$

where we set $h_{ii}^{(1)} = h_{ii}$. $\lambda(H)$ is the unique $(\min,+)$-eigenvalue of H.

Let for any $\alpha \in R, z \in R^n$, $\alpha\ o'\ z = (\alpha + z_1,\ \ldots,\ \alpha + z_m)^T$ and for any pair of matrices $C,\ D$ of equal size with elelments $c_{ik},\ d_{ik},\ (C \oplus' D)_{ik} = \min(c_{ik}, d_{ik})$.

Proposition 3. *([2] Theorem 4.2.4, p. 76) Let $H_\lambda = (-\lambda)\ o'\ H$.*

$$\Gamma(H_\lambda) = H_\lambda \oplus' H_\lambda^2 \oplus'\ \ldots\ \oplus' H_\lambda^m.$$

Then every column of $\Gamma(H_\lambda)$ with zero diagonal entry is a principal $(\min,+)$-eigenvector corresponding to $\lambda(H)$. Any $(\min,+)$-eigenvector of H can be expressed as a $(\min,+)$-linear combination of principal $(\min,+)$-eigenvectors of H.

4 Properties of set $M(A,B)$

Let us note that any solution x of system (1) must satisfy the inequalities

$$x_j \le (B\ o'\ x)_i - a_{ij},\ \forall i \in I, j \in J,$$

so that

$$x_j \le \min_{i \in I}((B\ o'\ x)_i - a_{ij}),\ \forall j \in J.$$

The maximum solution x of these inequalities must therefore satisfy the equation system (compare Proposition 1):

$$\overline{x}_j = \min_{i \in I}(\min_{k \in J}(b_{ik} + \overline{x}_k)) - a_{ij},\ j \in J. \tag{4}$$

By interchanging the min-operations we obtain

$$\overline{x}_j = \min_{k \in J}(\min_{i \in I}(b_{ik} + \overline{x}_k) - a_{ij}) = \min_{k \in J}(\min_{i \in I}(b_{ik} - a_{ij}) + \overline{x}_k),\ j \in J. \tag{5}$$

Let us set

$$q_{jk} = \min_{i \in I}(b_{ik} - a_{ij}), j, k \in J. \tag{6}$$

Thus, the maximum solution $\overline{x}$ will satisfy relation

$$Q\ o'\ \overline{x} = \overline{x}, \tag{7}$$

where $Q = B^T\ o'\ -A$.

It follows that the necessary condition for $\overline{x}$ to be the maximum solution satisfying (4) is that the $(\min,+)$-eigenvalue $\lambda(Q)$ of matrix Q is equal to zero and $\overline{x}$ is a corresonding $(\min,+)-$eigenvector of Q. The condition is not sufficient, since in general according to Proposition 1 only inequality $A\ o\ x(A,b) \leq b$ holds and it can happen that for vector $\overline{x}$ satisfying (7) only relations $A\ o\ \overline{x} \leq B\ o'\ \overline{x},\ A\ o\ \overline{x} \neq B\ o'\ \overline{x}$ hold. If $\overline{x}$ satisfies (7), we can easily find out by a direct computation whether at the same time equality $A\ o\ \overline{x} = B\ o'\ \overline{x}$ holds. If it holds, then $\overline{x}$ solves the system (1).

The following numerical example illustrates the theoretical results obtained above.

Example 2. Let $m = n = 2$,

$$A = (a_{ij}) = \begin{pmatrix} -2 & -1 \\ 0 & 2 \end{pmatrix},\ B = (b_{ik}) = \begin{pmatrix} 0 & -1 \\ 2 & 3 \end{pmatrix}.$$

We have

$$Q = (B^T\ o'\ -A) = \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix}.$$

Since

$$(Q)^2 = Q\ o'\ Q = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},$$

we have $\lambda(Q) = 0$ (see Proposition 2). Furthermore, we have $Q = Q_{\lambda(Q)}$, $\Gamma(Q) = Q\ \oplus'\ (Q)^2$ and

$$\Gamma(Q) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

It follows that $\overline{x} = (0,0)^T$ is the principal $(\min,+)$-eigenvector of Q corresponding to $\lambda(Q) = 0$. We will verify whether in our case $\overline{x}$ solves system (1). We have

$$A\ o\ \overline{x} = B\ o'\ \overline{x} = (-1,2)^T,$$

which gives $\overline{x} \in M(A,B)$.

If we have

$$A = (a_{ij}) = \begin{pmatrix} -2 & -1 \\ 0 & 2 \end{pmatrix},\ B = (b_{ik}) = \begin{pmatrix} 0 & -1 \\ 4 & 3 \end{pmatrix},$$

then similar calculations result in $\overline{x} = (1,0)^T$ and it can be easily verified that

$$A\ o\ \overline{x} = (-1,2)^T \neq B\ o'\ \overline{x} = (-1,3)^T.$$

Remark 1. We saw in the second part of Example 2 that in general the $(\min,+)$-eigenverctor $\overline{x}$ satisfies only the inequality $A\ o\ \overline{x} \leq B\ o'\ \overline{x}$. Let us note that it follows from the general theory of $(\min,+)$-eigenverctors that each $(\min,+)$-eigenverctor is a $(\min,+)$-linear combination of so called principal $(\min,+)$-eigenverctors (see Proposition 3). Therefore $M(A,B) \neq \emptyset$ if and only if at least one principle $(\min,+)$-eigenvector of Q is an element of $M(A,B)$. If it is not the case, then $M(A,B) = \emptyset$. It follows that we have to test at most n principle $(\min,+)$-eigenvectors. If none of the pricipal $(\min,+)$-eigenvectors is in $M(A,B)$, then $M(A,B) = \emptyset$ since for any $(\min,+)$-linear combination $\tilde{x}$ of any principal $(\min,+)$-eigenverctors relations $A\ o\ \tilde{x} \leq B\ o'\ \tilde{x},\ A\ o\ \tilde{x} \neq B\ o'\ \tilde{x}$ hold.

5 The case $\lambda(Q) \neq 0$

We will investigate further the case that $\lambda(Q) \neq 0$. We will show that a solution can be found in the form $(x, y) = (x, x - \lambda(Q))$, i.e. we can set $\alpha = 0$ and $y = x - \lambda(Q)$. Let $Q_{\lambda(Q)} = (q_{jk} - \lambda(Q))$ and $\lambda(Q_{\lambda(Q)})$ denote the $(\min, +)$-eigenvalue of $Q_{\lambda(Q)}$. Then $\lambda(Q_{\lambda(Q)}) = 0$. To simplify the notations, let us set $\tilde{Q} = Q_{\lambda(Q)}$, $\tilde{q}_{jk} = q_{jk} - \lambda(Q)$, $\tilde{A} = A + \lambda(Q) = \tilde{A} = (a_{ij} + \lambda(Q))$, $\tilde{B} = B - \lambda(Q) = (b_{ij} - \lambda(Q))$. Furthermore, we have

$$\tilde{q}_{jk} = q_{jk} - \lambda(Q) = b_{kj} - a_{ij} - \lambda(Q)\ \forall j, k. \tag{8}$$

Note that

$$\tilde{q}_{jk} = b_{kj} - (a_{ij} + \lambda(Q)) = b_{kj} - \tilde{a}_{ij},$$

where $\tilde{a}_{ij}$ denote the elements of $\tilde{A}$ and

$$\tilde{q}_{jk} = b_{kj} - \lambda(Q) + a_{ij} = \tilde{b}_{ik} - a_{ij},$$

where $\tilde{b}_{ik}$ denote the elements of $\tilde{B}$.

It follows that if $\tilde{x}$ is a $(\min +)$-eigenvector of $\tilde{Q}$ corresponding to its $(\min, +)$-eigenvalue, which is equal to 0, we have

$$\tilde{A}\ o\ \tilde{x} = B\ o'\ \tilde{x},\ \texttt{so that}\ (A + \lambda(Q))\ o\ \tilde{x} = B\ o'\ \tilde{x} \tag{9}$$

or

$$A\ o\ \tilde{x} = \tilde{B}\ o'\ \tilde{x},\ \texttt{so that}\ A\ o\ \tilde{x} = (B - \lambda(Q))\ o'\ \tilde{x} \tag{10}$$

Finally, (9), (10) are transformed to

$$A\ o\ (\tilde{x} + \lambda(Q)) = \tilde{B}\ o'\ \tilde{x},\ A\ o\ \tilde{x} = B\ o'\ (\tilde{x} - \lambda(Q)). \tag{11}$$

6 Conclusion

Using well known results of the theory of one-sided equation systems, we derived necessary conditions, which fulfill steady state solutions of the two-sided (max,plus)/(min,plus)-linear equation systems (1). Connection with (min,plus)-eigenvalues and (min,plus)-eigenvectors of matrix $Q = B^T - A$ was derived. The results can be used in synchronizing groups of activities with deterministic processing times with the aim to minimize the delay between the groups of activities which must follow one after the other. The subject of a further research will be to find necessary and sufficient conditions for the steady state solutions and methods.

Acknowledgements

Supported by the grant No. #14-02424S of the Grant Agency Czech Republic.

References

[1] Baccelli, F. L., Cohen, G., Olsder, J. P.: Synchronization and Linearity, Wiley, 1992.

[2] Butkovič, P.: Max-linear Systems: Theory and Algorithms, Monographs in Mathematics, Springer Verlag 2010.

[3] Cuninghame-Green, R. A.: Minimax Algebra, Lecture Notes **166** Springer Verlag Berlin, Heidelberg, New York 1979.

[4] Krivulin, N. K.: Methods of Idempotent Algebra in Modelling and Analysis of Complex Systems, S.-Peterburg 2009, (in Russian).

[5] Vorobjov, N. N.: Extremal Algebra of positive Matrices, *Datenverarbeitung und Kybernetik*, **3**, 1(1967), 39-71 (in Russian).

Slovak Day Surgery Beds Reduction on the Basis of Cost Frontier Function

Beata Gavurova[1], Samuel Korony[2]

Abstract. Objective of our paper was to gain excessive beds numbers in the field of both paediatric and adult day surgeries in Slovakia. We used available indicators of eight Slovak regions with one day surgery facilities for both paediatric and adult patients during 2009 – 2013 (number of beds was input cost variable, numbers of patients was output variable). We obtained linear cost functions as dependencies of the numbers of beds vs numbers of patients by ordinary least squares. Then by corrected least squares we found relatively efficient regions that lie on deterministic cost frontiers. Košice region in 2012 was cost efficient from the viewpoint of paediatric patients. In case of adult day surgery relatively the best region was Banská Bystrica region in 2012. This way we found that the overall number of excessive beds in paediatric (adult) day surgery during 2009 -2013 was 869 (2366). Total surplus of day surgery beds in Slovakia during 2009 – 2013 was 3 235 beds. The results of submitted analysis present a valuable platform for the creators of health and social policies as well as for other actors of health system in Slovakia.

Keywords: Day surgery, Beds Reduction, Efficiency, Corrected Least Squares, Cost Frontier Function.

JEL Classification: D24
AMS Classification: 91B32

1 Introduction

The issue of the effectiveness of healthcare system is a current one in all countries and a subject of many scientific activities of specialized research teams [1][2]. Their importance is enhanced by a variety of systemic disorders, which result in a lack of efficacy, efficiency and quality of healthcare [3][4][5]. Their solution requires an active approach of stakeholders in the healthcare system and setting up effective national policies [8][9]. We have focused on the issue of indebtedness of Slovak healthcare and its solution through a system setup of reduction of beds by introducing one day surgery. Since 2002 the Slovak healthcare system is constantly in rising indebtedness which was caused by inefficient public resources in the system and weak budget criteria for state contributory organizations. During this period, the question of the introduction and use of day surgery was getting more and more attention because of its convenience for both patients and health insurance companies. Day surgery addresses medical conditions that cannot be solved by a single clinic visit, but do not require hospitalization for a few days [12]. The basic principle is that the post-operative treatment does not require more than 24 hours duration [13]. It is implemented on the basis of expert guidance of Ministry of Health about medical care performances no. 06937/2006 dated 30.1.2006, vocational guidance of Ministry of Health no. 12225/2009 of medical care performances and professional guidance of the Ministry of Health of the Slovak Republic no. 08465/2010 amending the Professional guideline of Ministry of Health. Since 2007, with the development of day surgery is associated a process to reduce in-patient facilities, which was part of the restructuring process. It was criticized by the government mainly due to indiscriminate implementing of foreign models. On basis of the results of own research we can say that day surgery system in Slovakia is currently running but its components are not properly adjusted. This requires an effective regulation of the health system which should take particularly into account the health needs of the population. Their finding is very difficult especially from the methodological point (the scope, structure, representativeness and reliability of available data). There were research studies absenting in Slovakia till 2013 which would be dealing with the problems of day surgery. This is also the reason for the stagnation and inefficient process of reduction of beds that have become disposal for many hospitals in Slovakia [10][11]. Primary aim of this study is to highlight their importance in creating optimal strategy towards the development of day surgery, as well as in the drafting of state health policy towards increasing the efficiency of Slovak healthcare system.

[1] Faculty of Economics, Technical University of Košice, Němcovej 32, 040 01 Košice, Slovakia, e-mail: beata.gavurova@tuke.sk
[2] Faculty of Economics, Matej Bel University, Cesta na amfiteáter 1, 974 01 Banská Bystrica, Slovakia, e-mail: samuel.korony@umb.sk

2 Data and research methods

The data source for our analysis of day surgery surplus beds are data files offered to us kindly by the National Health Information Center of Slovakia (http://www.nczisk.sk). Slovak regions with day surgery facilities are our research objects. We analysed input and output indicators of both paediatric and adult day surgeries in Slovak regions during the years 2009 - 2013. Number of day surgery patients was output. Corresponding numbers of beds were used as input. We assume that the number of patients is already defined and it is not rational to increase it due to different medical, personal and technological requirements of day surgery therapy patients. That is why we used cost oriented approach. Besides descriptive statistics we used cost frontier analysis which is an parametric alternative to DEA methods [6][7]. Classic regression methods (e.g. least squares method) offer information about average cost not minimum cost. So they have to be modified. In our article we used corrected least squares method. We start from additive model of cost frontier:

$$C_i = f(Y_i, \beta) + e_i \tag{1}$$

where C_i is cost of i-th unit, Y_i is the vector of its outputs and $f(Y_i,\beta)$ is cost frontier. Vector β is the vector of estimated parameters and e_i is residual of regression model. It is measured by the difference between real cost and minimum possible cost. Residual with the largest vertical distance below the cost function represents efficient case. Thus cost frontier is obtained by adding minimum residual value to the intercept. In our case we do not consider stochastic errors. Value of residual is due only to inefficiency. For statistic and graphic data analysis we used statistical software IBM SPSS version 19. In next part we present briefly statistic parameters of analysed variables (day surgery indicators) and then results of surplus beds analyses.

3 Results

a) Descriptive statistics results

In Slovakia there are eight self-government regions at NUTS III level. In our paper we use their abbreviations from National Health Information Center of Slovakia: Banská Bystrica (BC), Bratislava (BL), Košice (KI), Nitra (NI), Prešov (PV), Trnava (TA), Trenčin (TC) and Žilina (ZI). Basic statistical parameters of analysed available variables (numbers of day surgery paediatric and adult patients, numbers of day surgery paediatric and adult beds) are in Table 1. We used them in cost frontier analysis of Slovak regions during five years interval (2009 -2013) (Mean = arithmetic mean, Std. Dev = standard deviation. Number of cases is 40 (eight Slovak regions during five years interval).

Variable	Mean	Std. dev.	Minimum	Region	Maximum	Region
Paediatric patients	1 317	895	201	TC9	3 374	PV13
Adult patients	13 817	6 572.80	4 233	TA9	28 337	KI13
Paediatric beds	39	23.4	5	TA9	115	NI13
Adult beds	151	72.5	30	TA9	303	KI13

Table 1 Statistical parameters of day surgery indicators in Slovak regions during 2009 – 2013

Mean number of day surgery paediatric patients is 1 317. Minimum number (201) was in 2009 in Trenčín region, maximum number was 3 374 (Prešov region in 2013). Mean number of day surgery paediatric beds is 39 with range from minimum value 5 again in Trnava region in 2009 to maximum value 115 (Nitra region in 2013). Mean number of day surgery adult patients is 13 817. Minimum number (4 233) was in 2009 in Trnava region, maximum number was 28 337 (Košice region in 2013). Mean number of day surgery adult beds is 151 with range from minimum value 30 again in Trnava region in 2009 to maximum value 303 (again Košice region in 2013).

b) Cost frontier results

Let us begin with paediatric day surgery. We want to optimize number of beds therefore we use cost function. Then number of beds is dependent variable and the number of patients is independent variable. Coefficients of linear regression model of its cost function are given in Table 2. Regression model of paediatric day surgery cost function is significant in both terms: constant and linear ($p < 0.001$). Coefficient of determination is rather small ($R^2 = 0.310$).

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Constant	18.972	5.516		3.439	.001
Paediatric patients	.015	.003	.573	4.307	.000

Table 2 Linear regression model of paediatric day surgery cost function

Constant term is significant what is surprising. As we can see in Fig.1 it is caused mainly by leverage effect of NI13 point (Nitra region in 2013). Residual of cost function was minimal in case of Košice in 2012 (label KI12). So Košice region in 2012 was cost efficient from all Slovak regions during 2009 - 2013 with zero surplus beds. But there were also other regions with negligible surplus beds (values less than 10): Banská Bystrica (2010, 2011, 2012 and 2013), Bratislava (2009 and 2010) and Trnava (2009, 2010, 2011 and 2012).
It is clear from the picture 1 that Nitra region in 2013 has got the largest beds surplus (label NI13). The second one case was again Nitra region in 2012 (NI12).

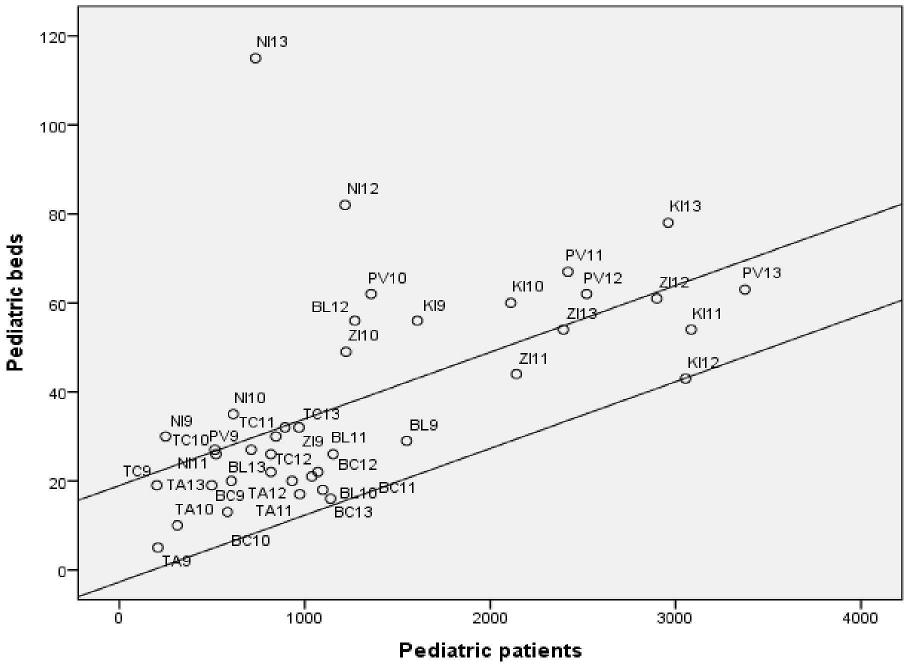


Figure 1 Paediatric day surgery linear cost function with corresponding cost frontier

Coefficients of linear regression model of adult day surgery cost function are in Table 3. Linear regression model of adult day surgery cost function is significant only in linear term ($p < 0.001$). Coefficient of determination is much better than in case of paediatric day surgery ($R^2 = 0.692$).

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Constant	23.782	14.948		1.591	.120
Adult patients	.009	.001	.837	9.421	.000

Table 3 Linear regression model of adult day surgery cost function

The variance of points around regression line is smaller in case of adult day surgery cost function (see Fig. 2). Banská Bystrica region in 2012 was efficient from all Slovak regions during 2009 – 2013 with minimum residual (zero surplus beds, label BC12). Other regions with small surplus beds (amounts less than 20) were: Žilina (2009 and 2010), Bratislava (2012), Trnava (2012) and Nitra (2013). The Prešov region in 2011 has got the largest beds surplus (label PV11). The second one case was Košice region in 2010 (KI10).

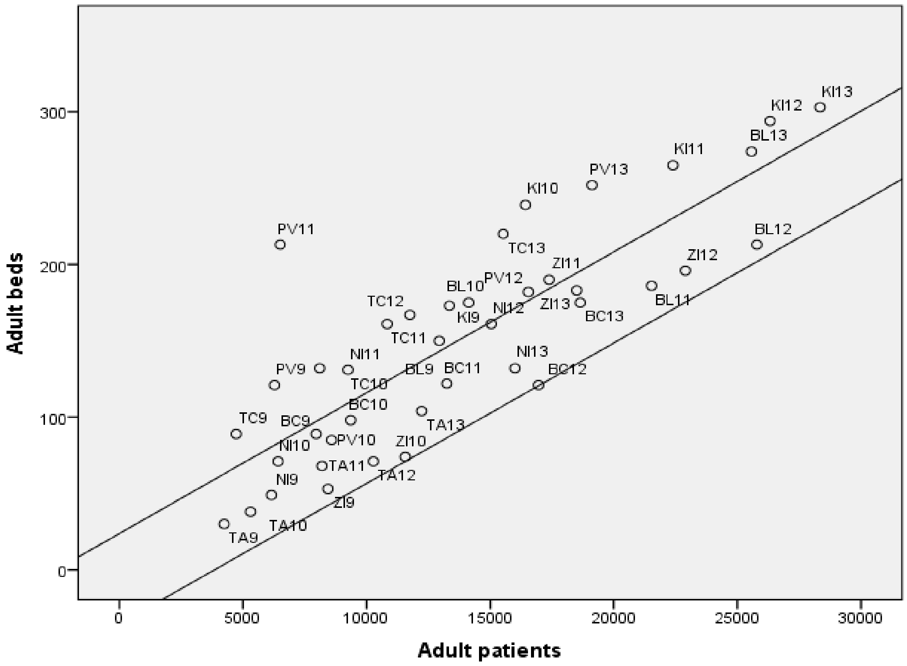


Figure 2 Adult day surgery linear cost function with corresponding cost frontier

	Beds surplus					Beds surplus			
Region	**Paediatric**	**Adult**	**Overall**	**Rank**	**Region**	**Paediatric**	**Adult**	**Overall**	**Rank**
BC2012	8	**0**	**8**	1	**PV2010**	44	41	86	21
TA2012	9	12	**20**	2	**KI2012**	**0**	87	87	22
TA2009	5	26	31	3	**BL2010**	4	85	90	23
ZI2009	21	11	32	4	**BL2013**	17	74	90	24
TA2010	8	24	32	5	**PV2012**	27	65	92	25
TA2011	5	28	33	6	**TC2009**	19	81	99	26
BL2011	12	23	34	7	**NI2011**	22	81	103	27
ZI2010	33	3	36	8	**KI2011**	11	94	104	28
BC2013	2	38	40	9	**TC2012**	13	94	106	29
ZI2012	20	20	40	10	**KI2013**	36	77	113	30
TA2013	14	27	41	11	**TC2010**	21	93	114	31
BC2011	9	35	44	12	**KI2009**	35	80	115	32
BL2012	40	11	50	13	**TC2011**	20	97	117	33
BC2010	7	47	54	14	**PV2009**	19	99	118	34
NI2009	29	28	57	15	**NI2012**	**67**	58	124	35
BC2009	14	51	65	16	**PV2013**	15	111	126	36
ZI2013	21	48	69	17	**NI2013**	**107**	20	126	37
BL2009	9	66	75	18	**TC2013**	20	112	132	38
NI2010	29	47	76	19	**KI2010**	31	123	**154**	39
ZI2011	15	65	80	20	**PV2011**	34	188	**222**	40

Table 4 Unnecessary day surgery beds in Slovak regions during 2009 -2013

Complete summary of all surplus beds in Slovak regions during analysed time interval is in the Table 4. Regions are ranked according to sum of both paediatric and adult day surgery surplus beds. From the table we can see that Banská Bystrica region in 2012 was the best of all regions in both paediatric and adult day surgery facilities from the viewpoint of surplus beds. The number of its surplus beds was only 8. The second one is Trnava region in 2012 (surplus beds 20). The last is Prešov region in 2011 with 222 surplus beds and the last but one is Košice region in 2010 (154 surplus beds).
In Fig. 3 and Fig. 4 are depicted overall sums of surplus beds in Slovak regions. The best region during whole time interval is Banská Bystrica with surplus beds value of only 39. Very similar is Trnava region with 41 surplus beds. The worst is Nitra region with 253 unnecessary paediatric day surgery beds. The second one is Prešov (139 beds). In the field of adult day surgery is the best one Trnava region (117 surplus beds). Prešov region is the worst one (504 beds). The overall number of excessive beds in paediatric day surgery was 869 and in the case of adult day surgery it was 2366. Total surplus of day surgery beds in Slovakia during 2009 – 2013 was 3 235 beds.

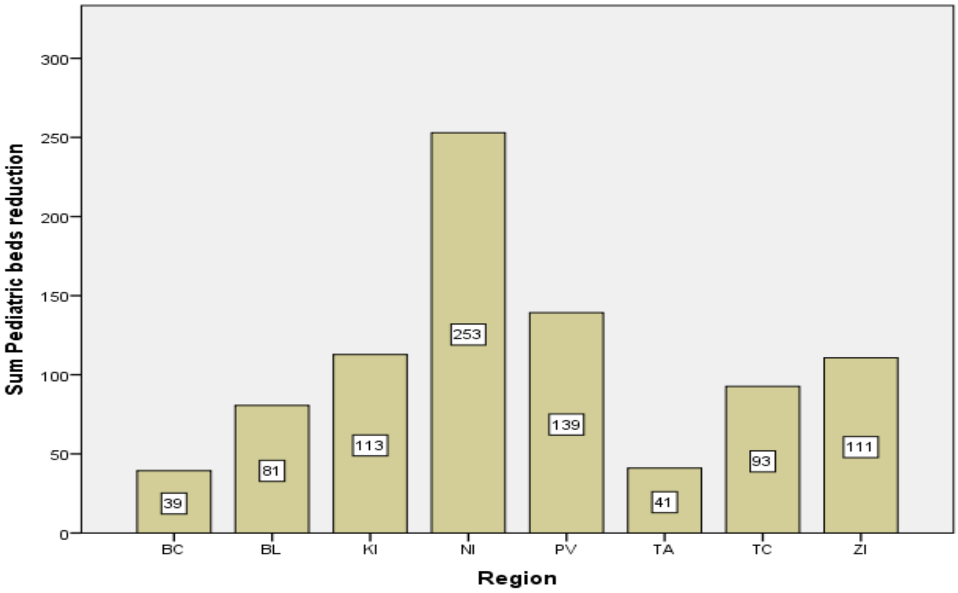


Figure 3 Bar graph of overall surplus beds of paediatric day surgery facilities in Slovak regions during 2009 - 2013

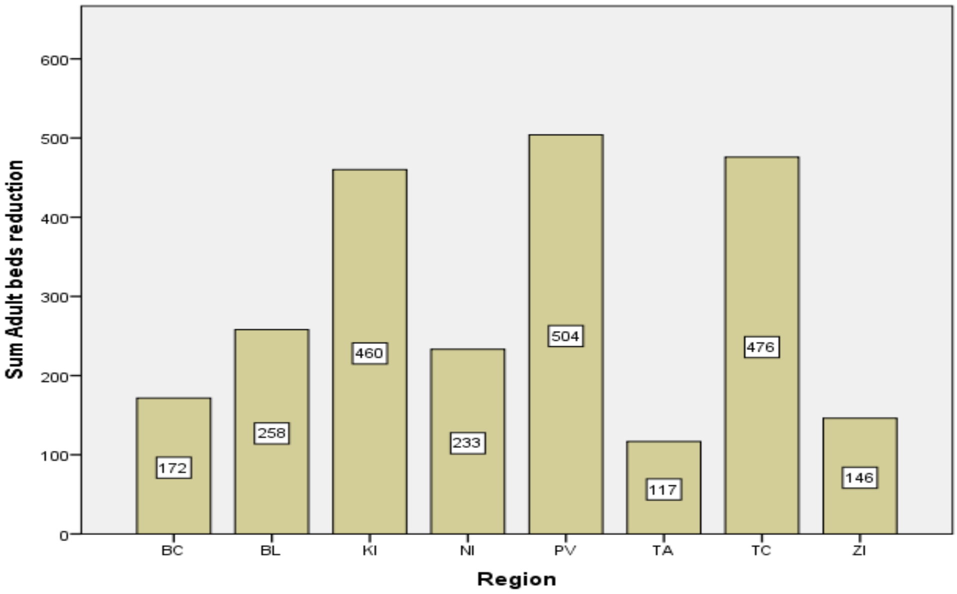


Figure 4 Bar graph of overall surplus beds of adult day surgery facilities in Slovak regions during 2009 – 2013

4 Conclusions

The health sector is increasingly starting to use demographic data to provide scope for the high-quality spatial analysis. Their special significance can be seen in solving the issues concerning on system of day surgery in which our country is lagging behind in comparison to other developed countries. Critical aspect of this situation is considered the 15-year period of its use in Slovakia. Despite significant legislative support of Ministry of Health, support of health insurance companies, as well as other actors of the healthcare system, there is no recorded positive progress in Slovakia in recent years. This relates to the incorrect setting of policy of financing of performances determined by the pricing of health insurance companies influenced by the form of ownership (state vs. two private). Objective of our paper was to gain excessive beds numbers in the field of both paediatric and adult day surgeries in Slovakia. By corrected least squares we found relatively efficient regions with minimum residuals from the viewpoint of deterministic cost linear frontiers. Košice region in 2012 was efficient from the viewpoint of paediatric patients. In case of adult day surgery relatively the best region was Banská Bystrica region in 2012. We found that the overall number of unnecessary beds in paediatric (adult) day surgery during 2009 - 2013 was 869 (2366). Total surplus of day surgery beds in Slovakia during 2009 – 2013 was 3 235 beds. Presented partial outcomes from analysis of day surgery in Slovakia reveals the importance of analysing data that are available from the national registry (National Health Information Center) and the importance of their links with demographical characteristics. The outputs of these analyses enable the appropriate system setup for day surgery in different regions of Slovakia, link them to policy financed through health insurance companies, support the improvement of information disciplines of healthcare and related improvement of the process of reporting on the national register which will also promote the formation of objectives in Slovak state health policy along the lines of increasing the efficiency of the healthcare system.

Acknowledgement
The work was supported by the VEGA Project No. 1/0986/15 *„Proposal of the dimensional models of the management effectiveness of ICT and information systems in health facilities in Slovakia and the economic-financial quantification of their effects on the health system in Slovakia“.*
Our thanks go to the National Health Information Center of Slovakia (http://www.nczisk.sk/) for providing the Slovak day surgery database as well as for a long-term support of our research activities within health policy.

References:

[1] Bem, A., Predkiewicz, K., Predkiewicz, P., and Ucieklak-Jez, P.: Determinants of Hospital’s Financial Liquidity. *Procedia Economics and Finance* **12** (2014), 27-36.
[2] Bem, A., and Michalski, G.: The financial health of hospitals. V4 countries case. *Sociálna ekonomika a vzdelávanie,* Banska Bystrica, (2014), pp. 1-8.
[3] Bem, A., and Michalski, G.: Hospital profitability vs. selected healthcare system indicators. *CEFE 2015 – Central European Conference in Finance and Economics*, Technical University of Kosice, Kosice, (2015), 52-61.
[4] Castoro, C., Bertinato, L., Baccaglini, U., Drace, C. A., and McKee, M.: *Policy Brief. Day Surgery: Making it Happen.* Brussels: WHO European Centre for Health Policy, (2007).
[5] Detmer, D. E., and Gellins, A. C.: Ambulatory surgery: A more cost-effective treatment strategy? *Archives of Surgery* **129,** 2 (1994), 123-127.
[6] Fried, H.O. et al. (eds.) *The Measurement of Productive Efficiency and Productivity Change.* Oxford University Press, Oxford, 2008.
[7] Jacobs, R., et al.: *Measuring Efficiency in Health Care*. Cambridge University Press, Cambridge, 2006.
[8] Szabo, S., and Sidor, J.: The Performance Measurement System – Potentials and Barriers for its Implementation in Healthcare Facilities. *Journal of Applied Economic Sciences* **9**, 4 (2014), 728-735.
[9] Szczygieł, N., Rutkowska-Podolska, M. and Michalski, G.: Information and Communication Technologies in Healthcare: Still Innovation or Reality? Innovative and Entrepreneurial Value - creating Approach in Healthcare Management. *5 th Central European Conference in Regional Science – CERS*, 2014, 1020-1029.
[10] Šoltés, V., and Gavurová, B.: The Functionality Comparison of the Health Care Systems by the Analytical Hierarchy Process Method. *E+M Ekonomie a Management* **17**, 3 (2014), 100-118.
[11] Soltes, M., Gavurova. B.: Quantification and comparison of avoidable mortality – causal relations and modification of concepts. *Technological and Economic Development of Economy* **21**, 6 (2016), 917-938.
[12] Šoltés, M., Pažinka, P., and Radoňak, J.: Laparoskopická hernioplastika TAPP v liečbe slabinovej prietrže – 10 ročné skúsenosti. *Rozhl Chir* **89**, 6 (2010), 384-389.
[13] Šoltés, M., Pažinka, P. and Radoňak, J.: Termické lézie v laparoskopickej chirurgii. *Endoskopie* **20**, 1 (2011), 14-16.

Analysis in a time series of milk-yield production

Tomas Haloun[1], Jaroslav Marek[2], Radko Rajmon[3], Martina Kršíková[4]

Abstract. Monitoring of milk yield is used as a tool for the operative management of individual cows as well as management at the herd level. Nevertheless, sometimes a general knowledge of the significance of a particular problem is not sufficient for appropriate managerial decisions and a need arises to weigh up the impact of particular disorder on a particular herd. Than it is necessary to evaluate animal data from a longer period, different lactation order etc. The aim of the introduced research is to present methods for comparing the milk yield of cows in different epochs. The data from selected Holstein dairy herds in the Czech Republic from an 8-year period was used as a model. Several statistical hypotheses (zero trend, absence of seasonal variation) were being tested. The results indicated a requirement to consider the relationship between labour date and milk production. Our statistical analysis enabled us to draw up a table of milk-yield indices showing the average changes in 305-day yield values among lactations and years.

Keywords: Statistical analysis of milk production, Wood function, Nonlinear regression, Estimates of unknown parameters, time series.

JEL classification: C13
AMS classification: 62J02, 62M10

1 The introduction

Milk yield has a substantial impact on the economic profitability of dairy farms. Therefore a great deal of attention is paid to the analysis of this indicator as well as the factors which influence milk production. There exists a host of models which evaluate the effects such as the success of genetic improvement, increasing quality of care, better feed composition and seasonal effect milk production at population level, cf. [6]. At farm level the milk production is influenced by factors which impact the whole herd but at the same time there are factors which affect individual animals. In order to make managerial decisions, it can be desirable to assess the yield of individual cows in the context of the herd at a concrete farm, which means working with seriously limited data volumes and different models. The basis for yield evaluation is modelling of a 305-day yield. Using a mathematical model for the description of the lactation curve leads to the need for finding a suitable regression function for fitting of measurements of daily yield, which are performed mostly only once a month. Several methods for approximation of a standard 305-day lactation cycle have already been proposed. Many studies have focused on modelling the 305-day yield. Cf. [2], [1], [3]. The Wood function is the most preferred method for solution of this nonlinear regression problem. See [5].

To evaluate phenomena observed in the long term, which have an individual impact and occur less frequently, it is desirable to standardize the estimated 305-day yields. A timeline may be convenient to capture correctly trend, seasonal, and cyclic components. The existence of these components needs to be tested at a chosen level of significance with suitable statistical tests. Subsequently, the productivity can be stripped off substantial factors. The main goal of this paper is to propose a suitable solution to the issue of yield correction in selected animals, which will make comparisons possible in the long term.

[1]Czech University of Life Sciences Prague, Faculty of Agrobiology, Food and Natural Resources, Department of Veterinary Science, Kamycka 129, 165 21 Praha 6, Suchdol, Czech Republic, tomas.haloun@email.cz

[2]University of Pardubice, Faculty of Electrical Engineering and Informatics, Department of Mathematics and Physics, 532 10 Pardubice, Czech Republic, jaroslav.marek@upce.cz

[3]Czech University of Life Sciences Prague, Faculty of Agrobiology, Food and Natural Resources, Department of Veterinary Science, Kamycka 129, 165 21 Praha 6, Suchdol, Czech Republic, rajmon@af.czu.cz

[4]University of Pardubice, Faculty of Electrical Engineering and Informatics, Department of Mathematics and Physics, 532 10 Pardubice, Czech Republic, st32357@student.upce.cz

1.1 Measurements

The data was obtained from the company ZD Zahori register of two farms. All recordings for all 3737 animals registered between the years 2001 and 2013 were exported. The data included 76724 measurements 2803, 1928, 1232, 1531 cycles for first, second, third, four and larger lactation, respectively. The data were exported directly from the register with subsequent attributes: Cow id, cowshed stall number, date of cows birth, day of cows disposal, date of labour, order of lactation, date of the control day, daily milk yield on the control day [in kilogram], protein content [in kilogram], number of somatic cells (Tab. 1).

43539	265	1	20.11.1997	14.05.2007	07.12.2004	6	25.01.06	11.40	3.92	3.65	240.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	23.03.06	41.20	4.52	3.40	28.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	24.04.06	42.00	3.34	3.08	237.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	24.05.06	39.60	3.35	3.14	67.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	26.06.06	35.20	4.07	2.83	183.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	25.07.06	26.20	3.74	3.11	391.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	23.08.06	30.80	3.53	3.17	327.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	25.09.06	30.60	3.00	3.07	115.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	24.10.06	28.40	2.80	3.39	44.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	23.11.06	23.20	3.19	3.43	102.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	20.12.06	20.60	5.13	3.74	637.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	23.01.07	23.00	4.31	3.60	313.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	22.02.07	22.40	4.32	3.58	311.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	26.03.07	20.40	4.27	3.43	531.00
43539	265	1	20.11.1997	14.05.2007	04.03.2006	7	23.04.07	19.00	4.89	3.74	546.00
43542	265	1	28.11.1997	20.03.2002	26.01.2000	1	23.02.00	20.40	5.69	3.25	34.00

Table 1 Samples of the data on monthly measurements of the daily milk yield of several cows

1.2 Estimation of lactation curve by Wood function

Several mathematical models for 305-day yield are at our disposal. For detail see [2], [1], [3]. The results in [3] indicated very good properties of Wood models. Estimation of this model will be described now. Let x be the day of lactation and Y is the daily yield. Let's consider that function f models how variable x explain variable $\mathbf{Y}$. Wood model for estimating of 305-day yield is given by function

$$f(\boldsymbol{\beta}, x) = \beta_1 x^{\beta_2} e^{-\beta_3 x}. \tag{1}$$

With nonlinear regression we expect that $\mathbf{Y}-\mathbf{Y}_0$ has distribution with mean value $X\boldsymbol{\beta}$, and covariance matrix where $\mathbf{Y}$ is vector of measured daily yields, $\mathbf{Y}_0 = f(\boldsymbol{\beta}_0; x)$, $\boldsymbol{\beta}_0$ is a suitable initial vector of unknown parameters. Ordinary squares estimate of unknown vector parameter is given by formula

$$\delta\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{Y} - \mathbf{Y}_0)', \tag{2}$$

where $\mathbf{X}$ is matrix of first partial derivatives of Wood function. Dimension of the matrix is $n \times 3$. After determining the partial derivatives of Wood function our matrix $\mathbf{X}$ has a form:

$$\mathbf{X} = \begin{pmatrix} x_1^{\beta_2} e^{-\beta_3 x_1}, & \beta_1 x_1^{\beta_2} e^{-\beta_3 x_1} \log(x_1), & -\beta_1 x_1^{\beta_2+1} e^{-\beta_3 x_1} \\ & \dots & \\ x_n^{\beta_2} e^{-\beta_3 x_n}, & \beta_1 x_n^{\beta_2} e^{-\beta_3 x_n} \log(x_n), & -\beta_1 x_n^{\beta_2+1} e^{-\beta_3 x_n} \end{pmatrix}. \tag{3}$$

Estimation of unknown parameters is determined by adding up the initial solution and correction given in formula (2):

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}_0 + \delta\widehat{\boldsymbol{\beta}}. \tag{4}$$

In nonlinear regression for functions with large Bates & Watts curvature and a small area of linearization a correction of the initial solution is given by relationship

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}_0 + \alpha\delta\widehat{\boldsymbol{\beta}}. \tag{5}$$

In process of numerical calculations we can choose equidistant step eg. $s = 1/1000$. So we will get values $\alpha_i = i/1000$, $i = 1, \dots, n$. In the cycle 1000 estimates of unknow parameters is obtained. An estimate with the smallest value of criterior for estimation of regression line

$$S_e = \sum_{i=1}^{n} (Y_i - \widehat{Y}_i)^2 \tag{6}$$

is considered as most satisfactory.

1.3 Example of approximation

This section will give a detailed description of individual steps in calculating unknown parameters of the approximate function in one milk cow. Data are selected from seven lactation cycles of the milk cow with ID 43539. The data for this cow in the table 2 was collected and calculated on the basis of data on date of the milk yield measurement, date of the cow's birthday and subsequent labors given in Table 1. Measured values of daily yields (in kilograms) are labelled y and the serial (order) number of the lactation cycle in measured days is marked as x.

x	19	51	81	114	143	172	205	214	244	271	325
Y	41.2	42.0	39.6	35.2	26.2	30.8	30.6	28.4	23.2	20.6	23.0

Table 2 Samples of monthly measurements of the daily milk yield with order of lactation day

We need an appropriate initial vector of unknown parameters of Wood function, which is close to the true value of unknown parameters. Suitable initial solution for approximation by Wood function can be set as (5; 0.8; 0.02), cf. [3]. Then vector Y_0 can be calculated by substitution of initial solution $\boldsymbol{\beta}_0$ and vector x of the order of days in lactation cycle, into Wood function. We need this vector to calculate the difference between estimated and measured milk yield. Calculated vector is $y_0 = (42.7, 39.6, 37.1, 34.6, 32.5, 30.6, 28.5, 26.8, 25.2, 23.8)'$. Numerical calculation provides the following subresults and results

$$\mathbf{X} = \begin{pmatrix} 0.94, & 125.58, & 810.36 \\ 0.87, & 155.76, & 2020.40 \\ 0.81, & 163.05, & 3005.43 \\ 0.76, & 163.74, & 3941.13 \\ 0.72, & 161.33, & 4648.61 \\ 0.67, & 157.40, & 5259.38 \\ 0.63, & 151.86, & 5848.33 \\ 0.59, & 146.45, & 6281.95 \\ 0.56, & 140.59, & 6656.25 \\ 0.52, & 135.20, & 6934.73 \end{pmatrix}, \qquad \begin{array}{c} \delta\widehat{\boldsymbol{\beta}} = (45.4125, 0.0080189, 0.0020595)', \ \alpha_{opt} = 3/1000, \\ \widehat{\boldsymbol{\beta}} = (5.1362375, 0.799976, 0.020006)', \ S_e = 79.62. \end{array}$$

Obtained estimate of unknown parameters is vector $\widehat{\boldsymbol{\beta}} = (5.1362375, 0.799976, 0.020006)'$. Estimate of lactation curve is presented in Fig. 1.

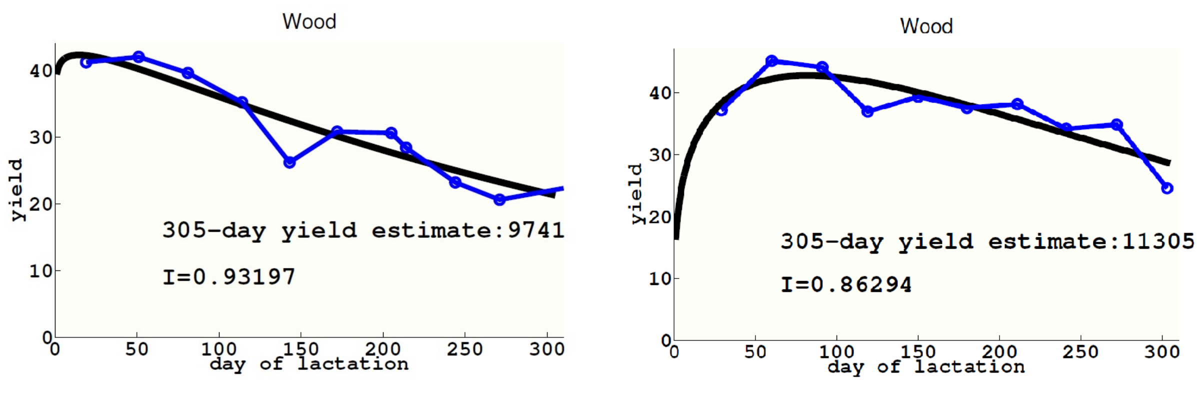


Figure 1 Estimates of 305 d yield of cow 43539 at 1th lactation and cow 54538 at 7th lactation

1.4 Applications to time series

In the process of finding a suitable methodology for data processing various test of statistical hypotheses about trend in years, quarters, months were performed. The results of these tests demonstrated a significant correlation between milk production and the years. Furthermore significant seasonal component appeared in the quarters of the year. Comparing months appeared to be problematic with regard to the different number of observations at a given interval and too many boundary observation. Coefficients (relative indices) for each quarter and each separate lactation were proposed as the optimal procedure. The final result of our study and our random sample is then a table of milk yield indices. This table facilitates comparison of milk yield of individual animals from different lactation and years.

So time series of milk yield can be described by model with trend and seasonal components

$$Y = a + bt + \alpha_1 q_1 + \alpha_2 q_2 + \alpha_3 q_3 + \alpha_4 q_4, \tag{7}$$

where q_1 is 1 only for first quarter, for another quarters of year is zero, where q_2 is 1 only for second quarter, for another quarters is zero, where q_3 is 1 only for third quarter, for another quarters is zero, where q_4 is 1 only for fourth quarter, for another quarters is zero. Analysis of milk yield time series show how trend and seasonal factor change 305-d milk yield. See Fig. 2 and Fig. 3.

Based on such time series we can directly make corrections of estimated yield for individual cows based on its classification into time interval. This leads to the possibility to relate yield to the selected period, which will be evaluated in series of milk yield indexes by value 1. Each additional period then has its own value index, which measures the relative change from baseline period in milk yield periods.

2 Results and Discussion

2.1 Trend in period 2005–2013

Obtained estimates of 305 day yield are depicted in Fig. 2. On axis x is date of calving. t-test for hypothesis of nonzero trend provide us information that trend is positive. The result obtained from ANOVA model simply provides information about existence of seasonal factors. See Table 3.

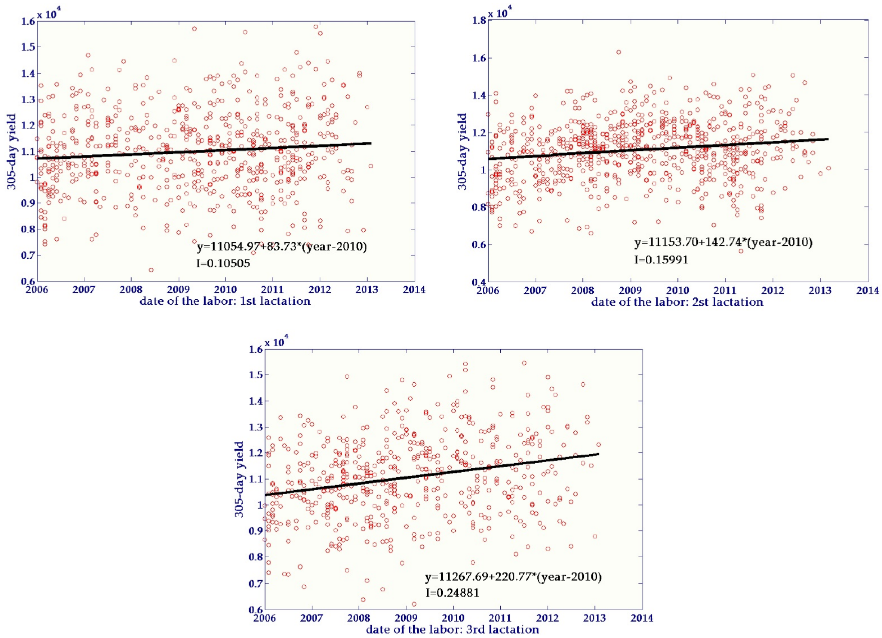


Figure 2 Time series of 305 d yield estimators between 2006 and 2013, 1st, 2nd, and 3rd lactation

A construction of graph in Fig. 3 is made on model (7). Intended milk yield indexes are presented in the Tab. 4. As an initial period with an index value 1 was chosen the first quarter of 2006.

first	second	third	fourth
10714	10005	10111	10743

Table 3 The mean values of milk-yield for quarters of year, 1st lactation

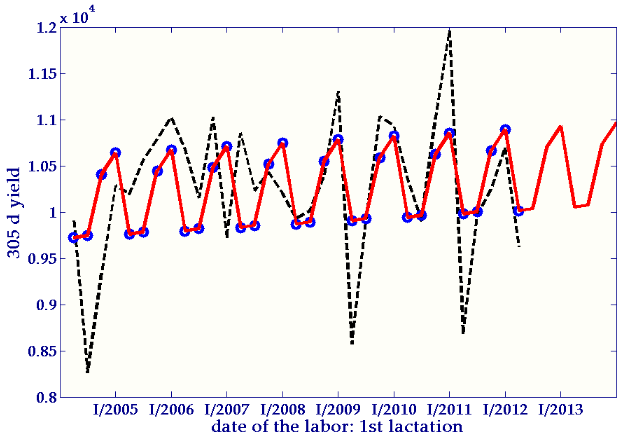


Figure 3 Estimates of trend and seasional component in 305 d yield in 2005 – 2012, 1st lactation

quarter/year	mean	n	s	index
III/2005	8258.238	180	522.13794	0.803045261
IV/2005	9340.996	204	282.69537	0.908334510
I/2006	10283.652	227	154.28539	1.000000000
II/2006	10193.088	186	156.32336	0.991193401
III/2006	10560.219	163	347.72766	0.991193401
IV/2006	10787.2	120	343.05917	1.048965873
I/2007	11027.059	138	665.71650	1.072290175
II/2007	10674.728	182	925.85370	1.038028903
III/2007	10154.052	184	662.67224	0.987397473
IV/2007	11024.403	212	389.96655	1.072031901
I/2008	9715.876	204	479.19580	0.944788486
II/2008	10853.934	184	506.42737	1.055455202
III/2008	10232.292	173	349.37335	0.995005665
IV/2008	10428.035	173	581.23500	1.014040051
I/2009	10198.219	151	825.39960	0.991692348
II/2009	9926.858	167	372.90256	0.965304738
III/2009	10016.76	169	461.18213	0.974046963
IV/2009	10400.896	182	406.96500	1.011401008
I/2010	11305.579	179	163.22684	1.099373938
II/2010	8572.864	159	1117.67100	0.833640034
III/2010	9900.303	198	164.36745	0.962722484
IV/2010	11031.356	145	315.29016	1.072708022
I/2011	10927.135	211	420.09340	1.062573393
II/2011	10353.197	184	274.22632	1.006762675
III/2011	9904.891	137	1138.0745	0.963168629
IV/2011	11021.393	176	848.06330	1.071739203
I/2012	11962.039	148	421.92883	1.163209237
II/2012	8680.581	103	981.57153	0.844114620
III/2012	9960.138	111	1048.89110	0.968540942
IV/2012	10243.146	78	250.45737	0.996061127

Table 4 The mean values of milk-yield for quarters of year, 1st lactation

3 Conlusion

When processing limited data sets or comparing individual milk cows, it is necessary to have the total 305-day yields stripped off selected factors. According to literature, a Wood function is considered as the most appropriate model for modelling a 305-day yield. Testing the statistical hypothesis indicated the existence of nonzero trend and seasonal variation. Gained knowledge allows to propose a method of 305-day yield corrections by factors of lactation and calving data. Corrections are made using the index table for years and quarters. The calculated milk yield indexes in a given cowshed can help producers when comparing cows and making decisions.

Acknowledgements

This research was supported by the institutional support of Czech University of Life Sciences and by Internal Grant Agency of University of Pardubice, the project SGS FEI_26_2016.

References

[1] Golebiewski, M., Brzozowski, P., and Golebiewski, L.: Analysis of lactation curves, milk constituents, somatic cell count and urea in milk of cows by the mathematical model of Wood. *Acta Veterinaria* Brno **1** (2011), 3-80.

[2] Leon-Velarde, C.,U., McMillan, I., Gentry, R. D., and Wilton, J. W.: Models for estimating typical lactation curves in dairy cattle. *Journal of Animal Breeding and Genetics* **112** (1995), 333-340. doi: 10.1111/j.1439-0388.1995.tb00575.x

[3] Marek, J., Rajmon, R., and Haloun, T.: Critical evaluation of seven lactation curve estimation models. In: *Intelligent Data Analysis and Applications. Proceedings of the Second Euro-China Conference on Intelligent Data Analysis and Applications, ECC 2015* (Abraham Ajith, Jiang Xin Hua, Snasel Vaclav, Pan Jeng-Shyang, eds.). Springer International Publishing, New York, 2015, 73-84. ISBN: 978-3-319-21206-7

[4] Wimmer, G., Palencar, R., Witkovsky, V.: *Spracovanie a vyhodnocovanie merania.* VEDA, Bratislava, 2002. p. 1-187. ISBN 80-224-0734-8.

[5] Wood, P.D.P.: Algebraic model of the lactation curve in References cattle. *Nature* **216** (1967), 164-165.

[6] Zavadilova, J.: Definition of subgroups for fixed regression in the test-day animal model for milk production of Holstein cattle in the Czech Republic. *Czech J. Anim. Sci.* **50**, vol. 1 (2005), 7-13.

[7] Zavadilova, J.: Genetic parameters for test-day model with random regressions for production traits of Czech Holstein cattle. *Czech J. Anim. Sci.* **50**, vol. 4 (2005), 142-150.

A two-stage Data Envelopment Analysis Model with application to banking industry in the Visegrad Group

Jana Hanclova[1]

Abstract. Data Envelopment Analysis (DEA) is widely used mathematical programming approach for comparing the inputs and outputs of a set of homogenous Decision Making Units (DMUs) by evaluating their relative efficiency. This paper focuses on this financing issue using a two-stage DEA model for 62 financial institutions in Visegrad Four countries in year 2013. The first stage measures fund collection efficiency and the second stage calculates operational efficiency of investment process. The estimated overall technical efficiency of the two-stage indicates that on average, deposits were under-produced in the first stage of production. Inefficient banks produce inside the production possibility set and should be able to simultaneously expand loans and securities given inputs and an appropriate choice of deposits if they become efficient. As a consequence, the underproduction of deposits resulted in fewer loans and securities investments. The results also found that the two-stage network system is superior to traditional DEA in the banking sector and confirmed that the traditional DEA approach overestimates efficiency scores.

Keywords: two-stage model, data envelopment analysis, banking sector.

JEL Classification: C44, D24, G20
AMS Classification: 90C05, 90C90

1 Introduction

Bank industry plays a critical role in the economic development of Visegrad Four (V4) countries. The economies of these countries have in common that after communism's collapse, there have been some changes - the transition to a market economy, joining the European Union in May 2004 and especially the transformation of banking system. The transition from centrally planned economy to market economy had been accompanied with restructuring and liberalization of the banking system. It had been associated with the privatization of some banks, the entry of foreign-owned banks, deregulation of interest rates and changes in legislation.

The improvements of management performances of financial institutions enable them to have a solid financial foundation and avoid the potential triggers of financial crisis. A number of approaches can be used to assess the banking industry including financial ratio analysis, data envelopment analysis (DEA) and the stochastic frontier analysis (SFA). Compared with ratio analysis and the SFA, DEA is considered a better means to analyse performance because it can produce multiple outputs using multiple inputs and requires no prior assumption regarding the specification of the best practice frontier. DEA is a linear programming based technique for measuring the relative efficiency of a set of homogenous decision making units (DMUs) and it is a non-parametric approach. The main DEA models used in banking industry can be classified into two categories: one is single-stage DEA approach and the other is two-stage or multi-stage DEA approaches. A good number of works adopt single-stage DEA approach (([1] [9], ([12], ([13] and [16]). However these single-stage models do not consider a banks internal structure relative to measures that characterize the operations performance of the bank. A bank is a black-box and that it ignores immediate production processes. In the two-stage DEA model each DMU uses its inputs to produce intermediate outputs in the first stage that become the inputs to the second stage from which final outputs are produced. The study of two-stage DEA approach is relatively active in banking industry in recent years and several important works have been reported ([2], [6], [7], [8]).

In this paper we investigate the efficiency status a banking system using the two-stage DEA model for Visegrad Four (V4) countries in the year 2013. We provide benchmarks for the financial institutions using four efficiency indexes for the input-oriented model. The overall technical efficiency is decomposed into the production efficiency in the first stage, the operational efficiency in the second stage. The last black-box efficiency is measured for the DEA model excluding intermediate measures. The rest of this paper is organized as follows. In the next section, we will present our two-stage data envelopment model for efficiency analysis of banking systems. Then, the application of our DEA model in evaluating the efficiency of 62 Visegrad Four financial institutions in

[1] VSB-Technical university of Ostrava, Faculty of Economics, Department of Systems Engineering, Sokolska tr. 33, 70200 Ostrava, Czech Republic, jana.hanclova@vsb.cz.

the year 2013 is presented in section 4. The last section summarizes our findings and suggests several directions for further research.

2 Two-stage DEA process

Charnes in [3] generalised Farrell's measure of single output efficiency into multiple outputs. Therefore, DEA has been widely used in various organizations and industries as the conventional model of one-stage production process. In practise, many production processes, such as those in the financial services sector, need more than one-stage to reflect sufficient management information. Wang pioneered intermediate production process as a two-stage DEA model in [18] to assess the impact of information technology on bank performance.

The two-stage DEA model is shown in Figure 1. This model includes all the outputs from the first stage (S1) as intermediate measures that make up the inputs to the second stage (S2). The resulting two-stage DEA model provides not only the overall efficiency score for the entire process, but as well yields the efficiency scores for each of the individual stages.

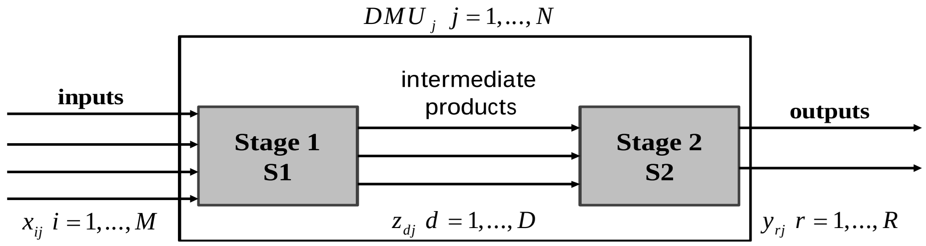


Figure 1 Two-stage process (source: own elaboration]

Seiford and Zhu developed DEA in [15] approach for evaluation of US commercial banks in a two-stage process characterized by profitability and marketability. They applied the standard DEA approach to each stage, which did not address potential conflicts between the two stages arising from the intermediate measures. The next study sought alternative ways to address the conflict between the two stages or to provide efficiency scores for both individual stages and the overall process. Liang, Cook and Zhu investigated the two-stage processes in in [11] via concepts adopted from non-cooperative and cooperative games. The resulting models were linear and the overall efficiency was a product of the efficiencies of the two individual stages. When there was only one intermediate measure connecting the two stages, both the non-cooperative (with leader-follower assumption) and cooperative models provided the same results as if the standard DEA model would be applied to the two steps separately and the efficiency decomposition was unique. Liang, Cook and Zhu developed in [11] **the centralized (C) or cooperative** DEA model under multiple intermediate measures where the overall efficiency decomposition was unique. This model provided a set of optimal weights on the intermediate factors that maximizes the overall efficiency scores, i.e. efficiencies of both stages are evaluated simultaneously. Chen, Cook and Zhu developed in [4] an approach for determining the DEA frontier or DEA projections for inefficient DMUs.

Consider a two-stage process, as it is shown in Figure 1, for each of a set of *N* DMUs. We assume each DMU_j $\left(j=1,2,\ldots,N\right)$ has *M* inputs $x_{ij},\left(i=1,2,\ldots,M\right)$ to the first stage and *D* outputs $z_{dj},\left(d=1,2,\ldots,D\right)$ from that stage. These *D* outputs then become the inputs to the second stage and are referred to as intermediate measures. The outputs from the second stage are $y_{rj},\left(r=1,2,\ldots,R\right)$. We denote for each DMU_j the efficiency for the first stage as e_j^{S1} and the second as e_j^{S2}. On the basis of the radial (CRS - constant return to scale) input-oriented DEA model we define:

$$e_j^{S1}=\frac{\sum_{d=1}^{D}w_d\,z_{dj}}{\sum_{i=1}^{M}v_i\,x_{ij}}\quad\text{and}\quad e_j^{S2}=\frac{\sum_{r=1}^{R}u_r\,y_{rj}}{\sum_{d=1}^{D}\tilde{w}_d\,z_{dj}},\quad j=1,2,\ldots,N, \tag{1}$$

where v_i, w_d, $\tilde{w}_d$ and u_r are unknown non-negative weights. It is assumed that $w_d=\tilde{w}_d$ for all *d*. If the *input-oriented DEA model* is used, then the two-stage **overall cooperative efficiency** is defined as decomposition of $e_o^{S1,coop}$ and $e_o^{S2,coop}$:

$$e_o^{S,coop} = e_o^{S1,coop} \cdot e_o^{S2,coop} = \max\left(e_j^{S1} \cdot e_j^{S2}\right) = \frac{\sum_{r=1}^{R} u_r\, y_{ro}}{\sum_{i=1}^{M} v_i\, x_{io}} \tag{2}$$
$$s.t.\; e_j^{S1} \le 1, \quad e_j^{S2} \le 1, \quad j = 1, 2, \ldots N.$$

As is indicated in [4], the centralized model (2) does not provide information on DEA projection. They developed models for determining the DEA projection for the inefficient DMUs. The input-oriented **projection frontier** (PF) model can be expressed as the dual model:

$$\begin{aligned} \min \quad & \tilde{\theta} \\ s.t. \quad & \sum_{j=1}^{N} \lambda_j\, x_{ij} \le \tilde{\theta} x_{io}, \quad i = 1, 2, \ldots, M, \\ & \sum_{j=1}^{N} \mu_j\, y_{rj} \ge y_{ro}, \quad r = 1, 2, \ldots, R, \\ & \sum_{j=1}^{N} \lambda_j\, z_{dj} \ge \tilde{z}_{do}, \quad d = 1, 2, \ldots, D, \\ & \sum_{j=1}^{N} \mu_j\, z_{dj} \le \tilde{z}_{do}, \quad d = 1, 2, \ldots, D, \\ & \tilde{z}_{do} \ge 0, \quad d = 1, 2, \ldots, D, \quad \tilde{\theta} \le 1, \lambda_j \ge 0, \mu_j \ge 0, \quad j = 1, 2, \ldots, N, \end{aligned} \tag{3}$$

where $\tilde{z}_{do}$ represents the set of new intermediate measures to be determined as "outputs" in the third set of constraints and also as "inputs" in the fourth set of constrains.

3 Empirical analysis

The scientific DEA literature that is dedicated to the efficiency evaluation of banks is based on **three approaches**: production approach, profitability approach and intermediate approach. These approaches are related to selection of inputs and outputs of the production units. The *production approach* is based on the assumption that banks provide services to customers. The outputs of the production units are services provided to customers that are represented by the number and the type of processed documents, transactions or special provided services over the considered period. The inputs are labour, materials, space and information systems expressed in terms of physical units or associated cost. The disadvantage of this approach is the focus only on the operating cost and ignorance of the interest expenses. The *intermediate approach* recognizes a bank as an intermediary in the capital market that transforms deposits, fixed assets and labour to create loans, interest and non-interest incomes (Chiu 2016). The *profitability approach* measures a bank profitability based on expenses as inputs and revenues as outputs. It is used the production and intermediation approaches in this study. Deposits are treated as outputs under the *production approach*, where banks want more and better deposits, whereas deposits are treated as inputs under the intermediation approach, where banks want fewer and better deposits. This research employs a two-stage network system to address this issue, in which deposits (*intermediate products*) are regarded as outputs for the first stage and inputs of the second stage according to [8].

The data samples in this empirical research were extracted from the Bank Scope – Bureau van Dijk database[2]. It was selected the sample of 62 financial institutions, mainly banks of V4 countries, according to the availability of all data for all institutions in the year 2013. Accordingly, for the first stage (S1), three inputs were chosen: personnel expenses (PE, salary expenses for all officers and employees of a bank), fixed assets (FA) and operational expenses (OE, excluded the salary and lease expenses for supporting the operation of the bank). Deposits (DEP) were as the intermediate output from the first stage and also as the input to the second stage (S2). The desirable outputs from the second stage were short-term, medium and long-term loans and advances to banks (LOA), securities investment revenues (SEC) and non-interest incomes on non-earning assets (NII). All variables were in thousand Euros.

[2] https://bankscope.bvdinfo.com/version-2016419/home.serv?product=scope2006

It was provided the two-stage performance measurement of the banks using **overall efficiency** (S_EFF) including intermediate measure. It was also estimated **black-box efficiency** (BB_EFF) for measuring bank efficiency of the two-stage process excluding intermediate measure i.e. as the one-stage process. It was also decomposed the overall efficiency into **production efficiency** (S1_EFF) for the first stage and **operational efficiency** (S2_EFF) for the second stage. All our optimization models were solved in GAMS software.

Efficiencies of 62 banks were estimated by both the one-stage DEA input- oriented model and also by the two-stage input-oriented DEA model in equation (2). The main statistics are given in Table 2 and the results are also illustrated in Figure 2 using boxplots.

	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25	50	75
BB_EFF	0.318	0.273	0.026	1.000	0.146	0.230	0.350
S1_EFF	0.296	0.250	0.058	1.000	0.151	0.192	0.317
S2_EFF	0.650	0.213	0.265	1.000	0.503	0.606	0.836
S_EFF	0.211	0.230	0.023	1.000	0.086	0.126	0.236

Table 1 Statistics of system efficiencies and efficiencies decomposition of two-stage processes.

The obtained results show that only two Czech banks have the best management performance within the V4 countries in 2013 i.e. B29 (PPF banka) and B31 (Czech Moravian Guarantee and Development Bank). It is marked these banks as the benchmark financial institutions in year 2013 in the Visegrad countries. The other financial institutions have the overall technical efficiency in the average score 0.21. In terms of production efficiency, there are only 4 financial institutions efficient (in addition to the two above mentioned banks – B30 (Stavební Sporitelna Ceské Sporitelny) and B36 (Raiffeisen stavební sporitelna)) however they have operational efficiency 0.76 which is higher than median 0.61. From the aspect of operational efficiency, the efficient banks excluding overall efficient banks include further 3 banks: B15 (Bank BPH), B53 (Privatbanka) and B60 (Evropsko-ruska banka), indicating that these banks perform better in intermediary services, but poorly in the production stage in range 0.04 – 0.33. As shown in Figure 2 operational efficiency (S2_EFF) is on average 0.65 with more efficient but with a little higher variability among banks in comparison with average productive efficiency 0.30.

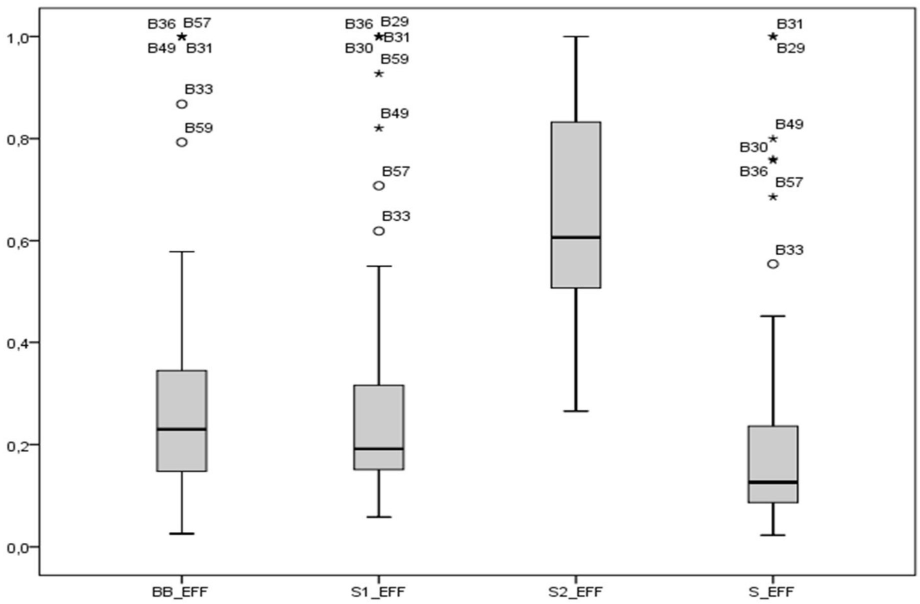


Figure 2 Boxplots of efficiencies of two-stage process.

This results show that we have only 2.3 % banks that are overall efficient in year 2013. The average efficiency of the overall stages is relatively very low (0.21). The sources of inefficiency arise from the first stage, deposit-generating stage, suggesting that the improvement of the intermediate output - raised fund (deposits) is necessarily. Figure 3 compares kernel density of deposits with kernel density of optimal intermediate measures. The results suggest to increase deposits in the production stage for the banks with the level of deposits up to median

(2 900 000 thousand Euro), focusing particularly on financial institutions, which is the value of deposit around 1st quartile (800 000 thousand Euro).

DEA became a main stream technology in bank studies on recent years. The methodology improvements include multi-dimensional and dynamic performance evaluation, dealing with non-homogenous financial institutions, relaxing the assumption of concavity for the production frontier, using undesirable outputs, using various optimization models etc. Fukuyama and Weber in [8] also found that two-stage network system is superior to traditional DEA in the banking sector and confirmed that the traditional DEA approach overestimate efficiency scores that is same conclusion as in paper [10]. The literature [2] investigated some Chinese commercial banks during the period 2008-2012 a new two-stage DEA approach. Their empirical results show that during the five years the average efficiencies of the banks were improved year by year because of performance improvement of their deposit-utilizing stage. This suggests that the managers of these banks should pay more attention to the relatively low efficient deposit-generating stage. The banks should improve their ability to absorb more deposits by applying more attractive policies. Moreover, the benchmarks for the banks prove that our approach is very useful for banks to reduce the gap between their intermediate measures in two stages. It means, the results provide a good alternative for well coordinating the activities of the two stages in order to reduce resource waste.

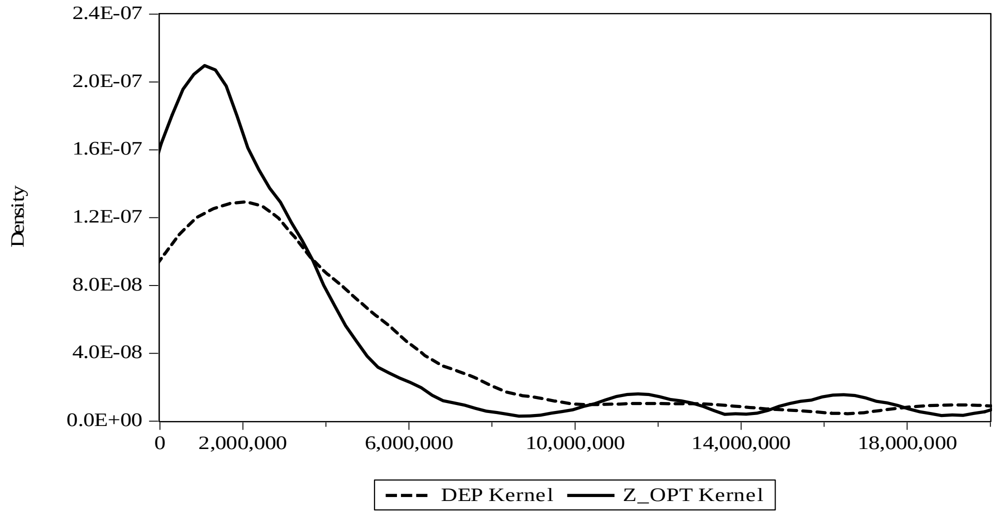


Figure 3 Kernel density of the deposits (DEP Kernel) and optimal intermediate measures (Z_OPT_Kernel).

4 Conclusions and remarks

In this paper a two-stage network technology is used for our Visegrad financial institutions where deposits and raised funds are produced as intermediate outputs in one stage of production and then used as input to produce final outputs of loans and securities investments in the second stage. The two-stage input-oriented DEA model yields an efficiency indexes that evaluate the benchmarking of performance of the financial institutions and also decompose the overall technical efficiency into the production and the operational efficiency for each units.

Based on the overall efficiency scores only two banks were identified as the best-practice. The estimated overall technical efficiency of the two-stage process in the financial institutions indicated that on average, deposits are under-produced in the first stage of production. Inefficient banks produced inside the production possibility set and should be able to simultaneously expand loans and securities given inputs and an appropriate choice of deposits if they were to become efficient. As a consequence, the underproduction of deposits resulted in fewer loans and securities investments. The results also found that two-stage network system is superior to traditional DEA in the banking sector and confirmed that the traditional DEA approach overestimates efficiency scores. The results are useful for efficiency improvement of inefficient financial institutions, because they help to identify inefficiencies arising from the internal system structure of each bank.

One direction for future research would be to extend the two-stage network model to a dynamic setting. In such a model, bank managers could choose to forego production of final outputs in a period and save resources for use in subsequent periods. For instance, rather than using deposits to make loans in the current period, those deposits might be carried over to the next period when economic conditions might allow even more loans to be made with fewer of those loans becoming non-performing. Additional future research directions could also include a more sophisticated performance measurement system as was suggested in the paper Toloo and Tichy [17]. Another extension of this research is to build open two-stage DEA model with undesirable variables using a slacks-based modelling.

Acknowledgements

This article was supported by the Czech Science Foundation (GACR project 16-17810S), European Social Fund within the project CZ.1.07/2.3.00/20.0296 and also through the Student Grant Competition project (SP2016/116) of Faculty of Economics, VŠB-Technical University of Ostrava.

References

[1] Akeem, U.O. and Moses, F.: An empirical Analysis of Allocative Efficiency of Nigerian Commercial Banks: A DEA Approach. *International Journal of Economics and Financial Issues* **4** (2014), 465-475.
[2] An, Q., Chen, H., Wu, J. and Liang, L.: Measuring slacks-based efficiency for commercial banks in China by using a two-stage DEA model with undesirable output. *Annals of Operations Research* **235** (2015), 13-35.
[3] Charnes, A., Cooper, W.W. and Rhodes, E.: Measuring the efficiency of decision-making units. *European Journal of Operational Research* **2** (1978), 429-444.
[4] Chen, Y., Cook, W.D. and Zhu, J.: Deriving the DEA frontier for two-stage processes. *European Journal of Operational Research* **202** (2010), 138-142.
[5] Chen, Y., Li, Y., Liang, L., Salo, A. and Wu, H.: Frontier projection and efficiency decomposition in two-stage processes with slacks-based measures. *European Journal of Operational Research* **250** (2016), 543-554.
[6] Chiu, C.R., Chiu, Y.H., Chen Y.C. and Fang, C.L.: Exploring the source of meta-frontier inefficiency for various bank types in the two-stage network system with undesirable output. *Pacific-Basin Finance Journal* **36** (2016), 1-13.
[7] Cook, W.D., Liang. L. and Zhu J. Measuring performance of two-stage network structures by DEA: A review and future perspective. *Omega* **38** (2010), 423-430.
[8] Fukuyama, H. and Weber, W.L.: Measuring Japanese bank performance: a dynamic network DEA approach. *Journal of Productivity Analysis* **44** (2015), 249-264.
[9] Hanclova, J. and Chytilova, L.: The impact of inclusion of non-traditional activities on efficiencies in European banking industry using the CCR-I model. In: *Proceedings of the 11th International Conference - Strategic Management and its Support by Information Systems 2015* (Nemec, R. and Zapletal, F., eds.). VŠB - Technical University of Ostrava, Ostrava, 2015, 208-220.
[10] Kao, C. and Hwang, S.N.: Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. *European Journal of Operational Research* **185** (2007), 418-429.
[11] Liang, L., Cook, W.D. and Zhu, J.: DEA models for two-stage processes: game approach and efficiency decomposition. *Naval Research Logistics* **55** (2008), 643-653.
[12] Melecky. L. and Stanickova, M.: The Competitiveness of Visegrad Four NUTS 2 Regions and its Evaluation by DEA Method Application. In: *Proceedings of the 29th International Conference on Mathematical methods in Economics* (Dlouhy, M. and Skocdopolova,V., eds.). Czech Society for Operations Research, Prague, 2011, 479-479.
[13] Pancurova, D. and Lyocsa, S.: Determinants of Commercial Banks' Efficiency: Evidence from 11 CEE Countries. *Finance a úvěr-Czech Journal of Economics and Finance* **63** (2013), 152-178.
[14] Repkova, I.: Efficiency of the Czech banking sector employing the DEA window analysis approach. *Procedia Economics and Finance* **12** (2014), 587-596.
[15] Seiford, L.M. and Zhu, J.: Profitability and marketability of the top 55 US commercial banks. *Management Science* **45** (1999), 1270-1288.
[16] Svitalkova, Z.: Comparison and evaluation of bank efficiency in selected countries in EU. *Procedia Economics and Finance* **12** (2014), 644-653.
[17] Toloo, M. and Tichy, T.: Two alternative approaches for selecting performance measures in data envelopment analysis. *Measurement* **65** (2015), 29-40.
[18] Wang. C.H., Gopal, R.D. and Zionts, S.: Use of Data Envelopment Analysis in assessing Information Technology impact on firm performance. *Annals of Operations Research* **73** (1997), 191-213.

Evaluation of Project Investments Based on Comprehensive Risk

Simona Hašková[1]

Abstract. Investment in a business project is an investment in the future, which is uncertain. Different risks are associated with uncertainty. This paper focuses on the investment risk. To minimize it the evaluation of projects is used based on the expected net present value of cash flows, which reflects two components of investment risk: 1) the possibility of different results; 2) the fact that an income of a particular project fluctuates randomly around its expected value. Both of these components relate to the economic side of things.

This contribution is an extension of the standard investment risk for an additional non-economic component having the character of a political or environmental risk, and the transition from the investment risk to the comprehensive risk. As a part of a comprehensive risk an investment becomes a bet in the investment lottery differentiating from the standard lottery by a low probability not to win anything. The aim of this paper is to show how the transition from investment risk to comprehensive risk changes the standard decision criteria for accepting projects.

Key words: Investment risk, comprehensive risk, adjusted decision criterion, project as an investment lottery.

JEL Classification: C02, C60
AMS Classification: 91B06, 91B50

1 An introduction to comprehensive risk

The extent to which the reality confirms the expected results of an investment in a business project will be known in the near or distant future, which is always *uncertain*. The *quantity* of this uncertainty relates to the density distribution of objective or subjective probabilities of possible outcome occurrences. The *quality* of uncertainty is usually associated with the degree of completeness and the veracity of its description. The key problem here is the adequacy of available information about the state of affairs of the real world and estimates of the future effects of today´s decisions.

The quantitative and qualitative aspect of uncertainty is subsequently reflected in the degree of investment risk occurring in connection with investments in business projects or plans, which is a part of a broader economic risk. In the field of managerial decision-making this issue has been solved by different methodologies applied in various fields – e.g. by stochastic networks shown in [6] and [7], sequential decision trees applied in [8] and [10], probabilistic decision trees demonstrated in [5].

Forlani and Mullins [2] were engaged with the project investment risk in detail; their approach was built on the work of Sitkin and Pablo [9] and enriched with the component of subjective risk perception, risk characteristics, contextual effects and expectations that play a key role in deciding on the acceptance or rejection of new projects.

Standard investment risk consists of two components: within the *first* one, it is expected that the project realization can take place according to one of several possible scenarios; however, we do not know in advance which one it will be. From the perspective of the company the sources of this investment risk are the *external* effects that the project implementer (entrepreneur) cannot exclude from the game. The *second* component of investment risk stems from the fact that an entrepreneur cannot control everything he could exclude from the game; its sources are predominantly *internal* effects associated with the substance of the project and the quality of its management. Therefore, the actual present value of the project income Π will randomly vary in time around its expected value E[PV(R)] (see Figure 1). The greater the degree of fluctuation the greater a component of investment risk originating from this inner source will be. The degree of fluctuation is reflected in the discount rate r of an analyzed project within the standard models (in more details in Brealey et al. [1] or Hašková et al. [3]).

[1] The Institute of Technology and Business / School of Expertness and Valuation, Okružní 517/10, 370 01 České Budějovice, CZ, e-mail: haskovas@post.cz

In this paper the standard investment risk is extended to another ingredient of a non-economic nature having the character of political or environmental risk (wars, expropriation, sanctions, natural disasters, etc.) due to which an investor can lose all the funds invested in the project. The inclusion of this component enables us move from the standard investment risk to comprehensive risk.

Within comprehensive risk the investment in the project becomes a bet in the investment lottery differentiating from the standard lottery only by the fact that the probability not to win anything is low (a general mathematical model of an investment lottery can be found in [11]). The aim of this paper is to show how the transition from investment risk to comprehensive risk changes the standard decision criteria for project acceptance.

2 A project as an investment lottery

By incorporating the non-economic component into the model of comprehensive risk we get an investment lottery presented graphically in Figure 1.

In it, $E[PV(R)] = \Sigma q_i \cdot \Sigma PV(R_{ij})$, $E[PV(I)] = \Sigma q_i \cdot \Sigma PV(I_{ij})$,
$E[NPV] = \Sigma q_i \cdot NPV_i = \Sigma q_i \cdot (\Sigma\, PV(R_{ij}) - \Sigma\, PV(I_{ij})) = \Sigma q_i \cdot \Sigma\, PV(R_{ij}) - \Sigma q_i \cdot \Sigma\, PV(I_{ij})$,

$$E[NPV] = E[PV(R)] - E[PV(I)] \tag{1}$$

In parameters $PV(R_{ij}) = R_{ij} / (1 + r)^j$ and $PV(I_{ij}) = I_{ij} / (1 + r)^j$, the variable Rij, or respectively I_{ij}, designates the current value of income R, or respectively of the investment I, in a scenario of a number i in the j-th year of the project lifetime, q_i is the probability of occurrence of a scenario of a number i, the fraction $1 / (1 + r)^j$ is the discount rate transforming the current value of the respective payment of the j-th year to the present time and NPV is net present value of the project.

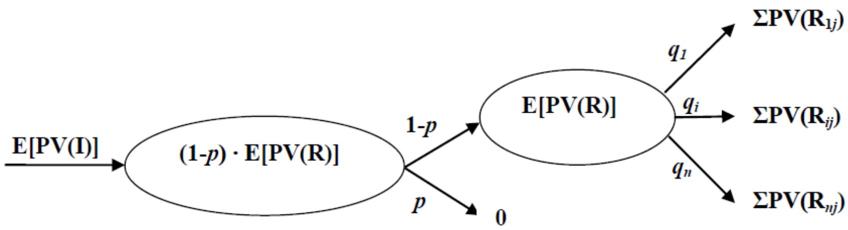


Figure 1 Scheme of an investment lottery based on comprehensive risk

In the model of investment lottery in Figure 1 the possible winnings, recorded in the leaves of the upper branch of the tree, are presented in current values of income ($E[PV(R_{ij})]$) generated by the respective scenarios of the project. The bottom branch shows the possibility of the investor not to win anything. If we subtract the current value of investment ($E[PV(I_{ij})]$) in every leaf of this tree that is substantially related to it, we get the tree in Figure 2.

Figure 2 shows a lottery model, in which an investor buys a lottery ticket whose price (investment) is I_{ij} within the current scenario i in the j-th year of the project lifetime according to the business schedule and obtains the income R_{ij} generated by the project Π. With this lottery ticket the investor either wins nothing with the probability p or wins the amount E[PV(R)] with the probability $1 - p$. As already mentioned, the probability value p in investment lotteries is substantially lower, unlike standard lotteries.

The project is successful if the expected present value of incomes generated by the project covers at least the expected present value of the investments spent on the project - the coverage of the expected capital input by the expected income output. For the project to be a success it then applies that the condition of project acceptability in form (2) has to be satisfied (see also Figure 2)

$$p \cdot 0 + (1 - p) \cdot E[PV(R)] = (1 - p) \cdot E[PV(R)] \geq E[PV(I)] \tag{2}$$

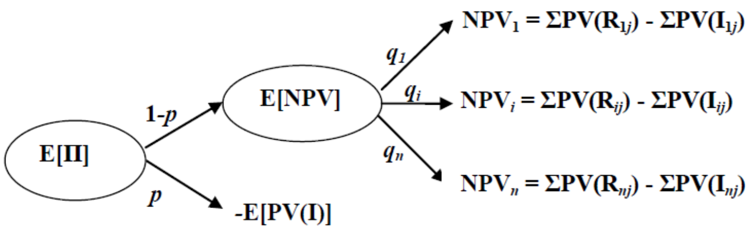


Figure 2 Modification of the tree from Figure 1 for the transition from the payouts to the profits

The variable

$$E[\Pi] = (1 - p) \cdot E[NPV] - p \cdot E[PV(I)] \tag{3}$$

at the root of the modified tree can be then interpreted as the expected present value of the investor´s profit from the project investment in terms of comprehensive risk.

3 Investment decision in terms of comprehensive risk

Multiplying the values of the expression on the left side of the condition for project acceptability (2) and after further adjustment we get E[PV(R)] – E[PV(I)] ≥ p · E[PV(R)], where E[PV(R)] – E[PV(I)] = E[NPV] is the expected present value of the project Π in terms of a standard investment risk (see Figure 3). From here

$$E[NPV] \geq p \cdot E[PV(R)] \tag{4}$$

or by substituting E[PV(R)] = E[NPV] + E[PV(I)] into the expression (4) and by further adjustment

$$E[NPV] \geq p \cdot E[PV(I)] / (1 - p). \tag{5}$$

Given that the value of expected returns E[PV(R)] is estimated with more difficulty than the value of the considered investment E[PV(I)], the latter relationship is preferable in practice.

Lemma:
Let us have two projects Π_1 and Π_2 with the same value E[NPV]. Let p_1 and p_2 be the probabilities of the occurrence of non-economic risk due to which the investor may lose everything he invested in the project. Then the following applies:

$$\text{if } \boldsymbol{p_1} \cdot E[PV(R)]_1 \leq \boldsymbol{p_2} \cdot E[PV(R)]_2, \text{ then } E[\Pi_1] \geq E[\Pi_2] \tag{6}$$

Proof:
$p_1 \cdot E[PV(R)]_1 \leq p_2 \cdot E[PV(R)]_2,$
$p_1 \cdot (E[NPV] + E[PV(I)]_1) \leq p_2 \cdot (E[NPV] + E[PV(I)]_2),$
$-p_1 \cdot E[NPV] - p_1 \cdot E[PV(I)]_1 \geq -p_2 \cdot E[NPV] - p_2 \cdot E[PV(I)]_2,$
$E[NPV] - p_1 \cdot E[NPV] - p_1 \cdot E[PV(I)]_1 \geq E[NPV] - p_2 \cdot E[NPV] - p_2 \cdot E[PV(I)]_2,$
$(1 - p_1) \cdot E[NPV] - p_1 \cdot E[PV(I)]_1 \geq (1 - p_1) \cdot E[NPV] - p_2 \cdot E[PV(I)]_2,$
$E[\Pi_1] \geq E[\Pi_2].$

From this it follows that the inclusion of a non-economic component to the comprehensive risk leads to the generalization of the standard decision criterion for the project acceptance from E[NPV] ≥ 0 to E[NPV] ≥ p · E[PV(R)], or respectively E[NPV] ≥ p · E[PV(I)] / (1 – p).

As we see we can get its original form (E[NPV] ≥ 0) if we completely disregard the non-economic component of risk (i.e. if p = 0). We also see that while the original criterion in the case of E[NPV] = 0 does not say anything about acceptance or rejection of the project, the generalized criterion in this case clearly refuses to accept it (the acceptability condition (2) cannot be met).

The rules for deciding between the two projects Π_1 and Π_2 are then modified as follows:

I. Every project that satisfies the condition (4) or (5) is acceptable, i.e. the project E[NPV] for which applies that E[NPV] ≥ p · E[PV(R)] or E[NPV] ≥ p · E[PV(I)] / (1 – p).
II. If both projects are acceptable priority is given to the project with the higher E[Π] value according to (3).
III. If both projects are acceptable with identical E[NPV] then (in accordance with Lemma (6)) priority is given to the project with a lower p · E[PV(R)] value.

4 Application of an approach based on comprehensive risk in the case study

The managerial study on „*Deciding between using the classical method of exploratory wells and seismic survey of a new oil field*“ presented in [4] builds on the discovery of a huge oil deposit along the African coast, confirmed in 2009 by a consortium of energy companies led by the US Anadarko Pertroleum. Within the framework project a simple model is derived of dependence of forecasted net present value (NPV) of deposit exploration and oil extraction on the choice of an exploration method used (exploratory wells - the project Π_V, seismic survey - project Π_S) and on the costs of mining rights TP. The graphic form of this dependency is shown in Figure 3.

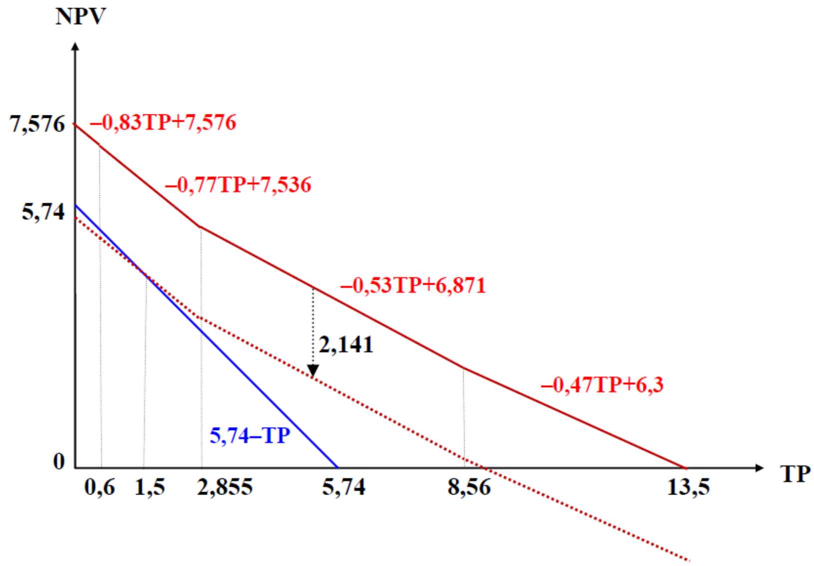


Figure 3 The graph of NPV dependence of projects Π_V and Π_S on the costs of mining rights TP (in mil. USD)

In it, the blue line (NPV = 5,74 – TP) corresponds to project Π_V, the red dotted line corresponds to project Π_S (for which in the interval $0{,}6 \leq TP \leq 2{,}855$ applies NPV = 7,536 – 2,141 – 0,77 · TP; the amount of 2,141 mil. USD is the seismic exploration cost). From the graph we see that in the case of TP = 1,5 the projects Π_V and Π_S have the same NPV = 5,74 – 1,5 = 7,536 – 2,141 – 0,77 · 1,5 = 4,24. Within the standard risk it is an impasse in which it is impossible to decide which of the projects to accept.

If we consider, however, that both the projects are intended to operate in the same socially unstable area with an evident non-economic risk $p_v = p_s = p$ reaching the level of units, then with respect to the amount of the contemplated investments $E[PV(I)]_V$ and $E[PV(I)]_S$, the meeting of criterion (5) can be expected and therefore to make a decision according to rule III. of section 3, i.e. according to the mutual ratio of the values $E[PV(R)]_V$ and $E[PV(R)]_S$. Then:

- $E[PV(R)]_V = E[NPV] + E[PV(I)]_V = 4{,}24 + TP + p_1 \cdot c_1 + p_2 \cdot c_2 + p_3 \cdot c_3 + C = 4{,}24 + 1{,}5 + 0.7 \cdot 2 + 0.2 \cdot 2{,}5 + 0.1 \cdot 3 + C = 7{,}94 + C$, where p_i is the probability for the requirement of oil well exploration at i thousand feet depth, c_i stands for the costs of the exploratory well, C is the price of the mining equipment (which is the same in both projects),
- $E[PV(R)]_S = E[NPV] + E[PV(I)]_S = 4{,}24 + TP + P + C = 4{,}24 + 1{,}5 + 2{,}141 + C = 7{,}881 + C$, where P is the cost of the seismic survey.

Since $E[PV(R)]_S \leq E[PV(R)]_V$, the seismic survey is preferable to the exploratory well.

5 Conclusion

The paper focused on the specification of investment risk and its extension to a non-economical component. This component was reflected in the decision criterion of the expected net present value E[NPV]. Each entrepreneur who adopts a project, the outcome which is uncertain, always undergoes an investment risk. The standard approach to its evaluation aimed at assessing the project profitability leans on the quantification of E[NPV]. This, in its standard form, takes into account both the possibility of different results due to the existence of different scenarios under which the project can take place, and the fact that the actual income of the project usually fluctuates randomly to a varying degree around its expected value. Both these components have an economic aspect.

The non-economic component of investment risk includes the political or environmental risk, due to which an entrepreneur may lose all funds invested in the project. The inclusion of this component led us to the transformation of the standard form of investment risk into the form of comprehensive risk, namely, to the modification of the decision criterion E[NPV]. The nature of the comprehensive risk lies in viewing the investment as a lottery. In it the entrepreneur buys a „lottery ticket“ at the expense of an investment with which, depending on the probabilities, he/she either loses everything or wins an amount equal to the expected net income (see conditions (1), (2), (3) graphically captured in Figure 2 and 3)), while the chance of losing everything is significantly lower compared to a classical lottery.

Given the fact that the value of the expected revenue is estimated with more difficulty than the value of the contemplated investment, practical decision-making taking into account the existence of comprehensive risk, leans primarily on condition (5). In the case of deciding between two projects with different expected profits Π_1 and Π_2, the existence of a non-economic risk p_1 and p_2 and with the same E[NPV] the project acceptability is defined by the conditions I., II. and III. They were derived from the proof of validity of the statement that if $p_1 \cdot \mathrm{E[PV(R)]}_1 \leq p_2 \cdot \mathrm{E[PV(R)]}_2$ is true, then also $\mathrm{E}[\Pi_1] \geq \mathrm{E}[\Pi_2]$ is true, where E[PV(R)] stands for the expected value of the income of project 1, or respectively project 2.

The procedure to the project evaluation based on comprehensive risk was applied in the managerial study, the aim of which was to decide on the choice of the optimal method of potential oil field exploration.

References

[1] Brealey, R. A., Myers, S. C., and Marcus, A. J.: *Fundamentals of Corporate Finance*. Mcgraw-Hill Education, New York, 2011.

[2] Forlani, D., and Mullins, J. W.: Perceived risks and choices in entrepreneurs' new venture decisions. *Journal of Business Venturing* **15**, 4 (2000), 305-322.

[3] Hašková, S., Chládek, I., and Kolář, P.: Contribution to the Formulation of Economically Efficient Subsidy Policy in the Area of Small Hydro Power Plants. *Littera Scripta* **7**, 1 (2014), 25-38.

[4] Hašková, S., and Kolář, P.: The mathematic modeling of process economization of natural resources. In: Proceedings of the 29th International Conference Mathematical Methods in Economics 2011 (Dlouhý, M., and Skočdopolová, V., eds.). Professional Publishing, Prague, 2011, 224-229.

[5] Kotsiantis, S. B.: Decision trees: a recent overview. *Artificial Intelligence Review* **39**, 4 (2013), 261-283.

[6] Neely, M. J.: Stochastic network optimization with application to communication and queueing systems. *Synthesis Lectures on Communication Networks* **3**, 1 (2010), 1-211.

[7] Puterman, M. L.: *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, New Jersey, 2007.

[8] Santos, J. R., Barker, K., and Zelinke I. P. J.: Sequential decision-making in interdependent sectors with multiobjective inoperability decision trees: application to biofuel subsidy analysis. *Economic Systems Research* **20**, 1 (2008), 29-56.

[9] Sitkin, S. B., and Pablo, A. L.: Reconceptualizing the determinants of risk behavior. *Academy of management review* **17**, 1 (1992), 9-38.

[10] Tan, B., and Srikant, R.: Online advertisement, optimization and stochastic networks. *IEEE Transactions on Automatic Control* **57**, 11 (2012), 2854-2868.

[11] Varian, H. R.: *Microeconomic Analysis*. W. W. Norton & Company Inc., New York, 1992.

On Comparing Various EWMA Model Estimators: Value at Risk Perspective

Radek Hendrych[1], Tomáš Cipra[2]

Abstract. The exponentially weighted moving average (EWMA) model is a particular modelling scheme supported by RiskMetrics that is capable of predicting the current level of financial time series volatility. It is designed to track changes in conditional variance of financial returns by assigning exponentially decreasing weights to the observed past squared measurements. Recently, several on-line (i.e. recursive) estimation techniques suitable for this class of stochastic models have been introduced. These methods undoubtedly represent attractive alternatives to the common identification and calibration procedures (i.e. off-line or batch); they can estimate and control the process behaviour in real time. The aim of the paper is to examine different EWMA model estimators by using financial data. For instance, one might consider the Value at Risk (VaR) backtesting approach since Value at Risk predictions are relevant outputs of the RiskMetrics EWMA modelling framework (especially from the practical point of view). Therefore, various VaR backtests can be used to study the adequacy of different EWMA model estimators.

Keywords: EWMA model, recursive estimation, RiskMetrics, Value at Risk.

JEL classification: C51
AMS classification: 62M10

1 Introduction

The exponentially weighted moving average (EWMA) model is a modelling scheme preferred by RiskMetrics to evaluate conditional heteroscedasticity (see e.g. [9]). This concept is frequently linked to investigating financial time series, more specifically to monitoring volatility (i.e. the conditional standard deviation of financial returns). The EWMA model has been primarily developed as a simple alternative to the GARCH models. The name of this concept originates from the fact that the conditional variance is an exponentially weighted sum of historical squared financial returns with the geometrically declining weights going backwards in time. Therefore, this model is easily capable of tracking changes in the conditional variance and volatility. Since its introduction in [9], it has been investigated from various theoretical and practical perspectives with many successful empirical applications. For example, one may employ the EWMA model to predict volatility, to calculate Value at Risk, or to define trading rules. Moreover, the RiskMetrics EWMA framework is regarded as the benchmark by many practitioners.

The only unknown parameter of the EWMA model that determines the geometrically declining weights is conventionally prescribed by experts or users. Alternatively, it can be estimated by employing common (i.e. off-line or batch) statistical identification procedures (e.g. the conditional maximum likelihood method). However, it is only exceptionally estimated recursively (i.e. on-line). On the other hand, it might be advantageous to apply a numerically effective technique that would estimate and control the model parameter (and consequently the model behaviour) in real time. Recently, several recursive estimation schemes suitable for this class of stochastic models have been introduced (see [4] or [5]). These methods undoubtedly represent attractive alternatives to the conventional approaches. In addition, they have demonstrated their adequacy (refer to [4]).

[1]Charles University in Prague, Faculty of Mathematics and Physics, Dept. of Probability and Mathematical Statistics, Sokolovská 83, 186 75 Prague 8, Czech Republic, hendrych@karlin.mff.cuni.cz.
[2]Charles University in Prague, Faculty of Mathematics and Physics, Dept. of Probability and Mathematical Statistics, Sokolovská 83, 186 75 Prague 8, Czech Republic, cipra@karlin.mff.cuni.cz.

The aim of this contribution is to examine different EWMA model estimators by using real financial data. Our attention is concentrated primarily on investigating the recursive estimation techniques. It is motivated by recently published works (see [4] and [5]), where the qualities of recursive estimation algorithms were studied from other points of view. In particular, we shall consider the Value at Risk (VaR) backtesting approach as was suggested in [10] to study the adequacy of the different EWMA model estimators. From the practical point of view, the Value at Risk predictions are relevant outputs of the RiskMetrics EWMA modelling framework (consult e.g. with [9]). Although the Value at Risk suffers from several disadvantages (see e.g. [6]), it still remains popular and widely applied.

This paper is organized as follows. Section 2 reviews the RiskMetrics EWMA modelling framework and its fundamental features. Section 3 briefly discusses the various estimation procedures investigated in this work: (i) the recursive estimator originally introduced in [4] and (ii) the rolling-window estimator based on the conditional log-likelihood criterion. Section 4 shortly introduces the VaR backtesting methodology, which is applied here for evaluating the adequacy of different estimation alternatives. Section 5 analyses these approaches by employing rigorous statistical tests based on the calculated Values at Risk corresponding to various financial datasets. Conclusions are summarized in Section 6.

2 RiskMetrics EWMA modelling framework

Formally, the RiskMetrics EWMA model of financial returns $\{y_t\}_{t\in\mathbb{Z}}$ is commonly defined as (see [9]):

$$y_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = (1-\lambda) y_{t-1}^2 + \lambda \sigma_{t-1}^2, \tag{1}$$

where the only modelling parameter λ lies in the interval $(0, 1)$, $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a sequence of i. i. d. random variables with the standard normal distribution $\mathrm{N}(0, 1)$, and σ_t^2 is $\mathcal{F}_{t-1}$-measurable. Remind that $\mathcal{F}_t$ denotes the smallest σ-algebra with respect to which y_s is measurable for all $s \leq t$. Apparently, the positivity of σ_t^2, i.e. the conditional variance of y_t given $\mathcal{F}_{t-1}$, is ensured by using (1).

The one-step ahead prediction of σ_t^2 is expressed as:

$$\sigma_{t+1|t}^2 := \mathbb{E}(\sigma_{t+1}^2 | \mathcal{F}_t) = (1-\lambda) y_t^2 + \lambda \sigma_t^2 = \sigma_{t+1}^2. \tag{2}$$

The one-step ahead conditional Value at Risk (VaR) measure for a nominal coverage rate $\alpha \in (0, 1)$ denoted by $VaR_{t+1|t}(\alpha)$ is generally defined as:

$$\mathbb{P}\left[y_{t+1} < -VaR_{t+1|t}(\alpha) | \mathcal{F}_t\right] = \alpha \text{ for all } t \in \mathbb{Z}. \tag{3}$$

In the RiskMetrics modelling framework, this relation is reduced to:

$$VaR_{t+1|t}(\alpha) = -\Phi^{-1}(\alpha) \cdot \sigma_{t+1} \text{ for all } t \in \mathbb{Z}, \tag{4}$$

where $\Phi^{-1}(\cdot)$ denotes the quantile function of the standard normal distribution $\mathrm{N}(0, 1)$.

3 Different estimation methods for the RiskMetrics EWMA model

To estimate (calibrate) the EWMA model (1) and to calculate the Value at Risk according to (4), we shall employ the following methods: (i) the value of λ is prescribed (e.g. the choice 0.94 is obviously recommended for daily data in [9]), (ii) λ is estimated by the rolling-window estimator based on maximizing the conditional logarithmic likelihood function assuming that y_0 and σ_0^2 are either defined or observed, and (iii) λ is estimated by the one-stage self-weighted recursive estimation algorithm introduced in [4]. The adequacy of these estimators will be compared in the VaR backtesting framework based on empirical financial data, see Sections 4 and 5.

Firstly, we shall briefly review the one-stage self-weighted recursive estimation algorithm that controls the EWMA modelling parameter in real time (refer to [4]). In many instances, this approach may be advantageous. Recursive estimation methods are generally very effective in terms of memory storage and computational complexity since the current parameter estimates are obtained by using the previous estimate and actual measurements. Moreover, they can be employed to detect structural changes.

Applying the general recursive prediction error method (see e.g. [8]), one can derive a recursive scheme for on-line estimating the parameter λ of the EWMA model (1). Note that the negative conditional log-likelihood function is supposed as the core criterion that is minimized recursively. The resulting algorithm can be formulated as follows:

$$
\begin{aligned}
\widehat{\lambda}_t &= \widehat{\lambda}_{t-1} + \frac{\widehat{p}_{t-1}(y_t^2 - \widehat{\sigma}_t^2)\widehat{\sigma}_t^{2'}}{\gamma_t(\widehat{\sigma}_t^2)^2 + (\widehat{\sigma}_t^{2'})^2\widehat{p}_{t-1}}, \\
\widehat{p}_t &= \frac{1}{\gamma_t}\left\{\widehat{p}_{t-1} - \frac{\widehat{p}_{t-1}^2(\widehat{\sigma}_t^{2'})^2}{\gamma_t(\widehat{\sigma}_t^2)^2 + (\widehat{\sigma}_t^{2'})^2\widehat{p}_{t-1}}\right\}, \\
\widehat{\sigma}_{t+1}^2 &= (1-\widehat{\lambda}_t)y_t^2 + \widehat{\lambda}_t\widehat{\sigma}_t^2, \\
\widehat{\sigma}_{t+1}^{2'} &= -y_t^2 + \widehat{\sigma}_t^2 + \widehat{\lambda}_t\widehat{\sigma}_t^{2'},\ t \in \mathbb{N},
\end{aligned}
\tag{5}
$$

where $\widehat{\lambda}_t$ denotes the recursive estimate of the parameter λ at time t. We recommend initializing the procedure by these values: (i) $\widehat{p}_0$ is a large positive number, e.g. $\widehat{p}_0 = 10^5$, (ii) $\widehat{\lambda}_0$ should be taken from the interval $(0,1)$, e.g. $\widehat{\lambda}_0 = 0.94$ as it is usually preferred for daily data, (iii) $\widehat{\sigma}_1^2$ is a positive number (e.g. the sample variance of several initial measurements) and $\widehat{\sigma}_1^{2'}$ equals zero, (iv) $\{\gamma_t\}$ is a deterministic sequence of real positive numbers smaller or equal to one that either accelerate convergence or allow tracking parameter changes (see below or [4]). It should be emphasized that $\widehat{\sigma}_{t+1}^2$ denotes the one-step ahead prediction of the conditional variance applying the on-line estimate $\widehat{\lambda}_t$.

At each time t, it is necessary to check whether the current recursive estimate belongs to the interval $(0,1)$ before evaluating other quantities in (5). If not, one should artificially use the same estimate of λ as the previous one to avoid eventual identification problems. This simple projection ensures positivity of the conditional variance since $\widehat{\lambda}_0$ lies in the interval $(0,1)$. The sequence $\{\gamma_t\}$, the so-called *forgetting factor*, may be selected as follows: (i) γ_t gradually grows to one as t goes to infinity, e.g. $\gamma_t = 0.99\gamma_{t-1} + 0.01$, $\gamma_0 = 0.95$, (ii) $\gamma_t \equiv \gamma$ for some $\gamma \in (0,1)$, e.g. $\gamma = 0.996$, and all t. The first option estimates the model (1) supposing time-invariant λ. The increasing forgetting factor improves the convergence speed of the algorithm during the transient phase. The second option is associated with the eventuality that λ can vary over time. The constant forgetting factor less than one progressively reduces the influence of historical measurements, and thus enables to detect parameter changes. See [8] for details.

Theoretical properties of the suggested recursive estimation algorithm coincide with the off-line case (as t goes to infinity), where the corresponding conditional log-likelihood criterion is maximized. Namely, convergence and asymptotic distributional properties are identical for a sufficiently large number of observations. Refer to [8]. The scheme (5) can be further robustified similarly as in [4].

Alternatively, we shall introduce the rolling-window estimation scheme based on the negative conditional logarithmic likelihood function corresponding to the model (1). It can be defined as follows:

$$
\widehat{\lambda}_t^M = \arg\min_{\lambda\in(0,1)} \sum_{\tau=t-M+1}^{t} \left(\frac{y_\tau^2}{\sigma_\tau^2} + \log(\sigma_\tau^2)\right),\ t \geq M,
\tag{6}
$$

where $M \in \mathbb{N}$ denotes the rolling-window width. At each time the minimum (6) is repeatedly calculated and $\widehat{\sigma}_{t+1}^2$ is evaluated using the most recent estimate $\widehat{\lambda}_t^M$ and M consecutive observations. The estimation is initialized similarly as above. It is obvious that the estimation can start only after observing at least M financial returns. This scheme or its alternatives are relatively frequently applied in practice. However, such an estimator can be computationally complex since the optimization task is obviously solved by an iterative procedure based on M consecutive measurements at each time.

4 Evaluation of estimators by backtesting

To investigate the empirical adequacy of the estimators introduced in Section 3, we shall consider the methodology suggested, e.g., in [10], which is based on the Value at Risk backtesting. Backtesting denotes statistical techniques designed in order to verify the VaR measures ex post, i.e. to check if the real losses correspond to the VaR predictions. See e.g. [3] or [6]. Consequently, one might compare the different EWMA model estimators by applying various VaR backtests as performance measures.

Almost all VaR backtests are based on exceedances, when one-step ahead losses exceed Values at Risk (see [6]). In particular, one can define the *exceedance process* $I_t(\alpha)$ as the corresponding quantitative binary indicator of such events: $I_t(\alpha) = 1$ if $y_t < -VaR_{t|t-1}(\alpha)$ and $I_t(\alpha) = 0$ otherwise. Traditional backtesting methods mostly fall into two main categories (see [6]): (i) ***coverage tests*** assess whether the frequency of exceedances is consistent with the quantile of loss the VaR measure should reflect, and (ii) ***independence tests*** assess whether results appear to be independent from one period to the next. However, other alternatives exist (refer to [6]).

In this paper, we shall concentrate on four different backtests. Firstly, we consider the Kupiec unconditional coverage test of proportion of exceedances introduced in [7] (denoted as Test 1). Secondly, the Christoffersen independence test is used (denoted as Test 2). It is a likelihood ratio test that looks for unusually frequent consecutive exceedances (see [1]). Thirdly, the extended variant of Kupiec's test presented in [2] is applied (denoted as Test 3). It combines two or more variants of VaR exceedances simultaneously (i.e. VaR for various nominal coverage rates are assumed). Finally, the Engle and Manganelli test is employed (denoted as Test 4). It simply verifies the null hypothesis that the intercept and coefficients in the autoregression of demeaned exceedance process on its lagged values (see [3]).

5 Empirical study: Three European stock indices

In this section, we shall analyse empirical performance of the different estimation methods discussed in Section 3. Their adequacy is evaluated by using four VaR backtests shortly reviewed in Section 4. The studied dataset consists of daily logarithmic returns of three important European stock indices, i.e. the CAC40 stock index (France), the DAX stock index (Germany), and the FTSE stock index (Great Britain). The daily adjusted closing prices were obtained from `http://finance.yahoo.com`. We investigate one thousand consecutive log-returns up to 31st March 2016 (i.e. approximately four years of daily data). Nonetheless, we observed longer history in order to initiate both the algorithms (5) and (6).

The following techniques for estimating the RiskMetrics EWMA model (1) were considered to predict VaR (one-step ahead): (i) $\lambda = 0.94$ as is recommended for daily data (denoted as $CONST$), (ii) the recursive algorithm (5) with $\gamma_t = 0.99\gamma_{t-1} + 0.01$, $\gamma_0 = 0.95$ (denoted as $RE \uparrow 1$), (iii) the recursive algorithm (5) with $\gamma_t \equiv 0.990$ (denoted as $RE990$), (iv) the recursive algorithm (5) with $\gamma_t \equiv 0.996$ (denoted as $RE996$), (v) the rolling-window estimator (6) with $M = 100$ (denoted as $RW100$), and (vi) the rolling-window estimator (6) with $M = 250$ (denoted as $RW250$). Point out that the nominal coverage rate was set as $\alpha = 0.05$ (and also as 0.01 in Test 3 discussed in Section 4). It is noteworthy that the proposed choices of the rolling-window widths M and the constant forgetting factors γ_t correspond to each other (see e.g. [8]), namely (iii) to (v) and (iv) to (vi). All computations were performed in `R`.

Figure 1 displays the CAC40 daily log-returns jointly with the selected one-day VaR predictions introduced above (the coverage rate $\alpha = 0.05$). At first sight, one can see that there are only minor discrepancies between the recursive and the rolling-window VaR predictions while the predictions derived by using the EWMA model with the prescribed λ are less volatile. Analyses of the other VaR predictions and stock indices result in analogous conclusions (figures are not presented due to the limited space).

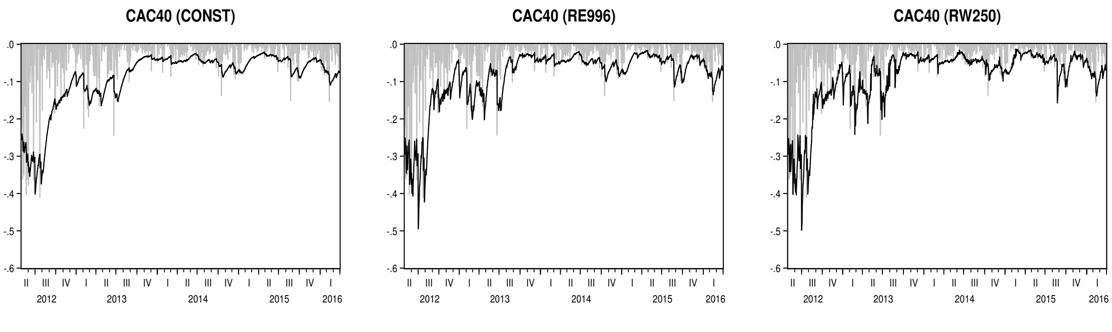


Figure 1 Negative CAC40 log-returns with selected one-day VaR predictions ($\alpha = 0.05$).

Tables 1, 2, and 3 show the results of the four VaR backtests considered in Section 4 (based on the VaR one-step ahead predictions). From these outputs, one can deduce two main conclusions. Firstly, one may identify that the analysed VaR modelling arrangement does not seem to be suitable for the investigated empirical datasets. It can be caused by the inappropriate model selection or by the arbitrary assumption of the standard normal distribution (compare with (4)). The results of backtests are not satisfactory in that sense that no estimator can be labelled as acceptable. On the contrary, the results demonstrate that the applied estimation methods are analogous in the sense of the calculated test statistics and the achieved p-values (there are no substantial discrepancies or even extraordinary excesses). Secondly, the computed exceedance proportions are comparable; however, in many cases they significantly differ from the declared 5% coverage rate (see the results of Test 1). Nevertheless, the main goal of this paper is to demonstrate that the computationally effective recursive estimation algorithm (5) is competitive with the computationally more complex rolling-window estimation scheme, which is commonly employed in practice. Consequently, the presented VaR backtesting outputs are not in contradiction to this statement.

6 Conclusion

The present paper investigated the qualities of recently introduced recursive algorithms for estimating the RiskMetrics EWMA model. The recursive estimation scheme (5) was compared with the commonly applied rolling-window estimator (6) by employing the VaR backtesting approach suggested, e.g., in [10]. It was shown that the computationally efficient recursive estimation algorithm (5) is competitive with the computationally more complex rolling-window estimation scheme based on various empirical financial datasets. However, the outputs of considered backtests indicated that the conventional VaR predictions are not adequate for the examined data. This should motivate further research of modifying the RiskMetrics EWMA modelling framework and the VaR real-time estimation.

Acknowledgements

This research was supported by the grant GA P402/12/G097.

References

[1] Christoffersen, P.: Evaluating interval forecasts, *International Economic Review* **39** (1998), 841–862.

[2] Colletaz, G., Hurlin, C., and Perignon, C.: The risk map: A new tool for risk management, *Journal of Banking and Finance* **37** (2013), 3843–3854.

[3] Engle, R. F. and Manganelli, S.: CAViaR: Conditional autoregressive Value at Risk by regression quantiles, *Journal of Business & Economic Statistics* **22** (2004), 367–381.

[4] Hendrych, R.: Robustified on-line estimation of the EWMA models: Simulations and applications. In: *Proceedings of the 33rd International Conference Mathematical Methods in Economics* (D. Martinčák, J. Ircingová, and P. Janeček, eds.), University of West Bohemia, Plzeň (CZ), 2015, 237–242.

[5] Hendrych, R. and Cipra, T.: Self-weighted recursive estimation of GARCH models, *Communications in Statistics - Simulation and Computation* (2015), in press (doi: 10.1080/03610918.2015.1053924).

[6] Holton, G. A.: *Value-at-Risk: Theory and Practice.* Academic Press, San Diego (CA), 2003.

[7] Kupiec, P. H.: Techniques for verifying the accuracy of risk measurement models, *Journal of Derivatives* **3** (1995), 73-84.

[8] Ljung, L. and Söderström, T. S.: *Theory and Practice of Recursive Identification.* MIT Press, Cambridge (MA), 1987.

[9] Morgan, J.: *RiskMetrics - Technical document.* (4th edition). Morgan Guaranty Trust Company, New York (NY), 1996.

[10] Sbrana, G. and Silvestrini, A.: Marginalization and aggregation of exponential smoothing models in forecasting portfolio volatility. In: *Mathematical and Statistical Methods for Actuarial Sciences and Finance* (C. Perna and M. Sibillo, eds.), Springer, Berlin (DE), 2012, 375–382.

CAC40	Proportion	Test 1	Test 2	Test 3	Test 4
$CONST$	0.060	1.98422	0.67799	**10.89947**	**10.35150**
		(0.15895)	(0.41028)	**(0.00430)**	**(0.03491)**
$RE \uparrow 1$	0.066	**4.91839**	0.17162	**17.95759**	**9.79110**
		(0.02657)	(0.67868)	**(0.00013)**	**(0.04410)**
$RE990$	0.067	**5.52377**	0.20236	**19.94164**	**11.23916**
		(0.01876)	(0.65282)	**(0.00005)**	**(0.02400)**
$RE996$	0.069	**6.83008**	1.21994	**18.28873**	**15.89149**
		(0.00896)	(0.26937)	**(0.00011)**	**(0.00317)**
$RW100$	0.071	**8.26095**	0.14786	**20.49056**	**12.56679**
		(0.00405)	(0.70059)	**(0.00004)**	**(0.01360)**
$RW250$	0.072	**9.02206**	0.82143	**20.74952**	**16.59069**
		(0.00267)	(0.36476)	**(0.00003)**	**(0.00232)**

Table 1 Backtests for CAC40: exceedance proportions, test statistics, and p-values (in brackets).

DAX	Proportion	Test 1	Test 2	Test 3	Test 4
$CONST$	0.040	2.25341	1.15578	**32.80609**	4.25393
		(0.13332)	(0.28234)	**(0.00000)**	(0.37273)
$RE \uparrow 1$	0.052	0.08317	1.87995	**39.08815**	2.66402
		(0.77305)	(0.17034)	**(0.00000)**	(0.61552)
$RE990$	0.059	1.61624	**5.16695**	**60.51582**	**9.59121**
		(0.20362)	**(0.02302)**	**(0.00000)**	**(0.04791)**
$RE996$	0.058	1.28428	**5.53229**	**39.36626**	**10.43186**
		(0.25711)	**(0.01867)**	**(0.00000)**	**(0.03375)**
$RW100$	0.059	1.61624	1.85057	**52.59733**	5.11321
		(0.20362)	(0.17372)	**(0.00000)**	(0.27588)
$RW250$	0.043	1.08068	0.10076	**38.36575**	1.75662
		(0.29854)	(0.75091)	**(0.00000)**	(0.78041)

Table 2 Backtests for DAX: exceedance proportions, test statistics, and p-values (in brackets).

FTSE	Proportion	Test 1	Test 2	Test 3	Test 4
$CONST$	0.067	**5.52377**	**8.27433**	**31.40050**	**23.26278**
		(0.01876)	**(0.00402)**	**(0.00000)**	**(0.00011)**
$RE \uparrow 1$	0.076	**12.36213**	**10.56582**	**38.43636**	**37.11047**
		(0.00044)	**(0.00115)**	**(0.00000)**	**(0.00000)**
$RE990$	0.071	**8.26095**	**13.45416**	**30.93453**	**39.38883**
		(0.00405)	**(0.00024)**	**(0.00000)**	**(0.00000)**
$RE996$	0.068	**6.16115**	**12.64470**	**36.50867**	**34.02678**
		(0.01306)	**(0.00038)**	**(0.00000)**	**(0.00000)**
$RW100$	0.078	**14.20448**	**7.48688**	**54.76527**	**31.12163**
		(0.00016)	**(0.00622)**	**(0.00000)**	**(0.00000)**
$RW250$	0.077	**13.26924**	**12.36007**	**33.96092**	**42.89799**
		(0.00027)	**(0.00044)**	**(0.00000)**	**(0.00000)**

Table 3 Backtests for FTSE: exceedance proportions, test statistics, and p-values (in brackets).

Detection of Feasible Domains of Interpolation Parameters within Generalized EOQ Type Inventory Models

Jiří Hofman[1], Ladislav Lukáš[2]

Abstract. The paper deals with some generalized inventory control models which are linked to the well-known EOQ model. There are dynamic periodic and deterministic represented by instantaneous replenishment in constant time periods. Opposite to EOQ models, the generalized models assume both demand rate and holding cost rate to be time-dependent along inventory cycle being normalized to unit length. The demand rate takes polynomial form, which is specified by particular interpolation conditions. Having particular model, we concern with feasible domain detection of interpolation parameters, which are used for construction of inventory level taking polynomial form of specified algebraic degree. Assuming instantaneous replenishment at the beginning of inventory cycle exclusively, the necessary condition of inventory level feasibility is expressed by condition that inventory level is non-increasing on inventory cycle globally. For various generalized inventory models, we investigate the feasibility domains of their interpolation parameters numerically using Mathematica. The results and some code snippets are presented and discussed in detail.

Keywords: generalized EOQ model, feasibility domain, interpolation parameters, inventory control model.

JEL Classification: C54, D29
AMS Classification: 90B05

1 Introduction

Many practical inventory control systems are still based upon the EOQ model regardless its simplicity and limits, in general. This model is classified as deterministic, dynamic and periodic one which seeks minimum of total cost under an objective to optimize lot size in order to reduce the cost of satisfying demand. Hence, the concept of the EOQ is that there is a trade off between the fixed order cost and the holding cost assuming demand rate being constant exclusively.

Since the EOQ business scenario is rather simple, it guides to various generalizations. First idea could be to eject an assumption that model is deterministic one, as the deterministic demand might sound too restrictive one in practice on opposite to stochastic demand. However, it needs quite different theoretical framework to apply. Another idea for generalization is induced by suspending constant assumptions being posed upon holding cost and rate of demand, respectively. In general, it stands exactly in correspondence with the concept of generalized EOQ type models which has been already given in [5].

There are many other various inventory models at disposal in literature, now. We are going to mention just a few ones. In [2], various lot sizing algorithms applied for inventory control and emphasizing the financial implications of corresponding inventory policies in supply-chain management are discussed. Deterministic models of perishable inventory with stock-dependent demand rate and nonlinear holding cost are presented in [3]. Inventory model given in [7] is based upon classical EOQ model with continuous reviewing of inventory level and facing deterministic constant demand rate, but assuming supplier side to undergo random supply disruptions causing uncertainty of replenishments. EOQ model for deteriorating items with exponentially dependent demand rate in which inflation and time value of money are taken into account is presented in [8]. Good survey of literature reviews in the area of lot sizing problems is presented thoroughly in [4]. There is also lot of textbooks covering the topic of inventory control problems, e.g. [1], and [6], being just two ones pursuant to our short subjective selection.

[1] University of West Bohemia/Faculty of Economics, Dept. of Economics and Quantitative Methods, Husova 11, 306 14 Pilsen, hofman@kem.zcu.cz.
[2] University of West Bohemia/Faculty of Economics, Dept. of Economics and Quantitative Methods, Husova 11, 306 14 Pilsen, lukasl@kem.zcu.cz

2 Generalized EOQ type models and feasible domains

Classical EOQ model assumes constant demand rate, infinite replenishment rate, zero lead-time, and minimization of total inventory cost per unit time $C(q)$. Let $z(\tau)$ denote an inventory level during a periodic inventory cycle t, with $\tau \in [0, t]$. As usual, let c_1, and c_3, denote unit holding cost, and fixed cost incurred per order (called ordering cost, too), respectively. Further, T gives total inventory control period, and Q gives aggregate product demand during T, thus providing the number of cycles by $n = Q/q = T/t$. Finally, let $N(q)$ stands for total inventory cost per inventory cycle being expressed as a function of unknown lot size q. There is well-known, that this quantity could be expressed equivalently as a function of unknown inventory cycle t, i.e. $N(t)$, as well.

There is well-known, the optimal replenishment lot size q_{opt} is solution of the optimization problem

$$\min C(q)\,, \quad q \in \,]0,\infty[$$
$$C(q) = N(q)/t\,, \;\; N(q) = c_3 + \int_0^t c_1 z(\tau)\mathrm{d}\tau\,, \;\; z(\tau) = q(1-\tau/t)\,, \;\; t = qT/Q, \tag{1}$$

thus yielding $q_{\text{opt}} = \operatorname{argmin} C(q) = \sqrt{(2c_3Q/(c_1T))}$, $C(q_{\text{opt}}) = \sqrt{(2c_1c_3Q/T)}$, with quantities c_1, c_3, T, and Q, given.

In [6], suitable general framework for large family of inventory models is presented in the following form

$$\text{Find } s(t) \in U, \;\; \text{extrem } \left(\int_0^T \varphi(z(\tau), \lambda)\mathrm{d}\tau\right), \;\; \forall\, z(t) \in V, \tag{2}$$

where U, and V denote set of feasible control, and set of feasible solutions, respectively.

Following [5], and [6] Ch.3.2.4, the generalized EOQ type models release restrictive assumptions of constant demand rate, and constant unit holding cost c_1. Hence, the total inventory cost per cycle $N(t)$ will take form

$$N(t) = c_3 + \int_0^t c_1(\tau)\, z(\tau)\mathrm{d}\tau, \tag{3}$$

provided a proper selection of functions $c_1(\tau)$ and $z(\tau)$ have been made. In [5], we have proposed algebraic polynomials, denoting $p_n(\tau)$ a polynomial of n-th degree, as usual. Let $c_1(\tau) = p_i(\tau)$, $i \geq 1$, and $z(\tau) = p_j(\tau)$, $j \geq 2$, whereas linear function $p_1(\tau)$ suits exactly to describe inventory level of the EOQ model.

Among advantages of such generalization mentioned in [5], the most important one is that higher degree of inventory level $z(\tau)$ enables us both to cope more flexibly with time dependent demand during inventory cycle, and also to use an additional intermediate information of inventory level being detected by inventory checks during inventory cycle course, prospectively.

However, an approximation $z(\tau) = p_j(\tau)$, $j \geq 2$, by interpolation polynomials could not be quite arbitrary, since $z(\tau)$ must be a non-increasing function on the whole inventory cycle, in general. Hence, crucial questions arise what are feasible domains of interpolation parameters of $p_j(\tau)$ in order to maintain that necessary condition, in particular. In **Fig. 1**, we depict normalized inventory levels $\zeta(\tau) = z(\tau)/q$, on normalized inventory cycle [0,1], with some of them being feasible ones, while the others not, at the first glance.

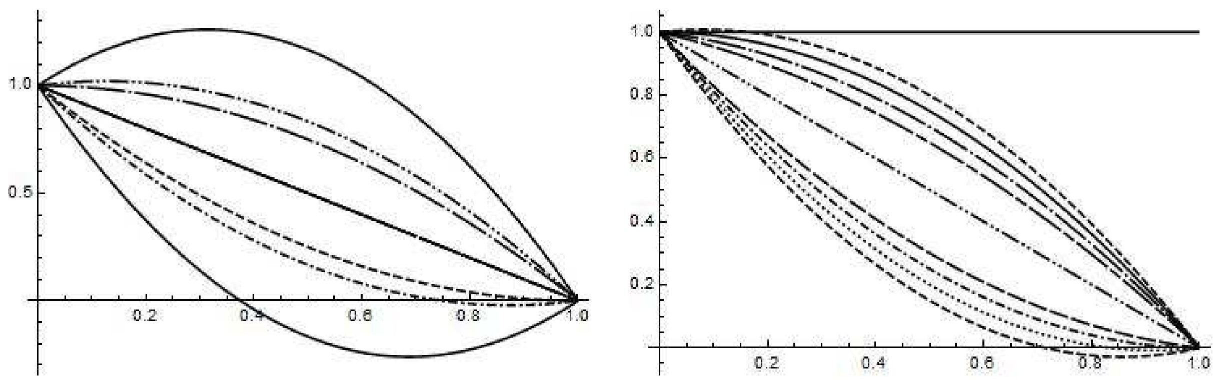


Figure 1 Normalized inventory levels $\zeta(\tau)$ under various quadratic approximations with $\tau \in [0,1]$

Let $p_n(\tau)$ approximating $\zeta(\tau)$ takes the usual form (4), with a_i, $i = 0, ..., n$, polynomial coefficients, which are to be determined by proper interpolation conditions uniquely

$$p_n(\tau) = a_0 + a_1\tau + \dots + a_n\tau^n = \sum_{i=0}^{n} a_i\tau^i,\ \tau \in [0,1]. \tag{4}$$

There are are two kinds of interpolation conditions – basic ones, and additional ones. Any $\zeta(\tau)$ must fulfill basic interpolation conditions

$$\zeta(0) = 1,\ \zeta(1) = 0, \tag{5}$$

which being applied upon (4) yields $p_n(\tau)$ in specific form

$$p_n(\tau) = 1 + a_1\tau + \dots + a_n\tau^n = 1+\sum_{i=1}^{n} a_i\tau^i,\quad \sum_{i=1}^{n} a_i = 1,\ \ \tau \in [0,1]. \tag{6}$$

Let assume the interpolation coefficients a_1, .., a_n are defined by set of quantities, e.g. arguments, function values, slopes, curvatures, and others, which form a vector ξ, with components ξ_j, $j = 1, .., m$. Hence, the expression (6) takes the following form

$$p_n(\tau;\xi) = 1 + a_1(\xi)\tau + \dots + a_n(\xi)\tau^n = 1+\sum_{i=1}^{n} a_i(\xi)\tau^i,\quad \sum_{i=1}^{n} a_i(\xi) = 1,\ \ \tau \in [0,1]. \tag{7}$$

Definition 1. *Within the GEOQ class models, a point $\xi \in \mathrm{R}^m$ is called feasible, if it satisfies condition* (8)

$$dp_n(\tau;\xi)/d\tau \le 0\ ,\quad \sum_{i=1}^{n} a_i(\xi) = 1,\quad \forall\ \tau \in [0,1]. \tag{8}$$

which in explicit form is

$$a_1(\xi) + \dots + n\, a_n(\xi)\tau^{n-1} = \sum_{i=1}^{n} i a_i(\xi)\tau^{i-1} \le 0\ ,\quad \sum_{i=1}^{n} a_i(\xi) = 1,\quad \forall\ \tau \in [0,1]. \tag{8a}$$

Definition 2. *Within the GEOQ class models, a set $V \subset \mathrm{R}^m$ is called feasible domain, if the condition* (8) *is satisfied by all its points, $\forall\ \xi \in V$.*

Though the **Def. 2** looks to be quantitative, it is qualitative in its natural sense, as it requires to check non-positivity of the first derivative of $p_n(\tau)$, representing actually the slope of inventory level $z(\tau)$, in uncountable number of points, i.e. $\forall\ \tau \in [0,1]$. In order to convert it into numerical plausible form, we relax the condition (8a) into a discrete form (8b), which is nothing else but simultaneous conjunction of point condition (9), for $k = 1, ..., K$, where K gives the total number of distinct points τ_k.

$$a_1(\xi) + \dots + n\, a_n(\xi)\tau^{n-1} = \sum_{i=1}^{n} i a_i(\xi)\tau^{i-1} \le 0\ ,\quad \sum_{i=1}^{n} a_i(\xi) = 1,\quad \forall\ \tau = \tau_k \subset [0,1]. \tag{8b}$$

A point condition, which provides a possibility to be checked numerically for given τ_k, takes the form

$$a_1(\xi) + \dots + n\, a_n(\xi)\tau^{n-1} = \sum_{i=1}^{n} i a_i(\xi)\tau^{i-1} \le 0\ ,\quad \sum_{i=1}^{n} a_i(\xi) = 1,\quad \tau = \tau_k \subset [0,1]. \tag{9}$$

3 Some numerical results

In this section, we describe three cases of numerical detection of feasible domains for inventory models of GEOQ(i, j) class, which was presented in [5]. In particular, we select GEOQ(0,2) model with two different kinds of additional interpolation conditions, i.e. Lagrangean type, and Hermitean type, too. Further, we will analyze GEOQ(0,3) model with Lagrangean type of additional interpolation conditions.

3.1 GEOQ(0,2)~L model

An inventory level $\zeta(\tau;\xi)$ is expressed by quadratic polynomial, which entails a linear demand rate, in particular

$$\zeta(\tau;\xi) = p_2(\tau;\xi) = 1 + a_1(\xi)\tau + a_2(\xi)\tau^2,\quad a_1(\xi) + a_2(\xi) = 1,\ \ \tau \in [0,1], \tag{10a}$$

and the additional inventory condition of Lagrangean type takes the following form

$$\zeta(\tau = \theta;\ \xi) = \omega,\quad \xi = (\theta, \omega)^{\mathrm{T}}\ ,\quad \theta, \omega \in\ ]0,1[\ ,\quad \text{i.e. } \xi \in\ ]0,1[\ \mathrm{x}\ ]0,1[\ \subset \mathrm{R}^2. \tag{10b}$$

Using Mathematica symbolic computation power we get explicit expressions of $a_1(\xi)$, and $a_2(\xi)$, by solving (10a), and (10b), in following form

$$a_1(\theta, \omega) = (1 - \omega - \theta^2)/(\theta\,(-1+ \theta))\ ,\quad a_2(\theta, \omega) = (-1 + \omega + \theta)/(\theta\,(-1+ \theta)). \tag{11}$$

Now, we are ready to detect numerically a feasible domain for this model, denoted ${}_{L}V_2$, where the sub-indices L, and 2, are pure symbolic ones thus expressing a link to the GEOQ(02)~L model, only. It can be defined in following way, just by adopting (9) for a set of discrete points $W = \{\tau_k\}$, $k = 1, ..., K$, where $W \subset [0,1]$

$$ {}_{L}V_2 = \{(\theta, \omega) \in]0,1[\text{ x }]0,1[\subset R^2 \mid ((1-\omega-\theta^2) + (-1+\omega+\theta)\tau) / (\theta(-1+\theta)) \le 0, \quad \forall \tau = \tau_k \in W\}, \tag{12} $$

as the condition $a_1(\xi) + a_2(\xi) = 1$, from (10a), yields a solution, for coefficients (11) being substituted therein, $\theta = 1$, and ω arbitrary, which is a degenerate solution only, since it does not obey the assumption $\theta \in]0,1[$, stated above.

In order to keep computer time elapsed for numerical analysis in a reasonable order, we select $W = \{0.1, 0.2, 0.3, ..., 0.9\}$. Mathematica provides a powerful plot command RegionPlot which can be used to detect ${}_{L}V_2$ p, as shown by the following code snippet, while the result is given in **Fig. 2**, on the left.

```
RegionPlot[Evaluate[(D[ς2L[v,θ,w],v/.→.1) ≤ 0 && (D[ς2L[v,θ,w],v/.→.2) ≤ 0 &&
(D[ς2L[v,θ,w],v/.→.3) ≤ 0 && (D[ς2L[v,θ,w],v/.→.4) ≤ 0 && (D[ς2L[v,θ,w],v/.→.5) ≤ 0 &&
(D[ς2L[v,θ,w],v/.→.6) ≤ 0 && (D[ς2L[v,θ,w],v/.→.7) ≤ 0 && (D[ς2L[v,θ,w],v/.→.8) ≤ 0 &&
(D[ς2L[v,θ,w],v/.→.9) ≤ 0], {θ,.01,.99},{w,.01,.99}]
```

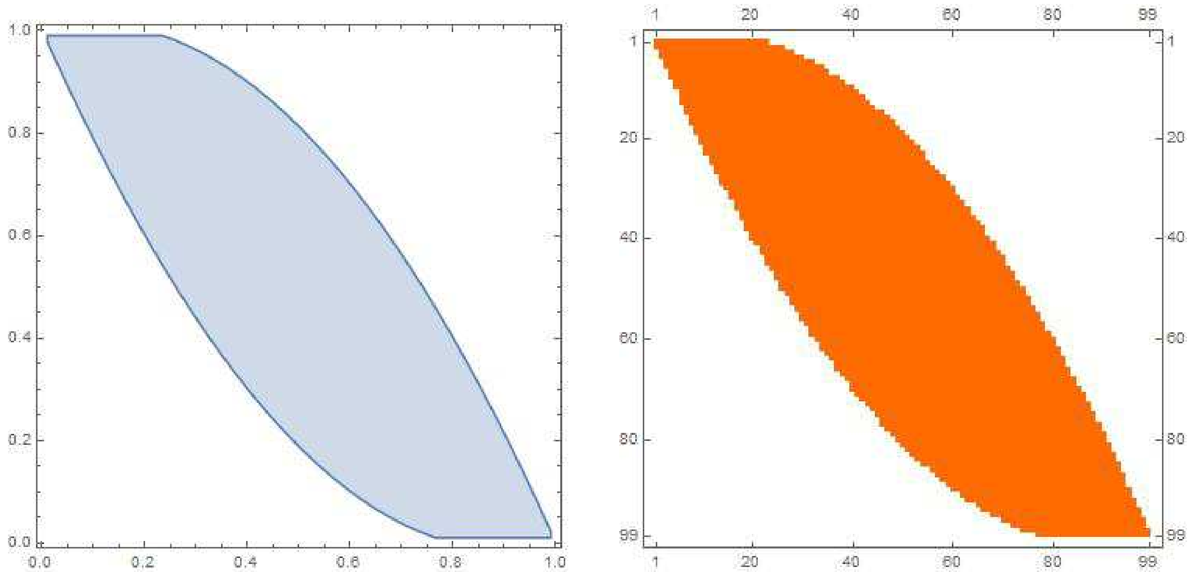


Figure 2 Feasible domain ${}_{L}V_2$ of GEOQ(0,2)~L model using RegionPlot ~(left), MatrixPlot ~(right)

In **Fig. 2**, the horizontal axis corresponds to θ, while the vertical one to ω, and ranging both parameters in $[0.01,0.99] \subset]0,1[$. In the code snippet, we use w for ω, and v for τ, simply. The plot on the right is produced by MatrixPlot, which is another very useful Mathematica command. It served us also as a key command in our general procedure for detecting feasible domains, as the RegionPlot is limited on 2-D regions. In spite of fact, that 3-D version exists. i.e. RegionPlot3D, we have developed a general procedure based upon idea to discretize an adequate subspace of R^m containing feasible domain V, prospectively. We will present it in **Sec. 3.3**.

In Fig. 3, the inventory levels $\zeta(\tau; 0.1, \omega_i)$, are depicted, for $\omega_i \in \{0.99, 0.96, 0.93, 0.90. 0.87, 0.84\}$.

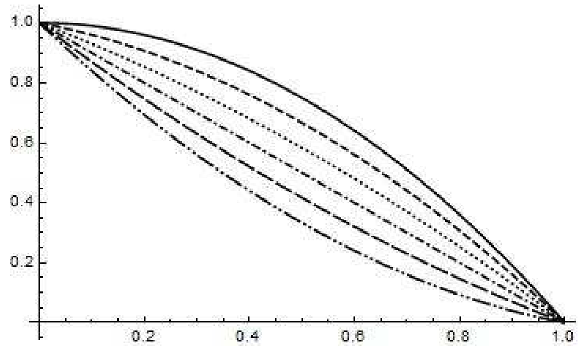


Figure 3 Some feasible inventory levels of GEOQ(0,2)~L model

3.2 GEOQ(0,2)~H model

An inventory level is given by (10a), again. However, the interpolation condition is of Hermitean type, and it can be applied either at a) $\tau = 0$, or b) $\tau = 1$, respectively

$$\text{a)}\quad d\zeta(\tau{=}0;\boldsymbol{\xi})/d\tau = -\varphi_h\,, \qquad \text{b)}\quad d\zeta(\tau{=}1;\boldsymbol{\xi})/d\tau = -\varphi_d\,, \qquad \boldsymbol{\xi} = \varphi \in \mathrm{R}. \tag{13}$$

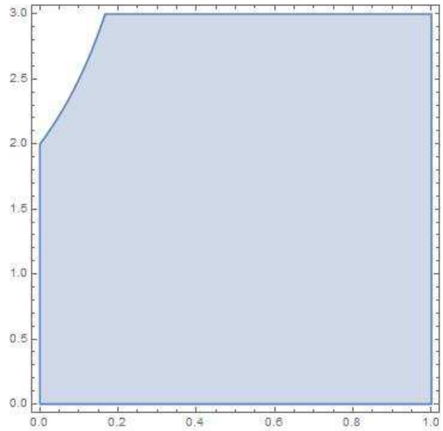
Within GEOQ(0,2)~H model, the inventory levels $\zeta(\tau;\varphi)$ can be derived in explicit forms analytically, for both cases, thus taking forms

$$\text{a)}\quad \zeta(\tau;\varphi) = 1 - \varphi\tau - (1-\varphi)\tau^2,\ \ \varphi = \varphi_h\,, \qquad \text{b)}\quad \zeta(\tau;\varphi) = 1 - (2-\varphi)\tau + (1-\varphi)\tau^2,\ \ \varphi = \varphi_d\,, \tag{14}$$

and yielding simple conditions of type (9), too

$$\text{a)}\quad -\varphi_h - 2(1-\varphi_h)\tau \le 0, \qquad \text{b)}\quad -(2-\varphi_d) + (1-\varphi_d)\tau \le 0, \qquad \forall\, \tau \in [0,1]. \tag{15}$$

We restrict ourselves to case a), since it has more practical usage. Adopting condition (15, a) into RegionPlot, we get RegionPlot[φ(1-2v) ≤ 2, {v,0,1},{ φ, 0,3}], and the result is plotted in **Fig. 4** left, which reveals ${}_{\mathrm{H}}V_2 = [0,2]$. On the right, there are demand rates of the model GEOQ(0,2)~H case a), for $\varphi_h \in \{0, 0.2, ..., 1.8, 2\}$.



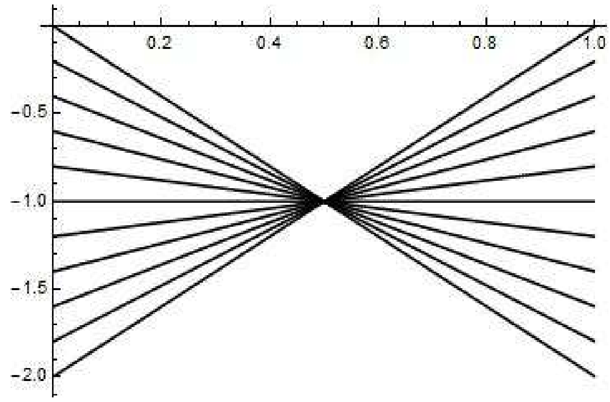


Figure 4 Feasible domain ${}_{\mathrm{H}}V_2 = [0,2]$ of model GEOQ(0,2)~H case a) model ~(left), feasible slopes ~(right)

3.3 GEOQ(0,3)~L model

An inventory level $\zeta(\tau;\xi)$ is expressed by cubic polynomial, which entails a quadratic demand rate, in particular

$$\zeta(k;\xi) = p_2(\tau;\xi) = 1 + a_1(\xi)\tau + a_2(\xi)\tau^2 + a_3(\xi)\tau^3\,,\ \ a_1(\xi) + a_2(\xi) + a_3(\xi) = 1,\ \ \tau \in [0,1], \tag{16a}$$

and two additional inventory conditions of Lagrangean type are

$$\zeta(\tau = \theta_h;\xi) = \omega_h,\ \ \zeta(\tau = \theta_d;\xi) = \omega_d,\ \ \xi = (\theta_h, \omega_h, \theta_d, \omega_d)^{\mathrm{T}},\ \ \theta_h, \omega_h, \theta_d, \omega_d \in\,]0,1[\,, \quad \text{i.e.}\ \ \xi \in (\,]0,1[\,)^4 \subset \mathrm{R}^4. \tag{16b}$$

Using Mathematica we get explicit expressions of $a_1(\xi)$, $a_2(\xi)$, and $a_3(\xi)$, by solving (16a), and (16b), again. Because of limited space of the paper, we do not list them here, on opposite to (11) of GEOQ(0,2)~L model, since they contain long algebraic expressions up to third order being composed from variables θ_h, ω_h, θ_d, ω_d, both in nominators and denominators. Formally, we express them simply

$$a_1(\xi),\ a_2(\xi),\ a_3(\xi),\quad \xi = (\theta_h, \omega_h, \theta_d, \omega_d)^{\mathrm{T}}. \tag{17}$$

Since $\xi \in (\,]0,1[\,)^4$ represents a point in 4-D cube, in general, we have to use another algorithm for detection a feasible domain than using RegionPlot type commands in Mathematica. Our general algorithm is based upon simple idea of full discretization both parameter region Ω, e.g. 4-D cube, and time domain [0,1], as well.

Discrete detection algorithm of feasible domain V of GEOQ class models – steps:

1. Build mesh on Ω, containing yet unknown V, by setting discrete points $\xi_i \in \Omega$, $i = 1, ..., I$.
2. Build mesh W by setting $\{\tau_k\}$, $k = 1, \ldots, K$.

3. Define binary function $f(\xi_i, W) = 1$, if (9) is fulfilled for $\forall\ \tau_k \in W$, at $\xi_i \in \Omega$,

   $= 0$, otherwise.

4. Point detection: if $f(\xi_i, W) = 1$ at $\xi_i \in \Omega$ then $\xi_i \in V$, else point ξ_i does not belong to V.

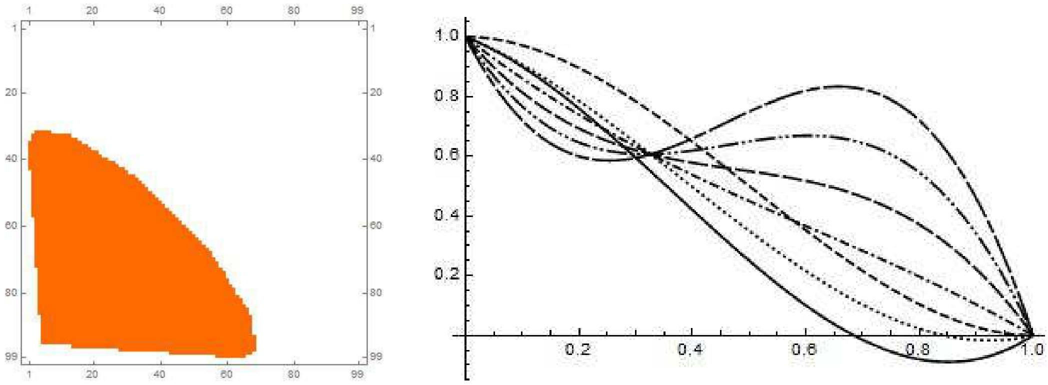


Figure 5 Feasible domain ${}_{L}V_3(\theta_h=.25,\ \theta_d=.75)$ of GEOQ(0,3)~LL model ~(left), $\zeta(\tau;\ \xi)$ ~(right)

In **Fig 5** on the left, we plot feasible domain ${}_{L}V_3(\theta_h=.25,\ \theta_d=.75)$ of GEOQ(0,3)~LL model using RegionPlot, which actually shows just a 2-D section of ${}_{L}V_3$ being 4-D, with $I = 99$ for both ω_h and ω_d, in general. The rows represents ω_h from top to down, whilst columns represents ω_d from left to right. On the right, there are several feasible and un-feasible inventory levels $\zeta(\tau;\ \xi)$ of GEOQ(0,3)~LL model with $\xi = (\theta_h = 0.25,\ \omega_h,\ \theta_d = 0.75,\ \omega_d)^{\mathrm{T}}$.

4 Conclusions

Framework of generalized EOQ-type models has been further developed by detection of feasible domains of admissible inventory levels for time dependent demand rates. Their descriptions are based upon additional interpolation conditions expressing intermediate information during inventory cycle.

General discrete detection algorithm is given in basic steps. Some numerical examples of feasible domain detections are presented. All calculations were performed in sw Mathematica.

Near future research will be focused on thorough numerical experiments with discrete detection algorithm, and its further numerical efficiency development. Next, we would like to concentrate ourselves also upon a role of inventory, its financial issues in particular, within broader framework of firm valuation, as an amount of firm asset allocated in inventory might cause awkward and unexpected effects thereon.

Acknowledgements

The research project was supported by the grant no. 15-20405S of the Grant Agency, Prague, Czech Republic.

References

[1] Axsaeter, S.: *Inventory Control.* 2-nd ed., Springer, New York, 2006.

[2] Beullens, P.: Revisiting foundations in lot sizing – Connections between Harris, Crowther, Monahan, and Clark. *Int. J. Production Economics* **155** (2014), 68–81.

[3] Giri, B.C., Chaudhuri, K.S.: Deterministic models of perishable inventory with stock-dependent demand rate and nonlinear holding cost. *EJOR* **105**(1998), 3, 467-474.

[4] Glock, Ch. H., Grosse, E. H., Ries, J. M.: The lot sizing problem: A tertiary study. *Int. J. Production Economics*, 2014, **155**, pp. 39-51.

[5] Hofman, J., Lukáš, L.: Generalized EOQ Type Inventory Models with Time Dependent both Demand Rate and Holding Cost Rate. In: *Conference Proceedings of 33rd Int. Conf. Math. Methods in Economics 2015* (Martinčík, D., Ircingová, J., and Janeček, P., eds.). University of West Bohemia, Pilsen, 2015, 249–254.

[6] Lukáš, L.: *Probabilistic models in management – Theory of inventory and statistical description of demand* (in Czech: *Pravděpodobnostní modely v managementu – Teorie zásob a statistický popis poptávky*). Academia, Praha 2012.

[7] Snyder, L.V.: A tight approximation for an EOQ model with supply disruptions. *Int. J. Production Economics* **155** (2014), 91–108.

[8] Tripathi, R. P., Singh, D., Mishra, T.: EOQ Model for Deteriorating Items with exponential time dependent Demand Rate under inflation when Supplier Credit Linked to Order Quantity. *Int. J. Supply and Operations Management* **1** (2014), 1, pp. 20-37.

Nonparametric regression via higher degree F-transform for implied volatility surface estimation

Michal Holčapek[1], Tomáš Tichý[2]

Abstract. Estimation of the implied volatility surface is a crucial step for pricing of illiquid or non-listed options. For its estimation we use market data of traded liquid options. In order to obtain smooth surface we mostly use some nonparametric regression approach, though some combination of market data can lead to the inconsistency of the estimated values (i.e. leading to arbitrage opportunities). In this contribution we focus on a novel nonparametric regression approach based on fuzzy transform technique. The original definition is extended into multivariate higher degree F-transform so that we are able to smooth (2+1)D discrete data. Proposed procedure allows us to calculate also partial derivatives of the estimated function easily. Extending the optimization problem by suitable constraint we can prevent most of the sources of arbitrage related to option valuation with implied volatility smiles.

Keywords: Fuzzy transform, fuzzy partition, implied volatitlity.

JEL classification: C44, G13
AMS classification: 41A10, 15A54

1 Introduction

The fuzzy transform (F-transform for short) is an approximation technique proposed by Perfilieva in [7] that has many applications in various fields like data analysis or image processing. The core of the F-transform technique consists in a fuzzy partition of the consider space by means of fuzzy sets satisfying a la Ruspini condition (or also the partition of unity condition). The fuzzy partition is then used to transform a (discrete or continuous) function into a vector of F-transform components. This step is called the *direct F-transform.* The approximation of the original function is provided by the *inverse F-transform*, where the linear combination of the F-transform components with respect to the weights derived from the fuzzy partition is applied. Note that the F-transform can fully reconstruct a function only in very special cases, e.g., when the function is constant. An ideal reconstruction of functions (Shannon sampling theorem) by the F-transform technique has been investigated in [10]. To improve the approximation ability, the F-transform technique was extended using polynomials to higher degree F-transform in [8]. A further development including two dimensional case and applications has been done in [9]. A summary of theoretical results on two dimensional continuous F-transform of higher degree can be found in [3]. A discrete version of multivariate F-transform of higher degree including a basic analysis of its properties can be found in [4].

One of the well-known technique of non-parametric regression is the kernel polynomial regression and its multivariate generalization. The estimation of a function f at a point x from sample data is determined by the value of a polynomial which is derived by the polynomial weighted regression, where the weights are specified by a kernel function centered at x. The respective derivatives of the mentioned polynomial at x are used for the estimation of possible derivatives of f at x. For details, we refer to the books [6] and [11]. Comparing the computation of function values by the kernel polynomial regression and the F-transform components of discrete higher degree F-transform, one can recognize that both techniques are very similar. The significant difference consists in the estimation (reconstruction) of the original function, where the higher degree F-transform uses only a limited (usually small) number of nodes of a

[1]Institute for Research and Applications of Fuzzy Modelling, NSC IT4Innovations, 30. dubna 22, 701 03 Ostrava 1, Czech Republic, michal.holcapek@osu.cz
[2]VŠB – Technical University of Ostrava, Department of Finance, Sokolská 33, 701 21 Ostrava, Czech Republic, tomas.tichy@vsb.cz

given space at which the polynomials are derived instead of the kernel polynomial regression with the polynomials that are computed at each point (or for a large number of points) of the space. A reduction of a computational effort of polynomial kernel regression motivates us to introduce a non-parametric regression with help of the higher degree F-transform.

In this contribution we extend the original definition of the F-transform into multivariate higher degree F-transform and apply it in order to smooth discrete 2D market data of option implied volatilities. This novel approach allows us to obtain derivatives of the smoothed function directly and in this way calculate several useful measures, such as a so called state price densities. While in Section 2 multivariate fuzzy partition is defined, Section 3 is devoted to the definition of discrete higher degree multivariate F-transform itself. In Section 4, the non-parametric regression is briefly mentioned and finally, in Section 5 an illustrative example is provided.

2 Uniform multivariate fuzzy partition

Let $\mathbb{N}_0$, $\mathbb{Z}$ and $\mathbb{R}^d$ ($\mathbb{R}^d_+$) denote the set of natural numbers including zero, integers and the set of real (positive) vectors, respectively. A uniform fuzzy partition, which is commonly considered in applications of the F-transform technique, can be defined using a univariate generating function K that is modified by the bandwidth parameter h (see, e.g., [5]). For the multivariate case, we have proposed in [4] a straightforward generalization of the definition of generating function.

Definition 1. A function $\mathcal{K} : \mathbb{R}^d \to [0, 1]$ is said to be a *d-dimensional generating function* if $\mathcal{K}$ is an even continuous function (fuzzy relation) for which

$$\mathcal{K}(\boldsymbol{x}) \begin{cases} > 0, & \text{if } x \in (-1, 1)^d; \\ = 0, & \text{otherwise.} \end{cases} \tag{1}$$

and the function of one variable $\mathcal{K}(c_1, \dots, c_{i-1}, x, c_{i+1}, \dots, c_n)$ is non-increasing in $[0, 1]$ for any choice of $i = 1, \dots, d$ and constant $c_j \in (-1, 1)$, $j = 1, \dots, d$ with $i \neq j$.

For $d = 1$, $\mathcal{K}$ is said to be a *univariate generating function* and denote it by K. In this paper, we use the standard approach to determine multivariate generating functions with help of the product of univariate generating functions. Particularly, if K is a univariate generating function, we define

$$\mathcal{K}(\boldsymbol{x}) = \prod_{i=1}^{d} K(x_i), \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{2}$$

Let $\mathcal{K}$ be a d-dimensional generating function, and let $\mathbf{H}$ be a $d \times d$ diagonal matrix with the elements from $\mathbb{R}_+$. In literature (see, e.g., [11]), the matrix $\mathbf{H}$ is called the *bandwidth matrix.* A *scaled d-dimensional generating function* $\mathcal{K}$ with respect to $\mathbf{H}$ is the function $\mathcal{K}_{\mathbf{H}} : \mathbb{R}^d \to [0, 1]$ defined by

$$\mathcal{K}_{\mathbf{H}}(\boldsymbol{x}) = \mathcal{K}(\mathbf{H}^{-1}\boldsymbol{x}). \tag{3}$$

Note that the bandwidth matrix can be introduced in a more general setting, where $\mathbf{H}$ is assumed to be symmetric and positively definite (see, e.g., [11]). The multivariate uniform fuzzy partition is defined as follows (cf., [4]). We use $\boldsymbol{X}^T$ to denote the transpose matrix to the matrix $\boldsymbol{X}$.

Definition 2. Let $\mathcal{K}$ be a d-dimensional generating function, and let $\mathbf{H}$ be a bandwidth matrix, $\mathbf{R}$ be a diagonal matrix of positive elements and $\boldsymbol{c} \in \mathbb{R}^d$. A family of fuzzy relations $\mathbb{A} = \{\mathcal{A}_{\boldsymbol{k}} \mid \boldsymbol{k} \in \mathbb{Z}^d\}$ defined by

$$\mathcal{A}_{\boldsymbol{k}}(\boldsymbol{x}) = \mathcal{K}_{\mathbf{H}}(\boldsymbol{x} - \boldsymbol{c} - \boldsymbol{k}\mathbf{R}), \quad \boldsymbol{x} \in \mathbb{R}^d,$$

is said to be a *uniform d-dimensional fuzzy partition of* $\mathbb{R}^d$ determined by the quadruplet $(\mathcal{K}, \mathbf{H}, \mathbf{R}, \boldsymbol{c})$ provided that *a la Ruspini condition* is satisfied, i.e., $\sum_{\boldsymbol{k}\in\mathbb{Z}^d} \mathcal{A}_{\boldsymbol{k}}(\boldsymbol{x}) = 1$ for any $\boldsymbol{x} \in \mathbb{R}^d$. The parameters $\mathbf{R}$ and $\boldsymbol{c}$ are called the *shift matrix* and the *central node*, respectively. The function $\mathcal{A}_{\boldsymbol{k}}$ is called the k-th *basic function.*

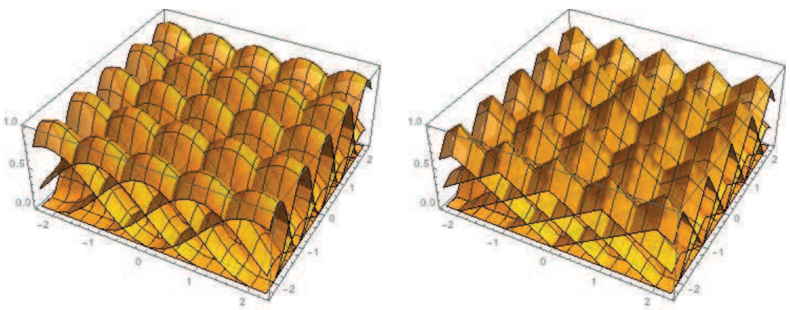


Figure 1 Parts of the 2D uniform triangle (left) and raised cosine (right) fuzzy partitions

For $d = 1$, we say that the quadruplet (K, h, r, k) determines a *uniform fuzzy partition of* $\mathbb{R}$. A deeper investigation of necessary and sufficient conditions for uniform fuzzy partitions can be found in [5]). The following theorem provides a necessary and sufficient condition for the multivariate fuzzy partitions determined by (2) (for the proof, we refer to [4]). We use $\boldsymbol{e}_j$ to denote the unit vector with 1 at the j-th coordinate.

Theorem 1. *Let $\mathcal{K} = \prod_{j=1}^{d} K$ be a multivariate generating function. The quadruplet $(\mathcal{K}, \mathbf{H}, \mathbf{R}, \boldsymbol{c})$ determines a uniform multivariate fuzzy partition of $\mathbb{R}^d$ if and only if the quadruplet $(K, \boldsymbol{e}_j\mathbf{H}\boldsymbol{e}_j^T, \boldsymbol{e}_j\mathbf{R}\boldsymbol{e}_j^T, \boldsymbol{c}\boldsymbol{e}_j^T)$ determines a uniform fuzzy partition of $\mathbb{R}$ for any $j = 1, \dots, d$.*

Let I be a compact subset of $\mathbb{R}^d$. The least subfamily $\mathbb{B}$ of a uniform multivariate fuzzy partition $\mathbb{A}$ of $\mathbb{R}^d$ satisfying a la Ruspini condition for any point of I is said to be a *uniform multivariate fuzzy partition of* I. We denote $\mathbb{B} = \mathbb{A}_I$ and put $\mathbb{A}_I = \{\mathcal{A}_{\boldsymbol{k}} \mid \boldsymbol{k} \in \mathbb{K}\}$ for $\mathbb{K} \subset \mathbb{Z}^d$.

Example 1 (Triangle and raised cosine 2D uniform fuzzy partitions). The functions $K^{tr}, K^{rc} : \mathbb{R} \to [0, 1]$ defined by

$$K^{tr}(x) = \max(1 - |x|, 0), \tag{4}$$

$$K^{rc}(x) = \begin{cases} \frac{1}{2}(1 + \cos(\pi x)), & -1 \leq x \leq 1; \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

for any $x \in \mathbb{R}$, are called the *triangle* and *raised cosine* univariate generating functions, respectively. Let $\star \in \{tr, rc\}$, and let $\mathcal{K}^\star(x, y) = K^\star(x)K^\star(y)$. In Fig. 1, one can see parts of the fuzzy partitions determined by $(\mathcal{K}^\star, \mathbf{H}, \mathbf{R}, \boldsymbol{c})$, $\star \in \{tr, rc\}$, where $\mathbf{H} = \operatorname{diag}\{2, 1\}$, $\mathbf{R} = \operatorname{diag}\{1, 1\}$ and $\boldsymbol{c} = (0, 0)$.

3 Discrete higher degree multivariate F-transform

3.1 Notation

In what follows, we use the multi-index notation. Let $\boldsymbol{i} \in \mathbb{N}_0^d$ and $\boldsymbol{x} \in \mathbb{R}^d$. Then, $|\boldsymbol{i}| = i_1 + \dots + i_d$ and $\boldsymbol{x}^{\boldsymbol{i}} = x_1^{i_1} \cdots x_d^{i_d}$. Let $m \in \mathbb{N}_0$. Put $D_m = \{\boldsymbol{i} \in \mathbb{N}_0^d \mid |\boldsymbol{i}| \leq m\}$ and define a linear ordering on D_m as follows. For any $0 \leq j \leq m$, denote by

$$N_j = \binom{j + d - 1}{d - 1} \tag{6}$$

the number of distinct d-tuples $\boldsymbol{i}$ with $|\boldsymbol{i}| = j$. Then, for $\boldsymbol{i}, \boldsymbol{j} \in D_m$, $\boldsymbol{i} \leq \boldsymbol{j}$ if $|\boldsymbol{i}| < |\boldsymbol{j}|$ or if $|\boldsymbol{i}| = |\boldsymbol{j}|$, then apply the reverse lexicographical ordering on $\boldsymbol{i}$ and $\boldsymbol{j}$. To emphasis the order of elements of D_m, we consider the ordered sequence $\boldsymbol{i}_1 < \dots < \boldsymbol{i}_N$, where $N = \sum_{j=0}^{m} N_j$. For example, $\boldsymbol{i}_1 = (0, 0, \dots, 0)$, $\boldsymbol{i}_2 = (1, 0, \dots, 0)$ and $\boldsymbol{i}_3 = (0, 1, 0, \dots, 0)$.

Any polynomial P of d variables and of degree m can be written in the form $P(\boldsymbol{x}) = \sum_{\ell=1}^{N} a_\ell \boldsymbol{x}^{\boldsymbol{i}_\ell}$, where $a_\ell \in \mathbb{R}$ for $\ell = 1, \dots, N$. We assume that a discrete function f of d variables is given at points $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$. Let $\{\mathcal{A}_{\boldsymbol{k}} \mid \boldsymbol{k} \in \mathbb{Z}^d\}$ be a uniform multivariate fuzzy partition of $\mathbb{R}^d$ determined by $(\mathcal{K}, \mathbf{H}, \mathbf{R}, \boldsymbol{t})$, and let $c_{\boldsymbol{k}} = \boldsymbol{k}\mathbf{R}$ denote the $\boldsymbol{k}$-th node. Denote by $\boldsymbol{Y} = (f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_n))^T$ the column vector of function values of f at points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$,

$$\boldsymbol{X}_{\boldsymbol{k}}^m = \begin{pmatrix} 1 & (\boldsymbol{x}_1 - \boldsymbol{c}_{\boldsymbol{k}})^{\boldsymbol{i}_2} & \cdots & (\boldsymbol{x}_1 - \boldsymbol{c}_{\boldsymbol{k}})^{\boldsymbol{i}_N} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (\boldsymbol{x}_d - \boldsymbol{c}_{\boldsymbol{k}})^{\boldsymbol{i}_2} & \cdots & (\boldsymbol{x}_d - \boldsymbol{c}_{\boldsymbol{k}})^{\boldsymbol{i}_N} \end{pmatrix}$$

the $n \times N$ matrix, where $N = \sum_{j=0}^{m} N_j$, and $\boldsymbol{A}_{\boldsymbol{k}} = \mathrm{diag}\{\mathcal{A}_{\boldsymbol{k}}(\boldsymbol{x}_1), \cdots, \mathcal{A}_{\boldsymbol{k}}(\boldsymbol{x}_n)\}$ which is the $n \times n$ diagonal matrix of weights computed as the function values of the basic function $\mathcal{A}_{\boldsymbol{k}}$ at points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. For the sake of simplicity, we omit m in $\boldsymbol{X}_{\boldsymbol{k}}^m$ and write only $\boldsymbol{X}_{\boldsymbol{k}}$ if no confusion can appear.

3.2 Discrete multivariate $\mathbf{F}_m$-transform

Let us assume that $I \subset \mathbb{R}^d$ is a compact subset and $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in I$.

Definition 3. Let $\mathbb{A}_I = \{\mathcal{A}_{\boldsymbol{k}} \mid \boldsymbol{k} \in \mathbb{K}\}$ be a uniform multivariate fuzzy partition of I. We say that the set of points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ is *sufficiently dense in* $\mathbb{A}_I$ if the matrix $\boldsymbol{X}_{\boldsymbol{k}}^T \boldsymbol{A}_{\boldsymbol{k}} \boldsymbol{X}_{\boldsymbol{k}}$ is invertible for any $\boldsymbol{k} \in \mathbb{K}$.

Now, we can proceed to the definition of the direct multivariate F-transform of higher degree that was introduced in [4].

Definition 4 (d-dimensional discrete $F_m^{\rightarrow}$-transform)**.** Let $\mathbb{A}$ be a uniform multivariate fuzzy partition of $\mathbb{R}^d$ determined by $(\mathcal{K}, \mathbf{H}, \mathbf{R}, \boldsymbol{c})$, let $I \subset \mathbb{R}^d$ be a compact subset, and let f be a discrete function given at points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in I$ that are sufficiently dense in $\mathbb{A}_I = \{\mathcal{A}_{\boldsymbol{k}} \mid \boldsymbol{k} \in \mathbb{K}\}$. A family $\{F_{m,\boldsymbol{k}}^{\rightarrow}[f](\boldsymbol{x}) \mid \boldsymbol{k} \in \mathbb{K}\}$ of polynomials of d variables of m degree in the form

$$F_{m,\boldsymbol{k}}^{\rightarrow}[f](\boldsymbol{x}) = \sum_{\ell=1}^{N} \beta_\ell^{\boldsymbol{k}} (\boldsymbol{x} - \boldsymbol{c}_{\boldsymbol{k}})^{\boldsymbol{i}_\ell}, \tag{7}$$

where $\boldsymbol{k} \in \mathbb{K}$ and $\boldsymbol{c}_{\boldsymbol{k}} = \boldsymbol{k}\mathbf{R}$, is called the *direct d-dimensional F-transform of m degree (F_m-transform for short) of f with respect to* $\mathbb{A}$ provided that

$$\boldsymbol{\beta}^{\boldsymbol{k}} = (\beta_1^{\boldsymbol{k}}, \ldots, \beta_N^{\boldsymbol{k}})^T = (\boldsymbol{X}_{\boldsymbol{k}}^T \boldsymbol{A}_{\boldsymbol{k}} \boldsymbol{X}_{\boldsymbol{k}})^{-1} \boldsymbol{X}_{\boldsymbol{k}}^T \boldsymbol{A}_{\boldsymbol{k}} \boldsymbol{Y} \tag{8}$$

for any $\boldsymbol{k} \in \mathbb{K}$. The polynomial $F_{\boldsymbol{k},m}^{\rightarrow}[f](\boldsymbol{x})$ is called the *$\boldsymbol{k}$-th component of F_m-transform of f with respect to* $\mathbb{A}$.

In [4], we proved that the discrete multivariate F_m-transform of a discrete function f given at points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, which is defined by

$$F_m[f](\boldsymbol{x}_i) = \sum_{\boldsymbol{k} \in \mathbb{K}} F_{m,\boldsymbol{k}}^{\rightarrow}[f](\boldsymbol{x}_i) \mathcal{A}_{\boldsymbol{k}}(\boldsymbol{x}_i), \tag{9}$$

well approximates f at points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. One could see that if f is a well-defined function on I and the values of f are known only for a sample $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in I$, then the function f can be approximated by formula (9), where $\boldsymbol{x}_i$ is now replaced by an arbitrary $\boldsymbol{x} \in I$.

4 Non-parametric regression

In this section, we restrict ourselves to two dimensional case, which is mostly used in practice including the implied volatility estimation. Let f be a function defined on a compact subset $I \subset \mathbb{R}^2$, let $Y_i = f(\boldsymbol{x}_i) + v^{1/2}(\boldsymbol{x}_i)\varepsilon_i$, $i = 1, \ldots, n$, be random variables, where $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in I$, $\mathrm{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = 1$, and $v(\boldsymbol{x}_i) \in \mathbb{R}_+$. Consider $Y = (Y_1, \ldots, Y_n)$ in formula (8), and assume that a uniform multivariate fuzzy partition $\mathbb{A}$ is given and the set of points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ is sufficiently dense in $\mathbb{A}_I = \{\mathcal{A}_{\boldsymbol{k}} \mid \boldsymbol{k} \in \mathbb{K}\}$. Then, a *non-parametric estimation of f* on I can be given by

$$\hat{f}(\boldsymbol{x}) = F_m[f](\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{K}} F_{m,\boldsymbol{k}}^{\rightarrow}[f](\boldsymbol{x}) \mathcal{A}_{\boldsymbol{k}}(\boldsymbol{x}), \quad \boldsymbol{x} \in I, \tag{10}$$

where $F^{\rightarrow}_{m,\boldsymbol{k}}[f](\boldsymbol{x})$, $\boldsymbol{k} \in \mathbb{K}$, are the F_m-transform components with respect to $\mathbb{A}$.

Let us assume that f is a continuous function, which is p_i times differentiable function in the i-th coordinate at each point $\boldsymbol{x} \in \operatorname{Int} I$, where $\operatorname{Int} I$ denotes the interior of I. Put $\mathbf{p} = (p_1, p_2)$, and let $\boldsymbol{j} \leq \mathbf{p}$. Then, a *non-parametric estimation of partial derivative* $\partial^{\boldsymbol{j}} f = \frac{\partial^{|\boldsymbol{j}|} f}{\partial^{j_1} x_i \partial^{j_2} x_2}$ on $\operatorname{Int} I$ can be given by

$$\widehat{\partial^{\boldsymbol{j}} f}(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{K}} \partial^{\boldsymbol{j}} F^{\rightarrow}_{m,\boldsymbol{k}}[f](\boldsymbol{x}) \mathcal{A}_{\boldsymbol{k}}(\boldsymbol{x}), \quad \boldsymbol{x} \in \operatorname{Int} I, \tag{11}$$

where $\partial^{\boldsymbol{j}} F^{\rightarrow}_{m,\boldsymbol{k}}[f](\boldsymbol{x}) = \frac{\partial^{|\boldsymbol{j}|}}{\partial^{j_1} x_i \partial^{j_2} x_2} F^{\rightarrow}_{m,\boldsymbol{k}}[f](\boldsymbol{x})$. Denote $s(\boldsymbol{x}, \boldsymbol{k}) = (1, (\boldsymbol{x} - \boldsymbol{c}_{\boldsymbol{k}})^{\boldsymbol{i}_2}, \dots, (\boldsymbol{x} - \boldsymbol{c}_{\boldsymbol{k}})^{\boldsymbol{i}_N})$, $\boldsymbol{k} \in \mathbb{K}$, and

$$D^{\boldsymbol{j}} = \operatorname{diag}\left\{ \frac{\partial^{|\boldsymbol{j}|} \boldsymbol{x}^{j_1}}{\partial^{j_1} x_i \partial^{j_2} x_2}(1,1), \frac{\partial^{|\boldsymbol{j}|} \boldsymbol{x}^{\boldsymbol{j}_2}}{\partial^{j_1} x_i \partial^{j_2} x_2}(1,1), \dots, \frac{\partial^{|\boldsymbol{i}|} \boldsymbol{x}^{\boldsymbol{i}_N}}{\partial^{i_1} x_1 \partial^{j_2} x_2}(1,1) \right\},$$

Then, (11) can be rewrite with help of matrices as follows:

$$\widehat{\partial^{\boldsymbol{i}} f}(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{K}} \left(s(\boldsymbol{x}, \boldsymbol{k}) D^{\boldsymbol{i}} \boldsymbol{\beta}^{\boldsymbol{k}} \right) \mathcal{A}_{\boldsymbol{k}}(\boldsymbol{x}), \quad \boldsymbol{x} \in \operatorname{Int} I, \tag{12}$$

where $\boldsymbol{\beta}^{\boldsymbol{k}} = (\boldsymbol{X}_{\boldsymbol{k}}^T \boldsymbol{A}_{\boldsymbol{k}} \boldsymbol{X}_{\boldsymbol{k}})^{-1} \boldsymbol{X}_{\boldsymbol{k}}^T \boldsymbol{A}_{\boldsymbol{k}} \boldsymbol{Y}$.

Example 2. Let us consider the continuous function $f(x, y) = (x + y)\sin(x/\pi)\cos(y/\pi)$, and assume that we know $Y_{ij} = f(i, j) + 16\,\xi(i, j)$, where $\xi(i, j) \sim N(0, 1)$ for any $1 \leq i, j \leq 15$. In Fig. 2, the estimation of f (left) and f'_x (right) from the sample (i, j, Y_{ij}), $0 \leq i, j \leq 15$, is depicted, where the blue dots represent the original discrete function $f(i, j)$ and its partial derivative $f'_x(i, j)$. We used the 2D triangle uniform fuzzy partition with $\mathbf{H} = \operatorname{diag}\{8, 8\}$, $\mathbf{R} = \operatorname{diag}\{4, 4\}$, and $\boldsymbol{c} = (0, 0)$. One can see that a portion of the white noise is suppressed and also the first derivative with respect to x are well estimated.

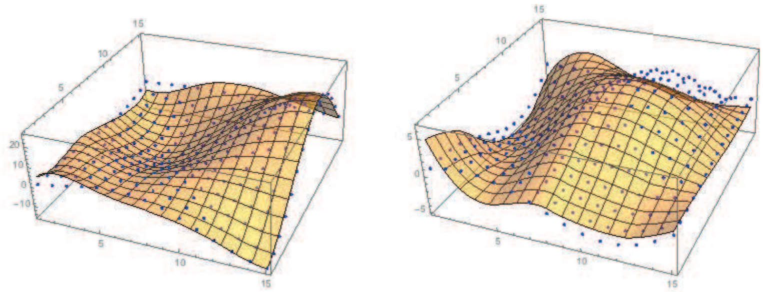


Figure 2 Estimation of the function f (left) as well as its partial derivative f'_x (right) from Example 2

5 Application to financial modeling

In this section, we will provide an example of the higher degree F-transform on real data within the classic problem of financial option pricing. Before we proceed to a real data, we briefly describe the motivation for the study. A standard approach to option pricing is based on Black-Scholes type (BS hereafter) models [2] utilizing the no-arbitrage argument of complete markets. However, there are several crucial assumptions, such as that the option underlying log-returns follow normal distribution, there is unique and deterministic riskless rate as well as the volatility of underlying log-returns. Since the assumptions are generally not fulfilled, the BS-type models provide false results. This is why market participants often use an artificial volatility, which match BS model with market prices (generally called implied volatility since it is implied by market prices). Obviously, such volatilities differ for various maturities and strike prices.[1]

In order to illustrate the procedure we have pick up implied volatilites of stock option traded at German market (Allianz) on a given day with various (normalized) strike prices (x) and maturities (y), see Figure 3. For the estimation, we used the F-transform of degree 3 w.r.t. the triangle 2-dimensional uniform fuzzy partition with $\mathbf{H} = \operatorname{diag}\{70, 1.2\}$, $\mathbf{R} = \operatorname{diag}\{35, 0.6\}$, and $\boldsymbol{c} = (2.5, 0)$. This estimation of respective partial derivatives can be simply obtained by formula (11) or (12).

[1] Recall, however, that when discrete observations of implied volatilities are smoothed to get a so called implied volatility surface we should take care about no-arbitrage conditions, where a state price densities play a crucial role, see e.g. [1] for more details.

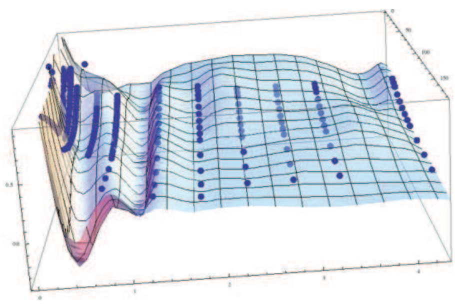


Figure 3 Estimation of implied volatility surface by multivalued F-transform of degree 3

6 Conclusion

The F-transform is an approximation technique applicable in various field of science. In this contribution its original definition was extended into multivariate higher degree F-transform so that we were able to smooth (2+1)D discrete data. A specific advantage of the approach is that one can simultaneously derive also derivatives of the smoothed function, which can has interesting application eg. in no-arbitrage estimation of option implied volatilities.

Acknowledgments

This work was supported by the project LQ1602 IT4Innovations excellence in science. The additional support was also provided by the Czech Science Foundation (GACR) through project No.16-09541S and moreover via SGS project SP2016/11 of VSB-TU Ostrava.

References

[1] Benko, M., Fengler, M., Hardle, W. and Kopa, M.: On extracting information implied in options, *Computational Statistics* **22** (2007), 543–553.

[2] Black, F. and Scholes, M.: The pricing of options and corporate liabilities, *Journal of Political Economy* **81** (1973), 637–659.

[3] Hodáková, P.: Fuzzy (F-)transform of functions of two variables and its applications in image processing, PhD thesis, University of Ostrava 2015.

[4] Holčapek, M. and Tichý, T.: Discrete multivariate F-transform of higher degree, in *In Proc. of 2015 IEEE International Conference on Fuzzy Systems*, Istanbul, 2015.

[5] Holčapek, M., Perfilieva, I., Novák, V., Kreinovich, V.: Necessary and sufficient conditions for generalized uniform fuzzy partitions, *Fuzzy sets and systems* **277** (2015), 97–121.

[6] Pagan, A. and Ullah, A.: *Nonparametric Econometrics*. Cambridge University Press, New York, 1999.

[7] Perfilieva, I.: Fuzzy transforms: Theory and applications, *Fuzzy sets and systems* **157(8)** (2006), 993–1023.

[8] Perfilieva, I. and Daňková, M.: Towards F-transform of a higher degree, in *In Proc. of IFSA/EUSFLAT 2009*, Lisbon, Portugal, 2009.

[9] Perfilieva, I., Hodáková, P. and Hurtík, P.: Differentiation by the F-transform and application to edge detection, *Fuzzy sets and systems* **288** (2016), 96–114.

[10] Perfilieva, I., Holčapek, M. and Kreinovich, V.: A new reconstruction from F-transform components, *Fuzzy sets and systems* **288** (2016), 3–25.

[11] Wand, M.-P. and Jones, M. *Kernel Smoothing*. Chapman&Hall/CRC Monographs on Statistics & Applied Probability, London, 1995.

A free software tool implementing the fuzzy AHP method

Pavel Holeček[1], Jana Talašová[2]

Abstract. The AHP method became very popular in multiple-criteria decision-making and it found its applications in diverse fields. Over the time, several modifications of the method for fuzzy environment have been devised. The paper introduces a new free software tool that implements one of these approaches. The elements of the pair-wise comparison matrix are allowed to be expressed by triangular fuzzy elements. The classical non-fuzzy methods based on the eigenvectors or the geometric means are also supported in the software. The presented software has been written as a web application, which means that it is available from any computer connected to the Internet without need to install any additional software. The presented tool makes it possible to design the (fuzzy) pair-wise comparison matrix in a user-friendly way, and to derive the priority vector from it. Various consistency indicators are also calculated.

Keywords: fuzzy, AHP, pair-wise comparison matrix, triangular fuzzy elements, software

JEL classification: C44
AMS classification: 90B50

1 Introduction

Methods based on the use of pair-wise comparison matrices represent an important group in the multiple-criteria decision-making. Probably the best-known representative of this group is Saaty's AHP [9]. Over the time, various authors devised modifications of the method suitable for fuzzy environment (e.g. [11, 4, 3, 7]). The aim of this paper is to introduce a new free software tool that can be used for calculations with the pair-wise comparison matrices used in the AHP, or their fuzzy versions. The matrices are allowed to contain not only real elements, but also triangular fuzzy elements. The current version of the software is able to derive (fuzzy) weights from these matrices and to measure the inconsistency of the data provided by the expert. In the future, another development and extension of the software is planned.

The paper is structured as follows. First, the mathematical methods used in the software will be described. As the software can work with both non-fuzzy and fuzzy pair-wise comparison matrices, methods for both of the cases will be described. Next, the new software tool will be presented. The functions and possibilities of this software will be discussed. Finally, possible directions for the future development will be outlined.

2 The used methods

In the presented software, the expert gives his/her preferences in form of a pair-wise comparison matrix. This matrix can contain non-fuzzy elements only (real numbers) or triangular fuzzy elements. In order to derive the weights from such a matrix and to measure the consistency of the information on expert's preferences encoded in the matrix, the following methods have been implemented.

[1] Palacký University in Olomouc, Faculty of Science, Department. of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/121, 77146 Olomouc, Czech Republic , pavel.holecek@upol.cz.
[2] Palacký University in Olomouc, Faculty of Science, Department. of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/121, 77146 Olomouc, Czech Republic , jana.talasova@upol.cz.

2.1 Pair-wise comparison matrix with non-fuzzy elements

Let us assume, that n criteria are taken into the account. In case of the classical AHP without any fuzziness involved, the expert provides the information on his/her preferences in form of a pair-wise comparison matrix, which has the following form:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \tag{1}$$

where, for all $i, j = 1, \dots, n$, the elements a_{ij} are positive and satisfy the condition that $a_{ij} = 1/a_{ji}$ (the matrix is reciprocal).

In the original AHP [9], the elements of the matrix A are taken from Saaty's scale $\{\frac{1}{9}, \frac{1}{8}, \dots, \frac{1}{2}, 1, 2, \dots, 9\}$. For example, by setting the value $a_{23} = 5$, the expert states that the second criterion is five times more important than the third one.

The weights of the individual criteria are derived in the original AHP using the *eigenvector method*. In this method, the eigenvector $\vec{v} = (v_1, \dots, v_n)$ corresponding to the maximum eigenvalue $\lambda_{\max}$ of the matrix A is determined. The weights w_k, $k = 1, \dots, n$, of the individual criteria are then determined as $w_k = v_k / \sum_{i=1}^{n} v_i$.

To propose the weights from a pair-wise comparison matrix, other methods are also used in the practice [6] and probably the best-known of them is the *geometric mean* method. In this method, the weights w_k, $k = 1, \dots, n$, are calculated as follows:

$$w_k = \frac{\left(\prod_{j=1}^{n} a_{kj}\right)^{1/n}}{\sum_{i=1}^{n} \left(\prod_{j=1}^{n} a_{ij}\right)^{1/n}}. \tag{2}$$

The weights obtained by this method are usually close to the ones calculated with the eigenvector method. Great advantages of this method are its simplicity and low computational demands.

In order to ensure the inconsistency of the pair-wise comparison matrix is on an acceptable level, Saaty proposed the following procedure [9]:

First, the *consistency index* is calculated:

$$CI = \frac{\lambda_{\max} - n}{n - 1}, \tag{3}$$

where $\lambda_{\max}$ is the maximum eigenvalue of the matrix A.

Subsequently, the *random index* RI for the given number of criteria n is considered. The RI is the average CI of randomly generated pair-wise comparison matrices of the same size as A with elements from the Saaty's scale. The RI is tabulated for different n in Table 1 [10]. Because of its random nature, the particular values may differ slightly in the literature (for comparison of values determined by different researchers, see [1]).

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49

Table 1 Values of the *random index* (RI) for different sizes (n) of the pair-wise comparison matrix.

Finally, the *consistency ratio* CR is calculated as follows:

$$CR = \frac{CI}{RI}. \tag{4}$$

According to the procedure proposed by Saaty, the matrix A is considered to be sufficiently consistent if $CR < 0.1$. Otherwise, the expert should revise the values in the pair-wise comparison matrix.

2.2 Pair-wise comparison matrix with triangular fuzzy elements

Sometimes, it can be difficult for the expert to compare certain pairs of criteria. In these cases, it would be more realistic to allow the expert to provide the data not only in form of real numbers, but also fuzzy numbers. Multiple methods that use pair-wise comparison matrices with fuzzy elements were proposed. From the vast range of the methods, the approach described by Ramík [7] has been selected and implemented in the presented software. The major advantages of this approach are its solid mathematical basis and suitability for computations. In this approach, the pair-wise comparison matrix is comprised of triangular fuzzy elements. Such a matrix $\tilde{A}$ is of the following form:

$$\tilde{A} = \begin{bmatrix} (a_{11}^L, a_{11}^M, a_{11}^U) & \dots & (a_{1n}^L, a_{1n}^M, a_{1n}^U) \\ \vdots & \ddots & \vdots \\ (a_{n1}^L, a_{n1}^M, a_{n1}^U) & \dots & (a_{nn}^L, a_{nn}^M, a_{nn}^U) \end{bmatrix}, \tag{5}$$

where for all $i, j = 1, \dots, n$:

- $a_{ij}^L, a_{ij}^M, a_{ij}^U$ are real numbers such that $1/\sigma \leq a_{ij}^L \leq a_{ij}^M \leq a_{ij}^U \leq \sigma$ for a chosen fixed $\sigma > 1$.
- $\tilde{a}_{ij} = \left(a_{ij}^L, a_{ij}^M, a_{ij}^U\right)$ implies that $\tilde{a}_{ji} = \left(\frac{1}{a_{ij}^U}, \frac{1}{a_{ij}^M}, \frac{1}{a_{ij}^L}\right)$. *(reciprocity)*

Besides the introduction of the fuzzy triangular elements, another difference compared to the classical AHP is that the preference intensities provided by the expert are not limited to the interval $[\frac{1}{9}, 9]$, but can be taken more generally form $[\frac{1}{\sigma}, \sigma]$ for a chosen value $\sigma > 1$.

The fuzzy weights $\tilde{w}_k = (w_k^L, w_k^M, w_k^U)$, $k = 1, \dots, n$, are then derived in this procedure as follows [7]:

$$w_k^L = C_{\min} \cdot \frac{\left(\prod_{j=1}^n a_{kj}^L\right)^{1/n}}{\sum_{i=1}^n \left(\prod_{j=1}^n a_{ij}^M\right)^{1/n}}, \quad \text{where } C_{\min} = \min_{i=1,\dots,n} \left\{ \frac{\left(\prod_{j=1}^n a_{ij}^M\right)^{1/n}}{\left(\prod_{j=1}^n a_{ij}^L\right)^{1/n}} \right\}, \tag{6}$$

$$w_k^M = \frac{\left(\prod_{j=1}^n a_{kj}^M\right)^{1/n}}{\sum_{i=1}^n \left(\prod_{j=1}^n a_{ij}^M\right)^{1/n}}, \tag{7}$$

$$w_k^U = C_{max} \cdot \frac{\left(\prod_{j=1}^n a_{kj}^U\right)^{1/n}}{\sum_{i=1}^n \left(\prod_{j=1}^n a_{ij}^M\right)^{1/n}}, \quad \text{where } C_{\max} = \max_{i=1,\dots,n} \left\{ \frac{\left(\prod_{j=1}^n a_{ij}^M\right)^{1/n}}{\left(\prod_{j=1}^n a_{ij}^U\right)^{1/n}} \right\}. \tag{8}$$

To measure the consistency of the pair-wise comparison matrix with triangular fuzzy elements, Ramík proposed the following index [7]:

$$NI_n^\sigma(\tilde{A}) = \gamma_n^\sigma \cdot \max_{i,j} \left\{ \max \left\{ \left| \frac{w_i^L}{w_j^U} - a_{ij}^L \right|, \left| \frac{w_i^M}{w_j^M} - a_{ij}^M \right|, \left| \frac{w_i^U}{w_j^L} - a_{ij}^U \right| \right\} \right\}, \tag{9}$$

where

$$\gamma_n^\sigma = \begin{cases} \frac{1}{\max\left\{\sigma - \sigma^{(2-2n/n)}, \sigma^2\left(\left(\frac{2}{n}\right)^{2/(n-2)} - \left(\frac{2}{n}\right)^{n/(n-2)}\right)\right\}} & \text{if } \sigma < \left(\frac{n}{2}\right)^{n/(n-2)}, \\ \frac{1}{\max\left\{\sigma - \sigma^{(2-2n/n)}, \sigma^{(2n-2/n)} - \sigma\right\}} & \text{otherwise.} \end{cases} \tag{10}$$

The value of the index ranges from 0 to 1, where 0 means that the matrix is fully consistent.

The mathematical theory described in this section has been implemented into a software tool whose introduction is the main goal of this paper. This software tool will be the topic of the next section.

3 The FuzzyAHP software tool

Because of the popularity of the AHP method, many software tools supporting this method have emerged (their list can be found for example in [2]). Although there are plenty of tools for the classical (non-fuzzy) AHP, there is a lack of software tools suitable for fuzzy problems. In [5], a fuzzy AHP software tool applied for a particular problem is described. However, the paper does not contain any download link and the software does not seem to be available on the Internet for the public. Another tool that works with pair-wise comparison matrices with fuzzy elements is FVK [8]. The method used in this tool makes it possible to take into account also the dependencies among the individual criteria. A great advantage of the tool is that it is available for free. The tool has been written as an add-on for Microsoft Excel. This can represent a limitation – the software is restricted to the computers with Windows operation system and Microsoft Excel only.

The fuzzy AHP tool presented in the paper is a web application. This means that it can be accessed from any computer connected to the Internet without need to buy or install any additional software. The tool is available for free.

3.1 The features of the introduced software tool

The presented software tool makes it possible to derive the weights and information on consistency from a pair-wise comparison matrix with either triangular fuzzy elements or with non-fuzzy elements (real number). Such calculation is performed in the software as follows. First, the user defines the number of criteria and their names. Then, the pair-wise comparison matrix has to be filled in (Figure 1). The user enters the values of the elements above the main diagonal. The other elements are calculated by the software automatically. This way the reciprocity of the pair-wise comparison matrix is ensured.

The expert enters each triangular fuzzy element as three numbers divided by a space. For simplicity, the software makes it possible to enter also a single number as an element of the matrix, which is then treated as a fuzzy element whose all three values are equal.

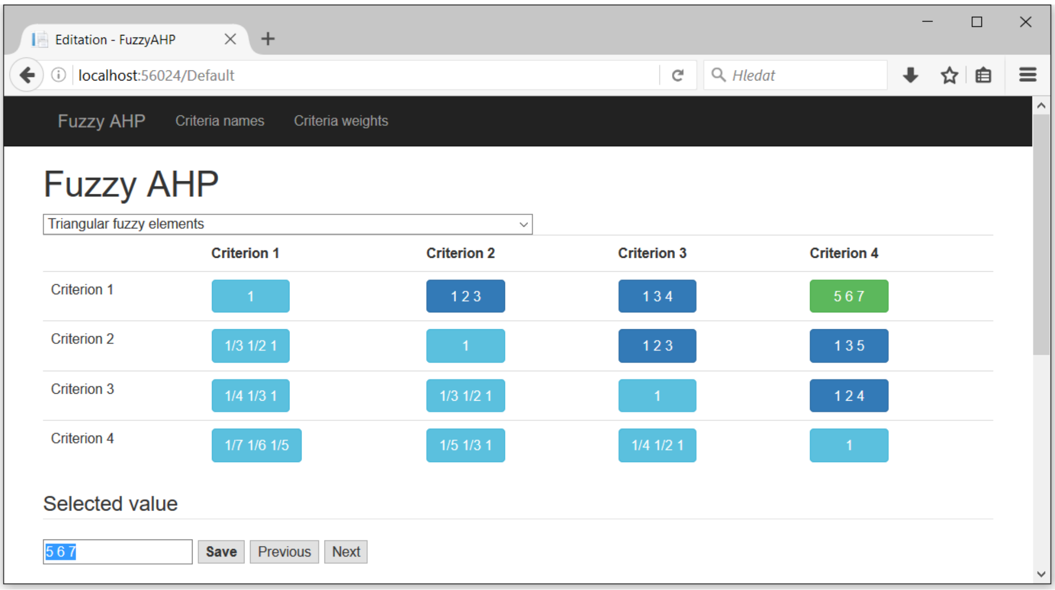


Figure 1 Editing a pair-wise comparison matrix with triangular fuzzy elements in the fuzzy AHP tool

As soon as all necessary data are filled in, the results are displayed. By default, the procedure described in Section 2.2 is used. The presentation of the results calculated by this method for the matrix in Figure 1 is depicted in Figure 2. The derived fuzzy weights are displayed in a numerical form and in form of a graph. Information on consistency through the NI index is also provided.

The software can be also used for calculation with a classical Saaty's matrix (without any fuzziness). The expert can define such matrix in the same way as before (the only difference is that single values instead of triplets are provided as elements of the matrix). The eigenvector method and the geometric

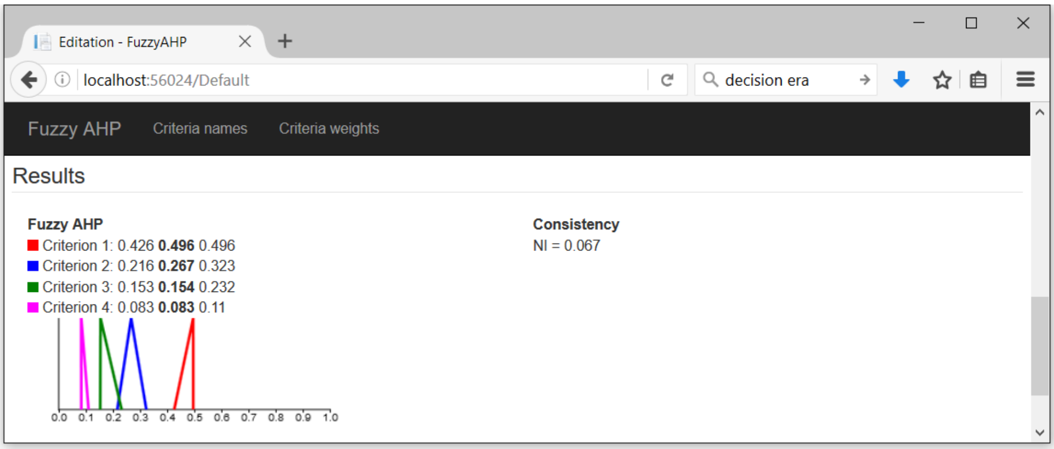


Figure 2 The results calculated from the pair-wise comparison matrix with triangular fuzzy elements

mean method can then be used for deriving the weights. The consistency indices mentioned in Section 2.1 are also calculated. The form, in which the results are presented in case of these non-fuzzy methods, can be seen in Figure 3.

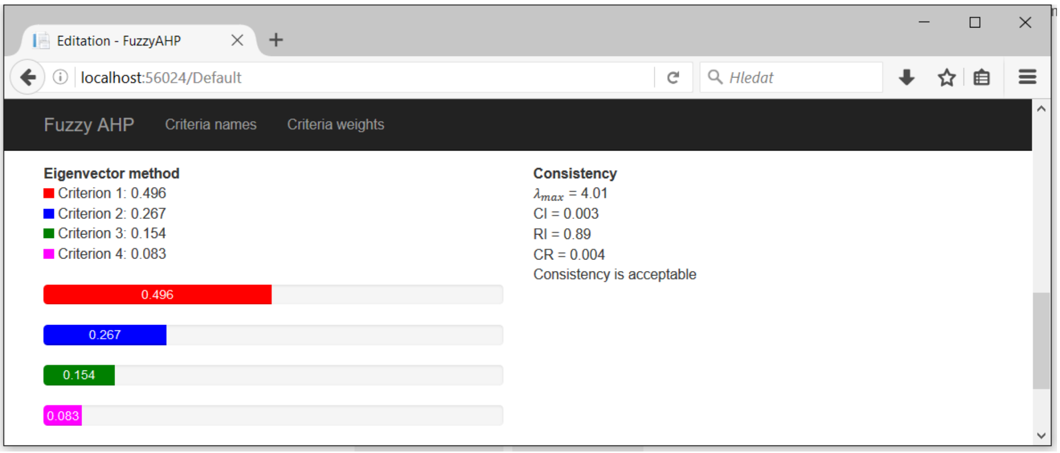


Figure 3 The results obtained by the eigenvector method for a pair-wise comparison matrix (the case without any fuzziness)

3.2 Plans for the future development

The further development of the software is planned in the future. In this section, some possible improvements that should be implemented in the later versions will be discussed.

The current version of the software is able to work with only one pair-wise comparison matrix at the time. This is sufficient for experimenting with the supported methods and studying their behavior. However, the future versions should support full hierarchical structure as it is used in AHP. This will make it possible for the expert to use the software for solving complex multiple-criteria decision-making problems.

Currently, three methods are supported in the software – two non-fuzzy methods and a single fuzzy method. In the literature, there is a whole range of other methods for fuzzy environment. Some of these methods could be implemented into the software in the future. This would give the expert an opportunity to compare the results obtained by different approaches.

4 Conclusion

The paper introduced a new software tool that works with pair-wise comparison matrices. The software makes it possible for the expert to define such matrix in a quick and comfortable way. Both classical (non-fuzzy) pair-wise comparison matrices and pair-wise comparison matrices with triangular fuzzy elements are supported. The software can derive the (fuzzy) weights and calculate the indices of consistency for the data provided by the expert. Multiple methods are supported for this task.

The advantage of the presented software is that it has been written as a web application so it is multiplatform (it can be used from any operation system) and does not require installation of any additional software. Another advantage is the price. The software is available for free.

This tool can be found at the following web address: `http://fuzzymcdm.upol.cz`.

Acknowledgements

The research has been supported by the grant GA14-02424S *Methods of operations research for decision support under uncertainty* of the Grant Agency of the Czech Republic.

References

[1] Alonso, J.-A., and Lamata, M. T.: Consistency in the analytic hierarchy process: a new approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **14**, 4 (2006), 445–459.

[2] Baizyldayeva, U., and Vlasov, O.: Multi-criteria decision support systems. Comparative analysis. *Middle-East Journal of Scientific Research* **16**, 12 (2013), 1725–1730.

[3] Buckley, J. J., Feuring, T., and Hayashi, Y.: Fuzzy hierarchical analysis revisited. *European Journal of Operational Research* **129**, 1 (2001), 48–64.

[4] Chang, D.-Y.: Applications of the extent analysis method on fuzzy AHP. *European Journal of Operational Research* **95**, 3 (1996), 649–655.

[5] Durán, O.: Computer-aided maintenance management systems selection based on a fuzzy AHP approach. *Advances in Engineering Software* **42**, 10 (2011), 821–829.

[6] Ishizaka, A., and Lusti, M.: How to derive priorities in AHP: a comparative study. *Central European Journal of Operations Research* **14**, 4 (2006), 387–400.

[7] Ramík, J., and Korviny, P.: Inconsistency of pair-wise comparison matrix with fuzzy elements based on geometric mean. *Fuzzy Sets and Systems* **161**, 11 (2010), 1604–1613.

[8] Ramík, J., and Perzina, R.: Solving decision problems with dependent criteria by new fuzzy multicriteria method in Excel. *Journal of Business and Management* **3**, 4 (2014), 1–16.

[9] Saaty, T. L.: *The Analytic Hierarchy Process.* McGraw-Hill, New York, 1980.

[10] Saaty, T. L., and Tran, L. T.: On the invalidity of fuzzifying numerical judgments in the Analytic Hierarchy Process. *Mathematical and Computer Modelling* **46** (2007), 962–975.

[11] van Laarhoven, P., and Pedrycz, W.: A fuzzy extension of Saaty's priority theory. *Fuzzy Sets and Systems* **11** (1983), 199–227.

Impact of Microstructure Noise on Integrated Variance Estimators: A Simulation Study

Vladimír Holý[1]

Abstract. Using high-frequency data for estimation of integrated variance of asset prices is beneficial but so-called microstructure noise occurs as the sampling frequency increases. This noise is caused by rounding of prices, bid-ask spread and overall discrete nature of transactions and can significantly bias estimates.

Microstructure noise is often modelled as i.i.d. variables independent of the efficient price process. However, real data indicate much more complicated structure than a white noise. In this paper we consider various models of microstructure noise. First, we model the noise as AR and MA time series. Next, we consider the noise dependent on asset prices. Another type of model is a noise created purely by rounding of prices.

We study the impact of microstructure noise on realized variance and other integrated variance estimators such as maximum likelihood estimator, the realized kernel estimator and the pre-averaging estimator. We simulate various structures of microstructure noise which violate the assumptions of estimators and we analyze the resulting integrated variance estimates.

Keywords: High-frequency data, microstructure noise, integrated variance estimation, realized volatility, simulations.

JEL classification: C53, C58, G17
AMS classification: 62M10, 91G70

1 Introduction

The main instrument for risk analysis of financial asset prices is variance estimation. To estimate daily or even intraday integrated variance, data sampled at high frequencies are required. As the sampling frequency increases so-called microstructure noise occurs. It is caused by rounding of prices, bid-ask spread and overall discrete nature of transactions and can significantly bias estimates. This noise is often modelled with assumptions that are not valid for real data. Some methods consider the noise independent of the efficient price process, some methods even assume simple white noise. However, microstructure noise can have much richer structure. For example, it was shown in [4] that microstructure noise in Dow Jones Industrial Average stocks returns is correlated with the efficient price, time dependent and substantially changes over time.

The goal of this paper is to illustrate the impact of some microstructure noise models on integrated variance estimators using simulations. In Section 2 we define some microstructure noise models. In Section 3 we introduce integrated variance estimators and analyze the impact of the noise. Finally, in Section 4, we summarize achieved results of simulations.

2 Microstructure Noise Setting

Let p_t^* be one-dimensional logarithmic price process of the asset and let's assume that it follows a continuous semi-martingale

$$p_t^* = p_0^* + \int_0^t \mu(s)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}W_s, \tag{1}$$

[1]University of Economics, Prague, W. Churchill Sq. 1938/4, 130 67 Prague 3, Czech Republic, vladimir.holy@vse.cz.

where W_s denotes the Wiener process, $\mu(s)$ is a finite variation càglàg drift process and $\sigma(s)$ is an adapted càdlàg volatility process. The process p_t^* is called efficient price. However, rather than p_t^*, we observe process p_t defined as

$$p_t = p_t^* + \varepsilon_t, \tag{2}$$

where ε_t is the mentioned market microstructure noise. This is common setting used for example in [5].

In our simulations, we consider efficient price p_t^* as standard Wiener process with unit variance. This assumption is not realistic for financial data, but it can be used as illustration for some basic behaviour as stated in [1]. In the following subsections we present models for microstructure noise.

2.1 White Noise

The most basic microstructure noise model is a white noise

$$\varepsilon_t \overset{i.i.d.}{\sim} \mathrm{N}(0, \sigma^2). \tag{3}$$

This noise is independent of p_t^*. We use unit variance for all noise models except the rounding noise model.

2.2 Autoregressive Noise

Next, we consider the noise as AR(1) time series

$$\varepsilon_t = \varphi \varepsilon_{t-1} + \chi_t, \quad \chi_t \overset{i.i.d.}{\sim} \mathrm{N}(0, \sigma^2). \tag{4}$$

In simulations, we use parameters $\varphi = 0.5$ (denoted in tables 1 and 2 as $\mathrm{AR}(1)_I$ noise model) and $\varphi = -0.9$ (denoted as $\mathrm{AR}(1)_{II}$).

2.3 Moving-average Noise

Another time series model is moving-average MA(1)

$$\varepsilon_t = \theta \chi_{t-1} + \chi_t, \quad \chi_t \overset{i.i.d.}{\sim} \mathrm{N}(0, \sigma^2). \tag{5}$$

In simulations, we use parameter $\theta = 0.7$.

2.4 Dependent Noise

Many integrated variance estimators assume microstructure noise independent of efficient price. However this may not always be the case and so we consider the microstructure noise ε_t correlated with p_t^*. In simulations, we use correlation coefficient $\rho = 0.9$.

2.5 Rounding Noise

Last model we consider is noise created purely by rounding of prices given by

$$\varepsilon_t = \lfloor p_t^* \rceil_k - p_t^*, \tag{6}$$

where $\lfloor p_t^* \rceil_k$ is p_t^* rounded to k digits. In simulations, we use $k = 1$. Unlike others the rounding noise is deterministic.

3 Impact of Microstructure Noise on Estimators

To quantify the risk of an asset, so-called integrated variance is used. It is defined as

$$IV(0,t) = \int_0^t \sigma^2(s)\mathrm{d}s. \tag{7}$$

There are many integrated variance estimators. In the following subsections we will briefly introduce a few of them and examine how they are affected by different types of noise.

The estimation of integrated variance using noisy data exhibits the following behaviour. If lower frequencies are used, the impact of the noise diminishes but the estimation is less precise. In Figure 1 we can see that the confidence intervals are wider and in Table 2 we can see that the standard deviations of mean errors are higher for greater period. Some papers recommend to use realized variance with 5 minute data[1], but most papers (e.g. [6]) recommend to use estimators robust to microstructure noise with data sampled at the highest possible frequency.

In simulations we estimate daily integrated variance using different frequencies ranging from 1 minute data to 60 minute data.

		Microstructure Noise Models					
Estimator	*Period*	White	$AR(1)_I$	$AR(1)_{II}$	MA(1)	Dep.	Round.
	1 min	**1.998**	**1.003**	**3.792**	**1.061**	**0.527**	**0.165**
Realized Variance	5 min	0.998	0.753	0.194	1.001	0.342	0.084
	20 min	0.670	0.584	1.148	0.661	0.275	0.054
	1 min	**-0.002**	**0.422**	**-0.340**	**0.844**	**0.149**	**0.000**
Maximum Likelihood	5 min	-0.005	0.131	0.150	-0.006	0.142	-0.005
	20 min	-0.011	0.043	-0.237	-0.011	0.139	-0.007
	1 min	**0.032**	**0.077**	**0.002**	**0.060**	**0.154**	**0.003**
Realized Kernel	5 min	0.059	0.083	0.103	0.060	0.158	0.003
	20 min	0.081	0.093	0.020	0.081	0.160	0.003
	1 min	**-0.035**	**-0.008**	**-0.058**	**-0.017**	**0.107**	**-0.034**
Pre-averaging	5 min	-0.058	-0.048	-0.009	-0.058	0.080	-0.058
	20 min	-0.079	-0.074	-0.097	-0.078	0.056	-0.079

Table 1 Means of relative errors of variance estimations calculated from 1000 simulations

3.1 Realized Variance Estimator

The natural approximation of integrated variance is realized variance given by

$$RV^n = \sum_{i=1}^{n} (p_{i\Delta_n} - p_{(i-1)\Delta_n})^2, \tag{8}$$

where Δ_n is the period corresponding to the number of observations n. This estimator is consistent only when there is no noise present. It is known that in the presence of white noise the realized variance diverges lineary to infinity with increasing number of observations. This is shown for example in [4]. As we can see in Figure 1, with other noise settings the divergence is nonlinear. When there is a strong negative autocorrelation in the noise, there is a danger of missing specific periodic parts of noise using some frequencies resulting in polyline in the third pair of plots in Figure 1. In the dependent noise setting we can see that even when using relatively lower frequencies the realized variance can still be significantly biased.

[1] Better a sparrow in the hand than a pigeon on the roof.

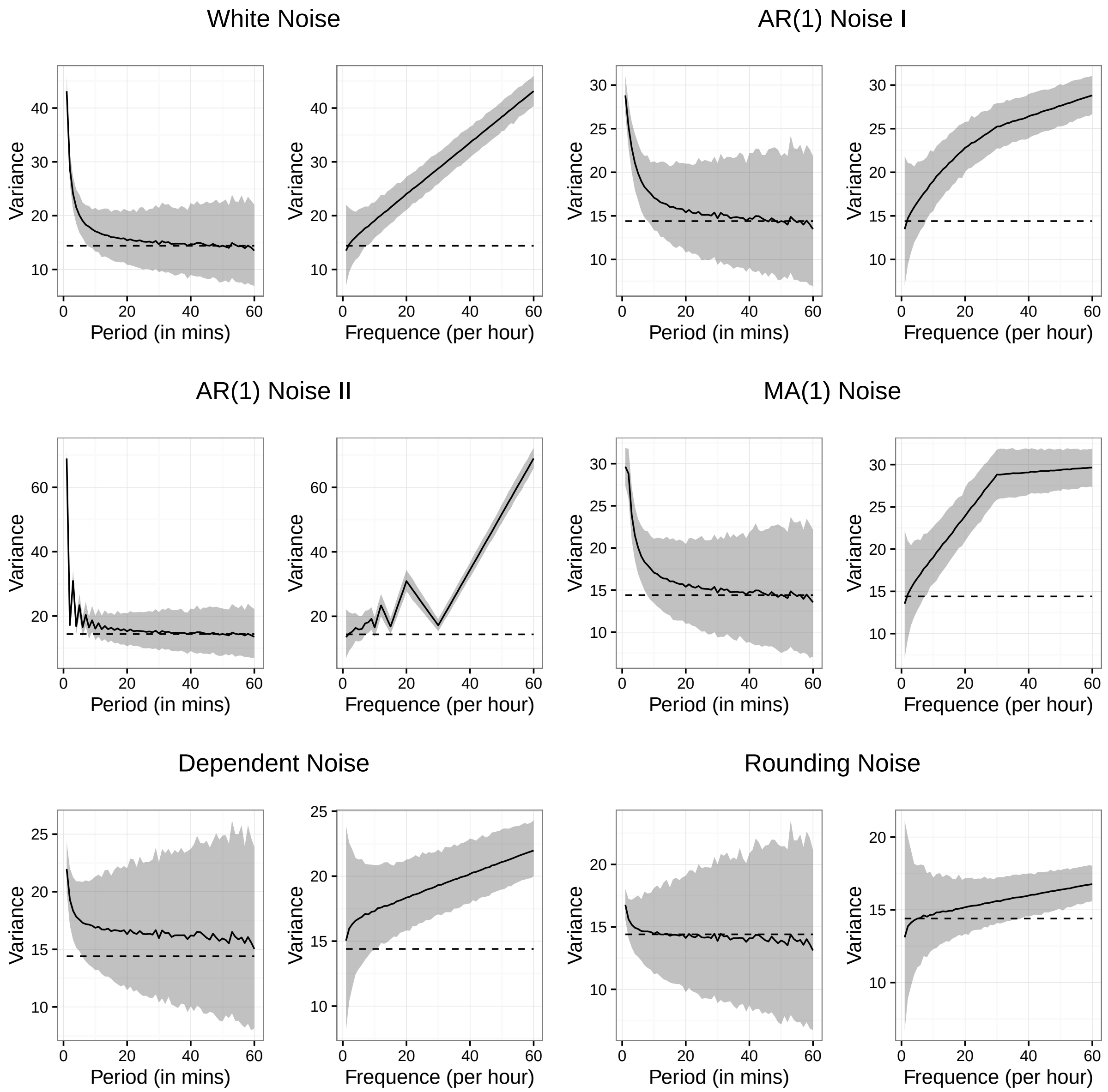


Figure 1 Means of realized variances (solid curves) with 95% confidence intervals (grey areas) and true integrated variances (dotted lines) for different microstructure noise models calculated from 1000 simulations

		Microstructure Noise Models					
Estimator	*Period*	White	$AR(1)_I$	$AR(1)_{II}$	MA(1)	Dep.	Round.
	1 min	**0.094**	**0.079**	**0.110**	**0.077**	**0.077**	**0.044**
Realized Variance	5 min	0.104	0.094	0.065	0.104	0.089	0.057
	20 min	0.109	0.103	0.123	0.109	0.098	0.068
	1 min	**0.089**	**0.109**	**0.057**	**0.164**	**0.096**	**0.067**
Maximum Likelihood	5 min	0.112	0.124	0.110	0.114	0.122	0.095
	20 min	0.129	0.138	0.099	0.130	0.143	0.117
	1 min	**0.119**	**0.123**	**0.118**	**0.121**	**0.138**	**0.118**
Realized Kernel	5 min	0.122	0.123	0.129	0.121	0.138	0.118
	20 min	0.123	0.124	0.119	0.124	0.139	0.118
	1 min	**0.142**	**0.144**	**0.141**	**0.142**	**0.159**	**0.141**
Pre-averaging	5 min	0.165	0.165	0.170	0.162	0.179	0.163
	20 min	0.178	0.177	0.176	0.177	0.192	0.176

Table 2 Standard deviations of relative errors of variance estimations calculated from 1000 simulations

3.2 Maximum Likelihood Estimator

Maximum likelihood estimator was proposed in [2]. It utilizes the fact that asset returns $r_t = p_t - p_{t-1}$ contamined by white noise follow MA(1) process. The daily variance as well as microstructure noise variance can then be estimated using maximum likelihood. As we can see in Table 1 when the noise is white this estimator gives the best results. However, when the white noise assumption is violated estimations are biased.

3.3 Realized Kernel Estimator

Realized kernel estimator proposed in [3] captures serial correlations induced by microstructure noise with a kernel. We use modified Tukey-Hanning kernel as recommended in [3]. This estimator deals well with all our noise settings as shown in Table 1.

3.4 Pre-averaging Estimator

Pre-averaging estimator was proposed in [7]. It is based on the idea to remove noise by locally averaging returns before computing realized variance. As we can see in Table 1 this estimator performs overall well even though it has a tendency to sligtly underestimate the daily variance. Compared to realized kernel it has higher standard deviations of relative errors as shown in Table 2.

4 Results

Our simulations show the impact of various microstructure noise models on realized variance and other integrated variation estimators. In the basic case of white noise, realized variance diverges lineary to infinity with increasing number of observations while other estimators are robust to this noise. Similar results apply to rounding noise. When the noise has AR(1) or MA(1) structure the divergence of realized variance is nonlinear. There is also a possibility of estimators missing some periodic parts of noise, which can lead to very different estimates while using relatively similar frequencies. The most problematic setting is dependent noise. All estimators have tendency to overestimate variance in the presence of noise correlated with efficient price.

Acknowledgements

The work on this paper was supported by the grant IGS F4/63/2016 of University of Economics, Prague.

References

[1] Ait-Sahalia, Y., and Jacod, J.: *High-Frequency Financial Econometrics*. Princeton University Press, 2014. ISBN 978-1-4008-5032-7.

[2] Ait-Sahalia, Y., Mykland, P. A., and Zhang, L.: How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise, *The Review of Financial Studies* **18** (2005), 351–416. ISSN 0893-9454.

[3] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N.: Designing Realized Kernels to Measure the ex post Variation of Equity Prices in the Presence of Noise, *Econometrica* **76** (2008), 1481–1536. ISSN 1468-0262.

[4] Hansen, P. R., and Lunde, A.: Realized Variance and Market Microstructure Noise, *Journal of Business & Economic Statistics* **24** (2006), 127–161. ISSN 0735-0015.

[5] Hautsch, N.: *Econometrics of Financial High-Frequency Data*. Springer Science & Business Media, 2011. ISBN 978-3-642-21925-2.

[6] Hautsch, N., and Podolskij, M.: Preaveraging-Based Estimation of Quadratic Variation in the Presence of Noise and Jumps: Theory, Implementation, and Empirical Evidence, *Journal of Business & Economic Statistics* **31** (2013), 165–183. ISSN 0735-0015.

[7] Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M.: Microstructure Noise in the Continuous Case: The Pre-Averaging Approach, *Stochastic Processes and their Applications* **119** (2009), 2249–2276. ISSN 0304-4149.

Recursive core of an OLG economy with production and comprehensive agreements

Milan Horniaček[1], Ľubica Šimková[2]

Abstract. We analyze an infinite horizon OLG economy with production and discounting of future utilities. Each agent lives for two periods. He is endowed with labor services in the first year of life. All goods are non-durable. Agents' utility functions have only consumption and provision of labor services as arguments. Agreements between members of non-singleton coalitions are comprehensive – they specify both labor and produced inputs contributed by members of a coalition, outputs, distribution of produced consumption goods between members of a coalition, and distribution of produced producer goods (that can be used in production only in the following period) between young members of a coalition. (Therefore, taking into account also discounting of future payoffs, the argument of Hendricks, Judd, and Kovenock showing emptiness of the core of OLG pure exchange economy does not apply here.) We prove (using one of Browder's fixed point theorems for a correspondence) that a strong recursive core is non-empty in our model. The latter is a subset of the strong core with the property that each stream of actions in it, for each period (with inputs of producer goods produced in the preceding period), induces a substream belonging to the strong core of the coalitional game starting at this date.

Keywords: OLG economy, recursive core, fixed point.

JEL Classification: C71, C73, D51
AMS Classification: 91A12, 91A13, 91A25, 91B50

1 Introduction

This paper has two goals. First, we develop a model of an OLG economy in which the strong core is nonempty, i.e., in which the argument in [3] showing the emptiness of the core cannot be used. We do it (keeping the usual simplifying assumptions that each agent lives for two periods) by adding production, discounting of future utilities, assuming that agents are endowed only by labor services (not by consumption goods) and they own producer goods (these are distributed between young agents and used in production in the following period when they are old), and comprehensive agreements on division of common outputs within non-singleton coalitions. The latter specify both labor and produced inputs contributed by members of a coalition, outputs, distribution of produced consumption goods between members of a coalition, and distribution of produced producer goods between young agents. (Producer goods can be used in production only in the following period when today's young agents will be old.) Thus, young agents cannot deprive old agents of consumption goods without giving up the use of producer goods owned by the old agents. Old agents, if they want to consume, have to make producer goods they own available to common production activities with young agents who contribute labor services. Moreover, discounting of future utilities implies that it is not possible to carry out forever the scheme in which utility of old agents in some period is increased at the expense of young agents and the latter are compensated in the next period at the expense of those who are young at it. (Compensations would grow in time without any limit.)

Second, we want to show that in the model fulfilling the first goal the strong recursive core is nonempty. The recursive core was introduced into the literature in [1] in the framework of capital accumulation model with infinitely lived agents. Its strong version used in this paper requires that each stream of actions in it, for each period (with inputs of producer goods produced in the preceding period) , induces a substream belonging to the strong core of the coalitional game starting at this date. That is, in the case of recursive strong core, we take into account also coalitions that some members joined only when they were old. Even such a coalition should not be able to increase (without cooperation with agents outside it) utility of some of its members (for members who joined it when they were old, utility in the second period of their lifetime) without decreasing utility of some

[1] Comenius University in Bratislava, Faculty of Social and Economic Sciences, Institute of Economics, Mlynské luhy 4, 82105 Bratislava, Slovakia, milan.horniacek@fses.uniba.sk
[2] Comenius University in Bratislava, Faculty of Social and Economic Sciences, Institute of Economics, Mlynské luhy 4, 82105 Bratislava, Slovakia, lubica.simkova@fses.uniba.sk

other member. Of course, a strong recursive core is a subset of the strong core. Therefore, by proving that the strong recursive core is non-empty in our model we fulfill both our goals.

Our results are important because they open the way to analysis of non-cooperative games in an OLG economy in which equilibria correspond to recursive core. Such games can shed light on formation of firms in an OLG economy with imperfect competition.

In the remainder of the paper, N denotes the set of natural numbers and $\Re$ is the set of real numbers. For $n \in N$, $\Re_+^n = [0,\infty)^n$. For a finite set X, $\#(X)$ is its cardinality. For $r \in \Re$, $[r]$ is the integer part of r. We endow each finite dimensional real vector space with the Euclidean topology and each infinite dimensional Cartesian product of finite dimensional real vector spaces with the product topology. $\Re^\infty$ endowed with product topology is metrizable. We denote metric on it by d. For a set X, $con(X)$ denotes its convex hull, 2^X is the set of all subsets of X, including the empty set, and (if X is a subset of a topological space) $cl(X)$ denotes the closure of X.

2 Model

The set of agents is N. Each agent lives for two periods. In the first period of his life he is young, in the second year he is old. The time horizon of the model is N. There is even $m \in N$ such that in each period exactly m agents live. The set of young agents living in period $t \in N$ is $I_t^- = \{0.5tm + i\}_{i=1}^{0.5m}$, the set of old agents in period 1 is $I_1^+ = \{1,\dots 0.5m\}$, and the set of old agents in period $t \in N \setminus \{1\}$ is $I_t^+ = I_{t-1}^-$. Function $\tau : I \to N$ assigns to each agent the first period of his life within the model. That is, $\tau(i) = 1$ for each $i \in \{1,\dots 0.5m\}$ and $\tau(i) = max\{n \in \{0\} \cup N | 0.5nm < i - 0.5m\} + 1$ for each $i > 0.5m$.

The finite set of goods in the model is $Q = L \cup V \cup Z$, where L (V, Z) is the set of labor services (consumption goods, producer goods). In order to deliver the message of the paper in the simplest possible way, we assume that all goods are non-durable, i.e. they can be used only in one period. Consumption goods can be used only in the period in which they are produced. Producer goods produced in period $t \in \{0\} \cup N$ can be used only in period $t+1$. For each $t \in N$ each agent $i \in I_t^-$ has endowment $\omega_i^- = (\omega_{i\ell})_{\ell \in L} \in \Re_+^{\#(L)}$ of labor services. We assume that

$$x_\ell^{sup} = sup_{t \in N}\, max_{i \in I_t^-}\, \omega_{i\ell} < \infty, \forall \ell \in L. \tag{1}$$

Old agents cannot provide labor services. They own producer goods produced in the preceding period. (Outputs of producer goods by each coalition in each period of its existence are distributed among young agents.) For each $t \in N$ and each $i \in I_t^+$, $\omega_i^+ = (\omega_{ik})_{k \in Z} \in \Re_+^{\#(Z)}$ is the vector of quantities of producer goods owned by agent i. For $i \in \{1,\dots 0.5m\}$ ω_i^+ is given (it originated before the beginning of the time horizon of the model), for $i > 0.5m$ it is the result of distribution of outputs of production activities in which i participated (by providing labor services) when he was young.

Taking into account (1) and the fact that nothing can be produced without using some labor service, it is reasonable to assume that for each $k \in V \cup Z$ there is $y_k^{max} > 0$ such that output of good k in the economy in any period cannot exceed y_k^{max}.

Production possibilities of any coalition in any period are described by the same production possibility set

$$Y = \left((x_k)_{k \in L \times Z}, (y_k)_{k \in V \times Z}\right) \subset \prod_{\ell \in L}\left[-x_\ell^{sup}, 0\right] \times \prod_{k \in Z}\left[-y_k^{max}, 0\right] \times \prod_{k \in V \cup Z}\left[0, y_k^{max}\right].$$

Elements of Y are vectors of gross inputs and outputs. Inputs have non-positive sign and outputs have non-negative sign. This approach to the production possibility set takes into account the fact that producer goods used as inputs should be obtained before the production starts.

Assumption 1. *(a)* $0 \in Y$, *(b)* Y *is closed, and* (c) *for each* $k \in L \times Z$ *there exists* $d_k > 0$ *such that, if* $x_\ell \geq -d_\ell$ *for each* $\ell \in L$ *or* $x_k \geq -d_k$ *for each* $k \in Z$, *then* $y_k = 0$ *for each* $k \in V \cup Z$.

Part (c) is a strengthened form of assumption that it is not possible to produce something from nothing. It says that it is not possible to produce something without using some minimal quantity of at least one labor service and some minimal quantity of at least one producer good. Note that we do not assume any form of returns to scale of production. If there are strictly decreasing returns to scale, it is efficient to produce in many small production units. (One coalition of agents can in one period run several production units.) Nevertheless, part (c) of Assumption 1 implies that these units cannot be infinitely small and their number is bounded from above by

$$min\left\{\Sigma_{\ell \in L}\left[\frac{0.5mx_\ell^{sup}}{d_\ell}\right], \Sigma_{k \in Z}\left[\frac{y_k^{max}}{d_k}\right]\right\}.$$

Each agent's utility from labor supplies and consumption over his lifetime is the sum of his utility when he is young and discounted utility when he is old. All agents use the same discount factor $\delta \in (0,1)$. (Our qualitative results would not change if we assumed different discount factors but with the supremum of them over all agents less than one.) Agent $i \in N$ has single period utility function $u_i : \prod_{\ell \in L}[0,\omega_{i\ell}] \times \prod_{k \in V}[0,y_k^{max}] \to \Re$. If $i > 0.5m$, his vector of supplies of labor services in the first period of his life is $\xi_i = (\xi_{i\ell})_{\ell \in L} \in \prod_{\ell \in L}[0,\omega_{i\ell}]$, his consumption vector when he is young is $\psi_i = (\psi_{ik})_{k \in V} \in \prod_{k \in V}[0,y_k^{max}]$ and his consumption vector when he is old is $\theta_i = (\theta_{ik})_{k \in V} \in \prod_{k \in V}[0,y_k^{max}]$, then his lifetime utility is $u_i(\xi_i,\psi_i) + \delta u_i(0,\theta_i)$. (For $i \in \{1,\ldots,0.5m\}$, his utility within the time horizon of the model is $u_i(0,\theta_i)$.)

Assumption 2. *For each* $i \in N$, u_i *is (a) continuous, (b) non-increasing in provision of each labor service* $\ell \in L$ *with* $\omega_{i\ell} > 0$, *(c) non-decreasing in consumption of each consumption goods,(d)* $u_i(0) = 0$, *and (e) there exists* $\overline{\rho} > 0$ *such that* $u_i\left(0,\left(y_k^{max}\right)_{k \in V}\right) \leq \overline{\rho}$ *for each* $i \in N$.

For each $i \in N$, vector a_i summarizes his participation in production activities and distribution of their results. We have $a_i = (\xi_i,\psi_i,\theta_i,\omega_i^+)$ for each $i > 0.5m$ and $a_i = (\theta_i,\omega_i^+)$ for each $i \in \{1,\ldots,0.5m\}$. For each $i \in N$ we set $a_i^+ = (\theta_i,\omega_i^+)$.

A distribution stream is a sequence $a = (a_i)_{i \in N}$. We denote by A the set of all feasible distribution streams. They are feasible with respect to endowments of young agents by labor services, ownership of producer goods by agents in $\{1,\ldots,0.5m\}$, and production possibility set Y. (In computation of components of elements of A in period t we use results of computations for previous periods and all feasible sets of production units in period t.) Taking into account (1), upper bounds on outputs, and part (b) of Assumption 1, A is a compact set.

A coalition is a non-empty subset of N. We denote the set of all coalitions by $\Im$. A distribution stream of coalition $C \in \Im$ starting in period $min_{i \in C}\,\tau(i) \leq t \leq max_{i \in C}\,\tau(i) + 1$ is a sequence

$$a_{Ct} = \left((a_i^+)_{i \in C:\tau(i)=t-1}, (a_i)_{i \in C:\tau(i) \geq t}\right).$$

For each $C \in \Im$ let $T(C) = \left\{t \in N \setminus \{1\} : C \cap I_t^+ \neq \emptyset, min_{i \in C \setminus I_t^+}\,\tau(i) \geq t\right\}$. We denote by $A_{Ct}\left((\omega_i^+)_{i \in C \cap I_t^+}\right)$ the set of feasible distribution streams of coalition $C \in \Im$ starting at period $t \in T(C)$. They are feasible with respect to endowments of young agents in C by labor services, ownership of producer goods by agents in C who are old in period t given by $(\omega_i^+)_{i \in C \cap I_t^+}$ and production possibility set Y when, starting from period t, agents in C do not cooperate with agents outside C. Taking into account (1), upper bounds on outputs, and part (b) of Assumption 1, $A_{Ct}\left((\omega_i^+)_{i \in C \cap I_t^+}\right)$ is a compact set. If $C \in \Im$ contains each of its members for his whole lifetime (which lasts only one period for each $i \in C \cap I_1^+$), we write only A_C. For each $a \in A$, each $t \in N \setminus \{1\}$, and each $i \in I_t^+$, we denote by $\omega_i^+(a)$ the ownership of producer goods by agent i generated by the previous activities recorded in a.

We denote by Γ the non-transferable utility coalitional game generated by the model described above, by $Core(\Gamma)$ its strong core, and by $Rcore(\Gamma)$ its strong recursive core.

Definition 1. *A distribution stream* $a \in A$ *belongs to* $Core(\Gamma)$ *if and only if there do not exist a coalition* $C \in \Im$ *and* $a^{(C)} \in A_C$ *such that* $u_i\left(\xi_i^{(C)}, \psi_i^{(C)}\right) + \delta u_i\left(0, \theta_i^{(C)}\right) \geq u_i(\xi_i, \psi_i) + \delta u_i(0, \theta_i)$ *for each* $i \in C \setminus I_1^+$ *and* $u_i\left(0, \theta_i^{(C)}\right) \geq u_i(0, \theta_i)$ *for each* $i \in C \cap I_1^+$, *with strict inequality for at least one* $i \in C$.

Definition 2. *A distribution stream* $a \in Core(\Gamma)$ *belongs to* $Rcore(\Gamma)$ *if and only if there do not exist* $C \in \Im$, $t \in T(C)$, *and* $a^{(C)} \in A_{Ct}\left(\left(\omega_i^+(a)\right)_{i \in C \cap I_t^+}\right)$ *such that* $u_i\left(\xi_i^{(C)}, \psi_i^{(C)}\right) + \delta u_i\left(0, \theta_i^{(C)}\right) \geq u_i(\xi_i, \psi_i) + \delta u_i(0, \theta_i)$ *for each* $i \in C \setminus I_t^+$ *and* $u_i\left(0, \theta_i^{(C)}\right) \geq u_i(0, \theta_i)$ *for each* $i \in C \cap I_t^+$, *with strict inequality for at least one* $i \in C$.

3 Non-emptiness of the recursive core

Proposition 1. *Game* Γ *has non-empty recursive core.*

Proof. Let

$$\alpha^* = \max\left\{\sum_{i \in I_1^+} u_i(0, \theta_i) + \sum_{i \in N \setminus I_1^+} \delta^{\tau(i)-1}\left(u_i(\xi_i, \psi_i) + \delta u_i(0, \theta_i)\right) \middle| a \in A\right\}. \tag{2}$$

Clearly,

$$0 \leq \alpha^* < 0.5 m\bar{\rho}\left(1 + \frac{1+\delta}{1-\delta}\right) = \frac{m\bar{\rho}}{1-\delta}. \tag{3}$$

Define

$$B = \left\{ \begin{array}{l} b = \left((b_i)_{i \in I_1^+}, \left((b_i^-, b_i^+)\right)_{i \in N \setminus I_1^+} \right) \in \prod_{i \in I_1^+} \left[0, \frac{m\bar{\rho}}{1-\delta}\right] \times \prod_{i \in N \setminus I_1^+} \left[0, \frac{m\bar{\rho}}{1-\delta}\right]^2 \\ \left| \sum_{i \in I_1^+} b_i + \sum_{i \in N \setminus I_1^+} \delta^{\tau(i)-1}\left(b_i^- + \delta b_i^+\right) \leq \alpha^* \right. \end{array} \right\}. \tag{4}$$

B is a non-empty, compact, and convex subset of $\Re^\infty$. For each $C \in \Im$ define function $\beta_C^{(0)} : con(A) \times B \to [0, \alpha^*]$ by

$$\beta_C^{(0)}(a,b) = \max\left\{ \begin{array}{l} \max\left\{ \begin{array}{l} \sum_{i \in C \cap I_1^+} \left(u_i\left(0, \theta_i^{(C)}\right) - b_i\right) + \sum_{i \in C \setminus I_1^+} \delta^{\tau(i)-1}\left[u_i\left(\xi_i^{(C)}, \psi_i^{(C)}\right) - b_i^- + \delta\left(u_i\left(0, \theta_i^{(C)}\right) - b_i^+\right)\right] \\ a^{(C)} \in A_C \end{array} \right\}, \\ 0 \end{array} \right\}. \tag{5}$$

For each $C \in \Im$ function $\beta_C^{(0)}$ is continuous and there exists $\bar{b} \in B$ such that $\beta_C^{(0)}(a, \bar{b}) = 0$ for each $a \in con(A)$. If $C \cap I_1^+ = \varnothing$, then (using part (c) of Assumption 1) coalition C, without cooperation with agents outside it, cannot produce anything and it cannot enable its members to achieve positive utilities. Thus, it is enough to take any $\bar{b} \in B$ with $\bar{b}_i^+ > 0$ for some $i \in C$. If there is $i \in C \cap I_1^+$, then it is enough to set $\bar{b}_i = \alpha^*$ and all other components of $\bar{b}$ equal to zero.

Define correspondence $W^{(0)} : con(A) \times B \to 2^{[0, \alpha^*]}$ by $W^{(0)}(a,b) = cl\left(\bigcup_{C \in \Im}\left\{\beta_C^{(0)}(a,b)\right\}\right)$. Clearly, $W^{(0)}$ is a nonempty-valued and compact-valued correspondence. From continuity of functions β_C, $C \in \Im$, and the use of discounting in (5) it follows that $W^{(0)}$ is a continuous correspondence.

For each $C \in \Im$ and each $t \in T(C)$ define function $\beta_C^{(t)} : con(A) \times B \to \left[0, \frac{m\overline{\rho}}{1-\delta}\right]$ by

$$\beta_C^{(t)}(a,b) = \max\left\{\begin{array}{l} \max\left\{\begin{array}{l}\sum_{i \in C \cap I_t^+}\left(u_i\left(0,\theta_i^{(C)}\right) - b_i^+\right) + \sum_{i \in C \setminus I_t^+} \delta^{\tau(i)-t}\left[u_i\left(\xi_i^{(C)},\psi_i^{(C)}\right) - b_i^- + \delta\left(u_i\left(0,\theta_i^{(C)}\right) - b_i^+\right)\right] \\ a^{(C)} \in A_{Ct}\left(\left(\omega_i^+(a)\right)_{i \in C \cap I_t^+}\right)\end{array}\right\}, \\ 0 \end{array}\right\}. \tag{6}$$

Function $\beta_C^{(t)}$ is continuous and there exists $\left(\overline{a},\overline{b}\right) \in A \times B$ such that $\beta_C^{(t)}\left(\overline{a},\overline{b}\right) = 0$. It is enough to take any $\overline{a} \in A$, $a^{(C)}$ solving the maximization program in (6) for $\overline{a}$, and $\overline{b} \in B$ with components corresponding to members of C in and after period t consistent with sum of discounted utilities generated by $a^{(C)}$.

Define correspondence $W^{(1)} : con(A) \times B \to 2^{\left[0, m\overline{\rho}(1-\delta)^{-1}\right]}$ by $W^{(1)}(a,b) = cl\left(\bigcup_{C \in \Im}\left(\bigcup_{t \in T(C)}\left\{\beta_C^{(t)}(a,b)\right\}\right)\right)$. For the same reasons as in the case of $W^{(0)}$, $W^{(1)}$ is a nonempty-valued, compact-valued, and continuous correspondence.

We define function $\beta_A : con(A) \to \left[0, \max\left\{\min\left\{d\left(\overline{a},a\right) \middle| \overline{a} \in A\right\} \middle| a \in con(A)\right\}\right]$ by

$$\beta_A(a) = \min\left\{d\left(\overline{a},a\right) \middle| \overline{a} \in A\right\}. \tag{7}$$

Function β_A is continuous. Of course, $\beta_A(a) = 0$ for each $a \in A$.

We define function

$$\beta_b : con(A) \times B \to \left[0, \max\left\{d\left(\left(\left(u_i(0,\theta_i)\right)_{i \in I_1^+}, \left(u_i(\xi_i,\psi_i), u_i(0,\theta_i)\right)_{i \in N \setminus I_1^+}\right), b\right) \middle| (a,b) \in con(A) \times B\right\}\right]$$

by

$$\beta_b(a,b) = d\left(\left(\left(u_i(0,\theta_i)\right)_{i \in I_1^+}, \left(u_i(\xi_i,\psi_i), u_i(0,\theta_i)\right)_{i \in N \setminus I_1^+}\right), b\right). \tag{8}$$

Taking into account continuity of functions u_i, $i \in N$, function β_b is continuous. We have $\beta_b\left(a, \left(\left(u_i(0,\theta_i)\right)_{i \in I_1^+}, \left(u_i(\xi_i,\psi_i), u_i(0,\theta_i)\right)_{i \in N \setminus I_1^+}\right)\right) = 0$ for each $a \in A$.

For each $C \in \Im$ let $\tau(C) = \min_{i \in C} \tau(i)$. Finally, we define function $\beta : con(A) \times B \to [0,1]$ by

$$\beta(a,b) = \frac{\beta_A(a) + \beta_b(a,b) + \max W^{(0)}(a,b) + \max W^{(1)}(a,b)}{\max\left\{\beta_A(a') + \beta_b(a',b') \middle| (a',b') \in con(A) \times B\right\} + \alpha^* + m\overline{\rho}(1-\delta)^{-1}}. \tag{9}$$

Taking into account continuity of functions and correspondences used in (9), β is a continuous function.

Let $b^{(1)}$ and $b^{(2)}$ be the unit vector in $\Re^\infty$ with the first and second component, respectively, equal to one. Define correspondence $G : B \to 2^{\left\{b^{(1)},-b^{(1)}\right\}}$ by $G(b) = \left\{-b^{(1)}\right\}$ if $b_1 \in \left[0, \frac{0.25m\overline{\rho}}{(1-\delta)}\right)$, $G(b) = \left\{-b^{(1)}, b^{(1)}\right\}$ if

$b_1 \in \left[\frac{0.25m\bar{\rho}}{(1-\delta)}, \frac{0.75m\bar{\rho}}{(1-\delta)}\right]$, and $G(b) = \{b^{(1)}\}$ if $b_1 \in \left(\frac{0.75m\bar{\rho}}{(1-\delta)}, \frac{m\bar{\rho}}{(1-\delta)}\right]$. Note that G is upper hemicontinuous. Define correspondence $F : con(A) \times B \to 2^{con(A) \times \Re^\infty}$ by

$$F(a,b) = \{a\} \times \left(\begin{array}{l} \left\{ \left(1 - \beta(a,b) \left(1 - \frac{min\{max\{b_1, b_2\}, 0.82m\bar{\rho}(1-\delta)^{-1}\}}{0.82m\bar{\rho}(1-\delta)^{-1}} \right) \right) b \right\} \\ - \left\{ \beta(a,b) max\{0.2m\bar{\rho}(1-\delta)^{-1} - b_2, 0\} b^{(2)} \right\} \\ + \left\{ \beta(a,b) max\{b_2 - 0.8m\bar{\rho}(1-\delta)^{-1}, 0\} b^{(2)} \right\} \\ + \beta(a,b) \left(\begin{array}{l} max\{0.8m\bar{\rho}(1-\delta)^{-1} - b_2, 0\} \\ + max(0.2m\bar{\rho}(1-\delta)^{-1} - b_1, 0) \end{array} \right) conG(b) \end{array} \right) \tag{10}$$

It follows from definition of F that

$$(a,b) \in F(a,b) \Leftrightarrow \beta(a,b) = 0. \tag{11}$$

Using properties of correspondence G and continuity of function β, correspondence F is non-empty-valued, closed-valued, convex-valued, and upper hemicontinuous. Its domain is a nonempty, compact and convex subset of $\Re^\infty \times \Re^\infty$, which, endowed with the product topology, is metrizable, and, hence, locally convex topological vector space. Using (10), for each $(a,b) \in con(A) \times B$ there exist $b' \in B$, $\lambda < 0$, and $h \in \Re^\infty$ such that $(a,h) \in F(a,b)$ and

$$(a,h) - (a,b) = \lambda[(a,b') - (a,b)]. \tag{12}$$

Thus, correspondence F satisfies all assumptions of Theorem 5 in [2, p. 288]. Therefore, by the latter theorem, it has a fixed point (a^*, b^*). Using (11), we have $\beta(a^*, b^*) = 0$. From this, (5)-(9), and Definitions 1 and 2, it follows that $a^* \in A$, b^* is a stream of utilities generated by a^*, $a^* \in Core(\Gamma)$, and $a^* \in Rcore(\Gamma)$. *Q.E.D.*

Acknowledgements

The research reported in this paper was financially supported by grant No. APVV-14-0020 of the Grant Agency APVV.

References

[1] Becker, R.A., Chakrabarti, S.K.: The recursive core. *Econometrica* **63** (1995), 401–423.
[2] Browder, F.E.: The fixed point theory of multi-valued mappings in topological vector spaces. *Mathematische Annalen* **177** (1968), 283–301.
[3] Hendricks, K., Judd, K., Kovenock, D.: A note on the core of the overlapping generations model. *Economics Letters* **6** (1980), 95-97.

The Invertibility of the General Linear Process and the Structures in its Background

Richard Horský[1]

Abstract. The modern conception of time series analysis is based on the Box Jenkins methodology. In this the area the statistics provides important tools for time series analysis in empirical level. However it is certainly useful to deal with the mathematical background of basic notions. It turns out that the functional analysis brings the efficient access to study structures, spaces and operators that are extensively used in Box Jenkins methodology.
The general linear process is defined by the help of the linear filter. This filter is an operator defined by a power series in the variable *B* (it denotes the lag operator). It is important to explain the convergence of this series and to define corresponded spaces in which these objects exist. The main goal is to describe the invertibility of the general linear process in general level, not only, as is usual, for polynomial or geometrical lag structures. It is done by the help of so called holomorphic functional calculus applied on the lag operator. The structure in background of this is a certain Banach algebra and its representation is given as well.

Keywords: stationary process, general linear process, lag operator, space of all bounded stochastic sequences, Banach algebra of linear filters, invertibility of linear filter.

JEL Classification: C44
AMS Classification: 90C15

1 The stationarity and the invertibility of the stochastic process

The stationarity and the invertibility are main properties of a stochastic process. The aim of this article is to define the problem of the invertibility of the general linear process on the general level and to solve it. We also perform the algebraic and analytical structures which are in background of the analysis of stochastic processes. We consider the stochastic processes the domain of which is the set of all integers.

It is suitable to recall some basic notions, concepts and facts to understand the sense of the next considerations.

1.1 The stochastic sequence and its basic characteristics

Stochastic process is a mapping $\mathbf{X}: T \to L_2(\Omega, \pi)$, where *T* is a time domain and $L_2(\Omega, \pi)$ is the space of all functions (random variables) defined almost everywhere on the measurable space Ω, functions, that are square integrable over the space Ω with respect to a probability measure π. This space is Hilbert space with the norm derived from the scalar product $E(XY) = \int_\Omega XY d\pi$. The mean and variance of any random variable $X \in L_2(\Omega, \pi)$, are finite numbers, particularly the mean $\mu = EX = \int_\Omega X d\pi$ and the variance $\sigma^2 = DX = EX^2 - \mu^2$. The time domain will be the set of all integers here. It means that the values of a stochastic process are random variables $X_t, t = 0, \pm 1, \pm 2, \ldots$ We will write $\mathbf{X} = (X_t)$ to emphasize that the stochastic process is a sequence and talk about the *stochastic sequence*.

The basic characteristics of the stochastic process are:

- the function of means $\mu_t = EX_t$;
- the function of variances $\sigma_t^2 = DX_t = EX_t^2 - \mu_t^2$;
- the autocovariance function $C(t,s) = \operatorname{cov}(X_t, X_s)$.

[1] University/of Economics, department of mathematics, Ekonomická 957, Prague 4, rhorsky@vse.cz

1.2 Stationary stochastic sequence and white noise

The concept of the stationary process is well-known (see [1], p. 12). Briefly said a stationary process is such that the functions of means and variances respectively are constant and the values of its autocovariance function are dependent only on the distance (time lag) of the random variables within the process: $C(t,s)=C(t-s,0)=\gamma_{t-s}$. The stationarity of a process ensures its stochastic stability.

A typical example of the stationary process is the white noise. The *white noise* $\mathbf{W}=(\varepsilon_t)$ is a stochastic process with $E\varepsilon_t=0$, $D(\varepsilon_t)=E(\varepsilon_t^2)=\sigma_t^2$ and the autocovariance $\gamma_k=\operatorname{cov}(\varepsilon_t,\varepsilon_{t-k})=E(\varepsilon_t\varepsilon_{t-k})$ for any positive integer k. In other words the white noise is an orthogonal system in the Hilbert space $L_2(\Omega,\pi)$. It can be transformed to be an orthonormal system, not necessary complete. The white noise may be used as a base for the definition of other processes, especially for the definition of the general linear process

1.3 General linear process and its invertibility

The *general linear process* (GLP) is the process of the form

$$X_t=\mu+\sum_{k=0}^{\infty}\psi_k\varepsilon_{t-k}\ , \tag{1}$$

where μ is a given scalar (mean of the process, w.l.o.g. mean is zero) and (ψ_k) is a given sequence of scalars (weights of the process), $\psi_0=1$. The convergence of the series is taken in the sense of the convergence in the mean square which is the same as the convergence in the norm topology in $L_2(\Omega,\pi)$. The convergence of the series in (1) is equivalent to the stationarity of the GLP. The necessary and sufficient condition for the convergence of (1) in the mean square is $(\psi_k)\in\ell_2$ (the space of all square summable scalar sequences). Another condition

$$\sum_{k=0}^{\infty}|\psi_k|<\infty \tag{2}$$

is only sufficient for the convergence in (1) since the space ℓ_1 of all absolutely summable scalar sequences is a proper subspace of the space ℓ_2.

The stochastic sequence (1) is said to be *invertible* if there exists a sequence of scalars (π_k) such that

$$\varepsilon_t=\sum_{k=0}^{\infty}\pi_k X_{t-k}. \tag{3}$$

This definition is given in [1], p. 86 and we explain this notion by means of the functional analysis in the last section 3.

2 Lag operator and its algebraic and analytical properties

The basic operator in the theory of stochastic sequences is the lag operator. Its role is not only to simplify formal writings but also it is an important linear operator with its own algebra and interesting analytical properties. Therefore we should specify the spaces which stand for its domain so that we might analyze its properties, especially its spectrum. Let us recall the definition of the lag operator which is given e.g. in [2] and in rather different manner in [3], where is also given its matrix representation in suitable spaces of stochastic sequences.

The *lag operator* is the operator

$$B:V\to V,\ \ B(X_t)=(X_{t-1})\ , \tag{4}$$

where V is the space of all stochastic sequences. In the following subsections it will be particularly specified.

2.1 Algebra of the lag operator

The set V in the definition (4) is in fact a linear space (of infinite dimension). The addition and scalar multiplication are defined in usual way, it means pointwise. It is seen that the mapping (4) is a linear operator. The space of all linear operators $\mathcal{L}[V]$ on the space V is the algebra (linear space with added operation, called multiplication or equivalently superposition of operators). The algebra $\mathcal{L}[V]$ contains all nonnegative integer powers of (4). They are defined by the principle of induction as follows

$$B^0 = I,\ B^{k+1} = B \circ B^k, \tag{5}$$

k is a positive integer. The symbol I in (5) denotes the identity operator on the space V. Thus by (5) we get $B^k(X_t) = (X_{t-k})$. The set of all (finite) linear combinations of the operators (5) is a subalgebra of the algebra $\mathcal{L}[V]$. Its elements are the polynomials in variable B. This algebra is obviously commutative. We denote it by the symbol. $\wp(B)$. This topics is also extensively studied in [2], page 19. As to the ":infinite" linear combinations of the nonnegative powers of B, i.e. the power series in B, it is necessary to introduce a convergence structure on the algebra $\mathcal{L}[V]$ besides the case we consider these series only in formal sense.

2.2 The convergence structures

The space V introduced in (4) is too wide for implementation a reasonable convergence structure. It seems to be appropriate to restrict ourselves to the space of all bounded stochastic sequences $\mathcal{B}(T)$. On one hand this space contains all stationary processes, on the other hand it has a nice analytical structure. Suppose that $\mathbf{X} = (X_t) \in V$. The function

$$\|\mathbf{X}\| = \sup_{t \in T} \sqrt{EX_t^2}. \tag{6}$$

has a finite value for any stochastic sequence $\mathbf{X} \in \mathcal{B}(T)$. It can be easily proved that (6) is a norm on the space $\mathcal{B}(T)$ and even that $\mathcal{B}(T)$ is a complete (Banach) space with respect to this norm (see e.g. [5], p.107). To avoid formal difficulties we will denote the lag operator by the same symbol B whichever space in the following considerations is its domain.

The lag operator B is bounded (continuous) operator on the space $\mathcal{B}(T)$. We can immediately get from (4) and (6) that $\|B\mathbf{X}\| \le \|\mathbf{X}\|$ and the equality is right, if we set $\mathbf{X} = \mathbf{E} = (1)$. Hence the norm of the lag operator is obviously

$$\|B\| = \sup_{\|\mathbf{X}\|=1} \|B\mathbf{X}\| = 1. \tag{7}$$

The norm (7) is the operator norm in the space $\mathcal{L}[\mathcal{B}(T)]$ of all bounded operators on the space $\mathcal{B}(T)$. The space $\mathcal{L}[\mathcal{B}(T)]$ is the Banach space since $\mathcal{B}(T)$ is Banach. The space $\mathcal{L}[\mathcal{B}(T)]$ is in fact algebra, since the product of two bounded operators is a bounded operator, i.e. it is a subalgebra of the algebra $\mathcal{L}[V]$. It obviously contains all nonnegative powers of the lag operator B and their finite linear combinations, i.e. $\wp(B)$ is a subalgebra of $\mathcal{L}[\mathcal{B}(T)]$. The power series in B and their convergence are analyzed in the section 3.

We may accept another point of view on the stochastic process and regard this process as a one sided stochastic sequence

$$\mathbf{X} = (X_t, X_{t-1}, \ldots, X_{t-k}, \ldots) \tag{8}$$

for any fixed integer t. Then we consider the space $\ell_\infty(L_2)$ instead of the space $\mathcal{B}(T)$. It is the space of all one sided bounded stochastic sequences of the type (8). The symbol $\ell_\infty(L_2)$ is derived from the symbol ℓ_∞ which denotes the space of all bounded scalar (one sided) sequences. The symbol L_2 in brackets is to emphasize the terms of the sequences are not scalars but functions from $L_2(\Omega, \pi)$. The space $\ell_\infty(L_2)$ is the Banach space with respect to the norm

$$\|\mathbf{X}\|_\infty = \sup_{k \ge 0} \sqrt{EX_{t-k}^2}\,. \tag{9}$$

If we extend the stochastic sequences (8) by zeros at time points t+1, t+2,... we get the stochastic sequence from the original definition given above in the section 1. This extension leads to an element of $\mathcal{B}(T)$ and thus the space $\ell_\infty(L_2)$ may be understood as a subspace of $\mathcal{B}(T)$, not only linear but also (closed) normed linear subspace, where the norm (9) is the norm (6) restricted to $\ell_\infty(L_2)$.

All the facts given above about the lag operator B are right if we replace $\mathcal{B}(T)$ by $\ell_\infty(L_2)$.

In [3] there is given the spectral analysis of the lag operator B in the space $\ell_\infty(L_2)$. The conclusions of this analysis can be formulated in the following proposition:

Proposition 1. (spectral properties of the lag operator)
The lag operator B defined in (4) is an element of the Banach space $\mathcal{L}[\ell_\infty(L_2)]$ *(of all bounded operators on the space* $\ell_\infty(L_2)$*) the norm of which is* $\|B\| = 1$. *Its spectral properties are as folows:*

i. *The resolvent set is* $\rho(B) = \{\lambda \in \mathbf{C} : |\lambda| > 1\}$.

ii. *The spectrum is* $\sigma(B) = \{\lambda \in \mathbf{C} : |\lambda| \le 1\}$. *All its elements are eigenvalues of B which means that the operator* $\lambda I - B$ *is not injective for any* $\lambda \in \sigma(B)$.

iii. *The range of the operator* $\lambda I - B$ *depends on the eigenvalue* λ *as follows:*

   a) $\lambda I - B$ *is surjective, i.e.* $R(\lambda I - B) = \ell_\infty(L_2)$ *for any* $|\lambda| < 1$.

   b) *The closure of the range of* $\lambda I - B$ *is a proper subspace of the space* $\ell_\infty(L_2)$ *for any* $|\lambda| = 1$.

3 Banach algebra of linear filters and its invertible elements

The basic notion in the theory of linear stochastic sequences is defined by the formula (1). We will suppose the GLP as the centered sequence, i.e. with zero mean. We can rewrite the formula (1) using the lag operator (4) in the form

$$\mathbf{X} = \psi(B)\mathbf{W}, \tag{10}$$

where

$$\psi(B) = \sum_{k=0}^{\infty} \psi_k B^k \tag{11}$$

The symbol $\psi(B)$ represents a formal abbreviation for the (formal) power series $\psi(B) = \sum_{k=0}^{\infty} \psi_k B^k$. We would like to get the convergence in (11). Then $\psi(B)$ will be a real object, an element of the space of all bounded operators on the space of all bounded stochastic sequences. The operator (11) is called the *linear filter* in Box Jenkins methodology. It transforms the white noise $\mathbf{W} = (\varepsilon_t)$ onto the process which is called GLP, more generally any bounded stochastic sequence onto (generally) another bounded stochastic sequence. In the Box Jenkins methodology one of the crucial problems is the problem of the invertibility of the given stochastic sequence. We recalled its definition, see (3). In fact this problem is the problem of the invertibility of the linear filter (11). In what follows our main goal is to describe the structure of the space of linear filters and give its representation. We notice that it is a Banach algebra and we describe the regular (invertible) elements of this algebra.

3.1 Lag structure of the GLP and Banach algebra of the linear filters

The notion of the Banach algebra is given for instance in [4], p. 59. The algebra W is called the *Banach algebra* if W is a Banach space and the multiplication is connected with the norm by the relation $\|fg\| \le \|f\|\|g\|$ for any $f, g \in W$. We can introduce some examples of Banach algebras.

Example 1a. The fields of all real or complex numbers are obviously Banach algebras. The norm is absolute value and the operation of multiplication is current number multiplication.

Example 1b. The space $\mathcal{A}(\mathcal{K})$ of all functions continuous on a compact set $\mathcal{K}$, analytical in int$\mathcal{K}$ (the interior of $\mathcal{K}$). The special case is that $\mathcal{K}=\mathcal{K}_r=\{\lambda \in \mathbf{C} : |\lambda| \le r\},\ r > 0$. The multiplication of functions is pointwise. The norm is *sup* (or *max*) norm.

Example 1c. The space ℓ_1 of all absolutely summable scalar sequences with Cauchy product of infinite series as multiplication. The norm is the sum of the series of the absolute values of the terms of sequence.

Example 1d. The space of all bounded linear operators $\mathcal{L}[W]$ on a Banach space W with multiplication as superposition of operators. The special case is the space of matrices of the given order n with multiplication of matrices. The norm is the operator norm.

In the previous section we introduced the Banach algebra $\mathcal{L}[\mathcal{B}(T)]$. On the contrary the algebra $\mathcal{L}[V]$ mentioned in the subsection 2.1 is not Banach. The algebra $\mathcal{L}[\ell_\infty(L_2)]$ is Banach and its subalgebra $\wp(B)$ of all polynomials in the variable B is an open subspace of the Banach space $\ell_\infty(L_2)$, so it is not a Banach space. We take its closure (with respect to the operator norm) $\overline{\wp(B)}$ to get the Banach algebra. What are the elements of this algebra? They are the operators of the form (11), i.e. the linear filters which are the sums of the convergent power series in B (the convergence in the operator norm topology of the space $\mathcal{L}[\ell_\infty(L_2)]$.

It is obvious that $\|B^k\| = \|B\|^k = 1$ for any non-negative integer k. It implies that

$$\|\psi(B)\| \le \sum_{k=0}^{\infty} |\psi_k| \|B^k\| = \sum_{k=0}^{\infty} |\psi_k| . \tag{12}$$

The inequality (12) shows that the condition (2) is sufficient for the linear filter (11) to be a bounded operator. The following proposition claims that there is a one-to-one correspondence between the scalar sequence $(\psi_k) \in \ell_1$ and the operator $\psi(B) \in \overline{\wp(B)}$ which preserves the norms. Thus we obtain the representation of the Banach algebra of linear filters by the Banach algebra from Example 1c.

Proposition 2. (the representation of the Banach algebra of linear filters)

The Banach algebra $\overline{\wp(B)}$ of the convergent power series (11) is isometrically isomorphic with the Banach algebra of scalar sequences ℓ_1.

Proof: It suffices to prove that in (12) there is the equality. We find a stochastic sequence $\mathbf{Z} \in \ell_\infty(L_2),\ \|\mathbf{Z}\|_\infty = 1$ such that $\|\psi(B)\mathbf{Z}\|_\infty = \sum_{k=0}^{\infty} |\psi_k|$. We define $\mathbf{Z} = (Z_{t-k})$ in this way: if $\psi_k \ne 0$ then $Z_{t-k} = |\psi_k|^{-1} \psi_k$ else $Z_{t-k} = 1$.

The sequence of scalars (weight of the process) is denoted as the *lag structure.* In the end of this section we remind some important examples of lag structures that define important types of stochastic sequences.

Example 2a. Suppose there exists a non-negative integer q such that $\psi_k = 0$ for any $k>q$, $\psi_q \ne 0$. Then the linear filter is a polynomial operator of the degree q. In Box Jenkins methodology the sequence expressed by (10) with a polynomial filter of degree q is denoted by the symbol MA(q), i.e. *the moving average process of the degree q.*

Example 2b. If $\psi_k = \lambda^k$ then $(\psi_k) \in \ell_1$ if and only if $|\lambda|<1$. The corresponded linear filter $\psi(B) = \sum_{k=0}^{\infty} (\lambda B)^k$ may be expressed in the form $\psi(B) = (I - \lambda B)^{-1}$. It follows from the Theorem on Neumann´s series (see [4], p.61.) The given sequence (ψ_k) is called the *geometric lag structure.* In this case the stochastic sequence $\mathbf{X}$ in (10) can be regarded as the unique solution of the operator equation $(I - \lambda B)\mathbf{X} = \mathbf{W}$, $|\lambda|<1$. This process is in Box Jenkins methodology called *autoregressive process of order 1* and denoted AR(1).

Example 2c. Let us consider an operator equation $\Phi(B)\mathbf{X} = \mathbf{W}$, where $\Phi(B)$ is a polynomial operator of the degree *p.* We ask the question whether there exists a unique solution of this equation in the space of bounded stochastic sequences. If so it is, this solution is called the *autoregressive process of the order p.* In other words, is the operator $\Phi(B)$ invertible? We may decompose $\Phi(B) = \prod_{j=1}^{r} \left(I - \lambda_j B\right)^{k_j}$, $\sum_{\mathrm{j}=1}^{r} k_j = p$. It means that the operator $\Phi(B)^{-1}$ (if it exists) is the product of some linear filters determined by the product of some geometric lag structures.

3.2 Regular (invertible) elements in the Banach algebra of linear filters

The lag structure (π_k) in (3) defines the linear filter $\pi(B) = \sum_{k=0}^{\infty} \pi_k B^k$ from $\overline{\wp(B)}$ if and only if it is the inverse to the lag structure $(\psi_k) \in \ell_1$ in (1). It means that it satisfies the infinite set of conditions: $\pi_0 = 1$, $\pi_k + \psi_1 \pi_{k-1} + \ldots + \psi_k \pi_0 = 0$, *k* positive integer. In this case the operator $\pi(B)^{-1} = \psi(B)$, where $\psi(B)$ is from (11).

However, we may formulate another equivalent condition for linear filter (11) to be invertible.

Proposition 3. (invertibility of the linear filter)

The linear filter $\psi(B) \in \overline{\wp(B)}$ *is an invertible operator if and only if all the roots of the complex function* $\psi(z) = \sum_{k=0}^{\infty} \psi_k z^k$ *are outside the unit circle* $\mathcal{K}_1 = \{z \in \mathbf{C} : |z| \le 1\}$.

Proof: The claim of the proposition follows from the spectral properties of the lag operator *B* (see Prop. 1. above). and the holomorphic functional calculus (see [4], p. 85).. The holomorphic functional calculus of the lag operator *B* is in fact a homomorphism defined on the algebra $\mathcal{A}$ (*B*) of all holomorphic functions $f : G \to \mathbf{C}$, where *G* is an open set in the complex plane which contain the spectrum $\sigma(B) = \mathcal{K}_1$ with values in the algebra $\overline{\wp(B)}$. Its domain contains for instance all elements of the algebras $\mathcal{A}(\mathcal{K}_r)$, $1 < r \le \infty$, see Example 1b. It is known that invertible are those operators which are the values for holomorphic functions with no roots in the spectrum of the original operator, i.e. *B* in our case (see [5], p. 275).

Conclusion

We presented one application of the holomorphic functional calculus to obtain the general description of the invertibility of the general linear process. We introduced the mathematical structure background for this application. It is the Banach algebra of linear filters. We gave also the representation of this algebra.

References

[1] Brockwell, P.J., Davis R.A.: *Time Series: Theory and Methods,* Springer-Verlag, New York, Berlin, Heidelberg, 1987

[2] Dhrymes, J.P. *Distributed Lags. Problems of Estimation and Formulation,* North Holland, Amsterdam, 1985.

[3] Horský, R..: The lag operator and its spectral properties. In: *Proceedings Part I of Mathematical Methods in Economics 2013*, 285–290.

[4] Lukeš, J.: *Zápisky z funkcionální analýzy,* Univerzita Karlova v Praze, Nakladatelství Karolinum, 2012

[5] Taylor, A.E.: *Úvod do funkcionální analýzy.* Academia, Praha, 1973.

DG Approach to Numerical Pricing of Local Volatility Basket Options

Jiří Hozman[1], Tomáš Tichý[2]

Abstract. The problem of determining the fair price of an option is a delicate issue arising from the assumptions made under a market model and the evaluation of such option prices is often relied only on approximations obtained by numerical schemes. It is well known that commonly used Black-Scholes type models can not capture all real market features. Therefore, one way to make these models suitable for real world pricing issues is to relax some of the model assumptions. In this paper we present one of the basket option models that try to remain consistent with the volatility smile. These multi-factor local volatility models incorporate the volatility as a function of maturity and moneyness into the pricing procedure. The presented numerical approach arises from the concept of the discontinuous Galerkin method and enables better resolving of occurred special properties of solutions of such types of options. Finally, the resulting schemes are demonstrated on practical experiments with real data.

Keywords: Option pricing, discontinuous Galerkin method, multi-factor Black-Scholes model, basket options, implied volatility, local volatility, Dupire formula, numerical solution.

JEL classification: C44, G13
AMS classification: 35, 90C15

1 Introduction

Options constitute a very important and no less interesting financial product not only because they can meet even very specific needs of market participants, but also through their challenging pricing procedures. Obviously, under the very restrictive assumptions of Black and Scholes [2] the pricing is quite simple, since it leads to analytical formula. On the other hand, assuming exotic options, ie. options with not so simple payoff function, see [5] for extensive review of (exotic) options, including many pricing formulas, and taking into account the real world behaviour more seriously, one often needs to rely on numerical procedures.

In this paper we develop a numerical scheme for pricing of basket options using discontinuous Galerkin approach, for more details see [9]. For the moment, we basically follow the assumptions of Black and Scholes, except that we consider the local volatility derived from the volatilities implied by the market prices of vanilla options. The model is defined in Section 2, while in Section 3 we provide its discretization. Numerical experiment follows in Section 4.

2 Two-Factor Local Volatility Model

The model that we focus on is the local volatility model of a basket option payoff of which is dependent on the value of a weighted sum of two assets. This is a flexible model which is widely used in practice, namely for multiple assets portfolios. Without loss of generality, we present the case of a basket put option consisting of a weighted sum of two correlated assets with payoff given by

$$\max(K - \alpha_1 S_1(T) - \alpha_2 S_2(T), 0), S_1 > 0,\ S_2 > 0, \tag{1}$$

[1]Technical University of Liberec, Department of Mathematics and Didactics of Mathematics, Studentská 2, 461 17 Liberec, Czech Republic, jiri.hozman@tul.cz
[2]VŠB - Technical University of Ostrava, Department of Finance, Sokolská 33, 701 21 Ostrava, Czech Republic, tomas.tichy@vsb.cz

where $K > 0$ is the strike price, $S_i(T)$ denotes the value of i-th underlying asset at maturity T for $i = 1, 2$ and positive weights satisfy $\alpha_1 + \alpha_2 = 1$. Note that the below-mentioned approach can be straightforwardly generalized for the basket call options using the put-call parity.

In the local volatility model the evolution of each underlying asset is governed by the stochastic differential equation (SDE)

$$dS_i(t) = rS_i(t)dt + \sigma_i(S_i(t), t)S_i(t)dW_i(t),\ i = 1, 2, \quad \text{with} \quad \langle dW_1(t), dW_2(t)\rangle = \rho, \tag{2}$$

where $r > 0$ is the risk free interest rate (here considered as constant), function $\sigma_i(S_i(t), t)$ is called the local volatility of the corresponding asset and $(W_1(t), W_2(t))_{t\geq 0}$ is a correlated two-dimensional Brownian motion with the constant correlation coefficient $\rho \in (-1, 1)$.

A common approach based on multidimensional Ito's lemma, construction of a risk-free portfolio and elimination of the random components leads to the wholly deterministic partial differential equation for pricing an option on two assets (see e.g. [1])

$$\frac{\partial V}{\partial t} + \mathcal{L}_t(V) = 0, \quad S_1 > 0,\ S_2 > 0,\ t \in (0, T] \tag{3}$$

where $V(S_1(t), S_2(t), t)$ denotes the value of a basket option and the differential operator $\mathcal{L}_t$ is defined as

$$\mathcal{L}_t(V) = \frac{\sigma_1^2(S_1,t)S_1^2}{2}\frac{\partial^2 V}{\partial S_1^2} + \rho\sigma_1(S_1,t)\sigma_2(S_2,t)S_1S_2\frac{\partial^2 V}{\partial S_1 \partial S_2} + \frac{\sigma_2^2(S_2,t)S_2^2}{2}\frac{\partial^2 V}{\partial S_2^2} + rS_1\frac{\partial V}{\partial S_1} + rS_2\frac{\partial V}{\partial S_2} - rV,$$

This backward differential equation (3) is solved on a time interval with the prescribed terminal condition at maturity date T given by the payoff function (1). In order to discretize (3) also in spatial coordinates it is desirable to restrict the problem on a bounded domain. For this purpose let S_i^{max} denote the maximal sufficient value of i-th asset and consider the domain $\Omega := (0, S_1^{max}) \times (0, S_2^{max})$. The boundary $\partial\Omega$ of the domain Ω can be decomposed into three parts: $\Gamma_1 = (0, S_1^{max}) \times \{0\}$, $\Gamma_2 = \{0\} \times (0, S_2^{max})$, $\Gamma_3 = \partial\Omega \cap I\!R_+^2$, and each part should be equipped with additional boundary values on it. These values are chosen compatible with the payoff function and using knowledge on the asymptotic behavior of options. For put options considered here we use mixed boundary conditions in the following form

$$\frac{\partial V}{\partial S_2}(S_1, S_2, t)\Big|_{\Gamma_1} = \lim_{\varepsilon\to 0+}\frac{\partial V}{\partial S_2}(S_1, \varepsilon, t),\ \frac{\partial V}{\partial S_1}(S_1, S_2, t)\Big|_{\Gamma_2} = \lim_{\varepsilon\to 0+}\frac{\partial V}{\partial S_1}(\varepsilon, S_2, t),\ V(S_1, S_2, t)\Big|_{\Gamma_3} = 0, \tag{4}$$

where the simple concept of extrapolated Neumann boundary conditions is used on both coordinate axes.

2.1 Degenerate Parabolic PDE

At first, it should be suitable to introduce the change of temporal variable t in order to transform studied problem (3) into the initial-boundary value one. Setting $\hat{t} = T - t$ the time to maturity and spatial substitutions $x = (x_1, x_2) = (S_1, S_2)$ lead to the degenerate-parabolic partial differential equation (written in divergence form) for the unknown price function $u(x, t) : \Omega \times (0, T) \to I\!R_0^+$ satisfying that

$$\frac{\partial u}{\partial \hat{t}} - \operatorname{div}\left(I\!D(x, \hat{t}\,) \cdot \nabla u\right) + \nabla \cdot \vec{f}(x, \hat{t}; u) + \kappa(x, \hat{t}\,)u = 0 \quad \text{in } \Omega \times (0, T), \tag{5}$$

where the degeneracy is captured by the common factor x_i appearing in the terms of the symmetric positive semi-definite matrix

$$I\!D(x, \hat{t}\,) = \begin{pmatrix} d_{11}(x, \hat{t}\,) & d_{12}(x, \hat{t}\,) \\ d_{12}(x, \hat{t}\,) & d_{22}(x, \hat{t}\,) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \sigma_1^2(x_1, \hat{t}\,)x_1^2 & \rho\sigma_1(x_1, \hat{t}\,)x_1\sigma_2(x_2, \hat{t}\,)x_2 \\ \rho\sigma_1(x_1, \hat{t}\,)x_1\sigma_2(x_2, \hat{t}\,)x_2 & \sigma_2^2(x_2, \hat{t}\,)x_2^2 \end{pmatrix}. \tag{6}$$

Then the vector-valued function can be expressed in terms of (6) as

$$\vec{f}(x, \hat{t}; u) = \left(\frac{\partial d_{11}}{\partial x_1}(x, \hat{t}\,) + \frac{\partial d_{12}}{\partial x_2}(x, \hat{t}\,) - rx_1,\ \frac{\partial d_{12}}{\partial x_1}(x, \hat{t}\,) + \frac{\partial d_{22}}{\partial x_2}(x, \hat{t}\,) - rx_2\right) \cdot u(x, \hat{t}\,)$$

This function represents a convection flux in (5) and indicates the direction of a propagation of information, see Figure 3 (left) with the example of the corresponding vector filed for specific real data. The last term on the left-hand side of (5) is the reaction term with factor κ defined as the scalar function

$$\kappa(x, \hat{t}\,) = 3r - \frac{\partial^2 d_{11}}{\partial x_1^2}(x, \hat{t}\,) - 2\frac{\partial^2 d_{12}}{\partial x_1 \partial x_2}(x, \hat{t}\,) - \frac{\partial^2 d_{22}}{\partial x_2^2}(x, \hat{t}\,),$$

which plays role as a variable reaction coefficient arising from the rewriting of original equation (3) into the divergence form (5).

Additionally, to close the system, the initial and boundary conditions in terms of a new value function are rewritten as

$$u(x,0) = \max(K - \alpha_1 x_1 - \alpha_2 x_2, 0) \quad \text{and} \quad \left(I\!D \cdot \nabla u(x,\hat{t}\,) \right) \cdot \vec{n} = 0 \text{ on } \Gamma_1 \cup \Gamma_2, \quad u(x,\hat{t}\,) = 0 \text{ on } \Gamma_3, \tag{7}$$

where $\vec{n}$ is the outer unit normal to Γ_i. Let us comment that homogeneous Neumann boundary condition from (7) prescribed on axes $x_1 = 0$ and $x_2 = 0$ is a priori fulfilled in the variational form and it corresponds to the so-called do-nothing boundary condition.

Finally, note that the pricing equation (3) is closely related to the convection-diffusion equation, which exhibits parabolic and hyperbolic behavior in dependency on a proportion of the convection and diffusion parts. Therefore, the numerical schemes for solving of such equation should be constructed with respect to these properties, see Section 3.

2.2 Local Volatility Treatment

Since the local volatility functions σ_i in (2) are typically unknown, it is necessary to determine their values in the way consistent with the observed market smile. We follow the approach from [3] with a few modifications and construct local volatilities using the three-step procedure. At first we calculate the implied volatilities θ_i for each underlying asset (and some fixed market observables) as the unique values satisfying the classical Black-Scholes pricing formula (for puts)

$$P_i(S_i, K, T, r, \theta_i) = Ke^{-rT}\Phi(-d_i + \theta_i\sqrt{T}) - S_i\Phi(-d_i), \; i = 1, 2, \tag{8}$$

where $d_i = \frac{\ln(S_i/K) + (r + \theta_i^2/2)T}{\theta_i\sqrt{T}}$, $i = 1, 2$, and Φ denotes the standard normal cumulative distribution function.

The next step is concerned with a suitable approximation of implied volatilities obtained from (8). We use one of the simplest ways and seek the volatility surfaces $\theta_i(K, T)$ as a globally continuous composition of piecewise quadratic functions in the space-time domain, which are constructed by the least squares method.

Finally, local volatilities $\sigma_i(S_i, t)$ can be expressed in the terms of implied volatilities according to the following Dupire formula

$$\sigma_i^2(K,T) = \frac{\theta_i^2 + 2T\theta_i \frac{\partial\theta_i}{\partial T} + 2rKT\theta_i \frac{\partial\theta_i}{\partial K}}{\left(1 + Kd_iT\frac{\partial\theta_i}{\partial K}\right)^2 + K^2T\theta_i\left(\frac{\partial^2\theta_i}{\partial K^2} - d_iT\left(\frac{\partial\theta_i}{\partial K}\right)^2\right)}, \; i = 1, 2, \tag{9}$$

Although the relation (9) requires a sufficiently smooth functions θ_i, the discontinuous approach allows us to use also piecewise smooth ones. The cuts of local volatility surfaces obtained after this calibration are plotted in Figure 1, together with the preceding two steps.

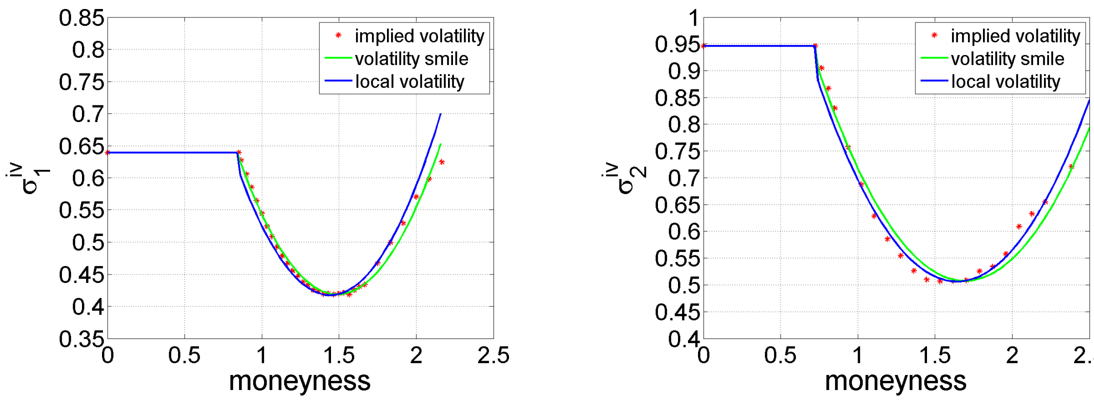


Figure 1 The cuts of the resulting volatility surfaces and local volatility functions together with the initial implied volatility observations (as red dots) for Allianz stock (left) and Deutsche Bank stock (right)

3 DG Numerical Scheme

We recall the DG framework from [8] with some modifications for the local volatility approach, cf. [4]. The approximate solution is sought in the finite dimensional space S_h^p consisting from piecewise polynomial, generally discontinuous, functions of the p-th order defined on the domain Ω. A simple example of a discontinuous piecewise linear function is shown on Figure 2 (left).

Similarly as in [7], we introduce the semi-discrete solution $u_h = u_h(\hat{t})$ represented by the system of the ordinary differential equations

$$\frac{d}{d\hat{t}}(u_h, v_h) + \mathcal{A}_h(u_h, v_h) = 0 \quad \forall v_h \in S_h^p, \forall \hat{t} \in (0,T) \tag{10}$$

where $u_h(0)$ is the initial condition, $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega)$ and the form $\mathcal{A}_h(\cdot,\cdot)$ stands for the DG semi-discrete formulation of the operator $\mathcal{L}_t$.

Consequently, we realize the discretization in time by an implicit Euler scheme (as in [1]) for the equidistant time partition $0 = \hat{t}_0 < \hat{t}_1 < \cdots < \hat{t}_s = T$ with the time step τ. We define the DG approximate solution of problem (5) as functions $u_h^m \approx u_h(\hat{t}_m)$, $\hat{t} \in [0,T]$, $m = 0,\ldots,r-1$, satisfying the following numerical scheme

$$\left(u_h^{m+1}, v_h\right) + \tau\mathcal{A}_h\left(u_h^{m+1}, v_h\right) = \left(u_h^m, v_h\right) \quad \forall v_h \in S_h^p,\ m = 0,1,\ldots,s-1, \tag{11}$$

with the starting data u_h^0 given by (7).

Further, if we rewrite the discrete DG solution as a linear combination of basis functions, i.e.

$$u_h^m(x) = \sum_{k=1}^{DOF} \beta_k^m \cdot \varphi_k(x),\ x \in \Omega, \qquad \text{where } S_h^p = \mathcal{L}\left(\varphi_1,\ldots,\varphi_{DOF}\right)$$

and set the vector of real coefficients $U_m = \left\{\beta_k^m\right\}_{k=1}^{DOF} \in I\!R^{DOF}$, then we can interpret the scheme (11) at each time level as the sparse matrix equation

$$(\mathbf{M} + \tau\mathbf{A})\,U_{m+1} = \mathbf{M}U_m, \tag{12}$$

where the matrix $\mathbf{M}$ is related to the mass matrix and the matrix $\mathbf{A}$ to the bilinear form $\mathcal{A}_h$, respectively. The existence and uniqueness of the solution of the discrete problem (11) is guaranteed by the invertibility of the corresponding system matrix in (12), i.e. the DG solution u_h^m is uniquely determined by the solution vector U_m of coefficients of basis functions. For illustration, the sparsity pattern of the system matrix is depicted in Figure 2 (right).

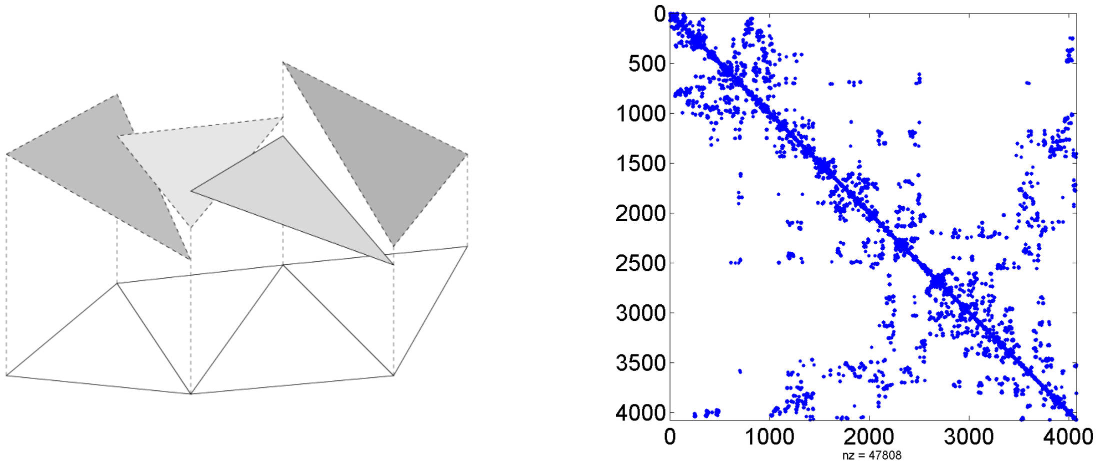


Figure 2 An example of the discontinuous piecewise linear approximation (left) and visualization of the sparsity pattern of the system matrix for the initial mesh, $\#\mathcal{T}_h = 1\,226$ and P_1 approximations (right)

4 Numerical Experiments on Real Datasets

In this section we intend to apply the DG method to the simplified market problem based on real datasets. We consider the basket put option consisting of a weighted sum of the Allianz ($\alpha_1 = 0.6$) and Deutsche Bank ($\alpha_1 = 0.6$) stocks with the correlation factor $\rho = 0.88$, the strike at 40 Euro and the expiration date in 94 calendar days, cf. [8]. The market data was observed on September 13, 2011, when the closing prices of both stocks were $S_1^{ref} = 59.79$ Euro and $S_2^{ref} = 23.40$ Euro, respectively. The value of risk-free interest rate is assumed to be constant and determined by the implied value $r = 0.01557$ p.a. and with respect to a given maturity, no dividend is expected ($q = 0$). On the other hand, the volatilities are variable and the sample dataset of implied volatilities for both underlying are given in Table 1 and Table 2. The corresponding local volatilities are constructed according to the Dupire formula (9) as functions of the moneyness and maturity, see also Figure 1.

K	θ_1^{impl}	K	θ_1^{impl}	K	θ_1^{impl}	K	θ_1^{impl}	K	θ_1^{impl}	K	θ_1^{impl}
0.01	0.6392	60	0.5445	72	0.4552	84	0.4194	96	0.4254	120	0.5703
51	0.6392	62	0.5246	74	0.4470	86	0.4198	98	0.4306	125	0.5980
52	0.6275	64	0.5086	76	0.4379	88	0.4184	100	0.4338	130	0.6245
54	0.6054	66	0.4921	78	0.4331	90	0.4198	105	0.4671		
56	0.5853	68	0.4786	80	0.4254	92	0.4212	110	0.4992		
58	0.5646	70	0.4667	82	0.4234	94	0.4184	115	0.5295		

Table 1 The implied volatilities observed for put options on the Allianz stock of September 13, 2011, over the different strikes for maturity $T = 94/365$

K	θ_2^{impl}	K	θ_2^{impl}	K	θ_2^{impl}	K	θ_2^{impl}	K	θ_2^{impl}	K	θ_2^{impl}
0.01	0.9461	20	0.8301	28	0.5856	36	0.5067	44	0.5337	52	0.6548
17	0.9461	22	0.7576	30	0.5547	38	0.5078	46	0.5581	56	0.7205
18	0.9052	24	0.6878	32	0.5267	40	0.5089	48	0.6090	60	0.7568
19	0.8670	26	0.6289	34	0.5100	42	0.5256	50	0.6328		

Table 2 The implied volatilities observed for put options on the Deutsche Bank stock of September 13, 2011, over the different strikes for maturity $T = 94/365$

The local volatility treatment is preprocessed in MATLAB and the rest of numerical experiments is carried out in the solver Freefem++ (see [6]) for the piecewise linear, quadratic and cubic DG approximations on adaptively refined domain $(0, 220) \times (0, 220)$ with the constant time step of 1 day. The Table 3 records the option values at the reference node $[S_1^{ref}, S_2^{ref}]$ for different polynomial orders in comparison with our previous approach, see [8]. One can easily observe that obtained results, to a certain extent, are comparable. Finally, Figure 3 (middle, right) captures the final mesh and the corresponding isolines of the piecewise linear solution in the zoomed domain.

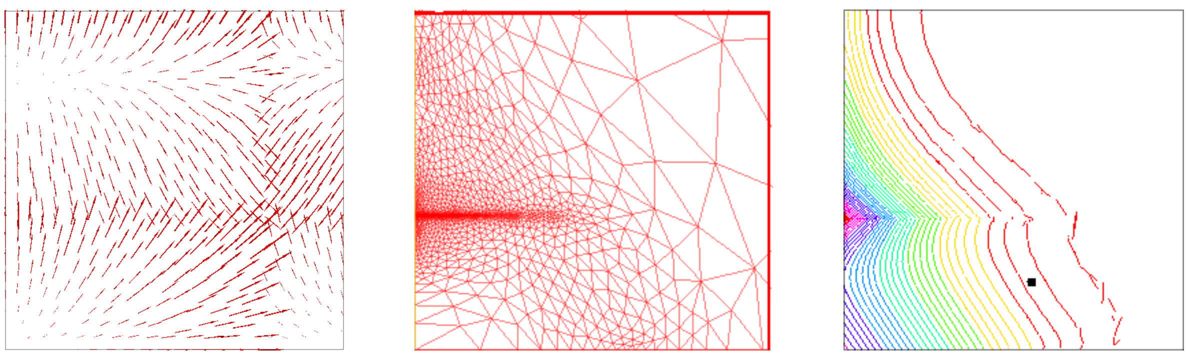


Figure 3 Zoomed plot of the vector field generated by physical fluxes (left), adaptively refined triangulation (middle) and detailed piecewise linear solution with the highlighted reference point (right); all is depicted on $[0, 110] \times [0, 110]$

polynomial order	P_1	P_2	P_3
option value (local vol., present method)	3.02215	3.01350	3.01078
option value (piecewise constant vol., [8])	2.99820	2.99906	2.99987

Table 3 The comparison of the computed option values at reference node for different polynomial degrees with respect to the volatility treatment

5 Conclusion

In this paper we have presented numerical scheme for pricing of 2-asset basket option with DG approach, which assumes local volatilities derived from market prices of relevant plain vanilla options and as a such should represent better estimation for the no-arbitrage prices. The presented scheme might be easily extended for basket options with more variables (stock prices).

Acknowledgements

Supported by the grant No. 16-09541S of the Czech Science Foundation and furthermore through SP2016/11, an SGS research project of VSB-TU Ostrava.

References

[1] Achdou, Y, and Pironneau, O.: *Computational Methods for Option Pricing.* Society for Industrial and Applied Mathematics, Philadelphia, 2005.

[2] Black, F., Scholes, M.: The pricing of options and corporate liabilities, *Journal of Political Economy* **81** (1973), 637–659.

[3] Dupire, B.: Pricing with a Smile, *Risk* **7**(1) (1994), 18–20.

[4] Du Toit, J., and Ehrlich, I.: Local Volatility FX Basket Option on CPU and GPU. *Numerical Algorithms Group, Technical report TR1/13*, High performance computing on graphics processing units, 2013.

[5] Haug, E.G.: *The Complete Guide to Option Pricing Formulas.* McGraw-Hill, New York, 2006.

[6] Hecht, F.: New development in FreeFem++, *Journal of Numerical Mathematics* **20** (2012), no. 3-4, 251–265.

[7] Hozman, J.: Analysis of the discontinuous Galerkin method applied to the European option pricing problem, *AIP Conference Proceedings* **1570** (2013), 227–234.

[8] Hozman, J., and Tichý, T.: A discontinuous Galerkin method for pricing of two-asset options. In: *Proceedings of the 33rd International Conference Mathematical Methods in Economics* (D. Martinčík, J. Ircingová, and P. Janeček, eds.), MME 2015, Plzeň, 2015, 273–278.

[9] Riviére, B.: *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation.* Society for Industrial and Applied Mathematics, Philadelphia, 2008.

Model based control of production flows in the serial logistic process

Bronislav Chramcov[1], Robert Bucki[2], Michał Trzopek[2]

Abstract. The paper highlights the problem of manufacturing flows in the system consisting of dispersed machines. Each machine is designed to carry out one operation with the use of the dedicated tool. Machines can be placed at different locations in the manufacturing area however, the manipulating arm delivers all semi-products. Moreover, the semi-products are moved to the required machine as well as to the storing zone with use of the same arm. Minimizing the total order realization time is the basic criterion as it brings the total manufacturing costs down. The analysis of the whole problem is supported by means of the simulator created exclusively for the purpose of this work. Implementing various initial data leads to obtaining certain results which are subsequently analyzed thoroughly in order to evaluate the simulation case study. Conclusions form the basis for the derivation of the significant assumptions about further development of information control of logistic systems.

Keywords: manufacturing system, heuristic algorithms, modelling, simulation

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

The Flexible Manufacturing System (FVS) is considered to be one of the most effective methods to minimize and eliminate management problems in industrial production systems. The aim of FVS is to enable the production (assembly) of several types of products in a single production system while maintaining minimum production time. To achieve this it is essential to address various alternatives of production at the design stage of the production system. The proposal for efficient production in logistic systems is one of the important tasks for managers of manufacturing companies. When designing a production system we need to think about several important issues which are interrelated [4]. The basic prerequisite for success in today's highly competitive industrial environment is therefore an efficient and flexible production, and especially the selection of appropriate management methodology [5]. The fundamental problem lies in the fact that any system downtime caused by blocking the production machine failure or maintenance may generate additional costs. These are primarily due to inactivity of at least one machine (workstation) and the impossibility to perform respective working operations in other machines arranged in series. Problems of this type are addressed in many works which primarily seek to increase the efficiency of such systems. Analysis and modeling of manufacturing systems (lines) were dealt with in many scientific works. Papadopoulos and Heavy did extensive search of the problem in the publication [8]. The problem of scheduling manufacturing of several types of products which is required for assembling various components or various manufacturing operations are among the most complex computational tasks. To optimize these production systems a variety of complex sophisticated methods can be used [7]. These methods also include the heuristic and meta-heuristic methods. These alternative methods are unable to determine the exact solution completely and in addition, this solution may sometimes not be found in a short time. However, these methods provide sufficient and fairly accurate solution in most cases. The methods, called heuristics, were originally based primarily on expert knowledge and experience in order to search for the space of possible solutions in a particularly comfortable way [9], [6]. Newer forms of these approaches are called meta-heuristics and were first introduced in the mid-80s. This type of search algorithms can solve complex optimization problems while using a set of multiple general heuristics [1]. These sophisticated methods are characterized by high demands on hardware and computational time. For this reason, simpler heuristics which can be used to effectively control the operation of a complex production system in real time were searched for [2]. The paper focuses on modeling the logistic system where each order can be made from a dedicated charge. The charge is either passed through a determined route in the manufacturing system equipped with machines performing specific operations on the semi-product or it is made by machines which are determined by the heuristic algorithm. It is also assumed that there is the buffer store zone serving storing needs. Moreover, the semi-products are moved to the required ma-

[1] Tomas Bata University in Zlín, Faculty of Applied Informatics, Nad Stráněmi 4511, 760 05 Zlín, The Czech Republic, chramcov@fai.utb.cz.
[2] Institute of Management and Information Technology, ul. Legionów 81, 43-300 Bielsko-Biała, Poland, rbucki@wsi.net.pl

chine as well as to the storing zone with the use of the same arm. In accordance with the contemporary lean approach to the problem of satisfying the needs of the customer certain algorithms are subject to validation to meet the criteria emerging from widely implemented logistic methods. Minimizing the total order manufacturing time remains the basic criterion as it brings the total manufacturing costs down. Orders are made either along the determined manufacturing route in case of defining a sequence of manufacturing operations or, in case of not defining such a sequence, according to the stochastically determined sequence of operations. No matter what sequence of operations is defined the proposed logistic formation is treated as a serial manufacturing system. The problem is represented by pre-programming time scaling manufacturing operations for the manipulating arm. The article presents the results of a specific simulation study. The simulator developed especially for this purpose was implemented [10].

2 Description of the manufacturing system

It is assumed that the manufacturing system consists of the charge zone, manufacturing hall, storing place for tools, storing place for semi-products and dispatching zone from which products are passed to customers. Let us assume that customers set orders in accordance with the vector of orders in the form $Z^k = \left[z^k_{m,n}\right]$ where $z^k_{m,n}$ is the number of units of the n-th production order set by the m-th customer at the k-th stage. The stage k, k=1,…,K is the moment of making the production decision. In the assumed real system machines are placed at different locations in the manufacturing hall in accordance with the manufacturing arrangement. Each machine performs only one type of manufacturing operation. In this case the specific tool is used. Charges are assumed to be supplied on the demand basis with no delay. Each type of product is associated with a certain type of charge. It is further assumed that there is a robotic arm which is responsible for the following operations: taking charge material from the entrance gate to the determined machine, transporting semi-products from a preceding machine to the determined subsequent one, transporting the ready product to the dispatching gate, transporting semi-products to the storing zone and back, manipulating with semi-products in the storing zone and replacing a used tool with a new one. It is assumed that the manufacturing system is equipped with a buffer zone where semi-products are stored for the determined period of time if required in case of the lack of manufacturing capacities for the certain semi-product. The buffer zone is located within the reach of the robotic arm.

2.1 Mathematical model of the manufacturing system

Let us assume that, from the information point of view, the manufacturing system is brought to its simplified form in which machines are to be associated with a certain defined point. This leads to creating the structure consisting of I rows and J columns. From this point of view it is possible to define the structure matrix of the system in the form $E = \left[e_{i,j}\right]$, where the element $e_{i,j}$=1 if the machine is placed in the i-th row of the j-th column, otherwise $e_{i,j}$=0. The base life matrix of tools used in the manufacturing system for a new brand set of tools used to manufacture units of the order is given in the form $G^{base} = \left[g^{base}_{i,j}\right]$, where $g^{base}_{i,j}$ is the base number of units which can be manufactured in the machine placed in the i-th row of the j-th column before the tool in this machine is completely worn out and requires immediate replacement. Let us introduce the matrix of the failure coefficient of order $z_{m,n}$ in the form $F = \left[f_{m,n}\right]$, determining how many units of the order $z_{m,n}$ can be made with the use of the each machine. Consequently, it is possible to define the life matrix for making the order $z_{m,n}$ in the form: $G(m,n) = \left[g(m,n)_{i,j}\right] = f_{m,n} \cdot G^{base}$, where the life matrix element $g(m,n)_{i.j}$ is the number of the order $z_{m,n}$ conventional units which can be made by means of the machine placed in the i-th row of the j-th column. If the order $z_{m,n}$ is not made by the machine placed in the i-th row of the j-th column, then $g(m,n)_{i.j} = -1$. Manufacturing operations on order units are carried out in various numbers of machines in different sequences. Let us introduce the route vector in the form: $D(m,n) = \left[d(m,n)_j\right]$. This vector specifies the manufacturing sequence for the n-th order set by the m-th customer (the sequence of manufacturing operations). The element $d(m,n)_j$ shows the machine (the number of the row) in the j-th column, which carries out the operation on the order $z_{m,n}$. If no machine from the j-th column is used to make the order $z_{m,n}$, then $d(m,n)_j = -1$. Similarly it is possible to define the vector of production times in the form: $T^{pr}(m,n) = \left[\tau^{pr}(m,n)_j\right]$, where $\tau^{pr}(m,n)_j$ is the manufacturing time of one conventional unit of the order $z_{m,n}$ with the use of specific machine in the j-th column.

2.2 Transportation and replacement times

It is possible to introduce matrices of transportation times for all types of product. The matrix in the form $T^{TR}(i,j)=\left[\tau^{TR}(i,j)_{i',j'}\right]$ specifies the transportation times between the machines in the manufacturing system where $\tau^{TR}(i,j)_{i',j'}$ is the transportation time (for any order) from the machine placed in the *i*-th row of the *j*-th column to the machine placed in the row *i'* of the column *j'*. At the same time: $\tau^{TR}(i,j)_{i',j'}>0$ if there is a transport operation taken into account, $\tau^{TR}(i,j)_{i',j'}=0$ otherwise. However, the above time is divided into the sum of two separate times as it is assumed that the semi-product is moved from the machine placed in the *i*-th row of the *j*-th column to the so-called central reference point and then from the central reference point to the machine placed in the in the row *i'* of the column *j'*. In accordance with the presented assumption the following assumptions can be taken for granted. Let us consider the matrix of transport times from machines to the central reference point in the form $T^{RefIN}=\left[\tau^{RefIN}{}_{i,j}\right]$, where $\tau^{RefIN}{}_{i,j}$ is the transport time of the product from the machine placed in the *i*-th row of the *j*-th column to the central reference point. Similarly, it is possible to consider the matrix of transport times from the central reference point to the machine placed in the in the row *i'* of the column *j'* in the form $T^{RefOUT}=\left[\tau^{RefOUT}{}_{i',j'}\right]$, where $\tau^{RefOUT}{}_{i',j'}$ is the transport time of the product from the central reference point to the machine placed in the row *i'* of the column *j'*.

Similarly, it is possible to define the matrix of transportation times from the charge zone as well as the matrix of transportation times to the dispatching zone. These matrices take the form: $T^{IN}=\left[\tau^{IN}{}_{i,j}\right]$ or $T^{OUT}=\left[\tau^{OUT}{}_{i,j}\right]$. The variable $\tau^{IN}{}_{i,j}$ specifies the transportation time from the charge zone to the machine placed in the *i*-th row of the *j*-th column and the variable $\tau^{OUT}{}_{i,j}$ specifies the transportation time from the machine placed in the *i*-th row of the *j*-th column to the dispatching zone. Moreover, there is a need to introduce the matrix of transportation times for new tools which are stored in the storing place for tools. The matrix in the form $T^{ToolIN}=\left[\tau^{ToolIN}_{i,j}\right]$ is defined where $\tau^{ToolIN}_{i,j}$ is the transportation time of a new tool from the storing place to the machine placed in the *i*-th row of the *j*-th column. Additionally, there is a need to introduce the matrix of transportation times for worn out tools to the storing place for used tools. This matrix is presented in the form $T^{ToolOUT}=\left[\tau^{ToolOUT}_{i,j}\right]$, where $\tau^{ToolOUT}_{i,j}$ is the transportation time of a worn out tool from the machine placed in the *i*-th row of the *j*-th column to the storing place for worn out tools.

It is assumed that the manufacturing system is equipped with a buffer zone where semi-products are stored for the determined period of time if required in case of the lack of manufacturing capacities for the certain semi-product. The buffer zone is located within the reach of the robotic arm. The matrix of transport times of semi-products to the determined place in the buffer zone is introduced in the form $T^{BufIN}=\left[\tau^{BufIN}{}_{i,j}\right]$, where $\tau^{TR-Buf}{}_{i,j}$ is the transport time of the semi-product to the determined place in the buffer zone from the machine placed in the *i*-th row of the *j*-th column. Additionally, the matrix of transport times of semi-products to the determined machine from its storing place in the buffer zone is introduced in the form $T^{BufOUT}=\left[\tau^{BufOUT}{}_{i,j}\right]$, where $\tau^{TR-Buf}{}_{i,j}$ is the transport time of the semi-product to the determined machine placed in the *i*-th row of the *j*-th column from its storing place in the buffer zone.

2.3 State of the manufacturing system

Throughout the course of manufacturing process the state of the system changes. The state of the manufacturing system changes after every operation on the order unit. Therefore, there is a need to analyze the state of orders, the state of tools in machines, charge materials, storing zones, etc. The state of the order changes after each production decision. The order matrix is modified at each *k*-th state according to the form (1) where $x^k_{m,n}$ is the number of units of the *n*-th order set by *m*-th customer to be made at the *k*-th state.

$$z^k_{m,n}=\begin{cases} z^{k-1}_{m,n}-x^k_{m,n} & \text{if a certain number of units of the } n-\text{th order is made at the } k-\text{th state} \\ z^{k-1}_{m,n} & \text{otherwise} \end{cases} \tag{1}$$

It is possible to present the state of the tool in machines by means of the matrix of state of the manufacturing system in the form $S^k=\left[s^k_{i,j}\right]$, where $s^k_{i,j}$ represents the value of relative tool wear ($0\le s^k_{i,j}\le 1$) of the tool in the machine placed in the *i*-th row of the *j*-th column. This element takes its value according to (2) where

$x^k(m,n)_{i,j}$ is the number of units of the order $z_{m,n}$ made by the machine placed in the i-th row of the j-th column and $g_{n,a}$ is the base number of units of the n-th order which can be manufactured before the a-th tool (the tool in the specified machine) is completely worn out and requires an immediate replacement.

$$s^k = \begin{cases} s_{i,j}^{k-1} + \dfrac{x^k(m,n)_{i,j}}{g_{i,j}} & \text{if a certain number of units of the order } z_{m,n} \text{ is made} \\ & \text{in the } i-\text{th machine in the } j-\text{th column at the } k-\text{th state} \\ s_{i,j}^{k-1} & \text{otherwise} \end{cases} \tag{2}$$

If the state of the tool should be exceeded ($s_{i,j}^k \geq 1$), no unit of any order can be made in the specified machine and it triggers the need to carry out the replacement process of the tool. If the tool in the machine placed in the i-th row of the j-th column has to be replaced with a new one, the state of this tool takes the value zero after carrying out the replacement procedure.

In addition, it is useful to monitor how many conventional units of any order can be made with the use of the specific machine. Hence, the matrix of the available capacity of the manufacturing system in the form $P_{i,j}^k = \left[p_{i,j}^k\right]$ is introduced where $p_{i,j}^k$ is the number of conventional units of any order, which still can be made in the i-th machine of the j-th column. If there is a remaining available capacity in the machine in the i-th row of the j-th column but the subsequent unit cannot be made fully in this machine, then the replacement process is carried out automatically. The available capacity of the machine in the i-th row of the j-th column is determined in the form (3) where expression $\lfloor x \rfloor$ represents the floor function and $g_{i,j}$ is the base number of units which can be manufactured before the tool (the tool in the specified machine) is completely worn out and requires an immediate replacement.

$$p_{i,j}^k = \left\lfloor g_{i,j} \cdot \left(1 - s_{i,j}^k\right)\right\rfloor \tag{3}$$

3 Case Study with the Use of the Simulator

The simulator of the general manufacturing system was created on the basis of the assumptions described above. The discussed simulator was created with the use of C# environment and it is used for simulating the behaviour of the dedicated manufacturing system [10].

3.1 Definition of the specific production system

The case study assumes that the complex production system consists of 11 machines. Each machine performs only one type of operation with the use of the specific tool. The structure of system is described by means of the matrix E and the base life matrix of the tool is defined in the form (5). Three customers set orders to be made by the production system. The number of conventional units of the orders for each specific customer is specified in the matrix (6) and the failure coefficient for the specific order is defined by the matrix (7).

$$E = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \quad (4) \qquad G = \begin{bmatrix} 2 & 3 & 3 \\ 3 & 3 & 2 \\ 2 & 2 & -1 \\ 3 & 2 & 1 \end{bmatrix} \quad (5) \qquad Z^0 = \begin{bmatrix} 5 & 1 & 6 & 1 \\ 8 & 4 & 3 & 4 \\ 9 & 9 & 7 & 4 \end{bmatrix} \quad (6) \qquad F = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 7 & 4 & 2 & 7 \\ 2 & 7 & 8 & 3 \end{bmatrix} \quad (7)$$

We assume that the orders are made in accordance with the route vectors presented in (8). The manufacturing times of one conventional unit of the order $z_{m,n}$ with the use of the specific machine in the j-th column (according to the route matrix) are defined in the vectors presented in (9).

$$D = \begin{bmatrix} \{2,1,4\} & \{3,1,1\} & \{1,3,2\} & \{2,1,1\} \\ \{1,3,4\} & \{1,4,4\} & \{4,2,4\} & \{2,3,1\} \\ \{2,3,4\} & \{4,2,1\} & \{3,2,1\} & \{1,4,4\} \end{bmatrix} \quad (8) \qquad T^{pr} = \begin{bmatrix} \{1,1,2\} & \{3,2,2\} & \{2,2,1\} & \{2,3,2\} \\ \{1,1,3\} & \{1,1,2\} & \{2,3,1\} & \{2,2,2\} \\ \{2,2,2\} & \{2,1,3\} & \{2,3,3\} & \{3,3,3\} \end{bmatrix} \quad (9)$$

The matrix of transport times of semi-products from machines to the central reference point is shown in (10) and the matrix of transport times of semi-products from the central reference point to the specific machine is shown in (11). Similarly, the matrix of transportation times from the charge zone is defined in (12) as well as the matrix of transportation times to the dispatching zone is defined in (13).

$$T^{RefIN} = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 2 & 1 \\ 2 & 2 & -1 \\ 3 & 2 & 2 \end{bmatrix} \quad (10) \quad T^{RefOUT} = \begin{bmatrix} 2 & 2 & 2 \\ 3 & 3 & 3 \\ 3 & 3 & -1 \\ 3 & 1 & 2 \end{bmatrix} \quad (11) \quad T^{IN} = \begin{bmatrix} 2 & 2 & 1 \\ 3 & 1 & 3 \\ 3 & 2 & -1 \\ 3 & 3 & 1 \end{bmatrix} \quad (12) \quad T^{OUT} = \begin{bmatrix} 3 & 1 & 3 \\ 3 & 3 & 1 \\ 1 & 3 & -1 \\ 2 & 3 & 2 \end{bmatrix} \quad (13)$$

Finally, the matrix of transportation times for new tools is presented in (14), the matrix of transportation times for worn out tools is presented in (15), the matrix of transport times of semi-products to the determined place in the buffer zone is presented in (16) and the matrix of transport times of semi-products to the determined machine from its storing place in the buffer zone is presented in (17).

$$T^{ToolIN} = \begin{bmatrix} 2 & 3 & 3 \\ 1 & 2 & 1 \\ 2 & 2 & -1 \\ 3 & 1 & 1 \end{bmatrix} \quad (14) \quad T^{ToolOUT} = \begin{bmatrix} 3 & 3 & 3 \\ 2 & 1 & 1 \\ 3 & 1 & -1 \\ 1 & 3 & 3 \end{bmatrix} \quad (15) \quad T^{BufIN} = \begin{bmatrix} 3 & 1 & 3 \\ 3 & 3 & 2 \\ 1 & 1 & -1 \\ 2 & 3 & 1 \end{bmatrix} \quad (16) \quad T^{BufOUT} = \begin{bmatrix} 3 & 1 & 2 \\ 3 & 1 & 3 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \quad (17)$$

3.2 Results of the simulation

The simulation was run for initial values of state of tools. In order to control the choice of the order and subsequently directing it to the manufacturing zone there is a need to implement one of available heuristic algorithms from a wide range of them [3]. The heuristic control algorithms of the maximal order [H(max)] and the minimal order [H(min)] are considered for demonstration purposes. These algorithms choose the order characterised by either the maximal or minimal number of units to be manufactured. Consequently, a random choice of orders as the simulation method was used 100 times. Moreover, there are two strategies implemented: making the following order begins when such a possibility comes into being or making the following order begins when the previous one has completed. Some manufacturing criteria are used for evaluation of the implemented control algorithms. In our case, the total manufacturing time (T_{TOTAL}^{pr}), the total replacement time of tools (T_{TOTAL}^{TR}), the lost capacity of tools (P^{LOST}) or the total value of the unused capacity of tools (P_{FINAL}^{unused}) are taken into account.

The results of the simulation experiments are shown in the following tables. The best results for each manufacturing criteria are highlighted. The first analysis of results is based on Table 1 where two heuristic were implemented to illustrate possible differences in results.

Control Algorithm	Strategy	T_{TOTAL}^{pr}	T_{TOTAL}^{TR}	P^{LOST}	P_{FINAL}^{unused}
H(max)	(1)	224	127	**1,46**	**12,67**
	(2)	282	115	**1,46**	**12,67**
H(min)	(1)	**212**	129	2,05	11,07
	(2)	271	**100**	2,05	11,07
Random (24)	(1)	212	**101**	2,89	12,24
Random (45)	(1)	212	111	2,95	10,18
Random (53)	(1)	**204**	**101**	0,73	13,40
Random (97)	(1)	211	123	**0,58**	**13,54**

Table 1 Results of the simulation for heuristic control algorithms

The best result in terms of minimizing the total manufacturing time is achieved by implementing the heuristic control algorithm of the minimal order and strategy (1) however, it does not minimize the replacement time of tools which is minimized by the use of the same heuristic control algorithm and the second strategy. Nevertheless, it brings one of the worst results in terms of minimizing the total manufacturing time. Should the lost capacity of tools be considered, there is a need to exclude the algorithm of the minimal order as it delivers the worst results, even if two strategies are used. Then it is worth considering all other criteria (e.g. minimizing the total manufacturing time). Additionally, the algorithm of the maximal order leaves the system with the biggest final flow capacity for further use.

The next step required choosing orders at random which was meant to find a satisfactory solution at the set number of simulation experiments. The random choice experiments were carried out 100 times. The best results are shown in the second part of Table 1. The initial data remain the same as in the previous experiment. As it can be seen, the best result in terms of minimizing the total manufacturing time are achieved in the 53rd experiment with the use of the strategy (1). The best time of replacing tools is also deliver by the same experiments and, additionally, in the 24th experiment which does not minimize the total manufacturing time. The minimal lost capacity was obtained in the 97th simulation experiment. Moreover, the same experiment leaves the system with the best final capacity of the system for further use. However, this experiment minimizes neither the total manufacturing time nor the replacement time of tools. It is worth noticing that strategy (2) seems not to be effective enough in the course of 100 simulations. These results shown in the graphic form give us a larger insight into the discussed problem. It is obvious that carrying out more simulation experiments (e.g. 10000) is likely to deliver better results.

4 Conclusion

The paper highlights the problem of searching for the satisfactory solution in the complex logistic system. The system consists of dispersed machines which are served by the manipulating arm. The system is described by mathematical modeling accompanied by all required equations. The main criterion is to minimize the total manufacturing time. Consequently, the simulator created on the basis of the assumptions enabled the case study to be carried out. Two heuristics are tested however, only the random choice of orders delivers the best solution in the course of the set number of simulation experiments. Moreover, it should be emphasized that other elements of the manufacturing process emerged for further analysis e.g. the total replacement time of tools, the lost capacity and the remaining capacity. If the number of simulation experiments increases, then the probability of finding a better solution increases. The approach presented hereby requires further development by carrying out a large number of simulation experiments.

Acknowledgements

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme project No. LO1303 (MSMT-7778/2014) and also by the European Regional Development Fund under the project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089

References

[1] Abraham, A., Grosan, C. and Pedrycz, W.: Engineering Evolutionary Intelligent Systems. *Studies in Computational Intelligence* **82** (2008). Springer Verlag.

[2] Bucki, R., Suchánek, P., and Vymětal, D.: Information Control of Allocation Tasks in the Synthetic Manufacturing Environment. *International Journal of Mathematics and Computers in Simulation* **6** (3, 2012), 324-332.

[3] Chramcov, B. and Bucki, R.: Decision Making Support of Logistics Tasks in the Manufacturing System. In *Proceedings of 32nd International Conference Mathematical Methods in Economics*, Olomouc, 2014, pp. 106–111.

[4] Heavey, C. and Browne, J.: A model management systems approach to manufacturing systems design. *International journal of flexible manufacturing systems* **8** (2, 1996), pp. 103-130.

[5] Katsanos, E. and Bitos, A.: Methods of Industrial Production Management: A Critical Review. In *Proceedings of the 1st International Conference on Manufacturing Engineering, Quality and Production Systems*, 1, pp.94-99, Brasov, 2009. WSEAS Press.

[6] Lee, K. Y. and El-Sharkawi, M. A.: *Modern heuristic optimization techniques: theory and applications to power systems*. John Wiley & Sons, 2008.

[7] Modrák, V.: Using the Simulation Technique for Planning the Industrial Production in Metallurgy. *Metalurgia International* **18** (2, 2013), pp. 17-20.

[8] Papadopoulos, H. and Heavey, C.: Queueing theory in manufacturing systems analysis and design: A classification of models for production and transfer lines. *European journal of operational Research* **92** (1, 1996), pp. 1-27.

[9] Russell, S. J., Norvig, P. and Davis, E.: *Artificial intelligence: A modern approach*. Prentice Hall, 2010.

[10] Trzopek, M.: *The Computer Simulator of the Manufacturing System with the Serial Structure.*, Institute of Management and Information Technology, Bielsko-Biała, 2015, p.66 (unpublished)

The Comparison of DEA and PROMETHEE method to evaluate the environmental efficiency of national steel sectors in Europe

Lucie Chytilová, František Zapletal[1]

Abstract. The aim of this paper is the environmental efficiency of national sectors of European steel industry. In particular, emissions trading with CO_2 is involved in analysis. The core factor of the environmental policy of the EU is the EU ETS system, whose idea is trading of emissions allowances for CO_2. Steel companies obtain some amount of emissions allowances for free which authorizes them to manufacture products (together with emissions). That is why that number of allowances is considered as the input in our model. On the other hand, the amount of CO_2 emissions released is regarded as the output. In addition, the structure of companies size and GDP per capita are other inputs and amount of production is the second output. Two methods are used to evaluate the efficiency of units and compared. Apart of the traditional DEA approach, the PROMETHEE analysis is also used. The both approaches are applied to data set of 22 national steel sectors in Europe for year 2014. The results of both models are identical, except for a few specified exceptions.

Keywords: DEA, PROMETHEE, efficiency, allowances, EU ETS, emissions trading

JEL classification: C61
AMS classification: 90C15

1 Introduction

In 2005, the EU ETS (Emissions Trading Scheme of the EU) has been established with the aim of reduction CO_2 emissions in Europe. This system force industrial companies to cover their CO_2 emissions by emissions allowances, thus it can increase costs of companies. One of the impacts of emissions trading duty is the pressure put on the companies and their environmental efficiency in production. The environmental efficiency is related with amount of emissions released per each unit of production. This paper is devoted to measure the environmental efficiency of whole EU countries for iron and steel sectors. In particular, two methods are used to evaluate the efficiency and the results are further compared.

The efficiency is the main criterion for a comprehensive assessment of activities of each industry sector and individual economic operators. It is considered to be one of the sources of wealth and it may show how well the economy allocates scarce resources to meet the needs and demands of consumers. There exist many methods which may be used for measuring the efficiency in field of steel industry. In this paper, two basic method are used. First, the PROMETHEE method is used. Some application of this method are in the research by Behzadiana et al. [3]. The second used method is the Data Envelopment Analysis (DEA). Many researchers have used various types of DEA, for example, Lenort et al. [8] applied DEA model to measure the efficiency of sectors manufacturing base metal in 25 European countries. Similarly, Morfeld and Silveira [11] used DEA to investigate the energy efficiency of European iron and steel production from 2000 to 2010.

The rest of the paper has the following structure: Section 2 provides information about the quantitative methods which are used. Input and output variables are defined in Section 3. Section 4 focuses on application and discussion the results of environmental efficiency analysis. Section 5 gives some conclusions and remarks.

[1]VSB - Technical University of Ostrava, Faculty of Economics, Sokolská 33, 702 00 Ostrava, Czech Republic, lucie.chytilova@vsb.cz, frantisek.zapletal@vsb.cz

2 Introduction to the quantitative methods

The first method used in this contribution is PROMETHEE which is, generally, more familiar as the method of multi-attribute decision-making. But, here, it is used to generate the input/output efficiency frontier depicting the efficient (non-dominated) units. The core idea of the method is a pairwise comparison using a preference function, see [1].

The Data Envelopment Analysis (DEA) has been used in efficiency analysis of many applications - banking industry, insurance industry, health care and so on. The three most known and widely used DEA models are the CCR model by Charnes et al. [4], the BCC model by Banker et al. [2] and the Additive model by Charnes et al. [5].These models were formulated for desirable inputs and outputs. However, in real applications, there are frequently needed undesirable inputs and/or outputs.

In scientific literature, there are known some approaches and models which deal with undesirable inputs and/or outputs. These models are usually a transformation of the basic DEA model. One of the most known and acceptable transformation is based on the ADD approach by Koopmans [6]. The transformation is given by $f(Y) = -Y + \beta$ and it was used for example in work of Seiford and Zhu [12]. However, the weakness of this transformation is that classification may depend on β as well as the final ranking can do so. Another transformation is the multiplicative inverse $f(Y) = 1/Y$ (used by Lovell et al., [10]). There are also approaches that can avoid data transformation. The most known type of this approach suggests that undesirable inputs are regarded as desirable outputs or vise versa for an initial attempt to formulate the model [9]. The disadvantage of this method is that it is good just if the research is done for the operational efficiency. In this work all three above mentioned approaches are used.

2.1 The methods of the PROMETHEE family

Primarily, PROMETHEE methods solve different problems of discrete multi-criteria decision-making. The key step of PROMETHEE is a pairwise comparison between all pairs of alternatives considered, see [1]. Those comparisons are done using a preference function $P : \mathbb{R} \to [0, 1]$ (i.e. the function which assigns a preference level $P_j(a, b)$ to the difference in values of compared alternatives a and b regarding the j-th criterion). Using (1) and (2), positive and negative flows of each alternative can be calculated. A positive flow can be considered as the index aggregating strenghts of the alternative and, similarly, a negattive flow can be regarded as the index aggregating weaknesses of the alternative in comparison with others. PROMETHEE I method defines a preference relationship between alternatives using (3) and equality by (4). In all other cases, a pair of alternatives is uncomparable. PROMETHEE II enables a decision-maker to compare also pairs which are uncomparable by PROMETHEE I. It is done by net flows $\phi(a)$ which aggregate positive and negative flows to the only one index ($\phi(a) = \phi^+(a) - \phi^-(a)$), see (5).

Except the ranking of alternatives, PROMETHEE can be also used to measure the efficiency of alternatives. This is done using the net flows of alternatives. At first, the set of criteria must be split into two subsets - input criteria $I = \{1, 2, \ldots, r\}$, and output criteria $O = \{1, 2, \ldots, s\}$. Then, a value of net flow regarding only the criteria from I, and, separately, a value of net flow regarding only the criteria from O are calculated, see (6) and (7) which were derived from (1) and (2). To avoid negative values of a net flow, the value of one is added.

Each alternative is characterized by the point $[\phi_I(a), \phi_O(a)]$. We say that a dominates b if and only if $\phi_I(a) \geq \phi_I(b)$ and $\phi_O(a) \geq \phi_O(b)$ (and one of these two inequalities must be satisfied as strict). The set of non-dominated (efficient) solutions (i.e. a set of efficient solutions) can be graphically denoted by the efficient frontier.

$$\phi^+(a) = \frac{1}{n-1} \sum_{b \neq a} \sum_{j=1}^{k} w_j \cdot P_j(a, b) \tag{1}$$

$$\phi^-(a) = \frac{1}{n-1} \sum_{b \neq a} \sum_{j=1}^{k} w_j \cdot P_j(b, a) \tag{2}$$

$$a \prec b \Leftrightarrow (\phi^+(a) \geq \phi^+(b) \wedge \phi^-(a) \leq \phi^-(b)) \tag{3}$$

$$a = b \Leftrightarrow (\phi^+(a) = \phi^+(b) \wedge \phi^-(a) = \phi^-(b)) \tag{4}$$

$$a \prec b \Leftrightarrow \phi(a) > \phi(b) \tag{5}$$

$$\phi_I(a) = \frac{1}{n-1}\sum_{b\neq a}\sum_{j=1}^{r} w_j \cdot P_j(a,b) - \frac{1}{n-1}\sum_{b\neq a}\sum_{j=1}^{r} w_j \cdot P_j(b,a) + 1 \tag{6}$$

$$\phi_O(a) = \frac{1}{n-1}\sum_{b\neq a}\sum_{j=1}^{s} w_j \cdot P_j(a,b) - \frac{1}{n-1}\sum_{b\neq a}\sum_{j=1}^{s} w_j \cdot P_j(b,a) + 1 \tag{7}$$

where $P_j(a,b)$ is a preference degree which expresses the power of preference of a over b regarding the j-th criterion, $\phi^+(a)$, $\phi^-(a)$ and $\phi(a)$ stands for positive, negative and net flow of the alternative a, w_j is a weight assigned to the j-th criterion (In the analysis of this contribution, it would not make sense to assign the (different) weights; their effect is excluded by setting the equal values to them), n stands for a number of alternatives considered, k stands for a number of criteria considered.

2.2 Basic Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a non-parametric approach for measuring relative efficiency of decision making units (DMUs) with multiple inputs and outputs. The basic model of DEA is the CCR model, see Charnes et al. [4].

Banker et al. in [2] extended the CCR model. The extended model is called the BCC model and considers variable returns to scale assumption. The model has convex envelope of data which leads to more efficient DMUs. The mathematical model of dual multiplier form of input-oriented BCC model is:

$$\begin{array}{lll} & \max\, e_Q = \sum_{j=1}^{S} v_j y_{jQ} - v_0, & \\ \text{s.t.} & \sum_{i=1}^{R} u_i x_{iQ} = 1, & \\ & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} - v_0 \leq 0, & k = 1, ..., T, \\ & u_i \geq 0, & i = 1, ..., R, \\ & v_j \geq 0, & j = 1, ..., S, \\ & v_0 \in (-\infty, \infty), & \end{array} \tag{8}$$

where v_0 is the dual variable assigned to the convexity condition $\mathbf{e^T}\lambda = \mathbf{1}$ of envelopment form of input-oriented BCC model, where λ would be the weight for each DMU. The BCC model can also be rewritten into the envelopment form or changed into the output orientation.

2.3 Data Envelopment Analysis for Undesirable Outputs

The models in previous subsection are basic models where the inputs and outputs are desirable. However, in real applications, there are frequently needed undesirable inputs and/or outputs. In this paper we deal with one undesirable output Also, according to the efficiency frontier of PROMETHEE method we use just the BCC model so the results could be compared.

Suppose the DEA data domain Y contains desirable (good) and undesirable (bad) outputs, represented as Y^g and Y^b, respectively. Obviously, Y^g should increase and Y^b should decrease to improve the performance of DMU. However, in the standard BCC model (8), both Y^g and Y^b are supposed to increase to improve the performance as Y. In order to increase the desirable outputs and to decrease the undesirable outputs, there were found many approaches. In this paper, we use three of them - translation transformation based on work by Koopmans [6], transformation of multiplicative inverse used by Lovell et al. [10] and basic method - change the meaning of the variable used by Liu and Sharp [9].

Translation
Based upon work of Koopmans [6], the basic model (8) is used for the calculation if: the Y^b is multiplied

by value (-1) and then the maximum value of $\max|Y^b|$ is add to $(-Y^b)$. Note: This model is known as Model A in this paper.

Inverse Multiplication
According to work of Lovell et al. [10] , the basic model (8) is used for the calculation if the Y^b is changed as inverse value $\frac{1}{Y^b}$. Note: This model is known as Model B in this paper.

Change of Variables
Based upon the idea of work by Liu and Sharp [9], Y^b is changed into X and the basic model (8) is used. Note: This model is known as Model C in this paper.

3 Variables

Following data have been involved in the analysis of efficiency. All data are of annual frequency and they are aggregated by all steel producing companies doing their business in a particular EU country.

- **Number of allowances allocated for free [pcs]** (*input*) - companies participating in the EU ETS system obtain certain amounts of allowances for free from the EU authorities. This factor influences costs on emissions trading as well as abatement costs. These data have been derived from carbonmarketdata.com database.
- **Energy consumption [GW]** (*input*) - the higher energy consumption by steel companies, the higher amount of emissions released and, also, the higher costs on emissions trading. Eurostat database is the source of these data.
- **Number of CO_2 emissions released [pcs]** (*output*) - emissions are considered as the undesirable output. The environmental efficiency weakens with increasing emissions. Data on emissions have been aggregated from carbonmarketdata.com database.
- **Level of iron and steel production [tons]** (*output*) - total amount of iron and steel products manufactured in a country. Here, a slight simplification has been done because products are supposed to be homogenous (no differences in envrionmental burden of different iron and steel products are taken into account). Data were taken from Eurostat database.

These data have been involved as a result of the close analysis of previous research. The classical production model for the DEA, such as the one in work of Kumar and Khanna [7] consider as inputs population and energy. GDP and CO_2 emissions are considered as outputs (desirable and undesirable, respectively). As we are looking for the special part of the industry, we changed the input population and we have used the number of allowances allocated for free.

4 Results of environmental efficiency analysis

4.1 PROMETHEE analysis

The PROMETHEE analysis has been performed using Visual PROMETHEE software. Two scenarios have been considered with respect to preference functions used. Usual preference function (preference function which assigns 1 for any non-zero difference in values) is used in Scenario 1, meanwhile, V-shape preference function is supposed for Scenario 2 (linear non-decreasing function whose function values increase with increase in diffrence of values and which aasign the preference value of 1 only when comparing the alternative with best value to the alternative with worst value). The results of the analysis for Scenario 1 are shown in Fig. 1. It can be seen that only five countries are environmentally efficient - CZ, PL, BEL, NED and ESP. All other 17 countries are dominated by the former quintuplet. However, it is important to emphasize that, due to the usual shape of preference functions, the impact of differences in values are disregarded. The second scenario was not possible to depict graphically because of very small differences in evaluatins of alternatives. Nevertheless, the set of efficient countries is similar - it consists of CZ, PL, BEL, NED and ITA. That means that only Spain switched its position with Italy.

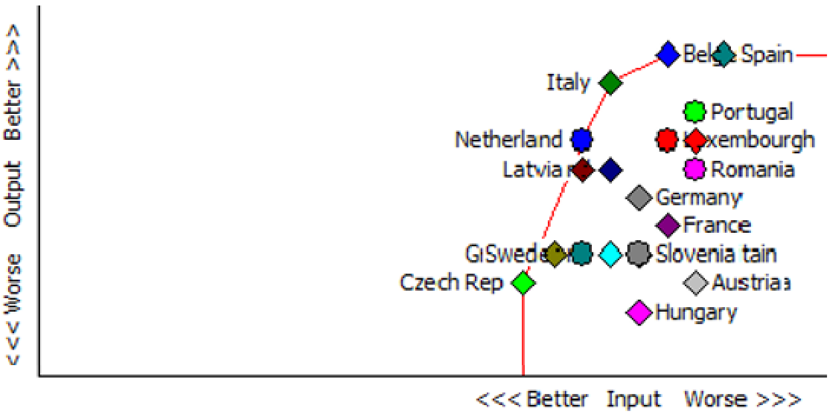


Figure 1 Efficiency frontier - PROMETHEE analysis

4.2 DEA analysis

All the DEA models have been performed using GAMS software. The results are shown in Table 1. It can be seen that all results for each models are similar. The biggest difference is between the Model A and the rest of the models. The difference is in the number of efficient countries. Seven countries are efficient - DK, ESP, ITA, LAT, LUC, NED, POR if the calculation have been done for the Model B and the Model C. Ten countries are efficient for the Model A - same countries as before and three more - F, SLO, SWE. Generally, the Model A gives same or highest efficient score then the rest of the models - for example: BEL has the efficient score for the Model A equal to 0.7681 and 0.6333 for the rest models. In case of GR, the efficiency score is different for all models - Model A (0.6419) > Model C (0.6185) > Model B (0.6006). Just for FIN and HU the highest efficiency score is in the Model C. If we have compared the Model B and Model C, we can see that the Model B gives the lowest efficient scores, for example in case of FIN the Model C gives the efficient score equal to 0.6182 and the Model B gives 0.5776. These differences are caused by different position of the undesirable output. We can see how each transformation influence the efficiency score.

DMU		Model A	Model B	Model C	DMU		Model A	Model B	Model C
Austria	AU	0.9871	0.9871	0.9871	Italy	ITA	1.0000	1.0000	1.0000
Belgium	BEL	0.7681	0.6333	0.6333	Latvia	LAT	1.0000	1.0000	1.0000
Bulgaria	BLG	0.8063	0.4364	0.4364	Luxembourg	LUC	1.0000	1.0000	1.0000
Czech Republic	CZ	0.4557	0.4557	0.4557	Netherlands	NED	1.0000	1.0000	1.0000
Denmark	D	0.6182	0.6182	0.6182	Poland	PL	0.8341	0.6176	0.6176
Finland	FIN	0.5776	0.5776	0.6182	Portugal	POR	1.0000	1.0000	1.0000
France	F	1.0000	0.5368	0.5368	Romania	RO	0.3254	0.3254	0.3254
Germany	DK	1.0000	1.0000	1.0000	Slovakia	SK	0.7540	0.3492	0.3492
Great Britain	GB	0.9073	0.8189	0.8189	Slovenia	SLO	1.0000	0.1409	0.1409
Greece	GR	0.6419	0.6006	0.6185	Spain	ESP	1.0000	1.0000	1.0000
Hungary	HU	0.5052	0.5511	0.5052	Sweden	SWE	1.0000	0.4367	0.4367

Table 1 DEA analysis - Efficiency of Model A, Model B and Model C

5 Conclusion

This paper deals with measuring the environmental efficiency of natioanl steel sectors in the EU. Two groups of methods have been used to do that - PROMETHEE methods and DEA methods. When comparing the results of both methods, significant differences have been found. The number of efficient units is similar (5 by PROMETHEE vs. 7 by DEA) but, only two of them (Italy and the Netherlands) are in common. Moreover, e.g. the Czech Republic has been considered to be the one of the least efficient countries by DEA and, also, as efficient by PROMETHEE. These facts can be a bit surprising. On the other hand, the algorithms of both methods are very different. Preference functions used in PROMETHEE help to prevent from overvaluing large differences in values according to some criterion. Unfortunately, no general recommendation can be given for a decision-maker when deciding on appropriateness of the methods. Maybe, it would be the best to use both approaches and then try to compare and justify possible differences in results. Anyway, when some unit is evaluated as efficient by both methods, it is reasonable to trust the results with high certainty.

For further research, we expect to revise (and maybe extend) the sets of inputs and outputs. Moreover, a dynamic (network) version of the model may be established in the future.

Acknowledgements

This paper was supported by grant No. GA 16-01298S of the Czech Science Foundation, SGS project No. SP2016/116 and the Operational Programme Education for Competitiveness Project CZ.1.07/2.3.00/-20.0296. The support is greatly acknowledged.

References

[1] Brans, J. P, and Mareschal, B: *PROMETHEE-GAIA: Une Methodologie d'aide a la Decision en Presence de Criteres Multiples.* SMA, Bruxelles, 2002.

[2] Banker, R. D., Charnes, A., and Cooper, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* **30** (1984), 1078–1092.

[3] Behzadian, M., Kazemzadeh, R. B., Albadvi, A., and Aghdasi, M. PROMETHEE: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research* **200** (2010), 198–215.

[4] Charnes, A., Cooper, W. W., and Rhodes, E. Measuring the efficiency of decision making units. *European Journal of Operational Research* **2** (1978), 429–444.

[5] Charnes, A., Cooper, W. W., Golany, B., Seiford, L., and Stutz, J. Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. *Journal of Econometrics* **30** (1985), 91–107.

[6] Koopmans, T. C. *Analysis of production as an efficient combination of activities.* In T. C. Koopmans (Ed.), Activity analysis of production and allocation, Cowles Commission, New York, 1951.

[7] Kumar, S., and Khanna, M. Productivity growth and CO2 abatement: A cross-country analysis using the distance function approach. *International Conference on Climate Change and Environmental Policy.* University of Illinois, 2002.

[8] Lenort, R., Baran, J., and Wysokinski, M. Application of Data Envelopment Analysis to Measure the Efficiency of the Metal Production Sector in Europe. *In Metal 2014: 23th International Conference on Metallurgy and Materials.* Ostrava: TANGER, (2014), 1795–1802.

[9] Liu, W. B., and Sharp, J. *DEA models via goal programming.* In G. Westerman (Ed.), Data envelopment analysis in the public and private sector. Deutscher Universitats-Verlag, 1999.

[10] Lovell, C. A. K., Pastor, J. T., and Turner, J. A. Measuring macroeconomic performance in the OECD: a comparison of European and non-European countries. *European Journal of Operational Research* **87** (1995), 507–518.

[11] Morfeld, J., and Silveira, S. Capturing energy efficiency in European iron and steel production-comparing specific energy consumption and Malmquist productivity index. *Energy Efficiency* **7** (2014), 955–972.

[12] Seiford, L. M., and Zhu, J. Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research* **42** (2002), 16–20.

Banking efficiency in the Visegrad Group according to Data Envelopment Analysis

Lucie Chytilová[1]

Abstract. The situation in the banking industry is very difficult all around the world. The Central Europe and especially countries of the Visegrad Group have to keep up with a lot of changes - historical development, new rules of the European Commission or the 2008 crisis. There are many studies which evaluate banking efficiency in the Visegrad Group. Almost all of these studies evaluate the efficiency under certainty and just for small sample. The problem of the studies are missing data so they chose just some banks. This may cause imprecise results. The aim of this paper is to determine the efficiency of the banking industry in Visegrad Group for a big group of banks (57 banks). The technical efficiency of banks in the Visegrad Group is estimated for years from 2009 to 2013. The special DEA models for calculations with the missing data are used. The development of calculated technical efficiency is analysed in the cross section between the units and in the time. Also the influence of the risk variable is measured as well.

Keywords: banking, Data Envelopment Analysis, efficiency, Visegrad Group, banking.

JEL classification: C61, G21, D24
AMS classification: 90C05

1 Intoduction

The situation in the banking industry is very difficult all around the world. The Central Europe and especially countries of the Visegrad Group have to keep up with a lot of changes - structural changes of banking regulation and financial market, development of security market, bigger stock market or new regulations by the European Union, see [4]. This all causes the big interest in the topic - efficiency in the banking sector of the Visegrad Group.

The Visegrad Group contains four countries - the Czech Republic, Hungary, Poland and Slovakia. These countries are connected not just geographically. The historical background is very similar in these countries. All four countries had been part of the Eastern Bloc which had fallen apart in 1989. This led them to integrate and cooperate. They started to create the market economy and have opened themselves to the rest of the world. The transformation of the banking system was an essential part of the transformation. In 2004 all countries of the Visegrad Group have jointed the European Union. Slovakia has joined the third stage of the European Monetary Union in 2009. Even there are some differences, it could be assumed that their financial and banking systems should not show major differences. Identification of the potential differences is one of the objectives of this paper.

This paper mainly focuses on the identification of the efficiency of commercial banks in the Visegrad Group during time period from 2009 to 2013 by using special model of DEA. Except of this paper, there are no papers which deal with so many banks (57 banks) for the region. The reason is the general problem with datasets. In this case, the dataset contained 57 banks from the Visegrad Group countries, but some variables were missing. To still be able to make the analysis, first the estimation of these missing variables was done and then the special DEA model was used. Model was calculated for the time period and close analysis of all results were done for each bank as well as for each country. All the calculation were done by GAMS software.

The rest of the paper has the following structure: Section 2 provides the formulations and the models

[1]VŠB-TU Ostrava, Faculty of Economics, Sokolská 33, 702 00, Ostrava, Czech Republic, lucie.chytilova@vsb.cz

of interval DEA to estimate efficiency bounds for the units with missing values. In Section 3, input and output variables are defined and information about the missing data and their estimation are given. Section 4 focuses on application and the efficiency analysis of a set of banks in Visegrad Group based on a dataset with missing values. The discussions about the results are provided in this section as well. Section 5 gives some conclusions and remarks.

2 Methodology

2.1 Classic DEA

Data Envelopment Analysis (DEA) is a non-parametric approach. It is widely used for measuring relative efficiency of decision making units (DMUs) with multiple inputs and outputs. Assume, there is a set of T DMUs (DMU$_k$ for $k = 1, ..., T$), let inputs and outputs data be $X = \{x_{ik}, i = 1, ...R; k = 1, ...T\}$ and $Y = \{y_{jk}, j = 1, ...S; k = 1, ...T\}$, respectively. Also, u_i for $i = 1, ...R$ and v_j for $j = 1, ..., S$ be the weights of the i^{th} input and the j^{th} output, respectively. Mathematically, the relative efficiency score of DMU$_k$ can be defined as:

$$e_k = \frac{\sum_{j=1}^{S} v_j y_{jk}}{\sum_{i=1}^{R} u_i x_{ik}}, \text{ for } k = 1, ..., T. \tag{1}$$

Charnes et al. [3] have proposed the following CCR model to measure the efficiency score of the under evaluation unit, DMU$_Q$ where $Q \in \{1, ..., T\}$:

$$\begin{array}{lll} & \max\, e_Q = \frac{\sum_{j=1}^{S} v_j y_{jQ}}{\sum_{i=1}^{R} u_i x_{iQ}}, & \\ \text{s.t.} & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} \leq 0, & k = 1, ..., T, \\ & u_i \geq 0, & i = 1, ..., R, \\ & v_j \geq 0, & j = 1, ..., S. \end{array} \tag{2}$$

The model (2) is non-linear. It is the model of linear-fractional programming. The model (2) could be transferred by Charmes-Cooper transformation to the standard linear programming problem:

$$\begin{array}{lll} & \max\, e_Q = \sum_{j=1}^{S} v_j y_{jQ}, & \\ \text{s.t.} & \sum_{i=1}^{R} u_i x_{iQ} = 1, & \\ & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} \leq 0, & k = 1, ..., T, \\ & u_i \geq 0, & i = 1, ..., R, \\ & v_j \geq 0, & j = 1, ..., S, \end{array} \tag{3}$$

where $Q \in \{1, ..., T\}$. DMU$_Q$ is CCR-efficient if and only if $e^* = 1$ and if there exists at least one optimal solution $(\mathbf{u}^*, \mathbf{v}^*)$ with $\mathbf{u}^* > \mathbf{0}$ and $\mathbf{v}^* > \mathbf{0}$ for the set $Q \in \{1, ..., T\}$. The inefficient units have a degree of relative efficiency that belongs to interval $[0, 1)$. Note: The model must be solved for each DMU separately.

The model (3) is called a multiplier form of the input-orient-CCR model. However, for computing and data interpretation, it is preferable to work with model that is dual associated to model (3). The model is referred as envelopment form of input-oriented CCR model, see [3]. There also exists the multiplier form and envelopment form of output-oriented CCR model. Both models give the same results, see [3].

Banker et al. [2] have extended the CCR model. The extended model is called the BCC model and considers variable returns to scale assumption. The model has convex envelope of data which leads to more efficient DMUs. The mathematical model of dual multiplier form of input-oriented BCC model is:

$$\begin{array}{lll} & \max\, e_Q = \sum_{j=1}^{S} v_j y_{jQ} - v_0, & \\ \text{s.t.} & \sum_{i=1}^{R} u_i x_{iQ} = 1, & \\ & \sum_{j=1}^{S} v_j y_{jk} - \sum_{i=1}^{R} u_i x_{ik} - v_0 \leq 0, & k = 1, ..., T, \\ & u_i \geq 0, & i = 1, ..., R, \\ & v_j \geq 0, & j = 1, ..., S, \\ & v_0 \in (-\infty, \infty), & \end{array} \tag{4}$$

where v_0 is the dual variable assigned to the convexity condition $\mathbf{e^T}\lambda = \mathbf{1}$ of envelopment form of input-oriented BCC model. Note: The BCC model can be rewritten into the envelopment form or changed into the output orientation.

The input-oriented BCC model will continue into the next section.

2.2 DEA for missing data

There exist many improvements of the classical DEA models. In this paper the main problems are the missing data. To deal with this issue the special model has to be use. This model is define below. As a inspiration, the model of Smirlis et al. [9] have been used. Their model have been done for the output-oriented DEA model, so the transformation was needed.

Assume, there are T DMUs, each using R inputs to produce S outputs. For any unit k ($k = 1, ..., T$), the level of its j^{th} output ($j = 1, ..., S$) is denoted by y_{jk} and by x_{ik} the level of its i^{th} input ($i = 1, ..., R$). Unlike the original DEA model, the interval DEA assumes that some of the crisp input x_{ik} and output y_{jk} values are not known and for them, it is only known that they lie within bounded intervals, i.e. $x_{ik} \in [x^L_{ik}, x^U_{ik}]$ and $y_{jk} \in [y^L_{jk}, y^U_{jk}]$, with the upper and lower bounds of the intervals x^L_{ik}, x^U_{ik}, y^L_{jk}, y^U_{jk} to be strictly positive constants.

To be able to introduce the intervals instead of exact data into the model (4), some transformation should be done.

The values x_{ik} and y_{jk} are expressed in terms of new variables s_{ik} and p_{jk}, respectively, to convert the non-linear model to a linear one. These new variables locate the level of inputs and outputs within the bounded intervals $[x^L_{ik}, x^U_{ik}]$ and $[y^L_{jk}, y^U_{jk}]$, respectively, as follow:

$$\begin{array}{ll} x_{ik} = x^L_{ik} + s_{ik}(x^U_{ik} - x^L_{ik}), & i = 1, ..., R;\ k = 1, ..., T \text{ with } 0 \le s_{ik} \le 1, \\ y_{jk} = y^L_{jk} + t_{jk}(y^U_{jk} - y^L_{jk}) & j = 1, ..., S;\ k = 1, ..., T \text{ with } 0 \le t_{jk} \le 1. \end{array}$$

Applying the above transformation to model (4), the following linear model is obtained:

$$\begin{array}{lll} & \max\ e_Q = \sum_{j=1}^{S} v_j\big(y^L_{jjQ} + t_{jkQ}(y^U_{jkQ} - y^L_{jkQ})\big) - v_0, & \\ \text{s.t.} & \sum_{i=1}^{R} u_i\big(x^L_{ikQ} + s_{ikQ}(x^U_{ikQ} - x^L_{ikQ})\big) = 1, & \\ & \sum_{j=1}^{S} v_j\big(y^L_{jj} + t_{jk}(y^U_{jk} - y^L_{jk})\big) - \sum_{i=1}^{R} u_i\big(x^L_{ik} + s_{ik}(x^U_{ik} - x^L_{ik})\big) - v_0 \le 0, & k = 1, ..., T, \\ & u_i \ge 0, 0 \le s_{ik} \le 1 & i = 1, ..., R, \\ & v_j \ge 0, 0 \le t_{jk} \le 1 & j = 1, ..., S, \\ & v_0 \in (-\infty, \infty). & \end{array} \tag{5}$$

It can be see that for inputs and outputs, there are new terms $u_i s_{ik}$ and $v_j t_{jk}$, respectively. These new terms may be replaced by new variables $q_{ik} = u_i s_{ik}$ and $p_{jk} = v_j t_{jk}$ which meet the needed conditions. The model (5) can be rewritten as follows:

$$\begin{array}{lll} & \max\ e_Q = \sum_{j=1}^{S} \big(v_j y^L_{jjQ} + p_{jkQ}(y^U_{jkQ} - y^L_{jkQ})\big) - v_0, & \\ \text{s.t.} & \sum_{i=1}^{R} \big(u_i x^L_{ikQ} + q_{ikQ}(x^U_{ikQ} - x^L_{ikQ})\big) = 1, & \\ & \sum_{j=1}^{S} v_j y^L_{jj} + \sum_{j=1}^{S} p_{jk}(y^U_{jk} - y^L_{jk}) - \sum_{i=1}^{R} u_i x^L_{ik} - \sum_{i=1}^{R} q_{ik}(x^U_{ik} - x^L_{ik}) - v_0 \le 0, & k = 1, ..., T, \\ & q_{ik} - u_i \le 0 & i = 1, ..., R, \\ & p_{jk} - v_j \le 0 & j = 1, ..., S, \\ & u_i, v_j \ge 0 & \forall i, j \\ & q_{ik}, p_{jk} \ge 0 & \forall i, j, k \\ & v_0 \in (-\infty, \infty). & \end{array} \tag{6}$$

In model (6), unknown variables under estimation are the weights u_i, v_j and the new variables q_{ik}, p_{jk} that denote the level of input and output values within the bounded intervals. For more details see [9].

3 Input and Output Variables

The important task for the efficiency measurements is to identify the right and relevant variables for the calculation.

There are known two main approaches for the bank efficiency evaluation as production unit. The main difference is the treatment of deposits. The production approach views deposits as output. Banks are producers of deposits, loans and other services. Inputs are define as physical variables. This approach was found by Benston [1]. Benston also found disadvantages - detailed database is required and it does not take into consideration the interest costs. The second approach was found by Sealy and Lindley [8] - intermediation approach. Banks are financial intermediaries between depositors and creditors. They collect deposits and other liabilities to apply them as interest-earning assets. Deposits are considered as input. In this case operating cost and interest cost are considered. It is the most common approach nowadays.

Even of the common use of the intermediation approach there is always question about variables. According to earlier analysis and research in this field (see [6], [5] and [7]), the input and output variables for this paper were selected according to intermediation approach. Plus the variable Equity is used as input and the Non-interest income as output, see Table 1.

	Variables	Description in the balance sheet	Units
Inputs	Physical capital (x_1 - FA)	Fixed assets	th Euro
	Labour (x_2 - LAB)	Number of employees	Number
	Loanable funds (x_3 - LF)	Deposits + Short term funding	th Euro
	Equity (x_4 - EQ)	Total equity	th Euro
Outputs	Advances (y_1 - ADL)	Loans + Advances to Banks	th Euro
	Investments (y_2 - INV)	Other Securities	th Euro
	Non-interest income (y_3 - NII)	Non-Earning Assets	th Euro

Table 1 Description of inputs and outputs

The required dataset of inputs and outputs was collected from the Bankscope[1]. In Table 1 is seen the description of inputs and outputs. This paper is based on period between 2009 to 2013. 57 banks were found for the analysis (14 - CZ, 21 - PL, 8 - SK, 14 - HU). There were even more banks for the interval, but for some of them there was not possible to estimate the missing values (usually case of one or two missing variables). The basic estimation techniques were used (linear regression and clustering of the units). For future, it would be good to improve the time period.

4 Results and Discussions

In Table 2 are seen the results for all banks during the whole time period. It is seen that during the all period, there are efficient: 7 banks in the Czech Republic, 4 banks in Hungary and Poland and just one bank in Slovakia. These results and also the geometric mean show that the most efficient banks are in the Czech Republic. The second place belongs to Hungary and third place is occupied by banks from Slovakia and Poland. According to efficiency score the Slovak banks are better but according to the relative number of over all efficiency the polish banks are considered as more efficient. These results are consistent with some previous researches which were done, see [6] and [7].

Table 2 shows that the bank with the biggest improvement for the time period is the Hungarian bank with DMU23 (KDB Bank Europe Ltd). Also another banks from Hungary improved a lot in this time period. So it can be concluded that this country have the biggest improvement during the period. The situation in the Czech republic is steady, as well as in Poland for the period from 2009 to 2013. The worst situation is in Slovakia - the biggest decrease during the time period. Overall it looks like that banks in these countries improved or at least stayed at same level, even after the financial crisis in 2008. Apart from the exception in the form of Slovakia and the crisis period.

[1] https://bankscope.bvdinfo.com/

DMU	Bank	Origin	2009	2010	2011	2012	2013
DMU01	Ceska Sporitelna a.s.	CZ	1.0000	1.0000	1.0000	1.0000	1.0000
DMU02	Ceskomorava Zarucni a ...	CZ	1.0000	1.0000	1.0000	1.0000	1.0000
DMU03	CSOB CZ	CZ	1.0000	1.0000	1.0000	1.0000	1.0000
DMU04	Evropsko-ruska banka	CZ	1.0000	1.0000	1.0000	1.0000	1.0000
DMU05	Expobank CZ	CZ	0.4316	0.4090	0.4689	0.3440	0.3580
DMU06	GE Money Bank	CZ	0.4417	0.4958	0.5158	1.0000	0.6582
DMU07	Hypotecni banka	CZ	1.0000	0.7537	0.5543	0.6037	0.5286
DMU08	J&T Banka	CZ	0.6206	0.5985	0.9018	1.0000	0.9085
DMU09	Komercni Banka	CZ	1.0000	1.0000	0.9508	0.9060	1.0000
DMU10	Modra pyramida stavebni spo.	CZ	0.7482	0.5248	0.4904	0.5179	0.5949
DMU11	PPF banka	CZ	1.0000	1.0000	1.0000	1.0000	1.0000
DMU12	Raiffeisen stavebni sporitelna	CZ	1.0000	1.0000	1.0000	1.0000	1.0000
DMU13	Stavebni Sporitelna CS	CZ	1.0000	1.0000	1.0000	1.0000	1.0000
DMU14	Unicredit Bank CR and SK	CZ	0.7573	0.7138	1.0000	1.0000	0.9707
DMU15	Bank of China	HU	1.0000	1.0000	1.0000	1.0000	1.0000
DMU16	Budapest Bank Nyrt-Buda.	HU	0.3799	0.3701	0.4703	0.5643	0.5686
DMU17	CIB Bank Ltd-CIB Bank Zrt	HU	0.3033	0.4557	0.7081	0.3893	0.5120
DMU18	DRB Del-Dunantuli Region. B.	HU	1.0000	1.0000	0.9359	1.0000	1.0000
DMU19	Erste Bank Hungary Nyrt	HU	1.0000	0.9847	1.0000	1.0000	0.7621
DMU20	FHB Kereskedelmi Bank Zrt	HU	1.0000	1.0000	0.7209	0.8936	0.9300
DMU21	Granit Bank Zrt	HU	1.0000	1.0000	1.0000	1.0000	1.0000
DMU22	K&H Bank Zrt	HU	1.0000	1.0000	1.0000	1.0000	1.0000
DMU23	KDB Bank Europe Ltd	HU	0.2616	0.3846	0.4172	0.6191	1.0000
DMU24	Magyar Cetelem Bank Rt	HU	0.1904	0.4580	0.8181	0.8966	0.8787
DMU25	MKB Bank Zrt	HU	0.9895	1.0000	1.0000	1.0000	1.0000
DMU26	OTP Bank Plc	HU	1.0000	1.0000	1.0000	1.0000	1.0000
DMU27	Raiffeisen Bank Zrt	HU	0.8123	0.6733	0.7206	0.9311	0.6405
DMU28	UniCredit Bank Hungary Zrt	HU	0.5967	0.3821	0.4076	0.6086	0.7493
DMU29	Alior Bank Spolka Akcyjna	PL	1.0000	0.5369	0.5215	0.5977	0.4204
DMU30	Bank Handlowy w Warszawie	PL	1.0000	1.0000	1.0000	1.0000	1.0000
DMU31	Bank Millennium	PL	0.6202	0.5881	0.4604	0.9364	0.8451
DMU32	Bank Ochrony Srodowia	PL	0.4293	0.4616	0.6055	0.7980	0.7821
DMU33	Bank Polska Kasa Opieki	PL	1.0000	1.0000	1.0000	1.0000	0.8123
DMU34	Bank Polskiej Spoldzielcuosci	PL	1.0000	1.0000	1.0000	1.0000	1.0000
DMU35	Bank SpOldzielczy w Brodnicy	PL	0.7280	0.8120	0.7612	0.7920	0.7147
DMU36	Bank Zachodni WBK	PL	0.6270	0.6448	0.7006	0.9236	1.0000
DMU37	BNP Paribas Bank Pola	PL	0.5953	0.6166	0.5470	0.8863	0.5454
DMU38	DNB Bank Polska	PL	0.4051	0.4670	0.2804	0.2156	0.2962
DMU39	Euro Bank	PL	0.1847	0.2892	0.2336	0.2955	0.3241
DMU40	FM Bank PBP	PL	0.9838	1.0000	0.9826	0.9255	0.7950
DMU41	HSBC Bank Polska	PL	1.0000	1.0000	1.0000	0.9740	0.9219
DMU42	ING Bank Slaski - Capital Gr.	PL	1.0000	1.0000	1.0000	1.0000	1.0000
DMU43	MBank Hipoteczny	PL	0.3922	0.3464	0.3246	0.6181	0.7603
DMU44	mBank	PL	1.0000	0.6585	0.6619	1.0000	0.8754
DMU45	Nordea Bank Polska	PL	0.3235	0.2471	0.4666	0.8202	0.3443
DMU46	Plus Bank	PL	0.5262	0.5008	0.4290	0.5208	0.4405
DMU47	RBS Bank (Polska)	PL	1.0000	1.0000	1.0000	1.0000	1.0000
DMU48	SGB Bank	PL	0.8710	0.8822	1.0000	1.0000	1.0000

DMU49	Volkswagen Bank Polska	PL	0.4127	0.3643	0.3732	0.3993	0.5053
DMU50	CSOB SK	SK	0.9732	0.7792	0.4912	0.6003	0.6115
DMU51	CSOB Stavebna Sporitelna	SK	1.0000	1.0000	1.0000	1.0000	1.0000
DMU52	OTP Banka Slovensko	SK	0.5479	0.4135	0.4258	0.4485	0.4067
DMU53	Prima banka Slovensko	SK	0.9894	0.9234	0.6978	0.8703	0.6339
DMU54	Sberbank Slovensko	SK	0.6748	0.3301	0.5051	0.4241	0.4788
DMU55	Tatra Banka	SK	0.7801	0.6312	0.5852	0.5918	0.6177
DMU56	Vseobecna Uverova Banka	SK	0.8559	0.8260	0.6772	0.5982	0.5903
DMU57	Post Bank JSC-Postova Banka	SK	1.0000	0.8151	1.0000	0.8955	0.9462
	average efficiency by country:	CZ	0.8246 (9)	0.7859 (8)	0.8143 (8)	0.8468 (10)	0.8229 (8)
	(number of efficient DMUs)	HU	0.6563 (7)	0.7052 (7)	0.7627 (6)	0.8196 (7)	0.8400 (7)
		PL	0.6525 (8)	0.6283 (7)	0.6201 (7)	0.7392 (7)	0.6812 (6)
		SK	0.8356 (2)	0.6723 (1)	0.6430 (2)	0.6491 (1)	0.6338 (1)

Table 2 Results of DEA

5 Conclusion

In this paper, the special input-oriented BCC DEA model for the calculation of banking efficiency in the Visegrad Group for the time period from 2009 to 2013 is used. First the missing values have been estimated by the regression analysis and clustering of the units. Then the special model have been used. The results have shown that the most efficient banking industry is in the Czech republic and the worst is in Poland (according to number of efficient banks which have been efficient for whole period) or in Slovakia (according to the efficiency score overall time). This gives similar results as the previous research (see 1. Chapter of [7]), where just 27 banks were used. The changes are due to the extension of dataset.

For further research, the sets of inputs and outputs or DMUs should be revised (and maybe extend). Moreover, a different version of the model may be established in the future.

Acknowledgements

This paper was supported by SGS project No. SP2016/116. The support is greatly acknowledged.

References

[1] Benston, G. J.: Branch Banking and Economies of Scale. *Journal of Finance* **20** (1965), 312–333.

[2] Banker, R. D., Charnes, A., and Cooper, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* **30** (1984), 1078–1092.

[3] Charnes, A., Cooper, W. W., and Rhodes, E. Measuring the efficiency of decision making units. *European Journal of Operational Research* **2** (1978), 429–444.

[4] Fang, Y., Hasan, I., Marton, K. and, Waisman, M.: Bank valuation in new EU member countries. *Economic Systems* **38** (2014), 55–72.

[5] Gulati, R., and Kumar, S.: Impact of non-traditional activities on the efficiency of Indian banks: an empirical investigation. *Macroeconomics and Finance in Emerging Market Economies* **4** (2011), 125–166.

[6] Hančlová, J., and Chytilová, L.: The impact of inclusion of non-traditional activities on efficiencies in European banking industry using the CCR-I model. In: *Proceedings of the Conference on Strategic Management and its Support by Information Systems 2015*, VŠB-TUO, Ostrava, 2015, 208–220.

[7] Hančlová, J., et al. *Optimization Problems in Economics and Finance.* VŠB-TUO, Ostrava, 2015.

[8] Lyroudi, K., and Angelidis, D.: Measuring Banking Productivity of the Most Recent European Union Member Countries. *Journal of Economics and Business* **9** (2006), 37–57.

[9] Smirlis, Y., G., Maragos, E., K., and Despotis D., K. Data envelopment analysis with missing values: An interval DEA approach. *Applied Mathematics and Computation* **177** (2006), 1–10.

Multi-period analysis and resource allocation among Czech economic faculties

Josef Jablonský [1], Vilém Sklenák [2]

Abstract. The paper deals with analysis of resource allocation among Czech economic faculties within the 7 years period since 2008 until 2014. The resources in the period *t* are divided among faculties according to their teaching and research performance in the period (*t* - 1). The teaching performance is measured by the number of students and number of graduated in the given period. The research performance is defined according to the national rules by the number of so called RIV points. The overall performance level in the given period is estimated using Data Envelopment Analysis model with constant returns to scale assumption. The input in the period *t* is partly based on the efficiency or super-efficiency scores from the preceding period. Super-efficiency score given by Andersen and Petersen model is used for discrimination among efficient units. The results of the allocation in the last period are compared to the ones calculated to the standard methodology. Numerical experiments are realized using own MS Excel macros and LINGO modelling and optimization system.

Keywords: data envelopment analysis, multi-period model, resource allocation

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Allocation of a limited amount of resources among the set of decision making units (DMUs) is an important task that finds its applications in various areas of human activities. One of the typical problems of this nature is allocation of funds among higher education institutions which is a problem that significantly influences all activities of the units under evaluation. The common practice in the Czech Republic consists in dividing of available resources among universities and then among faculties according to their performance in the past one or several periods. The overall performance of the universities (faculties) is given by their partial performances in various activities carried out. Among the main activities of higher educational institutions belong teaching, research, international relations, co-operation with partners, etc. Of course teaching and research are the most important ones in this context. That is why evaluation of teaching and research performance of the set of educational institutions (let us denote them further units or faculties) is a very important problem and the methodology used for this evaluation must reflect and respect many requirements.

The main aim of this paper is to contribute to the discussion in this field and try applying not so widely used techniques. The common practice in the Czech Republic uses basic multiple criteria decision making approaches, i.e. assigns expert weights to the defined criteria that describe various aspects of overall performance and then applies simple additive weighting. The given performance score is a basis for allocation of funds. In the last several tens of years a more sophisticated methodological approaches for efficiency and performance evaluation have been proposed and verified by various researchers. Data Envelopment Analysis (DEA) models introduced by Charnes et al. (1978) belong without any doubts among the most theoretically developed and applied techniques. Applications of DEA models in higher education are numerous. Let us mention at least three of them – (Beasley, 1995), (Thanassoulis et al., 2011) and (Jablonský, 2016). The time span of these publications shows that the research in this field is not finished yet.

The paper is organized as follows. Section 2 defines shortly basic DEA models and introduces multi-period DEA models which is a class of models that evaluates the performance of the set of units within several consecutive periods. Traditional models of this nature are usually oriented on the best period of the unit under evaluation and do not take into account interconnections among the periods. That is why their results can just hardly be used for allocation of funds among the units. An iterative approach that combines results of the performance evaluation in

[1] University of Economics, Prague, Department of Econometrics, W. Churchill Sq. 4, Praha 3, e-mail: jablon@vse.cz.
[2] University of Economics, Prague, Department of Knowledge Engineering, W. Churchill Sq. 4, Praha 3, e-mail: sklenak@vse.cz.

the period t and the amount of funds allocated to the unit in the period t (based on the evaluation in the period (t–1)) as the main input variable in the period (t+1) is introduced in Section 3. The proposed approach is illustrated on the real data set of 19 Czech economic faculties in multiple periods from 2008 until 2014. The given results are summarized in the final section of the paper.

2 Multi-period DEA models

Traditional DEA models analyse relative technical efficiency of the set of n DMUs that are described by m inputs and r outputs in one period. The efficiency score θ_q of the DMU_q is defined as the weighted sum of outputs divided by the weighted sum of inputs as follows:

$$\theta_q = \frac{\sum_{k=1}^{r} u_k y_{kq}}{\sum_{i=1}^{m} v_i x_{iq}}, \tag{1}$$

where u_k, $k = 1, 2, \ldots, r$ is the positive weight of the k-th output, v_i, $i = 1, 2, \ldots, m$ is the positive weight of the i-th input, and x_{ij}, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$ and y_{kj}, $k = 1, 2, \ldots, r$, $j = 1, 2, \ldots, n$ are non-negative values for the DMU_j of the i-th input and the k-th output respectively. Traditional DEA models maximize the efficiency score (1) under the assumption that the efficiency scores of all other DMUs do not exceed 1 (100%). This problem must be solved for each DMU separately, i.e. in order to evaluate the efficiency of all DMUs the set of n optimization problems must be solved. The presented problem is not linear in objective function but it can be converted into a linear optimization problem and then solved easily. The transformation consists in maximization of the nominator or minimization of the denominator in expression (1). The constraints of this LP optimization problem express the upper bound for efficiency scores of all DMUs except the unit q and the unit sum of the denominator/nominator in (1). The model that maximizes the nominator in (1) is referenced as DEA input oriented model, the model that minimizes the denominator is DEA output oriented model. In both cases the DMUs with $\theta_q = 1$ are lying on the efficient frontier estimated by the model and denoted as efficient units. Otherwise the units are inefficient and the efficiency score can be explained as a rate of input reduction or output expansion in order to reach the maximum efficiency. In some cases, it is more convenient working with dual problem to the linearized version of the model described above. This model, often referenced as input oriented envelopment CCR (Charnes, Cooper and Rhodes) model is formulated as follows:

$$\begin{aligned}
&\text{minimize} && \theta_q^{CCR} && \\
&\text{subject to} && \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- = \theta_q^{CCR} x_{iq}, && i = 1,2,\ldots,m, \\
& && \sum_{j=1}^{n} y_{kj}\lambda_j - s_k^+ = y_{kq}, && k = 1,2,\ldots,r, \\
& && \lambda_j \geq 0, && j = 1,2,\ldots,n.
\end{aligned} \tag{2}$$

The optimal objective function value of the model (2) is lower than 1 for inefficient units and equals 1 for the units weakly or fully efficient. In order to rank efficient units many models based on various principles have been proposed. An important group of these models are so called super-efficiency models. The first model of this nature was introduced in (Andersen and Petersen, 1993). Its input oriented formulation is very close to model (2). In Andersen and Petersen (AP) model the weight of the DMU_q λ_q is set to zero. It causes that the DMU_q is removed from the set of units and the efficient frontier changes its shape after this removal. Super-efficiency score θ_q^{AP}, that is greater than 1 for units CCR fully efficient, measures the distance of the unit DMU_q from the new efficient frontier. Its value expresses how many times the inputs may increase (in input oriented version of the model) in order the evaluated unit remains efficient.

Traditional DEA models (1) and (2) evaluate the efficiency or performance of the set of units in one time period only according to the values of input and output variables but it is clear that the performance in the period t can depend not only on the inputs in this period but on the inputs in one or several preceding periods. That is why various attempts how to solve this problem have been proposed in the past. This group of models is called multi-period DEA models. The first model of this class was introduced by Park and Park (2009). Unfortunately, their model (further as PP model) is not a real multi-period model because it returns just the best efficiency score of the unit under evaluation within all periods considered. It is the model oriented on the best period. It is clear that according to this property the PP model cannot be used as a tool for future allocation of resources. Its mathematical formulation can be found e.g. in (Park and Park, 2009) or (Jablonsky, 2016). Similarly to the PP model it is possible formulate models that are oriented on the worse period of the unit under evaluation or the model that returns average efficiency score over all periods. Their formulation is given in (Jablonsky, 2016). In addition, this paper contains an original SBM multi-period model that measures efficiency during all periods using undesirable slacks of inputs and outputs. Nevertheless, this model does not take into account interconnections among particular periods. In order to do it, the appropriate model is not linear and its solution is questionable. The DEA model that uses efficiency score given in the period (t–1) as the only input variable in the period t is as follows:

$$\begin{aligned}
&\text{minimize} && \sum_{q=1}^{n}\sum_{t=2}^{T}\theta_q^t && \\
&\text{subject to} && \sum_{j=1}^{n}\theta_j^{t-1}\lambda_{jq}^t \le \theta_q^t\theta_q^{t-1}, && t=2,3,\dots,T,\, i=1,2,\dots,m,\, q=1,2,\dots,n, \\
& && \sum_{j=1}^{n} y_{kj}^t\lambda_{jq}^t \ge y_{kq}^t, && t=1,2,\dots,T,\, k=1,2,\dots,r,\, q=1,2,\dots,n, \\
& && \theta_q^1 = 1, && q=1,2,\dots,n, \\
& && \lambda_{jq}^t \ge 0, && t=1,2,\dots,T,\ j=1,2,\dots,n.\ q=1,2,\dots,n,
\end{aligned} \tag{3}$$

where the DMUs are described by the same set of outputs and one input in T consecutive time periods $t = 1, 2, \dots, T$, and assume that y^t_{kj}, $k = 1, 2, \dots, r$, $j = 1, 2, \dots, n$ are the values of the k-th output in the t-th period of the DMU$_j$. The only input in the period t is the efficiency score of the units in the period (t–1) and its values for the first period are set to 1 for all units. θ_q^t is aggregative efficiency score of the DMU$_q$ in the period t. It is clear that model (3) is not linear in the first set of constraints (moreover, the number of variables and constraints can be very high) and that is why it cannot be solved easily. To overcome this problem we propose in the next section of the paper an iterative approach for analysis of multi-period systems.

Faculty	2007	2008	2009	2010	2011	2012	2013	Avg
FSV UK	1.4722	**1.7058**	1.1984	1.6965	1.1766	1.5307	1.3887	1.4527
EkF JČU	0.6730	0.6825	**0.9610**	0.6937	0.7666	0.6818	0.6175	0.7252
FSE UJEP	0.7713	0.6978	0.7977	0.7678	0.9260	0.8206	**0.9849**	0.8237
ESF MU	0.6921	0.6466	0.9762	1.0216	**1.0444**	1.0315	0.9765	0.9127
OPF SU	0.6206	0.6824	0.8831	**0.9331**	0.9072	0.8495	0.8372	0.8162
FE ZČU	0.9104	0.8333	0.8441	0.9530	**1.0022**	0.9381	0.8750	0.9080
HF TUL	0.4978	0.5363	0.6122	0.6490	0.5923	**0.8037**	0.6573	0.6212
FES UP	0.8354	0.6271	**0.8623**	0.8270	0.7358	0.6568	0.7726	0.7596
FP VUT	**1.0290**	0.9581	0.8996	0.8446	1.0065	1.0142	0.9721	0.9606
EkF VŠB	**0.9919**	0.9457	0.8883	0.8779	0.9000	0.9682	0.8525	0.9178
FME Zlín	1.3748	1.4282	1.3081	1.3163	**1.6378**	1.0406	1.4069	1.3590
FFU VŠE	**1.1239**	1.0924	0.8858	0.8394	0.8629	0.8370	0.7877	0.9184
FMV VŠE	0.7464	0.6835	0.6572	0.6340	0.7071	**0.7236**	0.6744	0.6895
FPH VŠE	0.9634	**0.9881**	0.8743	0.8047	0.8138	0.7390	0.6931	0.8395
FIS VŠE	0.7533	0.7983	0.8435	0.8182	0.8443	**0.9640**	0.8495	0.8387
NH VŠE	**1.5045**	1.4318	1.0453	0.9920	0.9135	1.0390	0.9321	1.1226
FM VŠE	0.7225	0.7049	0.8124	0.8441	0.8253	**0.9756**	0.9482	0.8333
PEF ČZU	0.9038	1.1285	1.3193	1.3149	1.3638	**1.4432**	1.3190	1.2561
PEF MZLU	0.8037	0.8461	**0.9368**	0.7982	0.8450	0.9348	0.8331	0.8568

Table 1 Efficiency (super-efficiency) scores in all periods.

3 Resource allocation: a case of Czech economic faculties

In this section we will test the hypothesis that the multi-period DEA analysis can generate results usable for resource allocation of the set of homogenous faculties. For this purpose the set of 19 Czech public economic faculties was investigated. Each of the faculties is described by two inputs - number of academic staff and labor costs. Three most important performance characteristics as outputs are taken into account – the number of students, the number of graduated, and the number of RIV points generated by the faculty in particular years (RIV points is an aggregated characteristic that measures the quality of publication outputs of individuals and/or educational units in the Czech Republic). The data set is available since 2007 until 2013. Table 1 contains efficiency scores θ^{CCR} and super-efficiency scores θ^{AP} for all faculties and periods. The best (highest) values are bolded.

Faculty	**Eff. score 2007**	**Eff. score' share**	**Labor costs 2007**	**Labor costs' share**	**Labor costs 2008 Original**	**Formula**	**+/–**
FSV UK	1.4722	8.47	51060	6.32	54914	60910	5997
EkF JČU	0.6730	3.87	58897	7.29	25549	45981	20431
FSE UJEP	0.7713	4.44	22016	2.73	25807	29498	3691
ESF MU	0.6921	3.98	40213	4.98	46559	36904	-9655
OPF SU	0.6206	3.57	38816	4.81	43255	34497	-8758
FE ZČU	0.9104	5.24	27365	3.39	27717	35520	7804
HF TUL	0.4978	2.86	26237	3.25	29261	25172	-4088
FES UP	0.8354	4.80	35365	4.38	38465	37822	-643
FP VUT	1.0290	5.92	27267	3.38	31958	38279	6321
EkF VŠB	0.9919	5.70	74416	9.22	78471	61449	-17022
FME Zlín	1.3748	7.91	27350	3.39	27803	46512	18709
FFU VŠE	1.1239	6.46	39407	4.88	40101	46717	6616
FMV VŠE	0.7464	4.29	61266	7.59	64748	48927	-15821
FPH VŠE	0.9634	5.54	40289	4.99	44219	43368	-852
FIS VŠE	0.7533	4.33	40342	5.00	39740	38418	-1322
NH VŠE	1.5045	8.65	22089	2.74	24231	46899	22668
FM VŠE	0.7225	4.15	17714	2.19	18368	26147	7779
PEF ČZU	0.9038	5.20	110854	13.73	115863	77948	-37916
PEF MZLU	0.8037	4.62	46559	5.77	46720	42782	-3937
Total	**17.3902**	**100.00**	**807525**	**100.00**	**823750**	**823750**	**0**

Table 2 Calculation of labor costs (2008)

The results presented in Table 1 show that only two faculties are efficient in all periods (FSV UK and FME Zlín). Application of the PP multi-period model (orientation on the best period) leads to the following ranking: FSV UK (1.7058), FME Zlín (1.6378), NH VŠE (1.5045), …, FMV VŠE (0.7236). The multi-period model oriented on the worse period (see Jablonský (2016)) generates the same ranking in first three places but the worse faculty is changed in this case, i.e. FSV UK (1.1766), FME Zlín (1.0406), NH VŠE (1.0390), …, HF TUL (0.4978). The same ranking (first three and last place) is given by average efficiencies over all periods (the last column of Table 1). The efficiency (super-efficiency) scores in the period *t* express the level of relative performance of the faculties according to the given model and variables (inputs/outputs) taken into account but they just hardly can be used as indicators for allocation of resources in the future periods because the results do not consider interconnections among the periods. In order to try overcoming this drawback we propose the following simple iterative procedure:

- The set of outputs contains the same three variables as in the previous experiments, i.e. the number of students, the number of graduated and the number of RIV points.
- The set of inputs includes two variables – the number of employees and the labor costs but this second variable is given in a different way comparing to the previous model. In 2007 the values of this variable correspond to the reality. In the other periods it is the result of a simple formula. The amount of 50 % of the labor costs in the period *t* is calculated according to the labor costs' share in the previous period and the

remaining 50 % according to the efficiency scores' share in this period. The principle of this calculation for year 2008 is illustrated in Table 2. E.g. for the FSV UK in 2008 the labor costs according to the formula are calculated as follows: [(1.4722/17.3902)*823750 + (51060/807525)*823750]/2. The results for 2008 show that the highest increase of the funds appears for EkF JČU but it is obviously given by very high fall of real labor costs in 2008. In the contrary, the highest decrease is evident for PEF ČZU which is easily explainable by a very high difference between both shares (13.73 % labor costs' share and 5.20 % efficiency scores' share).

- The efficiency scores are calculated using model (2) and, if necessary, super-efficiency scores using the AP model derived from model (2). Either two mentioned inputs are considered or just labor costs derived using the presented way are taken into account. The results of allocation of resources among faculties according to the given approach are presented in Table 3. The first 8 columns of the table contain differences in allocation of funds given by the model with two inputs and original allocation by government methodology. The positive/negative values indicate that the allocation by the model is higher/lower than the reality. The last column of Table 3 contain the same information but given by the model with one input variable (labor costs) only. The results show that there are faculties overfunded in all periods (e.g. PEF ČZU and FMV VŠE). In the contrary some of the faculties are underfunded in all or almost all periods (e.g. FME Zlín, NH VŠE, and others).

Faculty	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2014a**
FSV UK	5997	17544	10269	20844	-7166	-6503	-17364	-16567
EkF JČU	20431	14085	15480	10396	7587	1016	-1228	1642
FSE UJEP	3691	4762	4864	8954	8491	13237	15948	10085
ESF MU	-9655	-12901	-7325	-1177	2112	996	5945	3112
OPF SU	-8758	-9311	-4260	-8153	-6975	-5020	1934	6253
FE ZČU	7804	12856	15178	13311	13740	16804	16230	9831
HF TUL	-4088	-4645	-1970	-960	-199	5278	1094	1860
FES UP	-643	-2569	3251	5887	-4	6462	6626	7111
FP VUT	6321	8784	10445	7175	6786	6426	5138	5013
EkF VŠB	-17022	-17904	-19730	-15926	-12632	-14866	-15413	-10018
FME Zlín	18709	20030	19368	15750	23395	18230	22479	17873
FFU VŠE	6616	8052	3143	-327	83	-487	-212	-118
FMV VŠE	-15821	-18543	-24991	-30903	-26917	-30064	-34109	-29677
FPH VŠE	-852	-1517	-2978	-6524	-11828	-14165	-8428	-4523
FIS VŠE	-1322	113	-1668	-4084	-1161	-1731	-5411	-2075
NH VŠE	22668	25877	21707	20693	17412	17566	18870	11498
FM VŠE	7779	13082	16868	17407	20597	20076	19965	10512
PEF ČZU	-37916	-55918	-55652	-49018	-34498	-28167	-27135	-22003
PEF MZLU	-3937	-1876	-2001	-3346	1177	-5088	-4926	194

Table 3 Differences in real funding and funding by the model

4 Conclusions

Evaluation of efficiency and performance of higher educational institutions is of a high importance and the results of this evaluation can serve as starting points for allocation of limited resources among these institutions. The commonly used practice in the Czech Republic is based on weighting of several criteria among which the most important is the number of students and research performance of the institution in the last five years measured by RIV points. The paper compares this methodology with DEA approach – a traditional CCR input oriented model and its super-efficiency modification with five (or four) variables is applied. The given results show significant differences in funding of the faculties under evaluation according to the official methodology and the DEA models. Of course the proposed models do not cover all criteria considered by the official methodology but the differences in funding can just partly be explained by this drawback.

A future research in this field can be focused on extension of the number of variables in the model and application of other DEA models. SBM models, e.g. Tone's efficiency and super-efficiency models (see Tone (2002)), could be an interesting alternative to the traditional DEA models.

Acknowledgements

The research is supported by the Czech Science Foundation, project no. 16-01821S.

References

[1] Andersen, P. and Petersen, N.C.: A procedure for ranking efficient units in data envelopment analysis. *Management Science* **39** (1993), 1261–1264.

[2] Beasley, J.E.: Determining Teaching and Research Efficiencies. *Journal of the Operational Research Society*, **46** (1995), 441–452.

[3] Charnes, A., Cooper, W.W. and Rhodes, E.: Measuring the efficiency of decision making units. *European Journal of Operational Research* **2** (1978), 429–444.

[4] Jablonský, J.: Efficiency analysis in multi-period systems: an application to performance evaluation in Czech higher education. *Central European Journal of Operations Research* **24** (2016), 283–296.

[5] Park, K.S. and Park, K.: Measurement of multiperiod aggregative efficiency. *European Journal of Operational Research* **193** (2009), 567–580.

[6] Thanassoulis, E., Kortelainen, M., Johnes, G. and Johnes, J.: Costs and efficiency of higher education institutions in England: a DEA Analysis. *Journal of the Operational Research Society* **62** (2011), 1282–1297.

[7] Tone, K.: A slacks-based measure of super-efficiency in data envelopment analysis. *European Journal of Operational Research* **143** (2002), 32-41.

Selected Problems of a Performance Analysis Approach for networked Production Structures and their quantitative Solutions

Hendrik Jähn[1]

Abstract. Within economic environment the production of goods within networked production structures plays an increasingly important role. For the sustainability of a cooperation of enterprises from an economic perspective it is important to manage and control value-adding processes successfully. To support that objective a comprehensive Performance Analysis - Approach(PAA) has been developed. For the application and implementation of that approach in many cases mathematical methods represent a key role. In this contribution a choice of three specific problems within the context of networked value-adding is selected. By the application of mathematical methods quantitative solutions to the present problems are derived, presented and discussed. The focus of the explanation is on multi criterial decision-making and decision support systems. Here, however, primarily quantitative management methods in an economic environment stand in the focus of research.

Keywords: Production Net, Performance Analysis, Evaluation, Sanctions.

JEL classification: M14
AMS classification: 91E45

1 Motivation

For the management of networked production structures incentive and sanction mechanisms play an important role. They serve as instruments for the management and controlling of economic relations. The determination of sanctions is primarily based on control and supervision systems. This is founded on the assumptions of the New Institutional Economics [9], e.g. asymmetrical distribution of information among the actors, bounded rationality, individual maximisation of utility and opportunistic behaviour. Within that context a comprehensive approach for the value-added process-related performance analysis for enterprises operating in cooperation structures was developed [3]. In section 2 that approach is introduced in brief. The framework bases on a quantitative approach. That fact leads to specific problems where mathematical methods are applied in order to find appropriate solutions. A selection of three problems is introduced in section 2. Solutions are presented and discussed consecutively.

The performance analysis approach is founded on a variety of performance parameters. The problem discussed in section 3 addresses the possibility for considering different significances of the performance parameters for the reason of an individualised analysis. Additionally, the evaluation of performances is a major challenge. Ratings represented by credit points are one approach possible. Usually there exists an interval for allowed credit points. This contribution introduces in section 4 a methodology to construct a relationship between the actual performance and credit points granted. Another challenge is the quantification of sanctions. They reduce the profit share of an enterprise in case of an unsatisfactory performance. For the quantification of sanctions a relationship between the aggregated performance figure of an enterprise and the extent of sanctions is required. That problem is discussed in section 5.

[1]Chemnitz University of Technology, 09107 Chemnitz, Germany, hendrik.jaehn@wirtschaft.tu-chemnitz.de

2 Model Framework and Problem Formulation

There are frameworks for performance measurement in networked environments to be found in literature. However most are developed for long term cooperations [1], [7], [8]. That fact proves the necessity for a short term focused perspective. The framework of the performance analysis approach is designed for the application in value-adding specifically configured productions networks. It is realised by the application of an adapted value benefit analysis [10]. The approach comprises the analysis, measurement and evaluation of the enterprise-related performance including the derivation of consequences. That procedure guarantees a high level of automation because it is based on quantified data. The modular structure allows the exchange of components and individual related instruments of application. The coarse structure of that approach shows a classification into value-adding process-neutral (strategic) phases and value-adding process-specific (operational) phases [3]. Within that framework the performance provided by a selected enterprise within a specific value-adding process is focused for analysis.

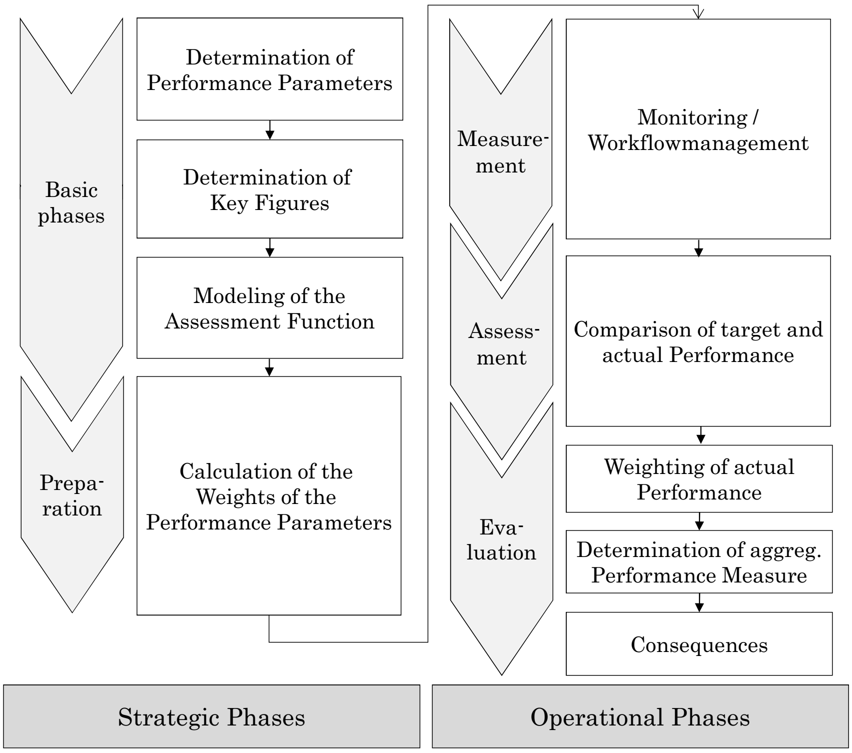


Figure 1 Strategic and operational Phases of the Analysis Approach

Figure 1 illustrates the exemplary sequence of the performance analysis approach. It is divided into stategic phases (left column, contains phases with tasks for a long term usage) and operational phases (right column, with value-adding process related- phases). These phases are divided into sub-phases which comprise of detailed process steps. An enterprise-specific aggregated performance figure represents the final result of the analysis. Based on that figure profit share affecting sanctions (income cuts) or bonuses (additional payments) are calculated. This allows the use of innovative steering mechanisms for the management and controlling of the network. It can be summarised as follows: "Performance analysis in the environment of networked production structures is the determination of the degree of target achievement in terms of the performance of an enterprise in a specific production process." Within the framework of that approach three problem fields are to be discussed in section 3 to 5: consideration of varying significances of performance parameters (preparation phase), determination of ratings (basic phase), and quantification of sanctions (evaluation phase).

3 Different significances of performance parameters

At first relevant performance parameters for the analysis must be identified. An adapted balanced scorecard-approach [5] serves as the basis. Out of four perspectives (finance, customer, internal business and enterprise) four hard-facts price, date of delivery, quality, response time and two soft-facts cooperation climate and confidence were identified as most relevant [3]. The information of soft-facts must be quantified first. For that purpose the Repertory Grid-methodology [4] is used. All performance parameters are represented by key figures. Based on that evaluation functions are constructed, see Figure 2.

Performance Parameter	Key Figure	Evaluation Function
Price	Discrepancy of price	Binary relation
Delivery date	Discrepancy of delivery date	Mathematical function
Response time	Required response time	Mathematical function
Quality	(Partial-) product quality	Binary relation / Value benefit a.
Cooperation	Quality of cooperation	Repertory Grid / Math. function
Confidence	Climate of confidence	Repertory Grid / Math. function

Figure 2 Fixing of key figures and evaluation function

In order to consider differing relevancies weightings for performance parameters are included. They can be determined by a method from decision theory [6]. Before beginning the determination of weightings a verification of differential and preferential independence of performance parameters is required. Each performance parameter gets a specific weighting. When using the trade-off-method for five (appropriate) key figures of performance parameters pairs of the specific trade-off rates must be determined. An appropriate forming results in five equations. Another equation which defines that the sum of weightings is one, completes the system of equations. This system consists of six equations with six unknown variables and hence is uniquely solvable.

Next an example is given. The trade-off-method is applied [2]. The order of key figures is at follows: discrepany of price (in money units), discrepancy of delivery date (in time units), required response time (in time units), product quality, cooperation climate, climate of confidence (in credit points, 10 points are the maximum achieveable). In that context, for example, subsequent indifference statements have been identified by decision makers:

$$
\begin{array}{lcl}
\{12,\ 0,\ *,\ *,\ *,\ *\} & \sim & \{0,\ 4{,}82,\ *,\ *,\ *,\ *\} \\
\{*,\ 0,\ 4{,}38,\ *,\ *,\ *\} & \sim & \{*,\ 1{,}03,\ 0,\ *,\ *,\ *\} \\
\{*,\ 0,\ *,\ 7,\ *,\ *\} & \sim & \{*,\ 2{,}00,\ *,\ 10{,}00,\ *,\ *\} \\
\{*,\ *,\ *,\ 10,\ 0,\ *\} & \sim & \{*,\ *,\ *,\ 7{,}00,\ 10,\ *\} \\
\{*,\ *,\ *,\ 10,\ *,\ 0\} & \sim & \{*,\ *,\ *,\ 7{,}85,\ *,\ 10\}
\end{array}
$$

To determine the target performance levels (credit point ratings) these values are inserted in utility functions of the parameters. The performance credit point ratings are determined based on the realised performance. These data serve as the basis for determining the weightings of the performance parameters. Following exemplary credit point ratings result and appropriate trade-off relations can be identified:

$$
\begin{array}{lcl}
\{0, 10, *, *, *, *\} & \sim & \{10, 5, *, *, *, *\} \\
\{*, 10, 6, *, *, *\} & \sim & \{*, 9, 10, *, *, *\} \\
\{*, 10, *, 8, *, *\} & \sim & \{*, 8, *, 10, *, *\} \\
\{*, *, *, 10, 0, *\} & \sim & \{*, *, *, 8, 10, *\} \\
\{*, *, *, 10, *, 0\} & \sim & \{*, *, *, 9, *, 10\}
\end{array}
$$

For each trade-off situation equations are formulated including weightings as unknown variables. For the first trade-off situation following equation arises:

$$w^P \cdot f_i^P(12) + w^L \cdot f_i^L(0) = w^P \cdot f_i^P(0) + w^L \cdot f_i^L(4,82) \tag{1}$$

Rearranging is as follows:

$$w^P = w^L \cdot \frac{f_i^L(4,82) - f_i^L(0)}{f_i^P(12) - f_i^P(0)} \tag{2}$$

At this point, corresponding credit point ratings are determined by using the evaluation function:

$$w^P = w^L \cdot \frac{5 - 10}{0 - 10} \tag{3}$$

Further shortening eventually leads to the first equation of the system of linear equations:

$$w^P = \frac{1}{2} \cdot w^L \tag{4}$$

All remaining equations are treated the same way so a system of linear equations results. Because six equations to calculate six weightings are necessary, one equation sets the value of one for the sum of all weightings. That system of equations then is dissolved, reulting six performance parameter-related weightings. Next, the consistency of weightings is checked. Finally the target performance as enterprise-related measure must be fixed in order to supply a criterion for comparing the realised performance of an enterprise with the targeted performance.

4 Determination of ratings

The operational phases especially include processes for measurement and evaluation of the performance. The measurement is realised by monitoring and data collection systems. The evaluation of the performance is based on a comparison of the already set target performance with the actual performance yielded. This is carried out for each performance parameter. An adapted form of the utility benefit analysis [10] has proved effective in that context. Here the performance provided is evaluated by granting credit points. For the formulation of a relationship between the performance provided and the scoring, the problem arises that credit point reviews are discrete, but the performance provided is continuously. An example for that problem is at follows: credit points are integers and usually within a range of 0 to 10. The quality of a product is good or medium or even very good, but it is scaled. That fact also applies to the price, which is never counted on whole Euro. This problem is solved by setting aside the condition for integrity. Then there exists a continuous distribution. Next a relationship between performance provided and credit points granted needs to be constructed. This is best achieved by means of a mathematical function. Many different forms of functions exist. For the determination of mathematical functions a variety of methods is available. These include, for example, the Lagrangian interpolation formula or Newton's algorithm or the algorithm of Neville-Aitken.

As an example, the evaluation function for the performance parameter "product quality" is to be determined. First, the shape of the function must be clarified. Primarily it is obvious that in case the quality does not reach a particular level (= target performance level 0) no credit points (target level = 0) will be awarded. It is also obvious that for the best possible quality (target performance level 10 or 100%) a maximum of credit points (target 10) is granted. In the simple case of a linear relationship an easy expressible linear function represents the evaluation function. In the present case, however, an S-shaped function is preferred. Although an average quality level (target performance level 5 or 50%) generates an average target value for credit points (5), however, no linearity is given in the lower and upper segment.

If the Lagrangian interpolation is applied in order to determine the evaluation function, supporting points are required. The outer supporting points are in that case (0,0) and (10,10). Now one point each to the left and right of center (5,5) must be determined. For right of the center for example (7,8) is suitable. With a product quality with a level of seven a target performance level of eight is assigned because the tolerances were still met in this case. The other value is represented by (3,2). For a product quality with a level of three only two credit points are granted because here already a strong violation of tolerance values is given. From the specified supporting points subsequently by means of Lagrangian interpolation an evaluation function can be determined. For the sample data, the following function results in order to calculate the competence cell-specific quality evaluation x_i^q:

$$f_i^Q(q_i) = x_i^q = \frac{1}{42}q_i^3 - \frac{5}{14}q_i^2 - \frac{4}{21}q_i \tag{5}$$

This utility function is appropriate as it assigns all values within the tolerance of seven to ten very high. Below the level of seven target performance levels are rated with a decreasing rate. Optionally, in some cases a modification is required. With a proper evaluation function, credit point ratings are calculated automatically very easily.

After completing the calculation of credits these figures are weighted according to their importance. Here the already determined weightings (see section 3) are used. Next, the performance parameter-related figures are aggregated to one total enterprise-related value. This is realised simply by adding all weighted credit points. Finally it has to be clarified what to do with the data.

5 Quantification of sanctions

Obviously, it appears useful to introduce sanction payments in case an enterprise has delivered an unsatisfactory performance. Based on the available figures it can be decided whether sanctions need to be paid by enterprises. Their extend must be calculated. As a precondition, a coherence between the total degree of fulfilment and possible sanctions must be developed.

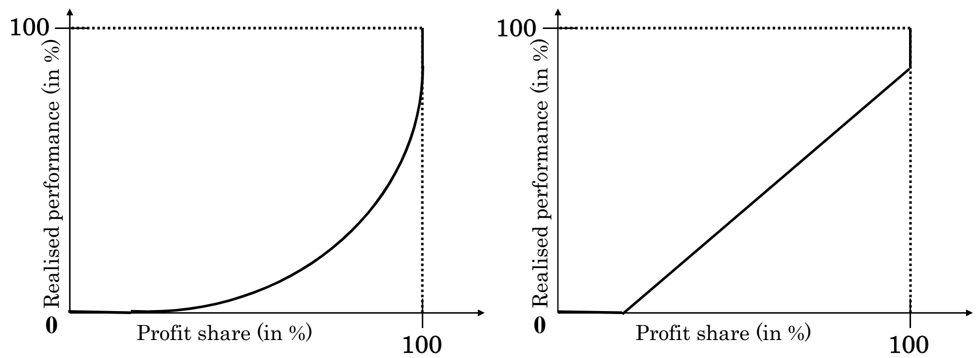


Figure 3 Example for sanction functions

Single credit point ratings are no integers but continuous values. Enterprise-specific aggregated credit point ratings are of this property, too. Also sanctions are not discrete. Thus the conditions are met to establish a relationship of the values in the form of a mathematical function. However, the major challenge is not to derive the mathematical function. This can be performed in analogy to section 4. However the shape of the function must be discussed. It is necessary to clarify whether a linear relationship is realistic or whether a different context is more descriptive for the situation. Thus, various constructs are also conceivable which are briefly discussed next.

In the presentation of functions, however, the profit share and not the sanction circumference is represented. Penalties diminish the profits. High sanctions for a poor performance must ensure a low profit share and vice versa. Consequently, functions for the relationship of profit share and realised performance must be constructed. Primarily, it is understandable that in case of a poor performance a large sanction amount reduces the profit share to a large scale. A good performance however ensures a higher profit share even by granting a bonus. Figure 3 takes that idea into account. Numerous further relationships are conceivable. By means of an appropriate procedure (see section 4) corresponding functions can be determined. At that stage all preconditions for the provision of enterprise-specific profit shares depending on the performance provided are fulfilled.

6 Conclusions

It has been proven that it is possible to develop an approach for the value-adding process-related performance analysis for enterprises operating in production networks that allows an almost completely automated procedure. That approach includes the value-adding process-related performance analysis of enterprises engaged in networked production structures. It closes a gap that wide-spread performance measurement systems leave behind. The performance analysis approach includes in addition to performance measurement also an evaluation procedure and as the result the quantification of consequences like sanctions or bonuses. Both soft-facts and hard facts are taken into consideration in order to receive an expressive result for the quantification of consequences. That model represents one example for the application of mathematical methods within an economical environment.

References

[1] Chalmeta, R., and Grangel, R.: Performance measurement systems for virtual enterprise integration, *Int. J. of Computer Integrated Manufacturing* **18** (2005), 1, 73–84.

[2] Eisenführ, F., Weber, M., and Langer, T.: *Rational Decision Making*. Springer, Berlin, 2010.

[3] Jähn, H.: Value-added process-related Performance Analysis of Enterprises acting in cooperative Production Structures, *Production Planning & Control* **20** (2009), 178–190.

[4] Jähn, H.: Integration of Soft factors in the Value-added Process-related Performance Analysis. In: *Proceedings of the 33rd International Conference Mathematical Methods in Economics (MME 2015)* (D. Martinčik, ed.), Cheb, 2015, 304–309.

[5] Kaplan, R.S., and Norton, D.P.: *The Balanced Scorecard: Translating Strategy into Action*. Harvard Business Review Press, Boston, 1996.

[6] Keeney, R.L., and Raiffa, H.: *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge University Press, Cambridge, 1993.

[7] Kulmala, H.I., and Lonnqvist, A.: Performance measurement of networks: Towards a non-financial approach, *Int. J. of Networking and Virtual Organisations* **3** (2006), 3, 178–190.

[8] Verdecho, M.-J., Alfaro-Saiz, J.-J., and Rodriguez-Rodriguez, R.: A Performance Management Framework for Managing Sustainable Collaborative Enterprise Networks. In: *Collaborative Systems for smart networked Environments* (L.M. Camarinha-Matos, and H. Afsarmanesh, eds.), 2014, 546–554.

[9] Williamson, O.E.: *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press, New York, 1975.

[10] Zangemeister, C.: *Nutzwertanalyse in der Systemtechnik*. 4th ed., Wittemannsche Buchhandlung, Munich, 1976.

Determinants of the shadow economy in the selected European Union member countries

Ondřej Jajkowicz[1]

Abstract. In this paper the influence of chosen, theory-based determinants of the shadow economy on the size of the shadow economy in EU member countries is tested. To test the effects of the selected determinants a panel data fixed effect regression model estimation was used. Problem of possible autocorrelation and heteroskedasticity, that are at the panel data fairly common, were resolved by model estimation using the "white period" estimation technique. Main theoretical assumptions refer to the fact, that development of variables capturing the tax and compulsory social security contributions burden of economic agents or over-regulation of official economy has a significant positive effect on the increase in the size of the shadow economy. The aim of this research is simply to confirm or refute a positive or negative effect of the above-mentioned and others selected determinants on the size of the shadow economy in the EU countries.

Keywords: Shadow Economy, Determinants of the Shadow Economy, Tax Burden, Regulatory Quality, Panel Regression.

JEL Classification: E26, O17, K42
AMS Classification: 62J05, 62M10

1 Introduction

The aim of this paper is to empirically verify the positive or negative effect of selected determinants on the size of the shadow economy in the countries of the Visegrad Group (V4). As determinants are meant determining parameters, factors or factors that are critical for desirable or undesirable activity (moving activities to the shadow economy) of man. The phenomenon of the shadow economy has been specified by economists almost 30 years ago. Period of exploration of the shadow economy has been accompanied by many attempts to find better and more precise definition, to find more accurate methods for capture and possibilities for accurate measurement of this phenomenon. Period of exploration of the shadow economy has been also accompanied by explanation of its implications for the official economy and the search for appropriate policies and measures for its elimination. The motivation to investigate the determinants of the shadow economy in these countries is the fact that this is still not a far explored issue. According to Schneider (2011) for the shadow economy can be considered all economic activities and incomes, whose main objective is to avoid government regulation, taxation or any capture. Already from this definition it is clear that the determinants that contribute to the existence, or increase in the size of the shadow economy are mainly excessive tax burden of economic agents and high level of regulation of the official economy. Besides these two major causes of existence and growth of the shadow economy size is possible to include other equally important determinants. For example, the proxies for the quality of the institutional environment, living conditions, and the quality and complexity of the legal and tax system. Despite the importance of the shadow economy a deeper analysis of this phenomenon in the V4 countries is missing. During the transformation of the centrally planned economies to market economies, these countries have undergone systematic changes accompanied by new economic, social and political relations in society. Individuals and society faced to a major challenge of adapting themselves to the new economic, political, legal and social environment, which in many cases meant respecting the new strict rules and restrictions imposed by the government. It is possible to assume that all of the facts mentioned above, including the development of private enterprise, gradual rise in rate of unemployment and the introduction of a standard market economy tax system influenced the development of the shadow economy in these countries.

2 Literature background

Based on the above facts are as the main cause for the existence of the shadow economy considered regulatory state intervention in the markets and its strong role in influencing the economy. Within economies operating on a market basis (EU member states) can be for a destabilizing element of the allocation process, considered the tax system. Especially the excessive tax burden of economic agents is a major problem in the case of the tax system. Johnson, Kaufmann and Shleifer (1997) in their study concluded, that countries with more regulation of official economy, higher tax burden and more corruption tend to have a higher size of the shadow economy. Johnson,

[1] Vysoká škola báňská – Technická univerzita Ostrava, Ekonomická fakulta, ondrej.jajkowicz@vsb.cz

Kaufmann a Zoido-Lobatón (1998) due to their findings states, that extent of regulatory discretion together with high levels of bureaucracy and weak rule of law are key driving forces to move activities to shadow economy. Also, the study of Schneider (2006) or Startiené and Trimonis (2010) confirm that the tax burden is the factor with the strongest influence on the existence and growth of the shadow economy. Marinov (2008) concludes that social and economic reasons which force economic agents to move their activities into the shadow economy are affected by government policy in the area of tax and regulatory measures. A correspondingly Schneider (2011) ranks grow of the tax burden and the level of compulsorily paid social security contributions among the most important driving forces of shadow economy existence. Another important driving force of existence of the shadow economy is the high level of the compulsorily paid social security contributions. Schneider (2013) states, that in the case of compulsorily paid social security contributions, it is necessary to realize that the shadow economy is made up of so-called "undeclared work". And in countries with high level of compulsory social contributions the "undeclared work" can be considered as a major problem. The greater the difference between total cost of labor in the official economy and the after-tax earnings from work, the greater is the incentive to work in the shadow economy. Development of the shadow economy is also dependent on the economic cycle, when in times of recession a tendency to shift part of the activities in the shadow economy is growing. In the case of an economic downturn occurs increased competition and for many companies in order to maintain their market position there is no other option than to shift from official to shadow economy activities (Fassmann, 2007). Very important factor affecting the existence of the shadow economy are the quality of the institutional environment and the quality of public services. Schneider (2013), in assessing the quality of the institutional environment as the driving force of the shadow economy came to the conclusion that it is important to undertake incentive oriented policy measures in order to make work in the shadow economy less attractive, to have policy institutions which work efficiently and as a constraint for selfish politicians and bureaucracy (to limit the role of the state apparatus) and to fight corruption with good governance and strict punishment. Within the quality of the institutional environment can be, in addition, according to Friedman, Johnson, Kaufman and Zoid-Lobatón (2000) and Mara (2011) among the factors affecting the existence of the shadow economy the corruption in the government sector included. Dreher and Schneider (2006) using cross-section analysis of 120 countries and panel data analysis of 70 countries found a strong relationship between corruption and shadow economy. Robust relationship between level of corruption and the shadow economy confirmed again Dreher and Schneider (2009), but this relationship depends on which indices of corruption were used in empirical analysis. Buehn and Schneider (2011) using a structural equation analyzed a sample of 51 countries around the world over the period 2000 to 2005 and they demonstrated positive relationship of corruption and the shadow economy size. According to Togler, Schneider and Schalteggara (2009), we cannot forgetting also the degree of urbanization, which reinforces the anonymity of the economic entities and thereby contribute to the development of the shadow economy activities. Conducted research of the empirical literature provides sufficient factual basis for the selection of the explanatory variables representing a major driving forces of the shadow economy in the empirical section of this article.

3 Model creation

In the empirical part of this paper a panel data multiple linear regression model is applied. This model is designed to explain factors influencing the size of the shadow economy in V4 group member countries. Panel data regression model ensure exploration of the relationship between selected explanatory variables including the proxies for tax burdens, quality of institutional environment, social security contribution burdens, regulation of official economy, control variables and so on and the size of the shadow economy in Czech Republic, Hungary, Poland and Slovakia. The dataset of the dependent variable for the estimation of the empirical model is composed of annual time series from 1999 to 2015. The model is based on the classical linear regression model with several explanatory variables. Using panel data, the classical linear multiple regression model is extended by a second dimension. The formal form of the panel data model with fixed effects is following:

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \beta_2 x_{2it} + \cdots + \beta_k x_{kit} + \mu_i + \varepsilon_{it}, \qquad (1)$$

Where y_{it} is tested dependent (explained) macroeconomic variable (The size of the shadow economy collected from Schneider (2011, 2013, 2015) estimates), x_{it} where $(i = 1,2, \dots, N)$; $(t = 1,2, \dots, T)$ are the independent (explanatory) variables. Namely the following independent variables:

- **Tax and social security contribution burdens**
1. The share of direct taxes in GDP (DT; %);
2. The share of indirect taxes in GDP (IT; %);
3. The share of social security contributions in GDP (SC, %).
- **Regulation of official economy**
1. The share of general government expenditures in GDP (GE; %).

- **Quality of institutional environment**

1. Quality of regulation (RQ);
2. Control of corruption (CC).

- **Control variables**

1. Urbanization rate, the share of urban population in total population in % (URB);
2. Unemployment rate (U; %).

To capture the tax burden on economic agents an indicators of the share of direct and indirect taxes in GDP were used. Out of taxes the economic entities are also burdened by compulsory payments of social security contributions. For this reason, an indicator of the share of compulsorily paid social security contributions in GDP is selected. The degree of regulation of the official economy is represented by the share of government spending relative to GDP. This variable in some way reflects on the state's participation in the official economy. Within the regulation of the official economy also an indicator of quality of regulation is selected. The quality of the institutional environment is in the model captured using two variables. Namely, it is the control of corruption indicator and quality of regulation indicator mentioned above. Among the independent variables belong also two control variables. The first of these is the urbanization rate and the second is level of unemployment. Time series of the share of direct taxes, indirect taxes and social contributions in GDP were drawn from the European Commission database relating to taxation in the EU (European Commission, 2015). Time series of indicators of the quality of regulation and control of corruption were drawn from the World Bank indicators database relating to the quality of public administration (i.e. Worldwide Governance Indicators); (The World Bank, 2015). These indicators, that are processed by the World Bank ranges from -2.5 (weak) to 2.5 (strong). Quality of regulation includes the ability and sensitivity of the government to implement appropriate policies and regulations to help develop the private sector. Control of corruption indicator shows the perception of the extent to which public power is exercised for private gain. Indicator takes into account both small and large forms of corruption, as well as "capture" of the state by elites and private interests. Time series of unemployment rate has been obtained from the Eurostat databases (Eurostat, 2015). The urbanization rate statistics from the World Bank are measured as a percentage of urban population to total population (The World Bank, 2015).

Symbol (μ_i), where ($i = 1,2, \dots, N$) denotes the effect of country-specific, which is constant in time and the symbol (ε_{it}), where ($i = 1,2, \dots, N$); ($t = 1,2, \dots, T$) then denotes error term with characteristics that are needed for proper parameter estimation. The final estimation of the regression model is performed by white period estimation technique. Using a panel data model is advantageous because it allows obtaining better estimates of the parameters. In particular, using them can be identified effects that would not be detectable only from time series or materially spatial data. Panel data allows us to eliminate the effect of hidden heterogeneity, if this effect is constant over time and significantly reduce the problems caused by the omission of some important variables that are difficult to measure, for example. This type of data provides a statistical advantage, because reality can be modeled with more data, and can solve the problem of collinearity between variables. **„**Panel data sets provide a rich environment for researchers to investigate issues which could not be studied in either cross-sectional or time series settings alone“ (Seddighi, 2012, s. 256)**.** The main importance of this type of data is that it provides a methodology for treating the problem of omitted variables inherent in cross-sectional data analysis.
The fixed effects model was chosen on the basis of Hausmann´s test performance. The main reason for fixed effects model selection is fact that we are in our regression analysis working with a larger number of countries. If we are using data for heterogeneous countries then the basic model does not distinguish between the various countries in analysis. In other words, by combining different countries we deny the heterogeneity that may exist among these countries. In other words each country is individual. Possible problem of heterogeneity resolves selected fixed effect model.

3.1 Model estimation and results

To avoid spurious regression time series were individually tested for the existence of a unit root. For this purpose a unit root tests especially designed for panels were used. Namely tests with common unit root process like Levin, Lin and Chu Test (LLC) that suggest a more powerful panel unit root test than performing individual unit root tests for each cross-section. The null hypothesis is that each individual time series contains a unit root against the alternative that each time series is stationary. On the other hand also Tests with individual unit root processes were used. One of this panel unit root tests is Im, Pesaran and Shin Test (IPS). Another test is for example combining p-value test, concretely Fisher-type tests proposed by Maddala and Wu (1999) and Choi (2001). For this paper, the above-described tests are sufficient. If the time series show a trend, it is appropriate to make them stationary. Non-stationary time series can be converted into stationary, for example through the differentiation. The time series stationarity testing pointed to non-stationarity of time series. Inserting these time series into the regression model would lead only to estimate the so-called “Spurious” regression. Identified non-stationarity was removed by logarithm and differences to the appropriate order. The fixed effect model was estimated via econometric program E-views version 7. Autocorrelation problem and heteroskedasticity problem, that are at the panel data fairly common, were resolved using the "white period" estimation technique. Changes in the variables representing the tax burden on businesses will not reflect in their behavior immediately. We can assume that in the real economy reactions of economic agents to change these variables comes with a lag. In order to know how late responding businesses to increase in the tax burden by moving their activities into the shadow economy, these variables would be delayed by several years. Another reason for the delay of these variables is the existence of a standard delay resulting from the formation of economic policy (for example, in approving the introduction of new taxes, changes in tax rates, etc.).

Independent variable	**Dependent variable**
Constant	-0.286628(-2.910501) ***
LOGRQ - Quality of regulation	-0.111560(-1.155549)
DLOGCC - Control of corruption	-0.085588(-1.819768)*
LOGU - Unemployment rate	0.063219(5.293310) ***
LOGGE - The share of general government expenditures in GDP	0.095435(2.840736) ***
LOGDT(-2) - The share of direct taxes in GDP	0.025772(6.736861) ***
DLOGIT (-2) - The share of indirect taxes in GDP	0.037442(1.487089)
DLOGSC(-2) - The share of social security contributions in GDP	0.051745(1.956043) *
DURB - Urbanization rate, the share of urban population in total population	6.856204(1.408025)
R^2	0.657370
Durbin-Watson stat.	2.129042
F-statistic	4.883717
Prob (F-stat.)	0.000332
The statistical significance on 1% (***), 5% (**) and 10% (*) level of significance.	

Table 1 Model estimation results
Source: Own estimate through the program E-views (version 7).
Note: In parentheses are the corresponding t-statistics.

Results of estimated model are shown in Table 1. Undelayed variables representing the tax burden and a same variables delayed by one period, were statistically insignificant. For this reason, they were omitted in the final model. Firstly, the influence of tax burden variables was tested with a one-year delay, but the results of this analysis were statistically insignificant. Statistically significant results of proxies for level of taxation in official economy, namely share of direct taxes in GDP, share of indirect taxes in GDP and share of social security contributions in GDP has been reached only in the case of a two periods (years) delay. Two periods (years in this case) delay in the development in tax burden of official economy can be considered sufficient for economic agents to adapt on changes in the tax burden and to make decisions about moving activities to shadow economy.

4 Conclusion

The results indicate a fairly close relationship (with respect to the use of panel data) between the size of the shadow economy and selected determinants. The relationship is represented by the coefficient of determination (R2 = 0.66). Almost 66% of the variation in the size of the shadow economy can be explained by the explanatory variables used. The findings confirmed theoretical predictions about selected determinants of the shadow economy in the V4 countries. Statistically significant are two of three variables expressing the tax burden on economic entities. So if the share of direct taxes in GDP (DT) is growing, the tax burden increases and ultimately the size of the shadow economy (share of GDP) grows. Businesses are trying to avoid taxation and shifting part of their activities into the shadow sector.

Another important variable determining the size of the shadow economy is variable representing share of the social security contributions in GDP (SC). Here it is necessary to realize again that the shadow economy is from the big part made up of so-called "Moonlighting" or “Illicit work”. And in countries with high levied compulsory social contributions may be “Illicit work” known as a major problem. In connection with this problem, we can mention for example the case of the Czech Republic, where the amount of social security contributions in the long term moves above the EU average. If these contributions levied too high, employers lose interest to recruit employees officially. Instead of official recruitment with the aim of avoiding such payments they finally recruit workers "without the proper registration". One percent increase in the size of direct taxes share in GDP and social security contributions share in GDP lead to grow of the change of the size of the shadow economy increase by around 0.026 and 0.052 percentage points. The only exception in the context of fiscal variables is a variable representing the share of indirect taxes in GDP (IT), which is in relation to the size of the shadow economy statistically insignificant. It was also found that the increase in unemployment (U) contribute to growth in the size of the shadow economy. One percent increase in the size of the unemployment rate leads due to results to an increase of the size of the shadow economy by around 0.063 percentage points. People who lose their jobs in the official economy and are not able to find the job in the official economy again move their activities into the shadow economy to compensate their lost income (wage). A statistically significant relationship was also demonstrated in the case of official economy regulation level represented as a share of government expenditure in GDP (GE). Due to the estimation results a one percent increase in the size of the government expenditures (proxy for level of regulation in official economy) leads to grow of the change of the size of the shadow economy increase by around 0.095 percentage points. Over-regulation of the official economy of the state, or excessive state's participation in the official economy contributes to growth in the size of the shadow economy in the countries surveyed. Without the statistically significant effect on the size of the shadow economy, according to the findings is indicator of regulatory quality (RQ) and the proxy for rate of urbanization (URB). But it is possible to assume that with a certain time delay of these variables we would achieve the statistically significant results. This assumption must be taken into account for further research of this issue. The last variable that affects the size of the shadow economy is an indicator of control of corruption (CC). Thus, if the perceived level of corruption by public is low (indicator is approaching its maximum 2.5) the size of the shadow economy decreases, and vice versa.

Acknowledgements

This paper was financially supported within the VŠB - Technical University SGS grant project No. SP2016/101 “Monetary, Fiscal and Institutional Aspects of Economic Policy in Selected Countries “

References

[1] Buehn, A., and F. Schneider. Corruption and the shadow economy: like oil and vinegar, like water and fire? International Tax and Public Finance, Volume 19, p. 172-19, 2011. ISSN 0927-5940

[2] Choi, In. Unit root tests for panel data. Journal of International Money and Finance 20, 2001,249–272.

[3] Dreher, A., and F. Schneider. Corruption and the shadow economy: an empirical analysis. CREMA, Center for Research in Economics, Management and the Arts.Working Paper No. 2006 – 01.

[4] Dreher, A., and F. Schneider. Corruption and the shadow economy: an empirical analysis. Public Choice, 2009, 144: p. 215–238.

[5] European Comission. Databases and indicators. [online database]. Washington, DC, [cit. 12.04.2016]. Available at: <http://ec.europa.eu/economy_finance/db_indicators/index_en.htm>

[6] Eurostat. General and regional statistics. [online database]. Washington, DC, [cit. 12.04.2016]. Available at: < http://ec.europa.eu/eurostat/data/database>

[7] Fassmann, M. Stínová ekonomika a práce na černo. Praha: SONDY, 2007, ISBN: 978-80-86846-21-7

[8] Friedman E., S. Johnson, D. Kaufman and P. Zoido-Lobatón. Dodging the grabbing hand: the determinants of unofficial activity in 69 countries. Journal of Public Economics, 2007, 76: 459-493. ISSN: 0047-2727.

[9] Johnson, S., D. Kaufmann and A. Shleifer. The Unofficial Economy in Transition. Brookings Papers Econ. Act, 1997.

[10] Johnson, S., D. Kaufman and P. Zoido-Lobatón. Regulatory Discretion and the Unofficial Economy. American Economic Review. Papers and Proceedings 88 (2), 1998, 387-392.

[11] Maddala, G. S., and S. Wu. A comparative study of unit root tests with panel data and a new simple test. Oxford bulletin of economics and statistics, special issue. 0305-9049, 1999.

[12] Mara, R. E. Causes and consequences of underground economy. MPRA Paper NO. 36438, 2011.

[13] Marinov, A. Hidden economy in the rural regions of Bulgaria. International Review on Public and Non - profit Marketing, 2008, 5: 71-80. ISSN: 1865-1992

[14] Schneider, F. Shadow economies and corruption all over the world: what do we really know? The Institute for the Study of Labour. Discussion Paper No. 2315, 2006.

[15] Schneider, F. The Shadow Economy in Europe, 2011. Johannes Kepler University of Linz, ATKEARNEY, VISA, 2011.

[16] Schneider, F. Size and Development of the Shadow Economy of 31 European and 5 other OECD Countries from 2003 to 2012: Some New Facts. Johannes Kepler University of Linz, 2011, [online]. [cit. 12.4.2016]. Available at: <http://www.economics.uni-linz.ac.at/members/Schneider/files/publications/2012/ShadEcEurope31.pdf>
[17] Schneider, F. The Shadow Economy in Europe, 2013. Johannes Kepler University of Linz. AT KEARNEY, VISA, 2013.
[18] Schneider, F. Size and Development of the Shadow Economy of 31 European and 5 other OECD Countries from 2003 to 2015: Different Developments. Department of Economics, Johannes Kepler University of Linz, 2015.
[19] Seddighi, H. R. Introductory Econometrics: A practical approach. Routledge, Milton Parg, Abingon, Oxon OX14 4RN, 2012, ISBN: 978-0-415-56687-2
[20] Startiené, G., and K. Trimonis. Causes and consequences of non-observed Economy. Economic and Management. (15), 2010, 275-80.
[21] Torgler, B., F. Schneider and Ch. A. Schaltegger. Local autonomy, tax morale, and the shadow economy. Public Choice, 2009, 144: 293–321. ISSN: 0048-5829.
[22] Worldbank. Worldwide Governance Indicators. [online database]. Washington, DC, [cit. 12.04.2016]. Available at: <http://data.worldbank.org/data-catalog/worldwide-governance-indicators>
[23] Worldbank. Indicators – Urban population. [online database]. Washington, DC, [cit. 12.04.2014]. Available at: <http://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS>

Designing a Robust Emergency Service System by Lagrangean Relaxation

Jaroslav Janáček,[1], Marek Kvet[2]

Abstract. Emergency service system design locates limited number of service centers at positions from a given set of possible locations to satisfy system users' demands for service. We study here the case of simple disutility perceived by an average user. The disutility is assumed to be proportional to the distance of a user location from the nearest service center and then, sum of distances from particular system users to the nearest located service center is minimized. A robust service system design is usually performed so that the design complies with specified scenarios by minimizing the maximal objective function of the individual instances corresponding with the particular scenarios. The min-max link-up constraints represent an undesirable burden in any integer programming problem due to bad convergence of branch-and-bound method. Within this paper, we try to overcome the drawback following from the link-up constraints by usage of the Lagrangean relaxation and the sub-gradient method. We provide the reader with a comparison of the original min-max approach to the suggested approach based on the Lagrangean relaxation of the troublesome constraints.

Keywords: emergency service system design, Lagrangean relaxation, radial formulation, robust design.

JEL Classification: C61
AMS Classification: 90C06, 90C10, 90C27

1 Introduction

When an emergency service system is designed, the designer must take into account that traversing time between service center and an affected user might be impacted by various random events following weather or traffic, the system designer must face the demand for system resistance to such critical events [2], [11], [13]. Most of the approaches to increasing the system resistance are based on making its design resistant to possible failure scenarios, which can appear in the road network as a consequence of random failures due to congestion, disruptions or blockages. An individual scenario is characterized by particular time distances between the users' locations and possible service center locations. A robust service system design has to comply with all the specified scenarios. The usual way of taking into account all scenarios is based on minimizing the maximal objective function of the individual instances corresponding with the particular scenarios. The min-max link-up constraints represent an undesirable burden in any integer programming problem due to bad convergence of branch-and-bound method, which is prevailing solving tool inside most available IP-solvers. Thus these approaches to the robustness constitute a big challenge to family of operational researchers and professionals in informatics.

An emergency service system design for a given road network in serviced area locates limited number of service centers at positions from a given set of possible locations to satisfy future system users' demands for service in case of emergency. Different objectives can be applied on the design. In this paper, the perceived disutility is assumed to be proportional to the distance of a user location from the nearest located service center and then, sum of distances from particular system users to the nearest located service center is minimized [3], [8], [9], [10].

Within this paper, we focus on the min-sum emergency service system design, which is robust considering given finite set of scenarios. Complexity of location problems with limited number of facilities to be deployed and necessity to solve large instances of the problem led to searching for a suitable algorithm. It was found that in contrast to original location-allocation formulation, the radial formulation of the problem can considerably ac-

[1] Prof. RNDr. Jaroslav Janáček, CSc., University of Žilina, Faculty of Management Science and Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, e-mail: jaroslav.janacek@fri.uniza.sk
[2] Ing. Marek Kvet, PhD., University of Žilina, University Science Park, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, marek.kvet@uvp.uniza.sk

celerate the associated solving process [1], [4], [5]. Simultaneously, an attention was paid to the radial formulation with homogenous system of radii [6], [7]. The later form of the radial formulation is used in this paper. Within this paper, we focus on ways to comply with bad convergence of the branch-and-bound method applied on the model with the link-up constraints linking the individual scenario objective functions up to their common upper bound. We present the original approach [12], [13] to the robust design using the radial formulation and compare it to the approach using the Lagrangean relaxation of the link-up constraints completed by sub-gradient method of Lagrangean multiplier improvement.
The remainder of the paper is organized as follows: Section 2 is devoted to the description of original robust design of the emergence system with min-sum objective including the radial formulation. The Lagrangean relaxation and Lagrangean multiplier adjustment is described in Section 3 and the associated numerical experiments are performed in Section 4. The results and findings are summarized in Section 5.

2 Radial formulation of the robust emergency system design

The robust emergency system design problem with radial formulation can be described using the following denotations. Let symbol J denote the set of users' locations and let symbol I denote the set of possible service center locations. We denote by b_j the number of users, which share the location j. To solve the problem, at most p locations must be chosen from I so that the maximal disutility perceived by the worst situated user be minimum. The value of user's disutility is given by the mutual positions of the user location and the location of the service center providing him with service. Let symbol U denote the set of possible failure scenarios. A particular scenario may correspond with some situation in underlying transportation network, e.g. weather, congestions and other critical events influencing travelling times. We assume that user's disutility grows with increasing distance between the user and the service center. Disutility following from the distance between locations i and j under a specific scenario $u \in U$ is denoted here as d_{iju}. The values of d_{iju} may be proportional to the network distances between the users' location j and the center location i under scenario u. Especially, the disutility perceived by a user of emergency system may correspond with estimated travelling time in minutes.The decisions, which determine the designed public service system, can be modeled by further introduced decision variables. The variable $y_i \in \{0,1\}$ models the decision on service center location at place $i \in I$. The variable takes the value of 1 if a service center is located at i and it takes the value of 0 otherwise.
In the robust problem formulation, the variable h is used as the upper bound of the objective function issues for individual scenarios. To formulate the radial model, the range $[d_0, d_m]$ of all possible disutility values $d_0 < d_1 < \ldots < d_m$ from the matrix $\{d_{iju}\}$ is partitioned into $v+1$ zones according to [6], [8]. The zones are separated by a finite ascending sequence of so called *dividing points* $D_1, D_2 \ldots D_v$ chosen from the sequence $d_0 < d_1 < \ldots < d_m$, where $0 = d_0 = D_0 < D_1$ and also $D_v < D_{v+1} = d_m$. The zone s corresponds to the interval $(D_s, D_{s+1}]$. The length of the s-th interval is denoted by e_s for $s = 0 \ldots v$. Further, auxiliary zero-one variables x_{jus} for $s = 0 \ldots v$ and $u \in U$ are introduced. The variable x_{jus} takes the value of 1, if the disutility of the user at $j \in J$ under scenario $u \in U$ from the nearest located center is greater than D_s and it takes the value of 0 otherwise. Then the expression $e_0 x_{ju0} + e_1 x_{ju1} + e_2 x_{ju2} + \ldots + e_v x_{juv}$ constitutes an upper approximation of the disutility d_{ju*} from user location j to the nearest located service center under scenario $u \in U$. If the disutility d_{ju*} belongs to the interval $(D_s, D_{s+1}]$, then the value of D_{s+1} is the upper estimation of d_{ju*} with the maximal possible deviation e_s. Let us introduce a zero-one constant a_{iju}^s under scenario $u \in U$ for each triple $[i, j, s]$, where $i \in I$, $j \in J$, $s \in [0..v]$. The constant a_{iju}^s is equal to 1, if the disutility d_{iju} between the user location j and the possible center location i is less or equal to D_s, otherwise a_{iju}^s is equal to 0. Then the radial-type min-sum robust emergency service system design problem can be formulated as follows:

$$\textit{Minimize} \quad h \tag{1}$$

$$\textit{Subject to:} \quad x_{jsu} + \sum_{i \in I} a_{iju}^s y_i \geq 1 \quad \textit{for } j \in J, \quad s = 0, 1, \ldots, v, \; u \in U \tag{2}$$

$$\sum_{i \in I} y_i \leq p \tag{3}$$

$$\sum_{j \in J} b_j \sum_{s=0}^{v} e_s x_{jus} \leq h \quad \textit{for } u \in U \tag{4}$$

$$y_i \in \{0, 1\} \quad \textit{for } i \in I \tag{5}$$

$$x_{jus} \geq 0 \quad \textit{for } j \in J, \; u \in U, \; s = 0, 1, \ldots, v \tag{6}$$

$$h \geq 0 \tag{7}$$

In this model, the objective function (1) represented by single variable h gives the upper bound of the all objective function values over the individual scenarios. The constraints (2) ensure that the variables x_{jus} are allowed to take the value of 0, if at least one center is located in radius D_s from the user location j and constraint (3) limits the number of located service centers by p. The link-up constraints (4) ensure that each perceived disutility is less than or equal to the upper bound h. As concerns the obligatory constraints (6), only values zero and one are expected in a feasible solution, but it can be seen that the model has integrality property regarding the variables x_{jus}. It can be noticed that in the optimization process all relevant values of x_{jus} are "pushed down" and the constraints (2) and (6) bound the variable x_{jus} from below by value of one or zero. It follows that the relevant values of x_{jus} stay at one of these values.

3 Lagrangean relaxation and robust emergency system design

As the min-max link-up constraints (4) represent an undesirable burden in any integer programming problem due to bad convergence of branch-and-bound method, we explore the possibility provided by Lagrangean relaxation to overcome this burden. The Lagrangean relaxation is applied on the constraints (4) only. Each of these constraints is associated with a nonnegative Lagrangean multiplier λ_u and the relaxed problem is formulated as follows.

$$Minimize\ \ h+\sum_{u\in U}\lambda_u(\sum_{j\in J}\sum_{s=0}^{v}b_je_sx_{jsu}-h)=h(1-\sum_{u\in U}\lambda_u)+\sum_{u\in U}\lambda_u\sum_{j\in J}\sum_{s=0}^{v}b_je_sx_{jsu} \tag{8}$$

Subject to (2), (3), (5), (6) and (7).

The problem (8), (2), (3), (5), (6) and (7) obviously has solution only for such setting of Lagrangean multipliers, where their sum is less than or equal to one. Having optimal solution of the relaxed problem with arbitrary multipliers meeting the stated rule, the value of optimal solution yields lower bound of the optimal solution of the original problem (1)-(7).
If the sum of Lagrangean multipliers is less than one, then the optimal value of h would equal to zero and it does not represents the upper bound at all. That is why we restrict ourselves on such setting of multipliers, where the sum equals to one exactly. In the next algorithm, we use the well-known proposition, which claims that the optimal solution of (8), (2), (3), (5), (6) and (7) is also optimal solution of (1)-(7) if and only if the resulting value of h meets each of constraint from the constraint set (4) and, in addition, if is satisfies the following complementary rules (9).

$$\lambda_u(\sum_{j\in J}\sum_{s=0}^{v}b_je_sx_{jsu}-h)=0\quad for\ u\in U \tag{9}$$

We try to reach or at least to approximate the optimal solution of the original problem (1)-(7) by an iterative algorithm, where the Lagrangean multipliers are readjusted at each iteration based on the optimal solution of (8), (2), (3), (5) and (6). The algorithm performs according to the following steps with the accuracy ε.

Step 0. Initialize $\lambda_u = 1/|U|$ for $u\in U$ and initialize the upper and lower bounds by LB=0 and UB=$+\infty$ respectively.

Step 1. Solve the following problem:

$$Minimize\ \ L=\sum_{u\in U}\lambda_u\sum_{j\in J}\sum_{s=0}^{v}b_je_sx_{jsu}$$

$$Subject\ \ to\ \ (2), (3), (5)\ and\ (6)$$

Perform Step 2 for the resulting L, $\boldsymbol{x}$ and $\boldsymbol{y}$.

Step 2. Define $H=\max\{\sum_{j\in J}\sum_{s=0}^{v}b_je_sx_{jsu}:u\in U\}$

If $UB>H$, then update the best found solution $UB=H$ and $\boldsymbol{y}^{best}=\boldsymbol{y}$.
If $LB<L$, then update the lower bound $LB=L$ and

$$\text{set}\quad g_u=\sum_{j\in J}\sum_{s=0}^{v}b_je_sx_{jsu}-H\quad for\ u\in U\ .$$

$$\text{set}\quad \alpha=0.9*\min\{(UB-LB)/\sum_{u\in U}(g_u)^2,\ \min\{-\lambda_u/g_u:u\in U,\ g_u<0\}\}$$

else set $\alpha=\alpha/2$.

Step 3. If UB-$LB<\varepsilon$, then terminate, otherwise

$$\text{set}\quad nar=\sum_{u\in U}(\lambda_u+\alpha g_u)\ ,$$

update $\lambda_u = (\lambda_u + \alpha g_u)/nar \quad for\, u \in U$ and go to Step 1.

4 Computational study

Within this section, we present the results of numerical experiments, which are aimed at ascertainment of the characteristics of the suggested approach from the viewpoint of computational time and the solution accuracy.

To compare the basic approach using the radial formulation (1) - (7) with the Lagrangean relaxation applied on the model (8), (2), (3), (5), (6) and (7), we performed the series of numerical experiments. To solve the problems described in previous sections, the optimization software FICO Xpress 7.9 (64-bit, release 2015) was used and the experiments were run on a PC equipped with the Intel® Core™ i7 5500U processor with the parameters: 2.4 GHz and 16 GB RAM.

The used benchmarks were derived from the real emergency health care system, which was originally implemented for the self-governing region of Žilina. Here, all cities and villages with corresponding number b_j of inhabitants were taken as the set of possible service center locations *I* and the set of system users *J* as well. The coefficients b_j were rounded to hundreds. The cardinality of *I* was 315. This sub-system covers the demands of all communities - towns and villages spread over the particular region by given number of ambulance vehicles. To enrich the pool of benchmarks, the parameter *p* limiting the number of located service centers varied so that the ratio of the cardinality of the set *I* to the value of *p* equals to 2, 3, 4, 5, 10, 15 and 20 respectively. The network distance from a user to the nearest located service center was taken as the user´s disutility. The achieved results are summarized in the Table 1.

An individual experiment was organized so that the model (1) - (7) was used first to obtain the basic design. Since the convergence of the objective function value was too slow, we limited the computational process by 15 minutes. After reaching this time limit, the computational process was terminated and the best found solution was taken as the result together with the associated lower bound. The computational time in seconds is denoted by *CT*. The value of the objective function (1) is reported in the column denoted by *h* and the lower bound is given in the column denoted by *LB*. To express the maximal difference between the best found solution and the optimal one, the value of *Gap* is used. It is defined as the difference between *h* and *LB* expressed in percentage of *LB*.

The same notation is used also for the second approach described by the model (8), (2), (3), (5) and (6) based on the Lagrangean relaxation. Here, several notes must be taken in consideration. This approach solves the problem iteratively with different settings of Lagrangean multipliers. Therefore, the value of *CT* contains the computational times of each iteration. The Lagrangean multipliers were computed according to the sub-gradient method. The total number of performed iterations is denoted by *It*.

Both approaches were aimed at the robust design of emergency service system. It means that the methods must concern all studied scenarios. Due to the lack of common benchmarks for study of robustness, the scenarios used in our computational study were created in the following way. We chose 25 percent of matrix rows so that these rows correspond to the biggest cities concerning the number of system users. Then we chose randomly from 5 to 15 rows and the associated disutility values in the individual rows were multiplied by the randomly chosen constant from the range 2, 3 and 4. This way, 10 different scenarios were generated for each self-governing region.

$\|I\|$	p	**BASIC DESIGN - 15 minutes**				**LAGRANGEAN RELAXATION**				
		CT [s]	*h*	*LB*	*Gap*	*CT* [s]	*h*	*LB*	*It*	*Gap*
315	158	189.8	2535	2535	0.0	74.2	2540	2518.0	7	0.9
315	105	236.9	5912	5912	0.0	50.2	5936	5849.7	6	1.5
315	79	261.9	9084	9084	0.0	50.3	9236	9001.1	6	2.6
315	63	902.5	12403	12354.6	0.4	50.6	12572	12161.3	6	3.4
315	32	900.8	24167	24064.4	0.4	91.4	24913	23673.8	5	5.2
315	21	901.5	33323	32787.9	1.6	43.4	33697	32164.7	5	4.8
315	16	914.2	41495	40315.8	2.9	44.0	41095	39489.8	5	4.1

Table 1 Results of numerical experiments for the self-governing region of Žilina with 315 possible service center locations and different number of centers to be located.

The achieved results reported in Table 1 show that the iterative approach based on the Lagrangean relaxation can be used to obtain a good feasible solution. It is much faster than the basic approach, but the accuracy of the resulting solution is a bit worse than the one obtained by the basic approach. From this point of view, the future research can be aimed at the initial settings of the Lagrangean multipliers and adjusting the suggested method to give better results in a short time.

5 Conclusions

The paper is focused on a robust design of the emergency service system with the min-sum quality criterion, where various scenarios are considered and robustness is ensured by minimizing the maximal objective function over individual scenarios. The considered user´s disutility is assumed to be proportional to the distance of the user from the nearest service center and then, sum of distances from all system users to the nearest located service center is minimized in the basic problem. The min-max link-up constraints represent an undesirable burden in any integer programming problem due to bad convergence of the branch-and-bound method. We have suggested an approximate approach based on the Lagrangean relaxation and the subsequent sub-gradient method, which enables to obtain the solution much faster. On the other hand, the accuracy of the result is not as good as we would like it to be and that is why we should focus our future research on the initial settings of the Lagrangean multipliers and adjusting the suggested method to give better results in a short time. Another topic may be aimed at finding relevant scenarios, which can significantly impact the performance of emergency service system.

Acknowledgement

This work was supported by the research grants VEGA 1/0518/15 "Resilient rescue systems with uncertain accessibility of service", VEGA 1/0463/16 "Economically efficient charging infrastructure deployment for electric vehicles in smart cities and communities", APVV-15-0179 "Reliability of emergency systems on infrastructure with uncertain functionality of critical elements" and by the project University Science Park of the University of Žilina (ITMS: 26220220184) supported by the Research & Development Operational Program funded by the European Regional Development Fund.

References

[1] Avella, P., Sassano, A., Vasil'ev, I..: *Computational study of large scale p-median problems*. In: Mathematical Programming 109, pp. 89-114, 2007.

[2] Correia, I., Saldanha da Gama, F.: Facility locations under uncertainty. *Location Science*, eds. Laporte, Nikel, Saldanha da Gama, pp. 177-203, 2015.

[3] Current, J., Daskin, M., Schilling, D.: *Discrete network location models*. In: Drezner Z (ed) et al. Facility location. Applications and theory. Berlin: Springer, pp. 81-118, 2002.

[4] García, S., Labbé, M., Marín, A.: *Solving large p-median problems with a radius formulation*. In: INFORMS Journal on Computing 23 (4), pp. 546-556, 2011.

[5] Janáček, J.: *Approximate Covering Models of Location Problems*. In: Lecture Notes in Management Science: Proceedings of the 1st International Conference ICAOR ´08, Vol. 1, Sept. 2008, Yerevan, Armenia, pp. 53-61, ISSN 2008-0050, 2008.

[6] Janáček, J., Kvet, M.: *Approximate solving of large p-median problems*. In: ORP3-Operational research peripatetic post-graduate programme: proceedings, Cádiz, Spain, September 13-17, 2011, Cádiz: Servicio de Publicaciones de la Universidad de Cádiz, 2011, ISBN 978-84-9828-348-8, pp. 221-225, 2011.

[7] Janáček, J., Kvet, M.: *Relevant Network Distances for Approximate Approach to Large p-Median Problems*. In: Operations research proceedings 2012: selected papers of the international conference on operations research: September 4-7, 2012, Hannover, Germany, ISSN 0721-5924, ISBN 978-3-319-00794-6, pp 123 – 128, 2012.

[8] Janáček, J., Kvet, M.: *Public service system design with disutility relevance estimation*. In: Proceedings of the 31st international conference: "Mathematical Methods in Economics", September 11-13, 2013, Jihlava, Czech Republic, ISBN 978-80-87035-76-4, pp 332 – 337, 2013.

[9] Janáček, J., Kvet, M.: Emergency system design with temporarily failing centers. *SOR 15: Proceedings of the 13th International Symposium on Operational Research*, Ljubljana: Slovenian Society Informatika, Section for Operational Research, ISBN 978-961-6165-45-7, pp. 490-495, 2015.

[10] Marianov, V., Serra, D.: *Location problems in the public sector*. In: Drezner Z (ed.) et al. Facility location. Applications and theory. Berlin: Springer, pp. 119-150, 2002.

[11] Marsh, M., Schilling, D.: Equity Measurement in Facility Location Analysis. *European Journal of Operational Research*, 74, pp. 1-17, 1994.

[12] Nash, J.: The Bargaining Problem. *Econometrica*, vol. 18, No. 2, pp. 155-162, 1950.

[13] Pan, Y., Du, Y., Wei, Z.: Reliable Facility System Design Subject to Edge Failures. *American Journal of Operations Research*, 4, pp. 164-172, 2014.

An importance of the population density for the location of the Emergency Medical Service stations

Marta Janáčková[1], Alžbeta Szendreyová[2]

Abstract. The paper deals with the location of the centers of the Emergency medical services (EMS). The problem belongs to the facility location problem. This location task is solved as a *p*-median problem and the optimization criterion is to maximize the system utility of the customers. The provision of the Emergency medical services is useful only in a limited period. This characteristic is ensured by the utility criterion. The paper describes the utility criterion, which we use in the model, and also explains its parameters. We test the optimal solution of the problem for two different sets of customers. In the first one, we regard municipalities as customers and we solve the task as a classic *p*-median problem. In the second one, inhabitants are considered to be customers and we solve the task as a weighted *p*-median problem. We also study the impact, which has the number of the inhabitants of the municipalities on the centers locations. We test the problems on the road network of two regions of Slovak Republic. The results are collected and shown in the tables.

Keywords: *p*-median, facility location problem, utility, network, public systems.

JEL Classification: C61
AMS Classification: 90B06, 90B80

1 Introduction

A providing a health care for the population is one of the services of the public service system. A part of this care is the provision of ambulance services in cases of an emergency. This is realized continuously and provided by the emergency medical care (ambulance crew with a doctor) or by the emergency health care (ambulance crew without a doctor). These crews are located in the Emergency medical services (EMS) stations. The EMS stations should be located so that to allow aid to each customer within 15 minutes of accepting the report about the urgent requirement.

In this paper, we deal with a design to place optimal the EMS stations. This task is solved as a classical *p*-median problem, as well as a weighted *p*-median problem. For these two approaches, we compare the availability of services in terms of usefulness, of kilometric distances and of number of differences in the location of EMS stations.

1.1 *P*-median problem

The *p*-median problem is used to decide how to locate the number of *p* centers (in rare cases, maximum *p* centers) to achieve the optimal value of the objective function. In our contribution, the system utility will be the optimality criterion. A formulation of this problem follows:

Let I denotes a finite set of the candidates for center locations (possible center locations) and let J denotes a set of the customers (dwelling places). The candidates for center locations and customers are located in the nodes of the network. Segments between nodes i and j are indicated by coefficients c_{ij} for each possible location $i \in I$ and for each dwelling place $j \in J$. Our task is to place the given number p of centers to some nodes from the set I and from them to serve each customers from the set J so that the value of the objective function reaches the maximum. The decision on placing or not placing a service center at a possible center location $i \in I$ is modeled by a variable y_i. The variable y_i takes the value of 1, if the center is located at place $i \in I$ and it takes the value of 0 otherwise. The decision on assigning of the customer from node j to the center at the place i is modeled by a variable z_{ij}. It takes the value of 1 if the customer j is served from the center i and takes the value of 0 otherwise. The model then can be constituted in the following form:

[1] University of Žilina, Department of Applied Mathematics, Univerzitná 1, 01026 Žilina, Slovak Republic, e-mail: marta.janackova@fstroj.uniza.sk.
[2] University of Žilina, Department of Mathematical Methods and Operations Research, Univerzitná 1, 010 26 Žilina, Slovak Republic, e-mail: Alzbeta.Szendreyova@fri.uniza.sk.

$$Maximize \quad \sum_{i \in I} \sum_{j \in J} c_{ij} z_{ij} \tag{1}$$

$$Subject\ to \quad \sum_{i \in I} z_{ij} = 1 \quad for\ j \in J \tag{2}$$

$$z_{ij} \le y_i \qquad for\ i \in I, j \in J \tag{3}$$

$$\sum_{i \in I} y_i \le p \tag{4}$$

$$y_i \in \{0,1\} \quad for\ i \in I \tag{5}$$

$$z_{ij} \in \{0,1\} \quad for\ i \in I, j \in J \tag{6}$$

The model's coefficients have the following meanings:

c_{ij} … evaluation of the segment between nodes *i* and *j*,
I … set of possible service centers locations,
J … set of customers (dwelling places),
p … required number of the location of the service centers.

In our paper, the coefficient c_{ij} is determined by the distance between nodes *i* and *j* in classical *p*-median problem and by the product of the distance between nodes *i* and *j* and the weight of the node *j* in the case of the weighted *p*-median problem. The weight of the node *j* is proportional to the number of inhabitants in the node *j*.

1.2 Utility function

The provision of the urgent medical care depends on the speed of its provision. According to the regulations in Slovak Republic, the medical services have to be provided within 15 minutes from the time the requirement was accepted. At serious health problem, the provision can be useless after exceeding this time limit. That is why we use maximization of the utility instead of minimization the distance (in kilometers or in minutes) when we design the placement of the EMS centers. To express the utility function, we need a function, which has a decreasing and "jumping" character. If we want easily compare the utility of the service for the customer, it is suitable to use a "normalized" form of the function, which causes that its maximum reaches always the same value. In our experiments, we use the function in the following form:

$$u(d) = \frac{1+e^{\frac{-d_{crit}}{T}}}{1+e^{\frac{d-d_{crit}}{T}}} \tag{7}$$

The variable d represents a distance between the EMS center and the customer (kilometric distance or time distance). The parameter d_{crit} represents the value, at which it occurs the jump in the function. The utility value is fundamentally changing when overstep the value of d_{crit}. The utility becomes negligible. The parameter T is the forming coefficient of the function. It affects the "steepness" of its course in a neighborhood of d_{crit}. The course of the function becomes shallower with increasing T. The significance of the jump softens and the course of the function gradually nears to linear character. For all values of parameters T and d_{crit}, the function $u(d)$ takes its maximum for $u(0) = 1$.

2 Experiments

According to the evaluation of the status at December 31st, 2013, 92 emergency medical care, 181 emergency health care and 7 stations of helicopter medical assistance are located in Slovak Republic at present. Their activities are managed by 8 regional operating EMS centers. In an effort to come closer to the reality, the required number of the location of the EMS centers in our design tasks is equal to the number of those municipalities, in which the EMS centers are really located. There are more EMS stations in some large municipalities. Our model cannot take into account this situation that is why we place less centers in our design as there are placed in current situation. The objective function is the utility function which was described in part 1.2.

2.1 Basic data of the task

We solve the experiments at the territory of the Slovak regions Žilina and Prešov. The testing is done on the road network of these regions. We solve two problems: classic *p*-median problem and weighted *p*-median problem. Each municipality is considered to be a customer in classic *p*-median problem. In weighted *p*-median problem, each hundred inhabitants of the municipality is considered to be a customer. The numbers of customers are rounded to the integer in this case. Each municipality is the possible location of the EMS center in both cases. We solve all tasks for the critical value $d_{krit} = 20$ (kilometers). The forming coefficient T takes the integer value of 1 to 22.

There are 12 emergency medical care and 24 emergency health care in the region of Žilina. They are located in 29 municipalities. According to available data, there are about 691 100 inhabitants in 315 municipalities in this region.

In the region of Prešov, there are 16 emergency medical care and 28 emergency health care. They are located in 32 municipalities. There are about 818 300 inhabitants in 664 municipalities in the region of Prešov.

2.2 Evaluation of the results

We locate the number of p centers so that the value of the utility reaches its maximum. The total utility of the optimal solution depends on the number of the customers. For relevant comparison, we also calculate the relative utility (an average per one customer). In addition to the utility, we monitor the traveled kilometers between the center and the customer assigned to it. Another monitored data is the maximum distance between the customer and the center which is assigned to him in the optimal solution.

For the region of Žilina, the required number of the service centers p is equal to 29 and the number of the possible service centers locations is 315. The number of the customers takes the value of 315. Table 1 shows the results of the classic p-median problem. The columns of the table contain step by step the values of the forming parameter T, the total utility of the optimal solution, the relative utility, the total distance among the centers and the assigned customers, the relative distance and the maximum distance (the worst availability for the customer).

T	Total utility	Relative utility	Total distance centers-municipalities	Relative distance per municipality	Maximum distance
1	314,96	1,00	2089	6,63	15
4	303,27	0,96	1949	6,19	16
7	289,61	0,92	1881	5,97	20
10	284,71	0,90	1876	5,96	20
13	284,19	0,90	1876	5,96	20
16	285,32	0,91	1874	5,95	20
19	286,99	0,91	1874	5,95	20
22	288,76	0,92	1874	5,95	20

Table 1 Evaluation of the utility for the Žilina region.
Number of candidates = 315, p = 29, d_{crit} = 20, number of customers = 315

T	Total utility	Relative utility	Distance centers-customers	Relative distance per customer	Maximum distance	Distance centers-municipalities	Relative distance per municipality
1	6910,63	1,00	30123	4,36	17	2125	6,75
4	6779,45	0,98	23363	3,38	26	2059	6,54
7	6609,49	0,96	22995	3,33	26	2067	6,56
10	6545,18	0,95	22995	3,33	26	2067	6,56
13	6537,12	0,95	22876	3,31	26	2102	6,67
16	6550,12	0,95	22876	3,31	26	2102	6,67
19	6570,05	0,95	22861	3,31	26	2138	6,79
22	6591,37	0,95	22861	3,31	26	2138	6,79

Table 2 Evaluation of the EMS centers design for the Žilina region.
Number of candidates = 315, p = 29, d_{crit} = 20, number of customers = 6911

The table 2 contains the results of the solution of the weighted p-median problem. The number of customers is 6911. The columns contain the following data: the value of T, the total utility, the relative utility, the total and relative distance among the centers and the customers assigned to them (inhabitants), the maximum distance, the total and relative distance among the centers and the assigned municipality, which the customer is from.

For the region of Prešov, the required number of the service centers p is equal to 32 and the number of the possible service centers locations is 664. The number of the customers takes the value of 664 in case of the classic p-median problem and it takes the value of 8183 in case of the weighted p-median problem. The results are in the tables 3 and 4.

T	Total utility	Relative utility	Total distance centers-municipalities	Relative distance per municipality	Maximum distance
1	662,00	1,00	6031	9,08	19
4	614,13	0,92	5681	8,56	31
7	577,42	0,87	5585	8,41	35
10	567,22	0,85	5583	8,41	35
13	568,23	0,86	5583	8,41	35
16	573,04	0,86	5583	8,41	35
19	578,84	0,87	5583	8,41	35
22	584,62	0,88	5583	8,41	35

Table 3 Evaluation of the utility for the Prešov region.
Number of candidates = 664, p = 32, d_{crit} = 20, number of customers = 664

T	Total utility	Relative utility	Distance centers-customers	Relative distance per customer	Maximum distance	Distance centers-municipalities	Relative distance per municipality
1	8174,81	1,00	59342	7,25	20	6167	9,29
4	7915,29	0,97	35409	4,33	24	6124	9,22
7	7677,08	0,94	34479	4,21	35	6181	9,31
10	7604,25	0,93	34195	4,18	35	6254	9,42
13	7604,95	0,93	34195	4,18	35	6254	9,42
16	7631,44	0,93	34144	4,17	42	6340	9,55
19	7665,42	0,94	34144	4,17	42	6340	9,55
22	7699,79	0,94	34134	4,17	42	6339	9,55

Table 4 Evaluation of the EMS centers design for the Prešov region.
Number of candidates = 664, p = 32, d_{crit} = 20, number of customers = 8183

In the table 5 we present the number of nodes (row H) in which the set of optimal solutions in case of the classic p-median problem and the set of optimal solutions in case of the weighted p-median problem differ from each other.

Region of Žilina

T	**1**	**4**	**7**	**10**	**13**	**16**	**19**	**22**
H	3	14	16	15	16	14	14	14

Region of Prešov

T	**1**	**4**	**7**	**10**	**13**	**16**	**19**	**22**
H	18	19	25	26	26	25	25	24

Table 5 Number of differences in the optimal locations of the EMS centers

As we expected, the results show that the solutions of the weighted p-median achieve better relative utility. More inhabitants gained an advantage, when the centers were moved to the municipalities with higher population. The reduction of the relative distance per customer shows this advantage. The relative distance per municipality became worse in weighted *p*-median problem, but this deterioration is not so significant (especially for low values of *T*). On the other side, the maximum availability became worse.

3 Conclusions

The model with small values of parameter *T* is most suitable for the urgent medical care. We have no relevant data on how the number of the EMS intervention correlate to the number of inhabitants. If the correlation is confirmed, the weighted *p*-median problem describes situation better as the classic *p*-median problem. The using of the classic *p*-median problem is suitable especially for the areas with even distribution of the population and for season, when the number of inhabitants changes (e.g. holiday).

Acknowledgements

This work was supported by the research grant VEGA 1/0518/15 “Resilient rescue systems with uncertain accessibility of service”.
We would also like to thank to “*Centre of excellence for systems and services of intelligent transport*” (ITMS 26220120050) for built up the infrastructure, which was used.

References

[1] Erlenkotter, D.: A Dual-Based Procedure for Uncapacitated Facility Location, *Operations Research* **26**, 6 (1978), 992–1009.
[2] Janáček, J., Emergency Public Service System Design Using IP-Solver. In: *Proceedings of the 1st WSEAS International Conference on Optimization Techniques in Engineering* (OTENG’13), Antalya, 2013, 97–101.
[3] Janáčková, M., and Szendreyová, A.: Properties of the cost matrix and the *p*-median problem. *International Journal of Applied Mathematics and Informatics* **8** (2014), 34–41.
[4] Xpress-MP Manual “*Getting Started*”. Dash Associates, Blisworth, 2005.

Fractional Brownian motion and parameter estimation

Josef Janák[1]

Abstract. The contribution introduces the concept of scalar fractional Brownian motion, which is an important stochastic process that does not have Markov property. The process is long–dependent on the past and its increments are not independent. For these qualities, fractional Brownian motion is used in hydrology [10], climatology [15] and also for modeling wind power resources [4]. It is also used for modeling of unemployment [3], inflation [5], volatility of assets [1] and trading volume [9]. Long memory models are also used in risk management [12] and for modeling of activity in computer networks [13].

The main topic of this article is parameter estimation for finite dimensional fractional Ornstein–Uhlenbeck process based on ergodicity (see [11]). Parameter estimation in the stochastic equation of the second order (which is basically the equation for the harmonic oscillation) is also discussed. The generalizations of this interesting topic (like infinite–dimensional case) are studied nowadays. The last section contains original author's results.

Keywords: Fractional Brownian motion, fractional Ornstein–Uhlenbeck process, parameter estimation.

JEL classification: C13
AMS classification: 60G22, 93E10

1 Introduction

In many applications of theory of probability, stochastic processes are used for modeling of input for some system. It is often assumed that these processes do have Markov property. But sometimes it can be observed that the real data have some dependency. The behavior of the process after time t does not depend only on the situation at the time t, but it depends on all the history before time t. One of the processes, which are convenient for modeling such data, is the fractional Browninan motion (fBm for short). FBm was introduced by Kolmogorov [8]. In the fifties of the twentieth century, Hurst studied the long–term characteristics of the flow of the Nile. From the data on the state of water levels recorded for last 800 years, he observed that long–term dependence of conditions of the water level is similar to the short–term. Hurst characterized this dependency by the parameter, which he estimated from the data, and the parameter is now called Hurst parameter.

In the following section of the paper, fBm is defined and some of its important properties are summarized. The third section introduces the space $\mathcal{H}$, that is the space of all deterministic functions, which are integrable with respect to the fBm. The stochastic integration with respect to fBm for the cases $H > \frac{1}{2}$ and $H < \frac{1}{2}$ is also mentioned.

Parameter estimation for finite dimensional fractional Ornstein–Uhlenbeck process based on ergodicity (see [11]) is discussed in Section 4. Parameter estimation for the stochastic equation of the second order can be found in Section 5.

2 Fractional Brownian motion

Consider the probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

[1]University of Economics, Department of Mathematics, Ekonomická 957, 148 00 Prague 4, janj04@vse.cz

Definition 1. A Gaussian process $(B^H(t), t \geq 0)$ on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in $\mathbb{R}$ is called the fractional Brownian motion (fBm) with Hurst parameter $H \in (0,1)$, if this process satisfies

(i) $\mathbb{E}B^H(t) = 0$ for any $t \in \mathbb{R}_+$,

(ii) $\operatorname{cov}(B^H(s), B^H(t)) = \frac{1}{2}(s^{2H} + t^{2H} - |t-s|^{2H}),$ for any $s, t \in \mathbb{R}_+$,

(iii) $(B^H(t), t \geq 0)$ has continuous paths $\mathbb{P} - a.s.$,

where $\operatorname{cov}(X,Y)$ denotes the covariance of random variables X and Y.

From (i) and (ii) it follows, that $B^H(0) = 0 \quad \mathbb{P} - a.s.$ If $H = 1/2$, then fBm is the standard Brownian motion. The existence of fBm and its some basic properties are stated in Theorem 1.

The proofs of Theorems 1, 2 and 4 can be found in [6]. The proof of Theorem 3 has been published in [14] and the proof of Theorem 6 (as well as more theorems based on ergodicity) has been published in [11].

Theorem 1. *Fractional Brownian motion exists, has stationary increments and is H-self-similar, i. e. for every $a > 0$ $(B^H(au), u \in \mathbb{R}_+) \overset{d}{=} (a^H B^H(u), u \in \mathbb{R}_+)$, where $\overset{d}{=}$ means equality in distribution.*

Instead of continuous process $(\Omega, \mathcal{A}, \mathbb{P}, X)$, its canonical representation $(C(\mathbb{R}^+), \mathcal{B}(C(\mathbb{R}^+)), X(\mathbb{P}), B^H)$ is often used, where $C(\mathbb{R}^+)$ is the space of all continuous functions on $\mathbb{R}^+$, $X(\mathbb{P})$ is the image of measure $\mathbb{P}$ using function $X : \Omega \to C(\mathbb{R}^+)$ and B^H is the projective process $B^H(t,\omega) = \omega(t)$ for $\omega \in C(\mathbb{R}^+)$ and $t \in \mathbb{R}^+$.

Characteristic properties of fBm are the following:

- From the Kolmogorov–Chentsov Theorem (see [7]) it follows, that the trajectories of fBm are locally Hölder continuous with exponent γ for any $\gamma \in (0,H)$, i. e. for any $T > 0$, there exist almost surely positive random variable $h_T : C(\mathbb{R}^+) \to \mathbb{R}$ and $\delta > 0$, so that

$$X(\mathbb{P})\left(\omega; \sup_{\substack{0<t-s<h_T(\omega)\\ s,t\in[0,T]}} \frac{|B^H(t,\omega) - B^H(s,\omega)|}{|t-s|^\gamma} \leq \delta\right) = 1,$$

  where the canonical representation is used. From this observation it follows, that on every interval $[0,T]$, almost all trajectories of fBm are γ-Hölder continuous with exponent $\gamma \in (0,H)$.

- For $H \in (1/2, 1)$ is fBm long-term dependent on the past, i. e. if $r(n)$ is defined by

$$\begin{aligned} r(n) &= \mathbb{E}B^H(1)(B^H(n+1) - B^H(n)) \\ &= \frac{1}{2}((n+1)^{2H} - 2n^{2H} + (n-1)^{2H}), \end{aligned}$$

  then

$$\sum_{n=1}^{\infty} r(n) = +\infty.$$

- FBm has finite variation of order $1/H$, i. e. for $p > 0$

$$\lim_{n\to\infty} \sum_{i=1}^{n} \left| B^H(t_i^{(n)}) - B^H(t_{i-1}^{(n)}) \right|^p \overset{\text{a.s.}}{=} \begin{cases} 0, & p > 1/H \\ \mathbb{E}|B^H(1)|^{1/H}, & p = 1/H \\ +\infty, & p < 1/H, \end{cases}$$

  where $\Delta_n(1) = \{0 = t_0^{(n)} < \cdots < t_n^{(n)} = 1\}, n \in \mathbb{N}$ is a sequence of divisions of interval $[0,1]$, such that $\Delta_n(1) \subset \Delta_{n+1}(1)$ and $\max|t_{i+1}^{(n)} - t_i^{(n)}| \to 0$.

- From property (ii) it follows, that

$$\mathbb{E}B^H(t)(B^H(t+h)-B^H(t))=\frac{1}{2}((t+h)^{2H}-t^{2H}-h^{2H}). \tag{1}$$

For $H<1/2$ is the expression (1) negative (increments are negatively correlated).
For $H>1/2$ is the expression (1) positive (increments are positively correlated).
For $H=1/2$ is known independence of increments of Brownian motion obtained.

There are many methods to estimate Hurst parameter H either from observation of whole trajectory of fBm or from discrete time series. The most commonly used R/S analysis as well as periodogram–based analysis in the frequency domain can be found in [16]. This (combined with the last point) is often used for testing of independency of increments of stock prices. (Where the estimator $\hat{H}$ evolves typically around 0,6.)

3 Introduction of stochastic integral

There are several problems with stochastic calculus for fBm. Since the trajectories of fBm does not have finite variation, Lebesgue–Stieltjes definition of integral cannot be used. FBm also does not have the semimartingale property, so Itô's stochastic calculus cannot be used as well. The difficulties in finding the proper definition of the stochastic integral was the cause, why stochastic differential equations with fBm are discovered only recently. Since this topic is quite technical and long, it is mentioned here only briefly.

Let $\mathcal{H}$ be the space of all $\mathbb{R}$–valued deterministic integrable functions by fBm.

Theorem 2. *Let $H>\frac{1}{2}$. Then the following holds true*

$$L^2([0,T])\subset L^{\frac{1}{H}}([0,T])\subset\mathcal{H}.$$

It means, that for any function $f\in L^2([0,T])$, the stochastic integral $\int_0^T f(t)\,dB^H(t)$ can be defined. First, the function f is approximated by step functions (which can be done, because the set of step functions is a dense set in $L^2([0,T])$), then the integrals for every approximating function are defined. Since the sequence of these stochastic integrals has Cauchy property in the space $\mathcal{H}$ (which is complete space), the stochastic integral is then defined as an appropriate limit of stochastic integrals obtained from step functions. For more in–depth analysis, see [14] and [6].

Theorem 3. *Let $H<\frac{1}{2}$. If $\gamma>\frac{1}{2}-H$, then*

$$C^\gamma([0,T])\subset\mathcal{H},$$

where $C^\gamma([0,T])$ is the space of all γ–Hölder functions on interval $[0,T]$.

In case of $H<\frac{1}{2}$, the set of integrable functions is different, but the stochastic integral is also defined using approximations of the integrand by step functions. (See [14], page 13 or [2].)

4 Parameter estimation for the fractional Ornstein–Uhlenbeck process

Consider the following linear stochastic differential equation

$$dX(t)=AX(t)\,dt+\Phi\,dB^H(t),\quad X(0)=x_0, \tag{2}$$

where $(B^H(t),t\geq 0)$ is $\mathbb{R}^m$-valued fBm, $A\in\mathbb{R}^{n\times n}$, $\Phi\in\mathbb{R}^{n\times m}$, $x_0\in\mathbb{R}^n$. Let $S(t):=e^{At}$ be a strongly continuous semigroup on $\mathbb{R}^n$, then the solution $(X^{x_0},t\geq 0)$ to (2) is defined by the mild form

$$X^{x_0}(t)=S(t)x_0+Z(t),\quad t\geq 0, \tag{3}$$

where $(Z(t),t\geq 0)$ is the convolution integral

$$Z(t)=\int_0^t S(t-u)\Phi\,dB^H(u). \tag{4}$$

The solution $(X^{x_0},t\geq 0)$ to the equation (2) is called the fractional Ornstein–Uhlenbeck process.

Theorem 4. *If the semigroup $(S(t), t \geq 0)$ is exponentially stable, i.e. there exist constants $M > 0$ and $\rho > 0$ such that for all $t \geq 0$*

$$|S(t)| \leq Me^{-\rho t},$$

then there is a Gaussian centered limiting measure $\mu_\infty = N(0, Q_\infty)$ for $(X^{x_0}(t), t \geq 0)$ such that

$$w^* - \lim_{t \to \infty} \mu_t^{x_0} = \mu_\infty$$

for each initial condition $x_0 \in \mathbb{R}^n$, where $\mu_t^{x_0} =$ Law $(X^{x_0}(t))$ and Law $(\cdot)$ denotes the probability distribution.

Remark 1. If $H > \frac{1}{2}$ then the covariance Q_∞ has the following form:

$$Q_\infty = \int_0^\infty \int_0^\infty S(u)\Phi\Phi^T S^T(v)\phi(u-v)\, du\, dv,$$

where $\phi(t) = H(2H-1)|u|^{2H-2}$.

If $H = \frac{1}{2}$, then

$$Q_\infty = \int_0^\infty S(t)\Phi\Phi^T S^T(t)\, dt.$$

Now consider the linear equation

$$dX(t) = \alpha AX(t)\, dt + \Phi\, dB^H(t), \quad X(0) = x_0, \tag{5}$$

which is the same equation as (2) with a real constant parameter $\alpha > 0$.

If $(S(t) = e^{At}, t \geq 0)$ is exponentially stable, then the semigroup $(S_\alpha(t) = S(\alpha t) = e^{\alpha At}, t \geq 0)$ is also exponentially stable and there is a limiting measure $\mu_\infty^\alpha = N(0, Q_\infty^\alpha)$.

For $H > \frac{1}{2}$, there is an useful formula for Q_∞^α:

$$\begin{aligned} Q_\infty^\alpha &= \int_0^\infty \int_0^\infty S_\alpha(u)\Phi\Phi^T S_\alpha^T(v)\phi(u-v)\, du\, dv \\ &= \frac{1}{\alpha^2} \int_0^\infty \int_0^\infty S(u)\Phi\Phi^T S^T(v)\phi\left(\frac{u}{\alpha} - \frac{v}{\alpha}\right) du\, dv \\ &= \frac{1}{\alpha^{2H}} Q_\infty, \end{aligned}$$

where Q_∞ corresponds to the case $\alpha = 1$ (see Remark 1). For $H = \frac{1}{2}$ is this equality obvious.

Based on the above results, some strongly consistent families of estimators of the parameter α may be proposed.

Theorem 5. *Let the semigroup $(S(t), t \geq 0)$ be exponentially stable and let $(X^{x_0}(t), t \geq 0)$ be a solution to (5). Let $\Phi \neq 0$. Define*

$$\hat{\alpha}_T := \left(\frac{Tr\, Q_\infty}{\frac{1}{T}\int_0^T |X^{x_0}(t)|^2\, dt} \right)^{\frac{1}{2H}}.$$

Then

$$\lim_{T \to \infty} \hat{\alpha}_T = \alpha.$$

The application is, that from the observation of the whole trajectory of the process $X^{x_0}(t)$, which is the solution to the equation (5), on some (long enough) time interval $[0, T]$, it is possible to compute the time–averages of $|X^{x_0}(t)|^2$, i.e. it is possible to compute the integral $\int_0^T |X^{x_0}(t)|^2\, dt$. If there is also a way to compute the trace of Q_∞, then Theorem 5 yields estimator of the parameter α, which is strongly consistent.

5 Parameter estimation for the stochastic equation of second order

Consider the following equation of second order

$$\ddot{x} + 2a\dot{x} + bx = \sigma\dot{\beta}_t^H, \tag{6}$$

with initial values $X(0) = x_0$ and $\dot{X}(0) = x_1$. Let $a > 0$, $b > 0$ be real parameters and $\sigma > 0$ is known.

This equation can be rewritten in the form (2) by setting

$$X(t) = \begin{pmatrix} X^0(t) \\ X^1(t) \end{pmatrix} \in \mathbb{R}^2,\ A = \begin{pmatrix} 0 & 1 \\ -b & -2a \end{pmatrix},\ \Phi = \begin{pmatrix} 0 & 0 \\ 0 & \sigma \end{pmatrix},\ B_t^H = \begin{pmatrix} B_1^H(t) \\ B_2^H(t) \end{pmatrix},$$

where $B_2^H = \beta^H$ and B_1^H is an independent fBm.

In the following part, let us suppose, that $H = \frac{1}{2}$ (due to simplicity). The eigenvalues of matrix A are $\lambda_1 = -a + \sqrt{a^2 - b}$, $\lambda_2 = -a - \sqrt{a^2 - b}$. If $a^2 < b$, then the real parts of eigenvalues of matrix A are negative and the semigroup $(S(t), t \geq 0)$ is exponentially stable. Let us denote $\alpha = -a$, $\beta = \sqrt{b - a^2}$. Then the eigenvalues of matrix A are in fact $\lambda_1 = \alpha + i\beta$, $\lambda_2 = \alpha - i\beta$. The fundamental system $(S(t), t \geq 0)$ for the equation (6) has the following form

$$S(t) = \begin{pmatrix} e^{\alpha t}\left(\cos(\beta t) - \frac{\alpha}{\beta}\sin(\beta t)\right) & \frac{1}{\beta}e^{\alpha t}\sin(\beta t) \\ e^{\alpha t}\left(-\beta - \frac{\alpha^2}{\beta}\right)\sin(\beta t) & \frac{1}{\beta}e^{\alpha t}(\alpha\sin(\beta t) + \beta\cos(\beta t)) \end{pmatrix}.$$

The covariance operator $Q_\infty^{(a,b)}$ of the limit measure $\mu_\infty^{(a,b)}$ equals to

$$\begin{aligned} Q_\infty^{(a,b)} &= \int_0^\infty S(t)\Phi\Phi^T S^T(t)\,dt \\ &= \int_0^\infty \frac{\sigma^2}{\beta^2}\begin{pmatrix} q_{11}(t) & q_{12}(t) \\ q_{21}(t) & q_{22}(t) \end{pmatrix} dt \\ &= \frac{\sigma^2}{\beta^2}\begin{pmatrix} \frac{\beta^2}{4ab} & 0 \\ 0 & \frac{\beta^2}{4a} \end{pmatrix} = \sigma^2\begin{pmatrix} \frac{1}{4ab} & 0 \\ 0 & \frac{1}{4a} \end{pmatrix}, \end{aligned}$$

where

$$\begin{aligned} q_{11}(t) &= e^{2\alpha t}\sin^2(\beta t), \\ q_{12}(t) &= q_{21}(t) = e^{2\alpha t}\sin(\beta t)(\alpha\sin(\beta t) + \beta\cos(\beta t)), \\ q_{22}(t) &= e^{2\alpha t}(\alpha\sin(\beta t) + \beta\cos(\beta t))^2. \end{aligned}$$

Then it is easy to obtain, that

$$\mathrm{Tr}\ Q_\infty^{(a,b)} = \sigma^2\frac{b+1}{4ab}.$$

Since the right hand side is decreasing in the variable $a \in \mathbb{R}_+$ (and also in $b \in \mathbb{R}_+$) on the set, where $a^2 < b$, the unique estimator of one of these parameters can be obtained if the value of the other one is known. Denote $Y(T) = \frac{1}{T}\int_0^T |\tilde{X}(t)|^2\,dt$, where $\tilde{X}(t)$ is the stationary solution to the equation (6) (which also exists), and set

$$a_T = \sigma^2\frac{b+1}{4bY(T)}.$$

Then according to Theorem 5 a_T converges $\mathbb{P} - a.s.$ to the real value of the parameter. For example if $b = 1$, then the estimator of parameter a is $a_T = \frac{\sigma^2}{2Y(T)}$.

6 Conclusion

The fractional Brownian motion has been introduced and its basic properties have been showed. There has also been introduced the space $\mathcal{H}$ of functions, which are integrable with respect to fBm. Notably if $H > \frac{1}{2}$, then $L^2([0,T]) \subset L^{\frac{1}{H}}([0,T]) \subset \mathcal{H}$, and if $H < \frac{1}{2}$, then $C^\gamma([0,T]) \subset \mathcal{H}$.

The second part of the paper presented the fractional Ornstein–Uhlenbeck process and some parameter estimates based on ergodicity. Parameter estimates for the stochastic equation of second order (i. e. stochastic oscillator) has also been derived.

7 Acknowledgement

The paper was supported by the GACR Grant no. 15-08819S entitled "Stochastic Processes in Infinite Dimensional Spaces".

The paper was supported by KMAT VŠE.

References

[1] Comte, F., Renault, E.: Long Memory in Continuous Time Stochastic Volatility Models, *Mathematical Finance* **8**, 4 (1998), 291–323.

[2] Duncan, T. E., Maslowski, B., Pasik–Duncan, B.: Linear stochastic equations in a Hilbert space with a fractional Brownian motion, *International Series in Operations Research & Management Science* **94** (2006), 201–222.

[3] Funke, M.: The Long Memory Property of the US Unemployment Rate, *Discussion Paper No. 19–98*, Hamburg University, Hamburg, 1998.

[4] Haslet, J., Raftery, A. E.: Space–Time Modelling with Long–Memory dependence: Assessing Ireland's Wind Power Resource, *Applied Statistics* **38**, 1 (1989), 1–50.

[5] Hassler, U., Wolters, J.: Long–memory in Inflation Rates: International Evidence, *Journal of Business & Economic Statistics* **13**, 1 (1995), 35–47.

[6] Janák, J.: *Geometrical Fractional Brownian Motion*, Bachelor Thesis, Faculty of Mathematics and Physics, Charles University in Prague, 2007.

[7] Karatzas, I., Shreve, S. E.: *Brownian Motion and Stochastic Calculus*, Springer Verlag, New York, 1994.

[8] Kolmogorov, A. N.: Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum, *C. R. Acad. Sci. URSS (N.S.)* **26** (1940), 115–118.

[9] Lobato, I. N., Velasco, C.: Long Memory in Stock Market Trading Volume, *Journal of Bussines & Economic Statistics* **18**, 4 (2000), 410–427.

[10] Mandelbrot, B. B., Wallis, J.: Noah, Joseph and Operational Hydrology, *Water Resources Research* **4**, 5 (1968), 909–918.

[11] Maslowski, B., Pospíšil, J.: Ergodicity and parameter estimates for infinite–dimensional fractional Ornstein–Uhlenbeck process, *Applied Mathematics and Optimization* **57**, 3 (2007), 401–429.

[12] Michna, Z.: Self–similar Processes in Collective Risk Theory, *Journal of Applied Mathematics and Stochastic Analysis* **11** (1998), 429–248.

[13] Norros, I.: On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks, *IEEE Journal on Selected Areas in Communications* **13**, 6 (1995), 953–962.

[14] Nualart, D.: Stochastic integration with respect to fractional Brownian motion and applications, *Stochastics and Stochastic Reports* **75**, 3 (2003).

[15] Pelletier, J. D., Turcotte, D. L.: Long–range Persistence in Climatology and Hydrological Time Series: Analysis, Modelling and Application to Drought Hazard Assessment, *Journal of Hydrology* **203**, 1–4 (1997), 198–208.

[16] Rose, O.: Estimation of the Hurst parameter of Long–Range Dependent Time Series, *Research Report No. 137*, University of Würzburg, Würzburg, 1996.

Some robust distances for multivariate data

Jan Kalina[1], Barbora Peštová[2]

Abstract. Numerous methods of multivariate statistics and data mining suffer from the presence of outlying measurements in the data. This paper presents new distance measures suitable for continuous data. First, we consider a Mahalanobis distance suitable for high-dimensional data with the number of variables (largely) exceeding the number of observations. We propose its doubly regularized version, which combines a regularization of the covariance matrix with replacing the means of multivariate data by their regularized counterparts. We formulate explicit expressions for some versions of the regularization of the means, which can be interpreted as a denoising (i.e. robust version) of standard means. Further, we propose a robust cosine similarity measure, which is based on implicit weighting of individual observations. We derive properties of the newly proposed robust cosine similarity, which includes a proof of the high robustness in terms of the breakdown point.

Keywords: multivariate data, distance measures, regularization, robustness, high dimension

JEL classification: C18
AMS classification: 62H30

1 Introduction

Various methods of multivariate statistics and data mining allow to reveal the multivariate structure of given data [5]. The task of classification analysis, clustering or dimensionality reduction have established their place in the analysis of continuous data in economic applications. Suitable measures of distance among the multivariate data (or between a pair of groups, clusters etc.) play a crucial role in numerous methods including linear discriminant analysis, classification trees, hierarchical clustering, partial least squares etc. However, standard distance measures are known to be too sensitive to the presence of outlying measurements in the data, which is true for the Euclidean distance, Mahalanobis distance, Manhattan distance, Pearson correlation coefficient, cosine similarity, and others.

In addition, some of standard distance measures encounter other problems for high-dimensional data with the number of variables p exceeding the number of observations n. In this so-called "small n-large p-problem" with $n < p$ or even $n \ll p$, it is especially the Mahalanobis distance which suffers from the phenomenon described as the curse of dimensionality. It remains computationally infeasible, which is a problem of numerical stability as the empirical covariance matrix is almost singular and thus practically not invertible. While the concept of a regularized Mahalanobis distance has been considered by some authors [5, 10], we are not aware of any efforts to propose a doubly regularized version, combining the regularization of the covariance matrix with a regularization of the means

In this paper, we recall the concept of a regularized Mahalanobis distance in Section 2 and propose its new doubly regularized version. It is based on a regularization of the covariance matrix as well as a regularization of the means, where the latter transform ensures a robustness against measurement errors or small errors in the data. We also investigate its efficient computation. Further, we propose and investigate a highly robust cosine similarity measure in Section 3. The method is based on a sophisticated implicit weighting of individual observations, which ensures a high robustness in terms of the breakdown point. Finally, Section 4 concludes the paper.

[1]Institute of Computer Science CAS, Pod Vodárenskou věží 2, 182 07 Praha 8, Czech Republic, kalina@cs.cas.cz
[2]Institute of Computer Science CAS, Pod Vodárenskou věží 2, 182 07 Praha 8, Czech Republic, pestova@cs.cas.cz

2 Regularized Mahalanobis distance

Mahalanobis distance is one of popular distance measures used for the analysis of multivariate data. As already mentioned in Section 1, it represents a popular tool in a variety of methods suitable for econometric applications; this is documented e.g. in [1, 13]. Various regularized versions of the Mahalanobis distance have been used for high-dimensional data. Most commonly, the regularized versions are obtained (only) by replacing the empirical covariance matrix $\mathbf{S}$ by such (regularized) counterpart, which is guaranteed to be a regular matrix [5]. Regularized estimators of the covariance matrix are either sparse or have the form of a linear combination of the empirical covariance matrix $\mathbf{S}$ and a regular (e.g. diagonal) target matrix [10]. In this section, we start by recalling principles of the regularized Mahalanobis distance and propose its generalization based on a double regularization, which modifies (transforms) the covariance matrix as well as the means of multivariate data.

We consider p-dimensional data to be observed in the total number of K ($K \geq 2$) groups. The Mahalanobis distance will be formulated in the following setting as a distance between two groups of data. Depending on the particular context, the Mahalanobis distance can be formulated also as a distance between two clusters, between two individual observations from one random sample, or between an observation and a group of data in an analogous way. Thus, we extend our results of [9].

We will use the notation $\bar{\mathbf{X}}_k$ and $\bar{\mathbf{X}}_l$ for the mean of the k-th and l-th group, respectively, where $k, l = 1, \ldots, K$ and $k \neq l$. The observations in the k-th group will be denoted as $(\mathbf{X}_{k1}, \ldots, \mathbf{X}_{kn_k})^T$. The overall mean is denoted by $\bar{\mathbf{X}}$.

Let us consider the regularization of the empirical covariance matrix $\mathbf{S}$ in the form

$$\hat{\mathbf{S}} = \lambda \mathbf{S} + (1-\lambda)\mathbf{T}, \quad \lambda > 0, \tag{1}$$

where a given target matrix $\mathbf{T}$ is used, which must be a regular symmetric positive definite matrix of size $p \times p$. Its most common choices include the identity matrix $\boldsymbol{\mathcal{I}}_p$, a diagonal matrix with identical element $s = \sum_{i=1}^{p} S_{ii}/p$, or a general diagonal matrix.

Definition 1. We assume the p-dimensional random variables observed in K groups, while a common covariance matrix $\boldsymbol{\Sigma}$ is assumed in each of the groups. Its pooled estimator across groups will be denoted by $\mathbf{S}$ and we consider $\mathbf{S}^*$ defined by (1) as its regularized version. The regularized Mahalanobis distance between the k-th and l-th group is defined as

$$d(\bar{\mathbf{X}}_k, \bar{\mathbf{X}}_l) = \Big((\bar{\mathbf{X}}_k - \bar{\mathbf{X}}_l)^T \left(\mathbf{S}^*\right)^{-1} (\bar{\mathbf{X}}_k - \bar{\mathbf{X}}_l) \Big)^{1/2}. \tag{2}$$

The idea of regularizing the mean of each of the K groups has been considered mainly in the classification task. Available approaches (e.g. [3, 9]) considered only in the form of L_1-regularization, where a robustness to small deviations of data has been reported. We will now generalize the concept of the regularized Mahalanobis distance using a double regularization:

- The matrix $\mathbf{S}$ is regularized as in (1);
- The means in (2) are replaced by regularized counterparts using various forms of regularization.

We formulate the definition of the regularized Mahalanobis distance in a general way, while explicit solutions will be expressed later for particular versions of regularization. We will also discuss tools for an efficient computation.

Definition 2. We assume the p-dimensional random variables observed in K groups, while a common covariance matrix $\boldsymbol{\Sigma}$ is assumed in each of the groups. Its pooled estimator across groups will be denoted by $\mathbf{S}$ and we consider $\mathbf{S}^*$ defined by (1) as its regularized version. Let $\bar{\mathbf{X}}'_k$ and $\bar{\mathbf{X}}'_l$ denote a regularized estimator of the expectation of the k-th group. where $\delta \in [0,1]$ is a given value and $k, l = 1, \ldots, K$ ($k \neq l$). The doubly regularized Mahalanobis distance between the k-th and the l-th group is defined as

$$d(\bar{\mathbf{X}}'_k, \bar{\mathbf{X}}'_l) = \Big((\bar{\mathbf{X}}'_k - \bar{\mathbf{X}}'_l)^T \left(\mathbf{S}^*\right)^{-1} (\bar{\mathbf{X}}'_k - \bar{\mathbf{X}}'_l) \Big)^{1/2}. \tag{3}$$

Explicit expressions for regularized means can be formulated for particular choices of the regularization. The following lemma represents an analog of results of [5], [3] and [15] from a regression context.

Lemma 1. 1. If $\bar{\mathbf{X}}'_k$ is defined as

$$\arg\min_{\mathbf{m}\in\mathbb{R}^p}\left\{\sum_{i=1}^{n_k}||\mathbf{X}_{ki}-\mathbf{m}||_2^2+\delta_1||\mathbf{m}||_1\right\}, \tag{4}$$

then

$$\bar{\mathbf{X}}'_k=\mathsf{sgn}(\bar{\mathbf{X}}_k)\left(|\bar{\mathbf{X}}_k|-\delta_1\right)_+=\mathsf{sgn}(\bar{\mathbf{X}}_k)\max\left\{|\bar{\mathbf{X}}_k|-\delta_1,0\right\}, \tag{5}$$

where $(z)_+$ denotes the positive part of z.

2. If $\bar{\mathbf{X}}'_k$ is defined as

$$\arg\min_{\mathbf{m}\in\mathbb{R}^p}\left\{\sum_{i=1}^{n_k}||\mathbf{X}_{ki}-\mathbf{m}||_2^2+\delta_2||\mathbf{m}||_2^2\right\}, \tag{6}$$

then

$$\bar{\mathbf{X}}'_k=\frac{\sum_{i=1}^{n_k}\mathbf{X}_{ki}}{n_k+\delta_2}. \tag{7}$$

3. If $\bar{\mathbf{X}}'_k$ is defined as

$$\arg\min_{\mathbf{m}\in\mathbb{R}^p}\left\{\sum_{i=1}^{n_k}||\mathbf{X}_{ki}-\mathbf{m}||_2^2+\delta_1||\mathbf{m}||_1+\delta||\mathbf{m}||_2^2\right\}, \tag{8}$$

then

$$\bar{\mathbf{X}}'_k=\frac{|\bar{\mathbf{X}}_k|-\delta_1/2}{1+\delta_2}\mathsf{sgn}(\bar{\mathbf{X}}_k). \tag{9}$$

We describe how to compute the regularized Mahalanobis distance in an efficient way avoiding computing the inverse of $\mathbf{S}^*$, which is expensive of order p^3 and numerically rather unstable. Suitable matrix manipulations allow to replace it by solving a set of linear equations using the following. We consider either the eigendecomposition of $\mathbf{S}^*$ in the form

$$\mathbf{S}^*=\mathbf{Q}_*\mathbf{\Theta}_*\mathbf{Q}_*^T, \tag{10}$$

where $\mathbf{\Theta}_*$ is diagonal and $\mathbf{Q}_*$ is an orthogonal matrix, or the Cholesky decomposition of $\mathbf{S}^*$ in the form

$$\mathbf{S}^*=\mathbf{L}_*\mathbf{L}_*^T, \tag{11}$$

where $\mathbf{L}_*$ is upper triangular matrix.

Lemma 2.

$$d(\bar{\mathbf{X}}'_k,\bar{\mathbf{X}}'_l)=||\mathbf{\Theta}_*^{1/2}\mathbf{Q}_*^T(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)||_2^{1/2}. \tag{12}$$

Proof. Starting with the right side of (12),

$$\begin{aligned}||\mathbf{\Theta}_*^{1/2}\mathbf{Q}_*^T(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)||_2 &= \left(\mathbf{\Theta}_*^{1/2}\mathbf{Q}_*^T(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)\right)^T\mathbf{\Theta}_*^{1/2}\mathbf{Q}_*^T(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)=\\ &= (\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)^T\mathbf{Q}_*\mathbf{\Theta}_*^{-1}\mathbf{Q}_*^T(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l),\end{aligned} \tag{13}$$

we come to the statement (12), because $(\mathbf{S}^*)^{-1}=\mathbf{Q}_*\mathbf{\Theta}_*^{-1}\mathbf{Q}_*^T$. □

Lemma 3.

$$d(\bar{\mathbf{X}}'_k,\bar{\mathbf{X}}'_l)=||\mathbf{L}_*^{-1}(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)||_2^{1/2}. \tag{14}$$

Proof. Starting with the right side of (14),

$$||\mathbf{L}_*^{-1}(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)||_2=\left(\mathbf{L}_*^{-1}(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)\right)^T\mathbf{L}_*^{-1}(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)=(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l)^T\mathbf{L}_*^{-T}\mathbf{L}_*^{-1}(\bar{\mathbf{X}}'_k-\bar{\mathbf{X}}'_l), \tag{15}$$

we come to the statement (14), because $(\mathbf{S}^*)^{-1}=\mathbf{L}_*^{-T}\mathbf{L}_*^{-1}$. □

Replacing the eigendecomposition of $\mathbf{S}^*$ with its Cholesky decomposition makes the computation cheaper, because the costs of Cholesky decomposition are about $1/3\cdot p^3$ floating point operations. On the other hand, Cholesky decomposition will suffer from instability when $\mathbf{S}^*$ is not positive definite.

3 Robust cosine similarity

The cosine similarity is another measure of distance commonly used for multivariate data. Although it is seemingly less popular compared to the Mahalanobis distance, the cosine similarity is especially common as a similarity measure between clusters of documents in text mining [12]. Numerous applications include recommending news articles, clustering of text documents (not only) in economic applications [13] and the measure also often serves as a benchmarking method in information retrieval. However, it suffers from sensitivity to outlying measurements (in both horizontal and vertical directions). Attention has been paid to robust versions of the cosine similarity in text mining applications. In this section, we propose a robust cosine similarity based on the LWS regression and study its properties including the breakdown point, which is a statistical measure of sensitivity against outliers in the data [6].

Commonly, the cosine similarity measure is used as a similarity measure between two variables. Therefore, let us consider a different setup compared to previous sections. We assume two random variables

$$\mathbf{X} = (X_1, \ldots, X_n)^T \quad \text{and} \quad \mathbf{Y} = (Y_1, \ldots, Y_n)^T \tag{16}$$

observed on n independent observations. The cosine similarity between $\mathbf{X}$ and $\mathbf{Y}$ is defined as

$$\cos\,\theta = \frac{\mathbf{X}^T\mathbf{Y}}{||\mathbf{X}||_2 \cdot ||\mathbf{Y}||_2} = \frac{\sum_{i=1}^n X_i Y_i}{\left(\sum_{i=1}^n X_i^2 \sum_{j=1}^n Y_j^2\right)^{1/2}}. \tag{17}$$

In the context of cluster analysis, the two random variables (16) may denote the centroids (i.e. averages values) across two clusters or some other prototypes (i.e. medians) of the clusters.

Before we obtain a robust counterpart of $\cos\,\theta$, let us recall the least weighted squares (LWS) estimator [14] of parameters in the linear regression model, which is one of highly robust regression estimators. It is based on assigning weights to individual observations, which are assigned after a permutation given implicitly during the computation of the estimator with the aim to down-weight less reliable outlying observations. Data-dependent adaptive weights of [2] can be recommended, because they ensure asymptotically a 100 % efficiency of the least squares estimator under Gaussian errors, while they guarantee a high breakdown point for data contaminated with outliers. Extensions of the idea of implicit weights assigned to individual observations turn out to yield promising results also in other models, e.g. in classification or correlation analysis [7, 8].

We will use the fact that the slope estimator in the linear regression model

$$Y_i = \beta X_i + e_i, \quad i = 1, \ldots, n, \tag{18}$$

has the form

$$b = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}, \tag{19}$$

i.e. $\cos\,\theta$ resembles the Pearson correlation coefficient r, except for the normalization. Before we arrive at robustifying $\cos\,\theta$, let us express

$$\cos\,\theta = b \cdot \frac{\sqrt{\sum_{i=1}^n X_i^2}}{\sqrt{\sum_{i=1}^n Y_i^2}}. \tag{20}$$

and denote

$$\bar{X}_{LWS} = \sum_{i=1}^n \tilde{w}_i X_i \quad \text{and} \quad \bar{Y}_{LWS} = \sum_{i=1}^n \tilde{w}_i Y_i, \tag{21}$$

where $\tilde{w}_1, \ldots, \tilde{w}_n$ are weights obtained by the LWS estimator in (18).

We define the robust cosine similarity $\tilde{r}_{LWS}(X, Y)$ as

$$\tilde{r}_{LWS}(X, Y) = \hat{\beta}_{LWS} \frac{\sqrt{\sum_{i=1}^n \tilde{w}_i X_i^2}}{\sum_{i=1}^n \tilde{w}_i Y_i^2}. \tag{22}$$

Further, we investigate robustness and asymptotic normality of the robust cosine similarity. We will investigate the asymptotic behavior of the robust cosine similarity computed with adaptive weights. Because its computation is based on an initial robust estimator of $\boldsymbol{\beta}$, the robustness properties of $\tilde{r}_{LWS}$ depend on the finite-sample breakdown point ϵ_n^0 of this initial estimator.

Theorem 1. *Let us consider a sequence of independent identically distributed two-dimensional random vectors* $(X_1, Y_1)^T, \ldots, (X_n, Y_n)^T$*, which are almost surely in a general position for* $n > 2$*. Let* ϵ_n^0 *denote the finite-sample breakdown point of an initial estimator of* $\boldsymbol{\beta}$ *in the model*

$$Y_i = \beta_0 + \beta_1 X_i + e_i, \quad i = 1, \ldots, n, \tag{23}$$

We use the LWS method with data-dependent adaptive weights [2] to estimate parameters in the linear regression model of $Y = (Y_1, \ldots, Y_n)^T$ *against* $X = (X_1, \ldots, X_n)^T$*. Then the finite-sample breakdown point of* $\tilde{r}_{LWS}(X, Y)$ *is larger than or equal to*

$$\min\{\epsilon_n^0, \ \{\lfloor (n-1)/2 \rfloor\}/n\}. \tag{24}$$

Proof. The breakdown point corresponds to the percentage of data needed to "break down" [6]. The result follows from [2] who derived the breakdown point of the LWS estimator of regression parameters and can be evaluated as

$$\min\{\epsilon_n^0, \ \{\lfloor (n+1)/2 \rfloor - (p+1)\}/n\}. \tag{25}$$

with $p = 2$. The robust cosine similarity inherits the breakdown properties of the robust estimator of the regression parameters. □

Other properties are derived for the estimator with given magnitudes of weights, considering general (but technical) assumptions $\mathcal{A}$ of [14]. Convergence in distribution will be denoted by $\xrightarrow{\mathcal{D}}$.

Theorem 2. *Let* $(X_1, \ldots, X_n)^T$ *and* $(Y_1, \ldots, Y_n)^T$ *be two sequences of non-random observations. We consider (18) as a special case of (23). Let Assumptions* $\mathcal{A}$ *be fulfilled. It holds for* $n \to \infty$ *that*

$$T = \frac{\tilde{r}_{LWS}(X, Y)}{\sqrt{1 - \tilde{r}_{LWS}(X, Y)}} \sqrt{n-1} \xrightarrow{\mathcal{D}} Z, \tag{26}$$

where Z *is a random variable with normal* $\mathsf{N}(0, 1)$ *distribution.*

Now we study the asymptotic distribution of the robust cosine similarity $\tilde{r}_{LWS}$. We consider the data $(X_1, Y_1)^T, \ldots, (X_n, Y_n)^T$ as a random sample from a bivariate normal distribution.

Theorem 3. *Under Assumptions* $\mathcal{A}$*, it holds that* $\tilde{r}_{LWS}$ *has asymptotically normal distribution for* $n \to \infty$*. For large sample size* n *and assuming* $\cos\theta \in (-1, 1)$*, the distribution of* $\tilde{r}_{LWS}$ *is approximated by*

$$\mathsf{N}\left(\cos\,\theta, \frac{1}{n}\right). \tag{27}$$

Proof. The asymptotic normality follows from the asymptotic normality of the least weighted squares estimator [14] and the relationship (20). The asymptotic expectation and variance of $\tilde{r}_{LWS}$ are equal to the expectation and variance of $\cos\,\theta$. □

4 Conclusions

This paper proposes new versions of distance measures applicable to multivariate continuous data contaminated by outliers. At the same time, the measures are reliable also for high-dimensional data with the number of variables largely exceeding the number of observations.

For high-dimensional data, the main obstacle of the traditional Mahalanobis distance is singularity of the empirical covariance matrix. Various regularized versions have been proposed, mainly with the aim to solve a classification task by means of a regularized linear discriminant analysis [10]. We propose a new concept of a doubly regularized Mahalanobis distance in Section 2, which combines a regularization of the covariance matrix with regularizing also the means. The classical means are replaced by their counterparts. We present explicit expressions for some important choices of regularization and discuss the possibility of an efficient computation of the new measure using two different tools of numerical linear algebra. The regularization is commonly understood as a tool for ensuring a (local) robustness against small departures in the observed data [4] and thus we can interpret regularized means as denoised versions of the classical means [11]. Nevertheless, the regularization itself cannot ensure robustness against severe outliers for continuous data in terms of the breakdown point.

A robust cosine similarity measure is proposed in Section 3. Its properties are derived including its asymptotic behavior for the number of observations going to infinity. The method is robust in a global sense, i.e. in terms of the breakdown point. This appealing robustness is ensured throughout the implicit weighting of individual observations, while reducing the influence of individual observations with the smallest contribution to the similarity between two variables. Compared to the classical cosine similarity, the novel highly robust cosine similarity measure may be more suitable for practical applications in cluster analysis (or text mining) replacing each of the variables (16) by the prototypes in the form of the means of two different clusters of p-dimensional data.

Acknowledgements

The work was financially supported by the grant 13-01930S of the Czech Science Foundation and by the Neuron Fund for Support of Science.

References

[1] Belloni, A. and Chernozhukov, V.: Least squares after model selection in high-dimensional sparse models. *Bernoulli* **19**, 2 (2013), 521–547.

[2] Čížek, P.: Semiparametrically weighted robust estimation of regression models. *Computational Statistics & Data Analysis* **55**, 1 (2011), 774–788.

[3] Guo, Y., Hastie, T. and Tibshirani, R.: Regularized discriminant analysis and its application in microarrays. *Biostatistics* **8**, 1 (2007), 86–100.

[4] Hansen, P.C.: *Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion.* SIAM, Philadelphia, 2007.

[5] Hastie, T., Tibshirani, R. and Friedman, J.: *The elements of statistical learning.* Springer, New York, 2009.

[6] Jurečková, J., Sen, P.K. and Picek, J.: *Methodology in robust and nonparametric statistics.* CRC Press, Boca Raton, 2012.

[7] Kalina, J.: Highly robust statistical methods in medical image analysis. *Biocybernetics and Biomedical Engineering* **32**, 2 (2012), 3–16.

[8] Kalina, J.: Implicitly weighted methods in robust image analysis. *Journal of Mathematical Imaging and Vision* **44**, 3 (2012), 449–462.

[9] Kalina, J. and Vlčková, K.: Robust regularized cluster analysis for high-dimensional data. *32nd International Conference MME 2014, Mathematical Methods in Economics*, Univerzita Palackého, Olomouc, 2014, 378–383.

[10] Pourahmadi, M.: *High-dimensional covariance estimation.* Wiley, Hoboken, 2013.

[11] Tibshirani, R., Hastie, T. and Narasimhan, B.: Class prediction by nearest shrunken centroids, with applications to DNA microarrays. *Statistical Science* **18**, 1 (2003), 104–117.

[12] Turchi, M., Perrotta, D., Riani, M. and Cerioli, A.: Robustness issues in text mining. *Proceedings of the 6th conference of Soft Methods in Probability and Statistics (SMPS)*, Konstanz, 2012, 263–272.

[13] Varian, H.R.: Big data: New tricks for econometrics. *Journal of Economic Perspectives* **28**, 2 (2014), 3–28.

[14] Víšek, J.Á.: Consistency of the least weighted squares under heteroscedasticity. *Kybernetika* **47**, 2 (2011), 179–206.

[15] Zou, H. and Hastie, T.: Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B* **67**, 2 (2005), 301-320.

A Note on Optimal Value of Loans

Vlasta Kaňková[1]

Abstract. People try to gain (in the last decades) own residence (a flat or a little house). Since young people do not posses necessary financial resources, the bank sector offers them a mortgage. Of course, the aim of any bank is to profit from such a transaction. Therefore, according to their possibilities, the banks employ excellent experts to analyze the financial situation of potential clients. Consequently, the banks know what could be a maximal size of the loan (in dependence on the debtor's position, salary and age) and what is a reasonable size of the installments. The aim of this contribution is to analyze the situation from the second side. In particular, the aim is to investigate the possibilities of the debtors not only in dependence on their present–day situation, but also on their future private and subjective decisions and on possible "unpleasant" events. Moreover, consequently according to these indexes, the aim of this contribution is to suggest a method for a recognition of a "safe" loan and simultaneously to offer tactics to state a suitable environment for future time. The stochastic programming theory will be employed to it.

Keywords: Loan, debtor, installments, multistage stochastic programming.

JEL classification: C44
AMS classification: 90C15

1 Introduction

Decisions of households, one of which taking of a mortgage is, are usually studied by microeconomics [2]. A standard approach to the analysis of such a decision would be to quantify subjective gains from the living in own estate and compare them with a discomfort incurred by the repayment of the mortgage. The analysis of risks associated with a such a decision, is less common, both in theory and, unfortunately, in practice. Neglecting the risks, nevertheless, can easily lead to situations, considered as catastrophic by the decision makers. The aim of the present paper is to outline a methodology, which could be used to a responsible analysis of risks associated with mortgage taking from the debtor's point of view.

In the paper, first, an example of a "classical" situation will be explained (Section 2), followed by stochastic programming models (Section 3). A simple stochastic programming problems will be constructed employing the original example (Section 4). Conclusion can be found in Section 5.

2 Problem Analysis – Example

Let us start with simple standard situation. A young married couple wants to gain own flat. Evidently, these young people have first to decide if they prefer flat or a little house. This decision depends on their nature, financial possibilities and conditions about loans (in this time). Young people are responsible and so they will try to analyze their possibilities. To this end let us assume that (in the start time) their a monthly income is

$$Z_0 = U_0 + V_0, \quad \text{where} \quad U_0 \quad \text{is an income of husband and} \quad V_0 \quad \text{is an income of wife.}$$

Evidently, this income can be divided into three parts Z_0^1, Z_0^2, Z_0^3, where Z_0^1 denotes means for a basic consumption, Z_0^2 denotes means that can be employed for a repayment of installments and Z_0^3 can be

[1]Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic, kankova@utia.cas.cz

considered as an allocation to saving. Consequently

$$Z_0 = Z_0^1 + Z_0^2 + Z_0^3, \quad Z_0^1,\, Z_0^2 > 0,\, Z_0^3 \geq 0. \tag{1}$$

Given the annuity repayments, which is the most standard way of repaying the loan and if we denote by a symbol M the value of the loan, by m number of identical installments and by ζ the loan interest rate, then the identical installments $b(M) := b(\zeta)$ in time points $t = 1, 2, \ldots, m$ (see, e.g., [7] or [9]) are given by

$$b(M) := b(\zeta) = \begin{cases} \frac{M\zeta}{1-v^m}, & \zeta \neq 0, \quad v = v(\zeta) = (1+\zeta)^{-1}, \\ \frac{1}{m}, & \zeta = 0. \end{cases} \tag{2}$$

It follows from the relations (1), (2) that (in the case when $\zeta \neq 0$) it is desirable (in "static" approach) the following inequality

$$\frac{M\zeta(1+\zeta)^m}{(1+\zeta)^m - 1} \leq Z_0^2 \tag{3}$$

to be fulfilled. Of course, this condition (in the extreme case) can be replaced by the inequality

$$\frac{M\zeta(1+\zeta)^m}{(1+\zeta)^m - 1} \leq Z_0^2 + Z_0^3. \tag{4}$$

If it is possible to assume that the relations (1), (2) will be fulfilled also in future, then the young people can take the loan equal to the maximal value M for which the inequality (3) (respective (4)) is fulfilled. However mostly it is necessary to assume that the financial situation of young married couple can change. For example: it is reasonable to assume that in some time period, say (m_1, m_2), $0 < m_1 < m_2 \leq m$ the married couple plan to have a baby. According to this fact and to the social politics of a state the young people can assume the less income in this time, approximately equal to

$$Z_1 = U_0 + V_1 = Z_0^1 + Z_1^2 + Z_1^3, \quad Z_1^2,\, Z_1^3 \geq 0,$$

where V_1 is the supposed income of wife in the time interval (m_1, m_2); Z_1^2 denotes the means, that can be employed for a repayment of installments (of course $Z_1^2 \leq Z_0^2$) and Z_1^3 saved amount in every year of this time interval (of course mostly $0 \leq Z_1^3 \leq Z_0^3$). Evidently without financial reserve the inequalities

$$Z_0^1 + Z_0^2 \leq U_0 + V_1$$

need to be fulfilled. Consequently, if

$$U_0 + V_1 < Z_0^1 + Z_0^2,$$

then a very serious trouble could arise. However, if the young couple saved every time point $t \in (1, \ldots, m_1 - 1)$ the amount Z_0^3 and if the inequality

$$\frac{(m_2 - m_1)M[\zeta(1+\zeta)^m]}{(1+\zeta)^m - 1} \leq (m_2 - m_1)[Z_0^2 - Z_1^2] + (m_1 - 1)Z_0^3 \tag{5}$$

is fulfilled, then they endure the time period (m_1, m_2) without financial troubles.

To construct the relation (5), it has been assumed that the amount Z_0^3 is deterministic, the same in every time point $t \in (1, \ldots, m_1 - 1)$ and that this amount can not be changed. However this situation can be a little different. To explain a new approach we suppose $m_1 = 3, \quad m_2 - m_1 = 2$ and one of the situations:

A 1. The deterministic value Z_0^3 (in the relation (1)) can be replaced by random values $Z_0^3(t)$; $Z_0^3(t)$, $t \in (1, m_1 - 1)$ with probability one positive. Consequently the deterministic income $Z_0 = Z_0^1 + Z_0^2 + Z_0^3$ is replaced by random $Z_{0,0} = Z_0^1 + Z_0^2 + Z_0^3(1)$ in the start point $t = 1$ and by $Z_{0,1} = Z_0^1 + Z_0^2 + Z_0^3(2)$ in the time point $t = 2$. Furthermore it is reasonable to assume that young people can these random amount invest (for example) into two assets to obtain:

a

in the first year the value	$\xi_{0,1}x_{0,1} + \xi_{0,2}x_{0,2}$
under the assumptions	$x_{0,1} + x_{0,2} \leq Z_0^3(1), \quad x_{0,1}, x_{0,2} \geq 0,$

- in the second year the value $\quad \xi_{1,1}x_{1,1} + \xi_{1,2}x_{1,2}$
  under the assumptions $\quad x_{1,1} + x_{1,2} \le Z_0^3(2), \quad x_{1,1}, x_{1,2} \ge 0$

  (under the assumptions that the profit in the time $t = 1$ can not influence the invested amount in the time $t = 2$). Evidently, it is desirable (for young people) the fulfilling of the relation

$$\frac{(m_2 - m_1)M[\zeta(1+\zeta)^m]}{(1+\zeta)^m - 1} \le (m_2 - m_1)[Z_0^2 - Z_1^2] + \sum_{i=0}^{1}[\xi_{i,1}x_{i,1} + \xi_{i,2}x_{i,2}], \tag{6}$$

  and of course the maximization of a possible profit, or

b.
  in the first year the value $\quad \xi_{0,1}x_{0,1} + \xi_{0,2}x_{0,2}$
  under the assumptions $\quad x_{0,1} + x_{0,2} \le Z_0^3(1), \quad x_{0,1}, x_{0,2} \ge 0,$

- in the second year the value $\quad \xi_{1,1}x_{1,1} + \xi_{1,2}x_{1,2}$
  under the assumptions $\quad x_{1,1} + x_{1,2} \le Z_0^3(2) + \xi_{0,1}x_{0,1} + \xi_{0,2}x_{0,2}, \quad x_{1,1}, x_{1,2} \ge 0.$

  (The profit obtained in the time $t = 1$ can be invested in the time moment $t = 2$).

Evidently, it is desirable (for young people) that the following relation holds:

$$\frac{(m_2 - m_1)M[\zeta(1+\zeta)^m]}{(1+\zeta)^m - 1} \le (m_2 - m_1)[Z_0^2 - Z_1^2] + \xi_{1,1}x_{1,1} + \xi_{1,2}x_{1,2} \tag{7}$$

and of course the maximization of a total profit.

Remark. $Z_0^3(1)$, $Z_0^3(2)$, $\xi_{0,1}$, $\xi_{0,2}$, $\xi_{1,2}$, $\xi_{1,2}$ are generally supposed to be random variables with "positive support". Consequently, it is necessary to "specify" the sense of relations in A.1. In details, it is necessary to "specify" when the operator of mathematical expectation, probability constraints, risk constraints or stochastic dominance constraints are employed in the optimization problems.

A.2 $Z_0^3(1)$, $Z_0^3(2)$ have a deterministic character. Let us assume that these amounts can be investigated into two assets (portfolio) with returns $\bar{\xi}_{0,1}$, $\bar{\xi}_{0,2}$, $\bar{\xi}_{1,1}$, $\bar{\xi}_{1,2}$. Mathematically saying, it is possible to determine $x_{0,1}$, $x_{0,2}$, $x_{1,1}$, $x_{1,2}$ fulfilling the relations

$$x_{0,1} + x_{0,2} \le Z_0^3(1), \quad x_{0,1},\ x_{0,2} \ge 0,$$
$$x_{1,1} + x_{1,2} \le Z_0^3(2), \quad x_{1,1},\ x_{1,2} \ge 0$$

to obtain random values
$$g_0 = \bar{\xi}_{0,1}x_{0,1} + \bar{\xi}_{0,2}x_{0,2},$$
$$g_1 = \bar{\xi}_{1,1}x_{1,1} + \bar{\xi}_{1,2}x_{1,2}.$$

Evidently, it is possible also to define random values Y_0, Y_1 by the following relation

$$\begin{aligned} Y_0 &= \tfrac{1}{2}\bar{\xi}_{0,1} + \tfrac{1}{2}\bar{\xi}_{0,2}, \\ Y_1 &= \tfrac{1}{2}\bar{\xi}_{1,1} + \tfrac{1}{2}\bar{\xi}_{1,2}. \end{aligned} \tag{8}$$

g_1, Y_1 are random values "depending" on $Z_0^3(1)$, and g_2, Y_2 "depending" on $Z_0^3(2)$. Employing the theory of a stochastic dominance [8] it is "reasonable" to determine $x_{0,1}$, $x_{0,2}$, $x_{1,1}$, $x_{1,2}$ such that

$$\begin{aligned} &F_{g_0} \succeq_1 F_{Y_0}, \qquad F_{g_1} \succeq_1 F_{Y_1}, \\ \text{or} \quad &F_{g_0} \succeq_2 F_{Y_0}, \qquad F_{g_1} \succeq_2 F_{Y_1}. \end{aligned} \tag{9}$$

(Symbols $\succeq_1, \succeq_2$ denote first and second order stochastic dominance; F_{g_0}, F_{g_1}, F_{Y_0}, F_{Y_1} distribution functions of g_0, g_1, Y_0, Y_1.) The definition of the stochastic dominance will be given in the next section. Moreover, we can evaluate the decision of $x_{0,1}, x_{0,2}, x_{1,1}, x_{1,2}$ for example by linear forms

$$c_{1,1}x_{1,1} + c_{12}x_{1,2}, \qquad c_{2,1}x_{2,1} + c_{2,2}x_{2,2}, \tag{10}$$

with $c_{0,1}$, $c_{0,2}$, $c_{1,1}$, $c_{1,2}$ considered generally to be random.

In the introduction we have tried to give a simple analysis of debtor's situation (for a time interval $(0,\, m_2)$) under very simple conditions. We have neglected many troubles and situations that can happen (e.g. illness, a loss of employment). We also omitted a possibility to gain "better" career or only increasing salary. In the next section we shall try to recall a survey of suitable mathematical models corresponding to introduced situations.

3 Stochastic Programming Problems

In this section we try to recall suitable types of the stochastic programming problems in static setting. To this end let $(\Omega, \mathcal{S}, P)$ be a probability space; $\xi\, (:= \xi(\omega) = (\xi_1(\omega), \ldots, \xi_s(\omega),\, \omega \in \Omega)$ an s–dimensional random vector defined on $(\Omega, \mathcal{S}, P)$; $F (:= F_\xi(z), z \in R^s)$ the distribution function of ξ; P_F the probability measure corresponding to F. Let, moreover, $g_0 (:= g_0(x, z))$ be a real-valued function defined on $R^n \times R^s$; $X_F \subset X \subset R^n$ a nonempty set generally depending on F, $X \subset R^n$ a nonempty "deterministic" set. If E_F denotes the operator of mathematical expectation corresponding to F and if for $x \in X$ there exists $\mathsf{E}_F g_0(x,\, \xi)$, then one-stage (static) "classical" stochastic optimization problem can be introduced ([6], [8]) in the form:

$$\text{Find} \qquad \varphi(F,\, X_F) = \inf\{\mathsf{E}_F g_0(x, \xi) | x \in X_F\}. \tag{11}$$

To our purpose we recall only special cases of X_F. We consider the case $X_F = X$ "deterministic" constraints; the case when there exist functions $\bar{g}_i (:= \bar{g}_i(x),\, x \in R^n)$, $i = 1, \ldots, s$ such that

- either

$$\begin{aligned} X_F (:= X_F(\alpha)) \quad &= \quad \bigcap_{i=1}^{s} \{x \in X : P_F[\omega : \bar{g}_i(x) \le \xi_i] \ge \alpha_i\}, \\ & \quad\quad \alpha_i \in (0,\, 1),\, i = 1, \ldots, s, \quad \alpha = (\alpha_1, \ldots, \alpha_s), \end{aligned} \tag{12}$$

- or

$$\begin{aligned} X_F (:= X_F(u_0,\, \alpha)) \quad &= \quad \bigcap_{i=1}^{s} \{x \in X : \min_{u^i}\{P_F[\omega : L_i(x,\, \xi) \le u^i] \ge \alpha_i\} \le u_0^i\}, \\ & \quad\quad u_0^i > 0,\, \alpha_i \in (0,\, 1),\, i = 1, \ldots, s, \\ & \quad\quad u_0 = (u_0^1, \ldots, u_0^s),\ \alpha = (\alpha_1, \ldots, \alpha_s), \\ L_i(x,\, z) \quad &= \quad \bar{g}_i(x) - z_i,\, i = 1, \ldots, s,\ z = (z_1, \ldots, z_s). \end{aligned} \tag{13}$$

Evidently, the case (12) corresponds to a special class of individual probability constraints and $L_i(x,\, z)$, $i = 1, \ldots, s$, in the case (13), can be considered as loss functions (for more details see, e.g., [5]).

Second order stochastic dominance constraints are the last considered type of constraints. To recall them let $g (:= g(x,\, z)) := g(x,\, \xi)$ be a function defined on $R^n \times R^s$, $Y (:= Y(z)) := Y(\xi)$ a random value with the distribution function F_Y. We can define first and second stochastic dominance constraints by:

- first order

$$X_F \;=\; \{x \in X : F_{g(x,\, \xi)}(u) \le F_Y(u) \quad \text{for every}\, u \in R^1\}, \tag{14}$$

- second order

$$X_F \;=\; \{x \in X : F^2_{g(x,\, \xi)}(u) \le F^2_Y(u) \quad \text{for every} \quad u \in R^1\}, \tag{15}$$

where $\quad F^2_{g(x,\, \xi)}(u) = \int\limits_{-\infty}^{u} F_{g(x,\, \xi)}(y)dy, \quad F^2_Y(u) = \int\limits_{-\infty}^{u} F_Y(y)dy, \quad u \in R^1.$

(For more information about stochastic dominance see, e.g., [8]).

4 Simple Mathematical Models

In this section we try to introduce simple optimization models in multiobjective (and multiperiod) setting corresponding to A.1. To this end we have to recall and generalize the notions mentioned firstly in the Introduction:

- M .... value of loan,
- m .... number of identical installments,
- ζ .... interest rate corresponding to loan,
- Z_t .... income of young married couple in time point $t \in \{0, 1, \dots, m\}$,
- U_t .... income of husband in time point $t \in \{0, 1, \dots, m\}$,
- V_t .... income of wife in time point $t \in \{0, 1, \dots, m\}$,
- Z_t^1 .... means determined for basic consumption in time point $t \in \{0, 1, \dots, m\}$,
- Z_t^2 .... means determined for repayment of installment in time point $t \in \{0, 1, \dots, m\}$,
- Z_t^3 .... allocation (maybe random) for saving in time point $t \in \{0, 1, \dots, m\}$,
- (m_1, m_2) ...... time interval in which income of wife is supposed to be smaller,
- $\xi_{t,j}$, $t = 0, 1, \dots, m$, $j = 1, 2$ random returns in time t and asset j in the approach A.1a,
- $\bar{\xi}_{t,j}$, $t = 0, 1, \dots, m$, $j = 1, 2$ random returns in time t and asset j in the approach A.1b,
- $x_{t,j}$, $\bar{x}_{t,j}$, $t = 0, \dots, m$, $j = 1, 2$, ... decision variables,
- F ...... a distribution function covering all random values occur that in the corresponding model.

First we generalize the approach of the situation A. 1a: Z_t^3 are for $t \in (1, m_1-1) \bigcup (m_2+1, m)$ supposed to be with probability one positive. Moreover, we assume that the corresponding amount can be investigated (of course in the case of positive value) in two assets with random returns $\xi_{t,1}$, $\xi_{t,2}$. If moreover we can assume that the profit obtained in the time point $t \in \{1, \dots, m\}$ can not be investigated in the time $t+1, \dots, m$, then evidently one of the possible corresponding stochastic optimization problem can be constructed as following:

$$\text{Find} \qquad \max M \tag{16}$$

under the system of constraints

$$\frac{M\zeta(1+\zeta)^m}{(1+\zeta)^m - 1} \le Z_t^2, \quad t = 0, 1, \dots, m_1 - 1, \tag{17}$$

$$P_F\{x_{t,1} + x_{t,2} \le Z_t^3\} \quad \ge \quad 1 - \varepsilon_t, \quad \varepsilon_t \in (0,1), \; x_{t,1}, x_{t,2} \ge 0, \quad t = 0, 1, \dots m_1 - 1, \tag{18}$$

$$P_F\{\frac{(m_2 - m_1)M[\zeta(1+\zeta)^m]}{(1+\zeta)^m - 1} \le \sum_{i=m_1}^{m_2} [Z_i^2 - Z_0^2] + \sum_{i=0}^{m_1-1})[\xi_{i,1}x_{i,1} + \xi_{i,2}x_{i,2}]\} \ge 1 - \varepsilon_0, \quad \varepsilon_0 \in (0, 1). \tag{19}$$

Evidently, in this case it is reasonable to add to an objective function (16) the second one

$$\mathsf{E}_F \sum_{i=0}^{m} [\xi_{i,1}x_{i,1} + \xi_{i,2}x_{i,2}] \tag{20}$$

with the corresponding constraints

$$\begin{aligned} P_F\{x_{t,1} + x_{t,2} \le \max(0, Z_t^3)\} \quad &\ge \quad 1 - \varepsilon_t, \; \varepsilon_t \in (0, 1), \; x_{t,1}, x_{t,2} \ge 0, \quad t = m_1, \dots m_2 - 1, \\ P_F\{x_{t,1} + x_{t,2} \le Z_t^3\} \quad &\ge \quad 1 - \varepsilon_t, \; \varepsilon_t \in (0, 1), \; x_{t,1}, x_{t,2} \ge 0, \quad t = m_2 + 1, \dots, m. \end{aligned} \tag{21}$$

Consequently, we have constructed two objective stochastic programming problem with objective (16) and (20) and constraints (17), (18), (19) and (21).

Starting with the situation A.1b we can obtain the problem:

$$\text{Find} \qquad \max M \tag{22}$$

under the system of constraints

$$\frac{M\zeta(1+\zeta)^m}{(1+\zeta)^m - 1} \le Z_t^2, \qquad t = 0, \ldots, m_1 - 1, \tag{23}$$

$$\begin{aligned} &P_F\{\bar{x}_{0,1} + \bar{x}_{0,2} \le Z_0^3\} \ge 1 - \varepsilon_0, \ \varepsilon_0 \in (0, 1), \ \bar{x}_{0,1}, \bar{x}_{0,2} \ge 0, \\ &P_F\{\bar{x}_{t,1} + \bar{x}_{t,2} \le \bar{Z}_t^3\} \ge 1 - \varepsilon_t, \ \varepsilon_t \in (0, 1) \quad \bar{x}_{t,1}, \bar{x}_{t,2} \ge 0, \quad t = 1, \ldots, m_1 - 1, \\ &\bar{Z}_t^3 = \max(0, Z_t^3) + \bar{\xi}_{t-1,1}\bar{x}_{t-1,1} + \bar{\xi}_{t-1,2}\bar{x}_{t-1,2}, \quad t = 1, \ldots, m, \end{aligned} \tag{24}$$

$$P_F\{\frac{(m_2 - m_1)M[\zeta(1+\zeta)^m]}{(1+\zeta)^m - 1} \le \sum_{i=m_1}^{m_2} [Z_i^2 - Z_0^2] + [\bar{\xi}_{m_2,1}\bar{x}_{m_2,1} + \bar{\xi}_{m_2,2}\bar{x}_{m_2,2}]\} \ge 1 - \varepsilon_0, \tag{25}$$

Evidently, in this case it is also reasonable to add to the objective function (22) the second one

$$\mathsf{E}_F[\bar{\xi}_{m,1}\bar{x}_{m,1} + \bar{\xi}_{m,2}\bar{x}_{m,2}] \tag{26}$$

and the corresponding constraints

$$P_F\{\bar{x}_{t,1} + \bar{x}_{t,2} \le Z_t^3 + \bar{\xi}_{t-1,1}\bar{x}_{t-1,1} + \bar{\xi}_{t-1,2}\bar{x}_{t-1,2},\} \quad \ge \quad 1 - \varepsilon_t, \ t = m_2 + 1, \ldots, m, \tag{27}$$

Remark. We have supposed (for simplicity) that a profit from the investigation in the time interval $(0, m_2)$ is included in the condition (25) and can not be employed in the time $t = m_2 + 1, \ldots, m$

5 Conclusion

In the last decades many people try to gain their own residence. Since they do not posses sufficient means, the bank sector offer them the loan. The aim of this contribution is to give a preliminary analysis of their situations and possible responsible behaviour. Three approaches have been analyzed in a very simple examples, two of them have been employed for a construction of stochastic optimization models. The results of [1], [3], [4] can be employed to investigate properties of these models. Employing these methodology a risk for young people can happen only with very small prescribed probability. However to deal with this new problem is over the possibilities of this contribution.

Acknowledgment This work was supported by the Czech Science Foundation under grant 15-10331S.

References

[1] Birge, J.R. and Louveaux, F.: *Introduction in Stochastic Programming.* Springer, Berlin 1999.

[2] Frank, Robert H., and Amy Jocelyn Glass: *Microeconomics and Behavior.* Mcgraw–Hill, New York 1991

[3] Kaňková, V. and Šmíd, M.: On Approximation in Multistage Stochastic Programs; Markov Dependence. *Kybernetika* **40** (2004), 5, 625–638.

[4] Kaňková, V.: Multistage stochastic programs via Autoregressive sequences and individual probability constraints. *Kybernetika* **44** (2008), 2, 151–170.

[5] Kaňková, V.: Risk measures in optimization problems via empirical estimates. *Czech Economic Review, Acta Universitatis Carolinea* **7** (2012), 3, 162–177.

[6] Kaňková, V. and Houda, M.: Thin and Heavy Tails in Stochastic Programming. *Kybernetika* **51** (2015), 3, 433–455.

[7] Luderer, B., Nollau, V. and Vetters, Klaus: *Mathematical Formulas for Economists.* Springer Science & Media, third edition, 2006.

[8] Shapiro, A., Dentcheva, D. and Ruszczynski, A: *Lectures on Stochastic Programming (Modeling and Theory).* Published by Society for Industrial and Applied Mathematics and Mathematical Programming Society, Philadelphia 2009.

[9] Šmíd, M. and Dufek, J.: Multi–period Factor Model of Loan Portfolio (July 10, 2016). Available at SSRN:http://ssrn.com/abstract=2703884 or http://dx.doi.org/10.2139/ssrn.2703884.

Google Trends and Exchange Rate Movements: Evidence from Wavelet Analysis

Svatopluk Kapounek[1], Zuzana Kučerová[2]

Abstract. Empirical exchange rate models that try to explain and forecast the movements of exchange rates do not serve as a good way for proper testing and forecasting of exchange rate movements nowadays because large exchange rate swings could be better explained by institutional, behavioural and other determinants and by market expectations than by economic fundamentals (such as interest parity condition, inflation differential etc.). The objective of the paper is to examine the short- and long-term co-movements between exchange rate of the US dollar and euro Euro and internet search data provided from Google Trends. Google Trends data provide monthly information about search intensity via search query indices on selected variables and they can serve as a proxy for the market expectations of these variables. We use the method of wavelet coherence analysis with phase shift to identify the causality in Granger sense. We apply Monte Carlo method to estimate the significance of results and edge effects. Our results indicate that data from Google Trends help to explain the movement of exchange rates.

Keywords: Google searches, exchange rates, investor attention, FX volatility.

JEL Classification: F31
AMS Classification: 65T60

1 Introduction

Exchange rates and its movements could be explained by established economic theories and by many empirical exchange rate models that try to explain and forecast the movements of exchange rates. However, these traditional empirical models do not serve as a good way for proper testing and forecasting of exchange rate movements nowadays because, in many cases, large exchange rate swings could be better explained by institutional, behavioural and other determinants and by market expectations than by economic fundamentals. Therefore, it is important to study other factors that might explain the exchange rate dynamics.

Meese and Rogoff [23] compare the out-of-sample forecasting accuracy of structural and time series exchange rate models from 1973 to 1981 to find whether these forecast are reliable or not. They came to surprising conclusion that a random walk model performs no worse than any estimated time series model of selected currency pairs in the time horizon of twelve months (the "Meese-Rogoff Puzzle"). As Frankel and Rose [12] claim this finding negatively influenced the modelling of exchange rates. Many authors tried to confirm or refute this conclusion, e.g. Bacchetta and Wincoop [1], Cheung et al. [4], Christoffersen and Diebold [6], Engel et al. [8], Engel and Hamilton [7], Engel and West [9], Faust et al. [10], Gourinchas and Rey [14], Leitch and Tanner [19], Mark [20], Meese [22], Tashman [26].

Moreover, we face another problem besides the problem of forecasting performance of models: substantial delays in the official release of statistical data. Therefore, policy makers are forced to make decisions when all up-to-date economic data are not available. According to Choi and Varian [5], there are sources of data providing real-time economic activity data available from private sector companies (Google, MasterCard, Federal Express, UPS, Intuit, etc.). As Castle et al. [3] state, economic policy makers have recently focused on the forecasting the current state (i.e. predicting the present) or more precisely "nowcasting". Koop and Onorante [17] define nowcasting as a method that uses currently available data to provide timely estimates of economic variables before their official estimates are produced.

We join the strand of research using the concept of nowcasting as a method based on using the data provided by social media (i.e. the usage of Big Data). Google Trends data provide monthly information about search in-

[1] Mendel University in Brno / Faculty of Business and Economics, Department of Finance, Zemědělská 1, Brno, Czech Republic, e-mail: kapounek@mendelu.cz.
[2] VSB-Technical University of Ostrava / Faculty of Economics, Department of National Economy, Sokolská třída 33, Ostrava, Czech Republic, e-mail: zuzana.kucerova@vsb.cz.

tensity via search query indices on selected variables and as such they can help us to measure the market expectations of these variables.

The objective of the paper is to examine the short- and long-term co-movements between exchange rate of the US dollar and euro and internet search data provided from Google Trends. We use the method of wavelet coherence analysis with phase shift to identify the causality in Granger sense. Then, we apply the Monte Carlo method to establish significance levels and confidence intervals for the wavelet power spectrum. The sample covers the period from 2004 to 2016.

The structure of the paper is as follow. Section 2 reviews the literature concerning the use of Google Trends data in forecasting and the problem of nowcasting. Section 3 introduces data and the methods used in the paper. Section 4 discusses our results. Section 5 concludes.

2 Literature Review

There is a growing body of research using Google Trends data for forecasting and nowcasting of selected economic variables. The use of Google Trends data for forecasting economic series was initially described by Choi and Varian [5]. Authors use Google Trends data and conclude that they are often correlated with various economic indicators and help to make short-term predictions.

In our paper, we focus solely on the foreign exchange markets. One of the first papers working with the Google Trends data to study the forecasting performance of exchange rate models is that of Bulut [2]. The author uses internet search data from Google Trends to capture the information set of decision makers and to assess the market expectations of selected macroeconomic fundamentals in a sample of twelve OECD countries for the period 2004-2014. The author concludes that the utilisation of the Google Search Data (i.e. *ex ante* variables) concerning current macroeconomic variables and nowcasting of these variables should be an alternative for proper testing of exchange rate determination models (with *ex post* variables) because of the existence of the lag in the availability of the official data to the market participants. Therefore, he suggests using the Google Trends Data to nowcast the future exchange rate movement.

Goddard et al. [13] study the relationship between investor attention and the dynamics of currency prices using a Google search volume index which measures the search intensity through Google for main currency pairs over the 2004-2011 period. According to the authors, Google Trends data can be used as a measure of acquisition of publicly-available information; compared to other professional trading platforms, Google gathers data from a lot of other sources and thus it can provide investors a highly diversified information set. The authors conclude that changes in investor attention are associated with changes in the holdings of the largest traders in foreign exchange markets when the causality runs mainly form investor attention to market volatility.

Seabold and Coppola [25] focus on foreign exchange markets and use Google Trends data to forecast price series in selected countries in Central America. The authors construct a new index for consumer search behaviour and find that the use of the Google Trends data improves the quality of forecasting in about 20 percent.

3 Data and Methods

To understand relations between the economic agents' attention and exchange rate movement we look at the frequency of appearances of selected words via Google. Therefore, we use searches of keywords received from the tool "Google Trends". Google Trends provide a time series index (from 0 to 100) of the volume of internet search queries on search keywords or phrases.[3] We assume that the indicator of searches represents sentiment of economic agents related to the demand for currency trading. In that sense we used keywords "United States Dollar" and "Euro" with particular emphasize on the searches in the category "Currency". The exchange rate movements were received from the world's trusted currency authority XE.com. We use data in the period from 2004M1 to 2016M3 (weekly frequency).

To understand nowcasting performance of the tool Google Trends we employ time-frequency domain analysis which enables differentiation between the short- and long-term co-movements and its changes in time. This approach is widely used in analysis of business cycle synchronization and dating (e.g. Pomenkova [24]; Kapounek and Pomenkova [16]; Fidrmuc et al. [11]; Kucerova and Pomenkova [18], Marsalek et al. [21]).We apply Continuous Wavelet Transform (CWT) as a band pass filter to time series ($x_n, n = 1, \dots, N$) with uniform time steps δt, where the time step is defined as the convolution of x_n with the scaled and normalized wavelet. We follow Grinsted et al. [15] and define the wavelet power as $|W_n^X(s)|^2$ and:

[3] The normalised search query index at a given point in time is a ratio of the total search volume for each query to the total number of all search queries.

$$W_n^X(s) = \sqrt{\frac{\delta t}{s}} \sum_{n'=1}^{N} x_{n'} \psi_0 \left[(n'-n) \frac{\delta t}{s} \right]. \tag{1}$$

where s represents scale in time. In practice, the complex argument of $W_n^X(s)$ can be interpreted as the local phase. To localize a function in frequency and time we use Morlet wavelet ψ_0 which provides an optimal trade-off between both time and frequency localization (Teolis [27]):

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-1/2\eta^2}, \tag{2}$$

where $\omega_0 = 6$ is dimensionless frequency and $\eta = s \times t$ dimensionless time by varying its scale s. To identify shocks in co-movements between the analysed time series x_n and y_n we apply the Cross Wavelet Transform (XWT):

$$W^{XY} = W^X W^{Y*}, \tag{3}$$

where * denotes complex conjugation (Grinsted et al. [15]). Additionally, we apply Wavelet Coherence (WTC) to identify common time-localized oscillations in nonstationary time series that can be interpreted as co-movement or correlation. Following Torrence and Webster [29] and Grinsted et al. [15] and we define the wavelet coherence of time series x_n and y_n as:

$$R_n^2(s) = \frac{\left| S\left(s^{-1} W_n^{XY}(s)\right) \right|^2}{S\left(s^{-1} \left| W_n^X(s) \right|^2\right) \times S\left(s^{-1} \left| W_n^Y(s) \right|^2\right)}, \tag{4}$$

where smoothing operator S is defined as $S(W) = S_{scale}(S_{time}(W_n(s)))$. S_{scale} represents smoothing operator along the wavelet scale axis and S_{time} smoothing operator in time, suitable or the Morlet wavelet (Torrence and Webster [29]).

Moreover, it is very important to identify a direction of causality which is given by the relative lag between the two time series. In this sense we applied phase shift to identify a time offset between the reflection and the maximum value on the waveform. Thus, we interpret phase shift as a lead or a lag between time series. We follow Grinsted et al. [15] and estimate the mean and confidence interval of the phase difference. The mean phase calculation is based on the circular mean of a set of angles $(a_i, i = 1, \ldots, n)$:

$$\bar{a} = \arg(X, Y) \text{ with } X = \sum_{i=1}^{n} \cos(a_i) \text{ and } Y = \sum_{i=1}^{n} \sin(a_i). \tag{5}$$

For a better understanding this issue, it is comparable to causality in Granger sense. However, the interpreting the phase as a lead or a lag have to be done relatively to the anti-phase, because a lead of 90° is also a lag of 270°.

Finally, we focused on the edge effects because wavelets are not completely localized in time in the case of very low frequencies. We follow concept provided by Torrence and Compo [28] who estimated statistical significance against an autocorrelation model with lag 1 and error term represented as white noise. The same approach was applied to identify significance levels of cross-wavelet power and wavelet coherence.

4 Results

Figure 1 shows time series representation in time and frequency domain. The upper left plot presents the time domain representations; it is apparent that the internet search queries (via Google) of the keyword "Euro" are more volatile than queries of the keyword "United States Dollar" during the analysed time period. We can identify also a few outliers that can cause biases in our results. The upper right plot and the both bottom plots show all analysed time series in frequency representation, after the Continuous Wavelet Transformation with the Morlet wavelet. In case of the "Euro" queries, we can identify two main periods of significant regularities in the time series at higher frequencies related to cycles of less than one year: in the years 2008 and 2012. We can interpret these two periods as periods with significant seasonality which are not particularly important for our results.

Similarly, the significant seasonality we can also find at the exchange rate movements. In this case, we can also identify significant cyclicality between the years 2008 and 2011 for periods longer than one year.

Additionally, all results in frequency domain show edge effects that cannot be ignored. This area is represented by V-shaped curve and lighter shade areas in the pictures. Thick contours show the 5% significance level of wavelet power spectrum.

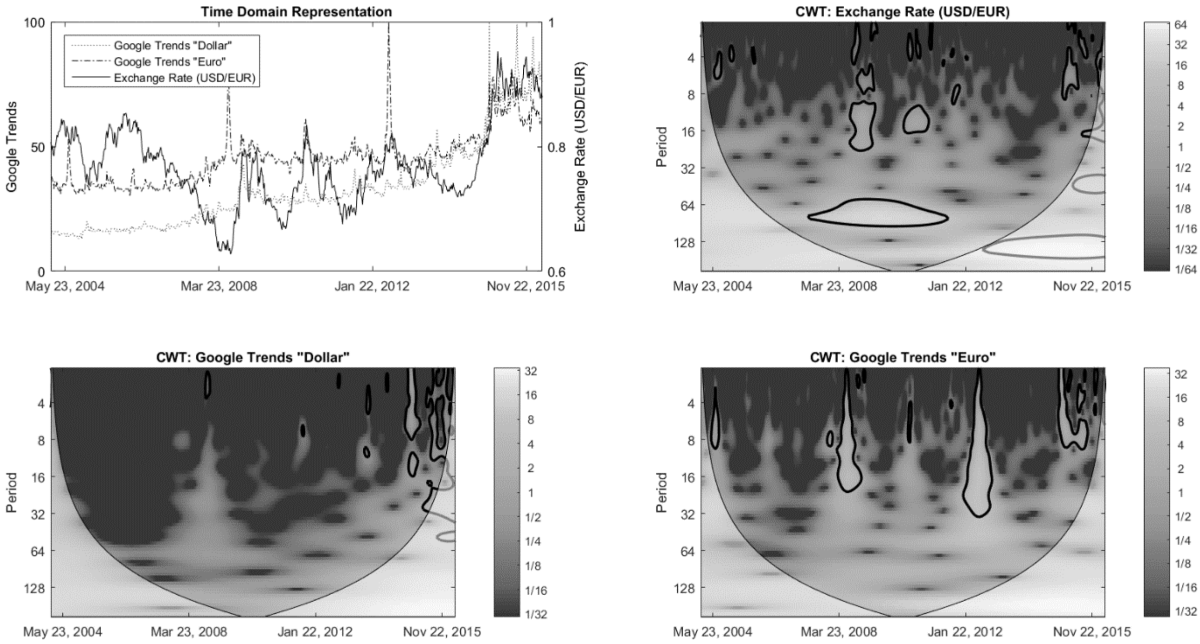


Figure 1 Time Domain Representation and Continuous Wavelet Transform (CWT)

Figure 2 shows the results of Cross Wavelet Transform (XWT) and Wavelet Coherence (WTC). Upper plots present the results of XWT where we can identify significant shocks inside the V-shaped area for the "Euro" currency in the period from 4 to 20. These shocks show peaks in the cross spectrum, which may have nothing to do with any relation of the analysed time series. Thus, it is very important for the robustness of our analyses that there are no significant shocks at longer than yearly periods. In case of the "Dollar" currency, there are almost no significant shocks.

Bottom plots illustrate the results of the time series co-movements. Except the significance tests, we can also identify a phase shift between the analysed time series, represented by arrows. Right arrows indicate that both time series are in-phase, left arrows indicate anti-phase. Thus, we can discuss a phase shift between 0° and 90° in the period from 64 to 128 where arrows turn to the bottom and to the bottom right corner. This effect is significant in the case of the Google searches of keyword "Euro" and the USD/EUR exchange rate (the bottom right plot). Thus, we can conclude that "Euro" Google search inquiries lead exchange rate movements particularly till 2009, after 2010 the character of this phase slightly changes because the arrows start turning to the right. Thus, there is shorter time lag in the co-movement of the Google inquiries and the exchange rate movement. Surprisingly, we didn't receive similar results for keyword "Dollar" (the bottom left plot). We assume that there are different determinants of the both of the analysed currencies.

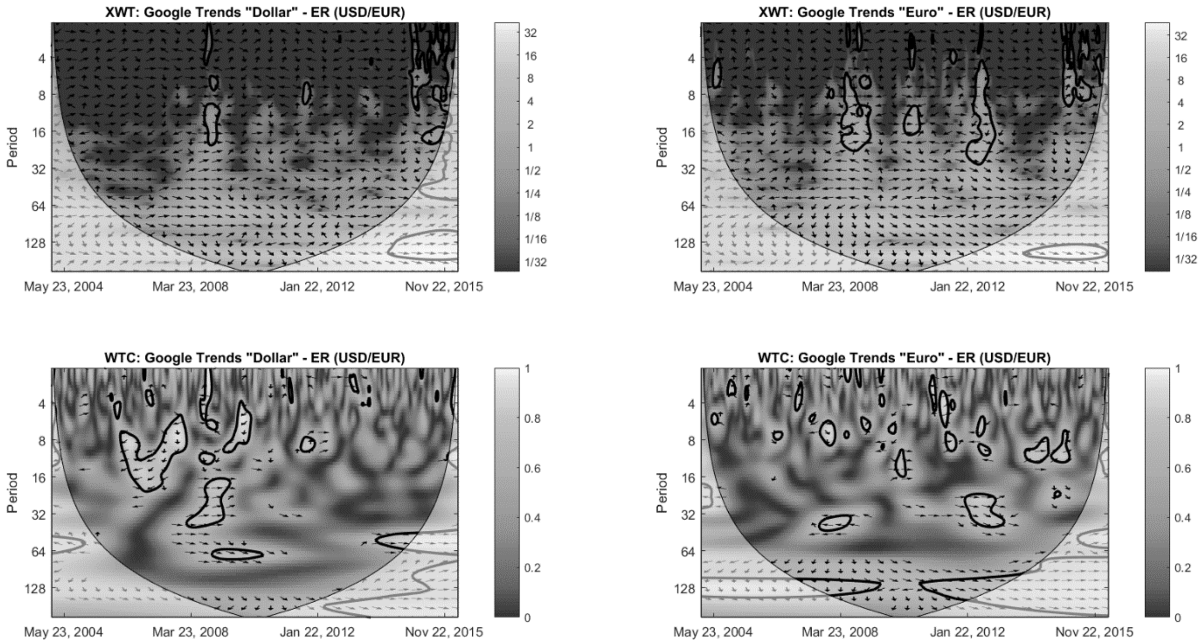


Figure 2 Cross Wavelet Transform (XWT) and Wavelet Coherence (WTC)

5 Discussion and Conclusions

In our paper, we tried to explain whether the Google Trends search inquiries able to predict (or nowcast) the exchange rate movements. Nowadays, there is a growing body of research studies using the Google Trends data for forecasting or, more precisely, nowcasting selected economic variables. The objective of the paper was to examine the short- and long-term co-movements between exchange rate of the US dollar and euro and internet search data provided from Google Trends. We used the method of wavelet coherence analysis with phase shift to identify the causality in Granger sense.

We confirmed the significant impact of the Google Trend search inquiries on the movement of the exchange rate of the euro currency. In this sense, we can conclude that the movement of the exchange rate of euro currency is responsive to the Google searches of the keyword “Euro”. Therefore, the Google Trends data could be used to nowcast the movement of this exchange rate, i.e. the increased economic agents’ attention is able to change this exchange rate. In case of the US dollar exchange rate, these results were not confirmed. In other words, the US dollar does not react to the sentiment of the economic agents. In this context, the US dollar proved to be a stable currency not influenced by the frequency of searches via Google.

Acknowledgements

This research was funded by the Czech Science Foundation, grant No. 16-26353S “Sentiment and its Impact on Stock Markets” and by the VSB-Technical University of Ostrava, Faculty of Economics under Grant SGS SP2016/101.

References

[1] Bacchetta, P., Wincoop van, E.: Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle? *American Economic Review* **96**, 3 (2006), 552–576.
[2] Bulut, L.: Google Trends and Forecasting Performance of Exchange Rate Models. *IPEK Working Paper*, 15–05 (2015). Available at: http://econpapers.ipek.edu.tr/IpekWParchives/wp2015/wp1505Bulut.pdf.
[3] Castle, J. L., Fawcett, N. W. P., Hendry, D. F.: Nowcasting is not just Contemporaneous Forecasting. *National Institute Economic Review* **201**, 1 (2009), 71–89.
[4] Cheung, Y.-W., Chinn, M. D., Pascual, A. G.: Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive? *Journal of International Money and Finance* **24**, 7 (2005), 1150–1175.
[5] Choi, H., Varian, H.: Predicting the Present with Google Trends. *The Economic Record* **88**, June (2012), 2–9.
[6] Christofferson, P. F., Diebold, F. X.: Cointegration and Long-Horizon Forecasting. *Journal of Business and Economic Statistics* **16**, 4 (1998), 450–458.

[7] Engel, C., Hamilton, J. D.: Long Swings in the Dollar: Are They in the Data and Do Markets Know It? *American Economic Review* **80**, 4 (1990), 689–713.

[8] Engel, C., Mark, N., West, K. D.: Exchange Rate Models Are Not as Bad as You Think. *NBER Working Paper* **13318** (2007).

[9] Engel, C., West, K. D.: Exchange Rates and Fundamentals. *Journal of Political Economy* **113**, 3 (2005), 485–517.

[10] Faust, J., Rogers, J. H., Wright, J. H.: Exchange Rate Forecasting: The Errors We've Really Made. *Journal of International Economics* **60**, 1 (2003), 35–59.

[11] Fidrmuc, J., Korhonen, I., Pomenkova, J.: Wavelet spectrum analysis of business cycles of China and G7 countries. *Applied Economics Letters* **21**, 18 (2014), 1309–1313.

[12] Frankel, J. A., Rose, A. K.: Empirical Research on Nominal Exchange Rates. In: Grossman, G. M., Rogoff, K. (eds): *Handbook of International Economics*. Amsterdam: Elsevier, 1995.

[13] Goddard, J., Kita, A., Wang, Q.: Investor Attention and FX Market Volatility. *Journal of International Financial Markets, Institutions and Money* **38**, September (2015), 79–96.

[14] Gourinchas, P. O., Rey, H.: International Financial Adjustment. *Journal of Political Economy* **115**, 4 (2007), 665–703.

[15] Grinsted, A., Moore, J. C., Jevrejeva, S.: Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics* **11** (2004), 561–566.

[16] Kapounek, S., Poměnková, J.: The endogeneity of optimum currency area criteria in the context of financial crisis: Evidence from the time-frequency domain analysis. *Agricultural Economics-Zemedelska Ekonomika* **59**, 9 (2013), 389–395.

[17] Koop, G., Onorante, L.: Macroeconomic Nowcasting Using Google Probabilities [online]. Available at: https://www.ecb.europa.eu/events/pdf/conferences/140407/OnoranteKoop_MacroeconomicNowcastingUsingGoogleProbabilities.pdf?e105896cfaba02ab33265ae4047d96be.

[18] Kucerova, Z., Pomenkova, J.: Financial and Trade Integration of Selected EU Regions: Dynamic Correlation and Wavelet Approach. *Ekonomický časopis* **63** (7), (2015), 686-704.

[19] Leitch, G., Tanner, J. E.: Economic Forecast Evaluation: Profits versus the Conventional Error Measures. *American Economic Review* **81**, 3 (1991) 580–590.

[20] Mark, N.C. Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability. *American Economic Review* **85**, 1 (1995), 201–218.

[21] Marsalek, R., Pomenkova, J., Kapounek, S.: A Wavelet-Based Approach to Filter Out Symmetric Macroeconomic Shocks. *Computational Economics* **44**, 4 (2014), 477–488.

[22] Meese, R. A.: Currency Fluctuations in the Post-Bretton Woods Era. *Journal of Economic Perspectives* 4, 1 (1990), 117–134. Messe, R. A., Rose, A. K.: Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, 14, 1–2 (1983), 3–24. Moosa, I., Burns, K.: Reappraisal of the Meese-Rogoff Puzzle. *Applied Economics* **46**, 1 (2014), 30–40.

[23] Meese, R., Rogoff, K.: Empirical Exchange Rate Models of the Seventies. *Journal of International Economics* **14** (1983), 3-24.

[24] Pomenkova, J.: Business Cycle Identification. *Ekonomicky Casopis* **60**, 9 (2012), 899–917.

[25] Seabold, S., Coppola, A.: Nowcasting Prices Using Google Trends: An Application to Central America. *World Bank Policy Research Working Paper* **3798** (2015).

[26] Tashman, L. J.: Out-of-Sample Tests of Forecasting Accuracy: An Analysis and Review. *International Journal of Forecasting* **16**, 4 (2000), 437–450.

[27] Teolis, A.: *Computational signal processing with wavelets.* Springer, 1998.

[28] Torrence, C., Compo, G. P.: A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society* **79** (1998): 61– 78.

[29] Torrence, C., Webster, P.: Interdecadal changes in the ESNO-Monsoon System. *Journal of Climate* **12** (1999), 2679–2690.

Economic efficiency of AOQL variables sampling plans

Nikola Kaspříková[1]

Abstract. Sampling inspection is a quality control tool used in industry to keep the quality of the products at satisfactory level while at the same time having the cost in control. When using acceptance sampling inspection, a decision on whether the lot of items is to be accepted or rejected is based on results of inspecting a sample of items from the lot. Rectifying acceptance sampling plans minimizing mean inspection cost per lot of the process average quality were designed by Dodge and Romig for the inspection by attributes. Plans for the inspection by variables were then proposed and such plans may be more economical than the corresponding attributes sampling plans. The recently proposed plans using EWMA statistic may lead to further improvements in the inspection cost. The design of the EWMA-based rectifying AOQL sampling plans minimizing the mean inspection cost per lot of process average quality is recalled. The measure for assessing the comparative economic efficiency of the plans, which may be used for getting a guidance for selecting the most appropriate type of the sampling plan to be used, is proposed and the economic characteristics evaluation of the plans is shown.

Keywords: acceptance sampling, inspection cost, optimization, AOQL, EWMA.

JEL classification: C44
AMS classification: 62D99

1 Introduction

Sampling inspection is one of the quality control tools used in industry to help keep the quality of the products at satisfactory level while at the same time having the cost in control. When using acceptance sampling inspection, a decision on whether the lot of items is to be accepted or rejected is based on results of inspecting a sample of items from the lot. There are several ways how acceptance sampling schemes may be classified. One such classification is according to whether an item is inspected by attributes, i.e. just classified as either good or defective (nonconforming) or by variables. Sampling plans for inspection by variables in many cases allow obtaining the same level of the protection as the corresponding sampling plans for the inspection by attributes while using a lower sample size. The basic notions of the variables sampling plans are addressed in [3].

The AOQL sampling plans minimizing the mean inspection cost per lot of process average quality when the remainder of rejected lots is inspected were originally designed by Dodge and Romig (see e.g. [2]) for the inspection by attributes. Plans for the inspection by variables and for the inspection by variables and attributes (all items from the sample are inspected by variables, the remainder of rejected lots is inspected by attributes) were then proposed and it was shown that these plans are in many situations more economical than the corresponding Dodge-Romig attribute sampling plans. The AOQL plans for inspection by variables and attributes have been introduced in [9], using approximate calculation of the plans. Exact plans, using non-central t distribution in calculation of the operating characteristic, have been reported in [10] and implemented in R package LTPDvar [6] The operating characteristics used for these plans are discussed by Jennett and Welch in [3] and by Johnson and Welch in [4]. It has been shown that these plans are in many situations superior to the original attribute sampling plans and similar results have been obtained for the LTPD plans, the analysis is provided in [7] and in [8]. Recent development

[1]University of Economics in Prague, Department of Mathematics,
Nám. W. Churchilla 4, Praha, Czech Republic, nb33@tulipany.cz

of acceptance sampling plans includes the work by Aslam et al. in [1] where the exponentially weighted moving average (EWMA) statistic is used for a design of the (p_1, p_2) sampling plans, i.e. sampling plans which satisfy the requirement to control the producer's risk and the consumer's risk. Using the EWMA statistic enables some savings in the cost of inspection as it allows using information on the quality in the previous lots. With the aim of obtaining further savings in the cost of inspection, the new AOQL plans for the inspection by variables and attributes, designed to use the EWMA statistic, have been proposed in [5]. Using an economic model similar to the model used in [7], a measure for assessing the comparative economic efficiency of the plans, which may be used for getting a guidance for selecting the most appropriate type of the sampling plan to be used, is proposed in this paper and an evaluation of the economic characteristics of the plans is shown. The structure of the paper is as follows: the AOQL plans for the inspection by attributes are recalled first, then the design of the recently introduced EWMA statistic based rectifying known sigma AOQL variable sampling plans minimizing the mean inspection cost per lot of the process average quality is recalled and finally the economic efficiency measure is introduced and the analysis of the economic performance of the plans is provided.

2 Attributes inspection plans

For the case that each inspected item is classified as either good or defective (the acceptance sampling by attributes), Dodge and Romig (see [2]) consider sampling plans (n, c) which minimize the mean number of items inspected per lot of process average quality, assuming that the remainder of the rejected lots is inspected

$$I_s = N - (N - n)\cdot L(\bar{p}; n, c) \tag{1}$$

under the condition

$$\max_{0<p<1} AOQ(p) = p_L. \tag{2}$$

The notation in equations (1) and (2) is as follows:
N is the number of items in the lot (the given parameter),
$\bar{p}$ is the process average fraction defective (the given parameter),
p_L is the average outgoing quality limit (the given parameter, denoted AOQL),
n is the number of items in the sample ($n < N$),
c is the acceptance number (the lot is rejected when the number of defective items in the sample is greater than c),
$L(p)$ is the operating characteristic (the probability of accepting a submitted lot with the fraction defective p).

The function AOQ is the *average outgoing quality, $AOQ(p)$* is the mean fraction defective after inspection when the fraction defective before inspection was p. The average outgoing quality (where all defective items found are replaced by good ones) is approximately

$$AOQ(p) = \left(1 - \frac{n}{N}\right)\cdot p \cdot L(p; n, c). \tag{3}$$

Therefore the condition (2) can be rewritten as

$$\max_{0<p<1} \left(1 - \frac{n}{N}\right)\cdot p \cdot L(p; n, c) = p_L. \tag{4}$$

The condition (2) protects the consumer against having an average outgoing quality higher than p_L (the chosen value), regardless of what the fraction defective p is before inspection.

3 AOQL variables inspection plans

The AQOL plans for the inspection by variables and attributes have been designed in [5] under the following assumptions: the measurements of a single quality characteristic X are independent, identically distributed normal random variables with unknown parameter μ and known parameter σ^2. For the quality characteristic X there is given either an upper specification limit U (the item is defective if its measurement exceeds U), or a lower specification limit L (the item is defective if its measurement is smaller than L).

The inspection procedure is as follows:

Draw a random sample of n items from the lot and compute the sample mean $\bar{x}$ and the statistic T at time t as $T_t = \lambda\bar{x} + (1-\lambda)T_{t-1}$, where λ is a smoothing constant between 0 and 1. The values of the smoothing constant over 0.5 give more weight to the sample in the current lot. Accept the lot if

$$\frac{U - T_t}{\sigma} \geq k \quad \text{or} \quad \frac{T_t - L}{\sigma} \geq k. \tag{5}$$

Suppose that c_s^* is the cost of inspection of one item by attributes and c_m^* is the cost of inspection of one item by variables and that the sample is inspected by variables. Then the inspection cost per lot with proportion defective p, assuming that the remainder of rejected lots is inspected by attributes (the inspection by variables and attributes), is $n \cdot c_m^*$ with probability $L(p,n,k)$ and $n \cdot c_m^* + (N-n) \cdot c_s^*$ with probability $1 - L(p,n,k)$. The mean inspection cost per lot of process average quality $\bar{p}$ is therefore

$$C_{ms} = n \cdot c_m^* + (N-n) \cdot c_s^* \cdot (1 - L(\bar{p}; n; k)) \tag{6}$$

Dividing (6) by c_s^* gives the objective function

$$I_{ms} = n \cdot c_m + (N-n) \cdot (1 - L(\bar{p}; n; k)), \tag{7}$$

where $c_m = c_m^*/c_s^*$ is the ratio of the cost of inspection of one item by variables to the cost of inspection of this item by attributes (this parameter has to be estimated in each real situation, it is usually $c_m > 1$). Note that both the function $I_{ms} = C_{ms}/c_s^*$ and the function C_{ms} have a minimum for the same acceptance plan (n,k). Therefore, we shall look for the acceptance plan (n,k) minimizing (7), instead of (6), under the condition (4).

Setting the value of c_m to 1 can be used in situations, when both the sample and the remainder of rejected lots are inspected by variables. Acceptance sampling by variables can thus be considered just as a special case of acceptance sampling by variables and attributes. Then instead of I_{ms} we may use notation I_m and setting $c_m = 1$ in (7) we obtain

$$I_m = N - (N-n) \cdot L(\bar{p}; n; k), \tag{8}$$

i. e. the mean number of items inspected per lot of process average quality, assuming that both the sample and the remainder of rejected lots is inspected by variables.

The task to be solved is to determine plan (n,k) minimizing (7) under the condition (4) for given values of input parameters N, c_m, p_L and $\bar{p}$. The operating characteristic is (see e.g. [1])

$$L(p,n,k) = \Phi((u_{1-p} - k)A), \tag{9}$$

where

$$A = \sqrt{\frac{n(2-\lambda)}{\lambda}}. \tag{10}$$

The function Φ in (9) is a standard normal distribution function and u_{1-p} is a quantile of order $1-p$ (the unique root of the equation $\Phi(u) = 1-p$).

4 Calculation and economic efficiency measure of the plans

Lets calculate the AOQL acceptance sampling plan for sampling inspection by variables when the remainder of rejected lots is inspected by attributes below. The task will be solved using the operating characteristic given by (9) and using the EWMA statistic with smoothing constant $\lambda = 0.9$. The resulting sampling plan will be compared with the corresponding Dodge-Romig plan available in [2] and evaluated with regard to the economic characteristics. We consider a lot of $N = 3500$ items considered in the acceptance procedure. The average outgoing quality limit is given to be $p_L = 0.015$. It is known that the average process quality is $\bar{p} = 0.01$. A cost of inspecting an item by variables is two times higher than the cost of inspecting the item by attributes, so the parameter c_m equals 2. We are to find the optimal AOQL acceptance sampling plan for sampling inspection by variables when the remainder of rejected lots is inspected by attributes.

The sampling plan can be calculated using a modified version of the code available in LTPDvar package [6] for R software [11]. The resulting plan is $n = 32$, $k = 1.942692$.

For the values of the input parameters given in our problem, there is plan (165, 4) for the acceptance sampling by attributes in [2]. Let us compare the plans ($n = 32$, $k = 1.942692$) and ($n = 165$, $c = 4$) with regard to the economic efficiency. For this purpose, let us introduce parameter e, which we define as

$$e = \frac{I_{ms}}{I_s} \cdot 100. \tag{11}$$

It may be worthwhile to provide a short comment on the meaning and interpretation of the parameter e. The parameter e as introduced in (11) is in effect based on the ratio of the mean inspection cost per lot of the process average quality of two sampling plans – the AOQL plan for the inspection by variables an attributes and the corresponding Dodge-Romig plan for the inspection by attributes (in fact, each cost may be multiplied by a positive constant of the same value, which has no significant impact on the resulting value of the quotient), and as such is a non-dimensional feature. A suggested interpretation of the resulting value of this parameter is as follows: if e is used for the comparison of the economic efficiency of the two plans, then the value 100 means, that the costs of these plans are the same; in case that we get e over 100, then it suggests that the plan for inspection by attributes is more efficient for the values of the input parameters considered; and finally if we get e below 100, then the sampling plan for the inspection by variables and attributes is preferable over the corresponding sampling plan for the inspection by attributes.

The expression $(100 - e)$ then represents the percentage of savings in the mean inspection cost per lot of the proces average quality when the sampling plan for the inspection by variables and attributes is used instead of the corresponding plan for the inspection by attributes. Let us denote the plan for the inspection by variables and attributes as (n_1, k) and the corresponding plan for the inspection by attributes as (n_2, c). Then the parameter e is defined as

$$e = \frac{n_1 \cdot c_m + (N - n_1) \cdot (1 - L(\bar{p}, n_1, k))}{N - (N - n_2) \cdot L(\bar{p}, n_2, c)} \cdot 100. \tag{12}$$

Since we get

$$e = 38.5,$$

it can be expected that approximately **61% savings in the inspection cost** can be made using the AOQL plan for the inspection by variables and attributes (32, 1.942692), in place of the corresponding Dodge-Romig plan $(165, 4)$.

Now consider the situation, when c_m equals 5. Then the solution based on the operating characteristic given by (9) gives plan (23, 1.924244) and the value of the parameter e is 71.7. In cases when c_m is even higher, the resulting value of the parameter e is greater. For example the optimal plan obtained for c_m equal to 8 results in plan (19, 1.9140) and leads to $e = 97.2$. When c_m is as high as 9, the parameter e is over 100.

The value of the parameter e depends on the values of the input parameters considered in each particular situation in practice. One of the characteristics with major impact on the resulting value of the e parameter is the c_m value. It may be expected that the value of e is generally rising in c_m, i. e. the comparative economic efficiency of the attribute inspection plans in comparison with the variables inspection plans is expected to be lower. For some cases in business practice, it may be useful to consider a break-even value of the c_m parameter. We define c_m^{BE} to be such value of the parameter c_m for which the parameter e just equals 100. For the situation considered in the example above, the value of the parameter e in response to the c_m values is shown in Figure 1. The c_m^{BE} value in our case equals 8.37, as shown in Figure 1. If c_m^{BE} is below 9, then for $c_m = 9$ we get e over 100 and then the attributes inspection plan is preferable. And similarly for the situations when c_m is below 8.37 (the value of c_m^{BE}), the plans for inspection by variables and attributes are more economically efficient (we get e below 100). Since the ratio of the unit cost of inspection by variables to the unit cost of inspection by attributes may not be known precisely in some cases in practice (the c_m value may be an outcome of more or less precise cost calculations), the break-even value of this parameter may be used to guide the decision about which plan to use. And then the guidelines for using such value in making the decision about which plan to use are simply as follows: in case that the c_m^{BE} value we get for the particular situation in practice is high, then the variables inspection sampling plans may seem preferable. Because we may expect that the real value

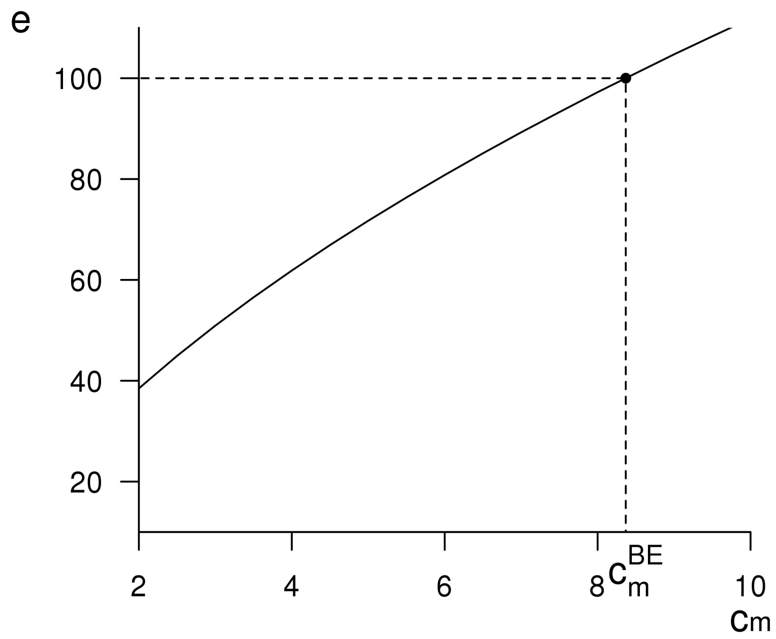


Figure 1 Economic evaluation of plans using e defined in (11) in response to c_m

of the c_m parameter is below c_m^{BE}. On the other hand, in case that we get a rather low value of the c_m^{BE}, which may be likely below the real value of c_m, then perhaps the sampling plan for the inspection by attributes is not worse than the plan for the inspection by variables and attributes.

5 Conclusion

It has been shown that the AOQL plans for the inspection by variables and attributes minimizing the mean inspection cost per lot of process average quality, which were designed to use the EWMA statistic in the decision procedure, may bring significant savings in the inspection cost. A measure for assessing the comparative economic efficiency of the plans, which may be used for getting a guidance for selecting optimal sampling plan to be used, has been discussed. The calculations of the plans may be performed easily using an R software extension package.

Acknowledgements

This paper has been produced with contribution of long term institutional support of research activities by Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] Aslam, M., Azam, M., and Jun, C.: A new lot inspection procedure based on exponentially weighted moving average, *International Journal of Systems Science* **46** (2015), 1392–1400.

[2] Dodge, H., and F., Romig, H. G.: *Sampling Inspection Tables: Single and Double Sampling.* John Wiley, New York, 1998.

[3] Jennett, W. J., and Welch, B. L.: The Control of Proportion Defective as Judged by a Single Quality Characteristic Varying on a Continuous Scale, *Supplement to the Journal of the Royal Statistical Society* **6** (1939), 80–88.

[4] Johnson, N. L., and Welch, B. L..: Applications of the Non-central t distribution, *Biometrika* **38** (1940), 362–389.

[5] Kasprikova, N.: Calculation and evaluation of new AOQL single sampling plans for inspection by variables. In: *Mathematical Methods in Economics 2015*, University of West Bohemia, Cheb, 2015, 349–353.

[6] Kasprikova, N.: *LTPDvar: LTPD and AOQL plans for acceptance sampling inspection by variables.* R package version 1.2. http://CRAN.R-project.org/package=LTPDvar. 2016.

[7] Kasprikova, N., and Klufa, J.: AOQL sampling plans for inspection by variables and attributes versus the plans for inspection by attributes, *Quality technology and quantitative management* **6** (2015), 133–142.

[8] Kasprikova, N., and Klufa, J.: Calculation of LTPD single sampling plans for inspection by variables and its software implementation. In: *International Days of Statistics and Economics*, University of Economics, Prague, 2011.

[9] Klufa, J.: Dodge-Romig AOQL single sampling plans for inspection by variables, *Statistical Papers* **38** (1997), 111–119.

[10] Klufa, J.: Dodge-Roming AOQL plans for inspection by variables from numerical point of view, *Statistical Papers* **49** (2008), 1 – 13.

[11] R Core Team: *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria, 2015. URL http://www.R-project.org

Mathematical model for sustainable value added calculation

Edward Kassem, Oldřich Trenz, Jiří Hřebíček, Oldřich Faldik [1]

Abstract. Sustainability Value Added (SVA) is an effective method for sustainability assessment. It plays a strategic role in decision making. It encourages the companies to deal with resources more effectively and efficiently. Sustainable Value Added represents the extra value created as a result of using economic, environmental and social resources, compared to a benchmark. It demonstrates the link between the organization's strategy and commitment to a sustainable global economy.

This paper aims to propose an improved method of sustainability assessment. It employs important and widely used financial value (e.g EVA) and new data oriented analysis (e.g Data Envelopment Analysis) for evaluating the sustainability of number of producers. This model reflects the specific requirements of the country and industry in which the company operates. This can be implemented by calculating the weights and benchmark values for each sector. The suggested model will be a backbone of WEBRIS – a web information system which is a combination of different information and communication technologies for quick and efficient data aggregation and assessment. Finally, the results visualization will be presented in the case study for breweries sector.

Keywords: DEA, efficiency, EVA, sustainability assessment, SVA, WEBRIS.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Sustainability performance can be said to be an ability of an organization to remain productive over time and hold on to its potential for maintaining long-term profitability. It shouldn't be acted only on the basis of economic results, whereas it should take into consideration non-financial factors [4, 18, 24, 25]. Because of that, all the current trends and frameworks depend on more comprehensive sustainability pillars; environmental, social, economic and governance performance. Therefore by following these frameworks and integrating associated activities, companies will be able to achieve long-term benefits [5, 9]. This will be done by engaging different companies in disclosure of their overall economic, environmental, social, and governance (ESG) impacts and helping them in understanding, measuring and communicating their sustainability performance.

One of the most famous and common methods for measuring corporate sustainability which incorporates three dimensions (economic, environmental and social) is called the triple bottom line approach [12, 24]. It calculates the value by using not only financial but also non-financial resources. This value is called sustainable value added (SVA). This approach simplifies the measurements and enables sustainable performance to be measured in monetary terms depending on the data availability on the enterprise level as well as on the benchmark [19]. It shows how much value or damage is created as a result of using economic, environmental and social resources, compared to a benchmark [13]. In order to use this method, the benchmark company should be chosen. In this research, Data Envelopment Analysis (DEA) is used for this purpose.

Data Envelopment Analysis (DEA) which is a relatively new "data oriented" approach for evaluating the efficiency of a number of producers. It combines the measurements of multiple inputs into any satisfactory overall measure of efficiency. In DEA the producers are usually referred to as the Decision Making Units (DMUs) which convert multiple inputs into multiple outputs. This approach refers to assessment of sustainability performance methodology [10] for Czech breweries which is based on the optimization algorithm. Relative efficiency is defined as a ratio of the total weighted output to the total weighted input. It can be achieved, if DMU performance can be enhanced by improving some inputs or outputs, without worsening others [8]. DEA can be used as a very powerful service management and benchmarking technique to evaluate nonprofit and public sector organizations [3, 10]. The advantage of this approach is that it provides the information about the sustainable management and efficiency of manufacture systems under different conditions. It also allows us to use a complex set of indicators for all the different sustainability dimensions [16].

[1] Mendel University in Brno, Faculty of Business and Economics, Department of Informatics, Zemědělská 1, 613 00 Brno, Czech Republic e-mail: edward.kassem@mendelu.cz.

The objective of the paper is to propose a modified and more accurate model for measuring a corporate sustainability performance. The model integrates environmental, social, economic and corporate governance indicators. It aggregates many indicators from different frameworks and allows the enterprises to compare their performance effectively. As an example, the results visualization is presented for six Czech breweries companies.

2 Materials and methods

2.1 Economic added value (EVA)

Economic Added Value is one of the most important and useful financial Indicators. It is considered as a backbone of enterprises' performance monitoring. EVA is used to determine the company's value. It designs compensation purposes, for interconnecting the strategic and the operations management of companies. It can be used not only for large corporation but also for small business and entrepreneurial endeavors [1]. Its construction at the end of the last century was an important moment in corporate theory and practice. Corporate results under the situation of the current economic theory and practices most often are measured using the Economic Value Added (EVA) indicator [21, 23]. EVA indicator from the perspective of financial management combines all the basic components required to describe the economic situation of the company. It is calculated according to the annual economic companies' reports by using Eq. (1).

$$EVA = (ROE - r_e).E \tag{1}$$

where E describes the market value of the firm's equity, ROE=*Net Income/E* and r_e is Return of Entity which is computed using the Eq. (2).

$$r_e = \frac{WACC.\dfrac{C}{A} - (1-t).\dfrac{I}{D}.\left(\dfrac{C}{A} - \dfrac{E}{A}\right)}{\dfrac{E}{A}} \tag{2}$$

where A, C, t, I, D and WACC are assets, Invested Capital, corporate tax rate,interest expenses, market value of the company's debt, and Weighted Average Cost of Capital, respectively. Note that, a positive number tells us that Company XYZ more than covered its cost of capital. A negative number indicates that the project did not make enough profit to cover the cost of doing a business. The calculation of EVA is not easy, but it can be achieved using different prepared spreadsheets, programs generated in different programming languages (e.g. MATLAB, Maple etc.) or using information and communication system like web information system WEBRIS [15] which uses HTML5 and PHP5 technologies with XBRL support.

2.2 Data envelopment analysis (DEA)

As the field of Data Envelopment Analysis has grown, varieties of models and analyses are implemented. Since DEA was introduced, researchers in a number of fields have quickly recognized that it is an excellent and easily used methodology for modeling operational processes and performance evaluation. Various forms of DEA models have been used for entities, such as hospitals, universities, and companies.

One of the famous models is called CCR model [6] used for relative effeciency calculation. It produces multiple optimal solutions. This non-uniqueness of solution hampers the use of cross-efficiency evaluation. Therefore, [11, 22] suggested introducing a secondary goal to optimize the input and output weights while keeping unchanged the CCR- efficiency of the target DMU. The formulated model is known as the aggressive formulation for cross- efficiency evaluation. Other different models as: minimizing the total deviation from ideal point, minimizing the maximum deviation, and minimizing the mean absolute deviation [20] were also implemented. These models are all established on the basis of an unrealistic ideal point, which defines the best relative efficiency of one as the target efficiency of each DMU. This target efficiency, however, is only realizable to DEA efficient DMUs, but not realizable to non-DEA efficient units. In order to solve this problem, set of models suggested to define the CCR-efficiencies of the n DMUs as their target efficiencies,which are all realizable to DEA efficient units and non-DEAefficient units. These models are mentioned as minimizing or maximizing the total deviation from the ideal point, minimizing or maximizing the squared sum of deviations from the ideal point, and minimizing the mean absolute deviation from the ideal point [26]. The newest models to measure the sustainability management and performance based on advanced DEA model are approaches combine dual-role factors and a cross-efficiency technique [7]. It assesses the conflicts and trade-off among environmental, economic and social interests by using three continuous multi-criteria approaches and a set of different weights.

2.3 Sustainable value added (SVA)

Sustainability Value Added (SVA) is an effective method for sustainability assessment. It plays a strategic role in decision making. It encourages the companies to deal with resources more effectively and efficiently. Sustainable Value Added represents the extra value created as a result of using economic, environmental and social resources, compared to a benchmark. It is expressed in absolute monetary terms. According to the method published by [9] the SVA value calculation can be expressed as follows: The gross value added of the company should be calculated (in unit €). After that, the amount of each environment or social resources should be determined (e.g t, m^3, ..etc). Then efficiency computed by dividing the gross value added on the amount of resources (unit €/t, €/m^3). The same steps should be done for the benchmark. Finally, the last two values are subtracted from each other and the result multiplied by the amount of considered indicator. This process is depicted in Fig.1.

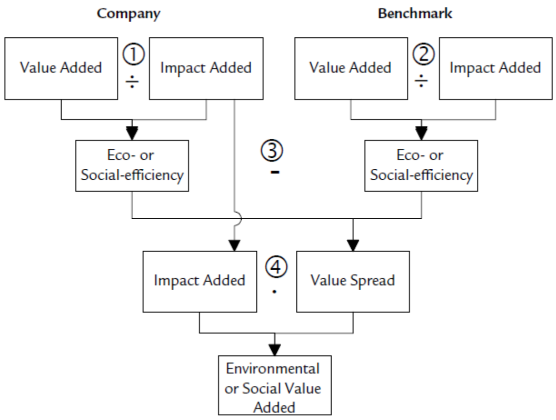


Figure 1 Evaluation steps of the Environmental or Social Value Added. Source: [9]

The calculation of Sustainable Value Added for the company in two different times t_1 and t_0 is presented in Eq. (3). where *n* is the number of relevant environmental indicators. $EIA_{i,t0}$, $EIA_{i,t1}$ describe the eco-effectiveness of environmental impacts in t_0 and t_1. $EE_{i,b}$ is eco-efficiency of the benchmark for *i* environmental resource. This value calculated using Eq. (4) where *b* refers to benchmark. $EG=(VA_{t1}-VA_{t0})$ representing economic growth. *VA* is the gross value added.

$$SVA = EG - \frac{1}{N}\sum_{i=1}^{N} EE_{i,b}.(EIA_{i,t1} - EIA_{i,t0}) \tag{3}$$

$$EE_b = \frac{VA_b}{EIA_b} \tag{4}$$

3 Proposed method of sustainability assessment

As mentioned above, the main objective of this method is improving the benchmark between the enterprises and providing an active participation in decision-making. The main calculation described in Fig. 1 remains the same. Whereas the improvements include several modifications, in order to achieve the following factors:

- comprehensive sustainability assessment: we focused our efforts on developing a comprehensive sustainability assessment. Therefore, environmental, social, economic and corporate governance indicators should be integrated. In this case, the proposed model won't only deal with financial indicators but should also include non-financial ones;
- suitability: The assessment should be done for different companies in the Czech Republic. However, the model can't be universal, because the indicators should reflect the specifics of the industry in which the company operates. Therefore, different available sustainability frameworks are used and specific set of indicators are chosen for each sector (e.g. agriculture, manufacture ...);
- simplicity and applicability: The modified model should be easy, simple, suitable and accurate. It reflects not only three dimensions (economic, environment, and social), but also the corporate governance pillar is add-

ed. As mentioned above, EVA is the most important and measured indicator which combines all the basic components required to describe the economic situation of the company. For this reason, the gross value added (VA) is replaced by Economic Value Added to describe the financial situation of the companies more efficiently;

After applying all mentioned modifications on the Eq. (3), it is presented in Eq. (5).

$$SVA = 3EVA_c - \frac{1}{N}\sum_{i=1}^{N}(\frac{EI_{i,c}}{EI_{i,b}}EVA_b) - \frac{1}{M}\sum_{j=1}^{M}(\frac{SI_{j,c}}{SI_{j,b}}EVA_b) - \frac{1}{K}\sum_{l=1}^{K}(\frac{GI_{l,c}}{GI_{l,b}}EVA_b) \tag{5}$$

where *EI, SI, GI,* w_{Ei}, w_{Ej}, w_{Gl} are values and the weights of environment, social and governance indicators, respectively. Symbol *b* refers to benchmark, while symbol *c* refers to the studied company. According to Eq. (5), increasing the value of environment indicator "for example the amount of hazardous waste is increased", will effect negatively on SVA. Increasing the economic value added of the company, in turn increases SVA value. In order to improve our proposed model we supposed that different indicators don't effect equally on enterprises score. That means, each indicator should have a different weight on sustainability calculation. This weight differs according to the country, size and sector of the company. The implementation of this improvement can be done using Eq. (6), where w_{av} is the average weight for *i's* indicator after applying DEA model.

$$SVA = 3EVA_c - \frac{1}{N}\sum_{i=1}^{n}(w_{i,av}\frac{EI_{i,c}}{EI_{i,b}}EVA_b) - \frac{1}{M}\sum_{j=1}^{m}(w_{j,avg}\frac{SI_{j,c}}{SI_{j,b}}EVA_b) - \frac{1}{K}\sum_{l=1}^{k}(w_{l,av}\frac{GI_{l,c}}{GI_{l,b}}EVA_b) \tag{6}$$

In our model, the benchmark value and the weight of each indicator are calculated using DEA model. By applying this model, the most efficient company can be determined which is taken as a benchmark company. More information about these calculations can be found in [17].

In order to apply this method, an example of six Czech brewery enterprises A, B, C, …, F is used. The important data is extracted from Amadeus database [1]. After that, without loss of generality, few KPIs are chosen. We considered KPIs as the amount of hazardous waste (ENV1) in tones, the amount of other waste (ENV2) in tones, number of employees (SO1) and percentage of women in supervising the company (GOV1) as organizations' inputs. The organizations' output is Economic Value Added (EC1) in CZK. In additional, a dual-role factor is considered as the average employees' salary and bonuses (SO2).

Company	Efficiency Score	w_{ENV1}	w_{ENV1}	w_{SOC1}	w_{SOC2}	w_{GOV1}
A	1	0.00296	0	0.00935	0.00905	0
B	1	0	0.00275	0.0096	0.01017	0.00087
C	0.42	0	0.01762	0.04278	0.05078	0.01447
D	1	0	0.00271	0.00944	0.01	0.00086
E	1	0	0.00725	0.0176	0.02089	0.00595
F	1	0	0	0.0294	0	0

Table 1 Efficiency score and input weights

As mentioned before, the weight of each indicator and the benchmark values will be calculated using DEA model. This model computes efficiency score of selected organizations (DMUs) using the linear programing in Maple [14]. Table 1 summarizes the output of DEA model which consists of several parameters. These parameters are the efficiency scores of breweries and the weight of each indicator used in their sustainability assessment.

The above mentioned DEA model is solved six times, ones for each target brewery. As a result, there are six sets of input and output weights. The six efficiency values are then averaged as the overall performance of the brewery sector. The highest efficiency score, company B, will be used as benchmark values in Eq. (5) or Eq. (6) for sustainability calculation. By applying Eq. (5) or Eq. (6) the sustainability value added of each enterprise is calculated.

Table 2 presents the final assessment of studied companies. These values make us understand which company has better contributed to sustainability development. The results present the sustainability value assessments with and without applying the averaged weights, Eq. (6) and Eq. (5), respectively.

Company	SVA	Modified SVA
A	239489.3	568282.7
B	0	69441.2
C	-113994	1952.749
D	-35667.5	20485.45
E	-57742	-1564.21
F	649930.8	677183.5

Table 2 Sustainability value added

According to presented results company F is the most sustainable company, while C or E is the less sustainable one using SVA or SVA modified model, respectively. Figure 2 depicts two types of curves. The red presents the sustainability value added without applying the averaged weight Eq. (5), whereas the blue one presents the proved version of sustainability value added which takes into consideration the weight of each indicator.

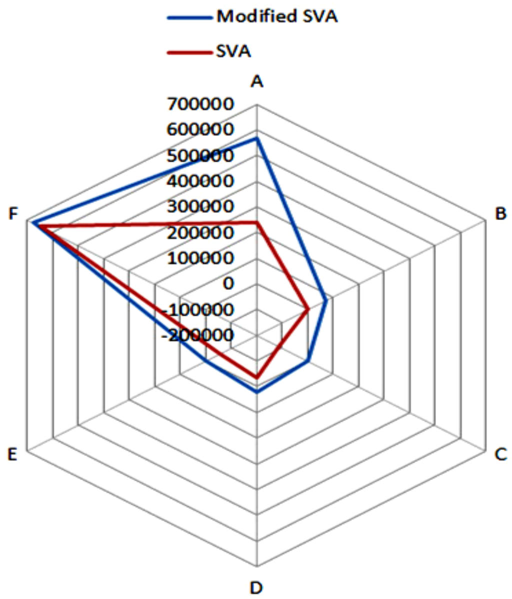


Figure 2 The level of SVA and modified SVA of the six Czech brewery enterprises.

4 Conclusion

Sustainability assessment is a comprehensive process to achieve the best performance and determine the weak points of the studied organization. An expensive data collecting and managing, difficulty of determining the appropriate sustainability indicators and capturing reliable data-information are the main barriers that face different organizations. In order to overcome them, WEBRIS system is suggested. WEBRIS is a combination of different information and communication technologies which can be used for quick and efficient data aggregation and assessment. The backbone of this system is a model used for sustainability assessment.

This paper aims to propose an improved method of sustainability assessment. It employs important and widely used financial value (e.g SVA, EVA) and new data oriented analysis (e.g Data Envelopment Analysis) for evaluating the efficiency of number of producers. This work is supported by make it reflects the specifics requirements of the country and industry (Czech breweries). The weights and benchmark values are calculated and used in sustainable value added equation. The results visualization are presented in the case study for breweries sector. According to our proposed method each company from different sectors can assess their sustainability in easy fast way. Then compare their results with other companies in the same sector.

Acknowledgements

This paper is supported by the Czech Science Foundation. Project name: Measuring corporate sustainability performance in selected sectors [Nr. 14-23079S].

References

[1] Amadeus Database. [Online]. Available on: http://www.eui.eu/Research/Library/ResearchGuides/Economics/Statistics/DataPortal/Amadeus.aspx

[2] Baxendale, S. J., and Bowen, L.: Economic Value Added for New Ventures and Small Business. *Journal of Small Business Strategy* **12** (2001), 41–51.

[3] Callens, I. and Daniel T.: Towards indicators of sustainable development for firms: a productive efficiency perspective. *Ecological Economics* **28** (1999): 41–53.

[4] Carroll, A. B.: Ethical challenges for business in the new millennium: Corporate social responsibility and models of management morality. *Business Ethics Quarterly* **10** (2000), 33–42.

[5] Chabowski, B. R., Mena, J. A., and Gonzalez-Padron, T. L.: The structure of sustainability research in marketing, 1958–2008: a basis for future research opportunities. *Journal of the Academy of Marketing Science* **39** (2011), 55–70.

[6] Charnes, A., Cooper, W.W., and Rhodes, E.: Measuring the efficiency of decision making units. *European Journal of Operational Research* **2** (1978), 429–444.

[7] Cook, W.D., Green, R.H., Zhu, J.: Dual-role factors in data envelopment analysis. *IIE Transactions* **38** (2006), 105–115.

[8] Cooper, William W. and Seiford, Lawrence M. and Zhu, J.: Data Envelopment Analysis, "Handbook on Data Envelopment Analysis, Springer US, Boston, MA, 2004, 1–39.

[9] Cruz, L., Avila Pedrozo, E., and de Fatima Barros Estivalete, V.: Towards sustainable development strategies: a complex view following the contribution of Edgar Morin. *Management Decision* **44** (2006), 871–891.

[10] Dong, F., Mitchell, P. D., and Colquhoun, J.: Measuring farm sustainability using data envelope analysis with principal components: The case of Wisconsin cranberry. *Journal of environmental management* **147** (2015), 175–183.

[11] Doyle, J., Green, R.: Efficiency and cross-efficiency in DEA: derivations, meanings and uses. *Journal of the Operations Research Society* **45** (1994), 567–578.

[12] Figge, F., and Hahn, T.: Sustainable value added-measuring corporate contributions to sustainability beyond eco-efficiency. *Ecological economics* **48** (2004), 173–187

[13] Hart, S. L., and Milstein, M. B.: Creating sustainable value. *The Academy of Management Executive* **17** (2003), 56–67.

[14] Hřebíček, J., Trenz, O., Chvátalová, Z., and Soukopová, J.: Optimization of corporate performance using data envelopment analysis with Maple. *In Aurelio Araujo et al. Engineering Optimization*. London, CRC Press, 2014, 763–767.

[15] Hřebíček J., Faldík O., Kasem E., Trenz O.: Determinants of Sustainability Reporting in Food and Agriculture Sectors. *Acta universitatis agriculturae et silviculturae Mendelianae Brunensis* **63** (2015), 539–552.

[16] Jablonský, J.: *Operations research: Quantitative models for economic decisions,* Czech, Praha, 2002.

[17] Kasem, E., Trenz, O., Hřebíček, J., and Faldík, O.: Key Sustainability Performance Indicator Analysis for Czech Breweries. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis* **63** (2015), 1937–1944

[18] Kaplan, R. S., and Norton, D. P.: Transforming the balanced scorecard from performance measurement to strategic management: Part I. *Accounting horizons* **15** (2001), 87–104

[19] Kuosmanen, T., and Kuosmanen, N.: How not to measure sustainable value (and how one might). *Ecological Economics* **69** (2009), 235–243

[20] Liang, L., Wu, J., Cook, W.D., Zhu, J.: Alternative secondary goals in DEA cross efficiency evaluation. *International Journal of Production Economics* **113** (2008), 1025–1030.

[21] Qi, L.: A review of economic value added (EVA) survey-From the aspects of theory and application. In *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference,* 2011, 507–509.

[22] Sexton, T.R., Silkman, R.H., and Hogan, A.J.: Data envelopment analysis: critique and extensions. In: Silkman, R.H. (Ed.), Measuring Efficiency: An Assessment of Data Envelopment Analysis. Jossey-Bass, San Francisco, CA, 1986, 73–105.

[23] Sharma, A. K., and Kumar, S.: Economic value added (EVA) literature review and relevant issues. *International Journal of Economics and Finance* **2** (2010), 200–220.

[24] Siew, R.: A review of corporate sustainability reporting tools (SRTs). *Journal of Environmental Management* **164** (2015), 180–195.

[25] Waddock, S., and Smith, N.: Corporate responsibility audits: Doing well by doing good. *Sloan Management Review* **41** (2000), 75–83.

[26] Wang, Y.M., and Chin, K.S.: Some alternative models for DEA cross- efficiency evaluation. *International Journal of Production Economics* **128** (2010), 332–338.

A single-stage approach for selecting inputs/outputs in DEA

Esmaeil Keshavarz[1], Mehdi Toloo[2]

Abstract. Data Envelopment Analysis (DEA) is an optimization-based methodology to evaluate the performance of Decision Making Units (DMUs), each one consumes multiple inputs to produce multiple outputs. From the statistical and empirical points of view, if the number of inputs/outputs is high in comparison with the number of DMUs, then a large percentage of DMUs will be identified as efficient and efficiency discrimination among DMUs is questionable. It also implies that the selection of relevant inputs/outputs is very crucial for successful theoretical and applied purposes. In order to deal with selective input/output measures in DEA, some two-stage approaches are proposed where selective measures are chosen in the first stage and then the DMUs are evaluated in the second stage based on the selected measures. The current study suggests a common weights DEA model which selects inputs/outputs as well finds the efficiency of DMUs in a single-stage approach. An illustrative example of the banking industry in Iran is provided to validate the proposed approach. Obtained Results emphasize that the suggested approach correctly deals with selective measures.

Keywords: Data envelopment analysis, the rule of thumb, selective measures, integrated single-stage model, banking industry.

JEL Classification: C44
AMS Classification: 90C05, 90C11, 90C90

1 Introduction

Consider the problem of evaluating performance of n homogeneous Decision Making Units (DMUs), which consume m inputs to produce s outputs. Data Envelopment Analysis (DEA) is a well-organized optimization-based methodology to do this evaluation. In the basic DEA models, the relative efficiency of a DMU is computed by maximizing the ratio of weighted sum of its outputs to weighted sum of its inputs, based on the condition that this ratio is less than or equal to one for all DMUs. In fact, to compute the efficiency of each DMU a fractional programming model should be solved. Charnes, et al. [4] in their pioneering work reformulated this fractional programming as a linear programming problem and so along with the speedy advances of linear programming and operations research, DEA models has been also rapidly developed. In the course of this development, some critical challenges have occurred [6]. One of these challenges occurs when the number of inputs and outputs is high in comparison with the number of DMUs; in this situation, most of the DMUs are evaluated efficient and hence the obtained results are not reliable. On the other hand, awkward deleting of some inputs or outputs from considerations can seriously affect the efficiency scores of DMUs.

Empirically, there is a rough rule of thumb [5], which expresses the relation between the number of DMUs and the number of performance measures:

$$n \geq \max\{3(m+s), m \times s\} \tag{1}$$

In some applications, the number of performance measures, which reflects the manager's interest, and the number of DMUs do not satisfy the rule of thumb. In such setting, the problem of selecting some inputs/outputs, in a manner which complies (1) is an important issue. A variety of researchers attempted to tackle this issue. Nataraja and Johnson [9] analyzed four most widely-used approaches (Efficiency Contribution measure (ECM), Principal Component Analysis (PCA-DEA), a regression-based test, and bootstrapping) for inputs/outputs selection in DEA.

Morita and Avkiran [8] considered an input and output selection method based on discriminant analysis using external evaluation. The authors used a 3-level orthogonal layout experiment to find an appropriate combination of inputs and outputs, where experiments are independent of each other. Jenkins and Anderson [7] described a

[1] Islamic Azad University, Sirjan Branch, Department of Mathematics, Hejrat Blvd., IAU campus, Sirjan, Iran, e-mail: ekeshavarz@iausirjan.ac.ir.
[2] VŠB-Technical University of Ostrava, Department of Systems Engineering, Sokolska třida 33, 701 21 Ostrava, Czech Republic, e-mail: mehdi.toloo@vsb.cz.

systematic statistical method for deciding which of the original correlated inputs/outputs can be omitted with least loss of information, and which should be retained. Amirteimoori and Emrouznejad [2] developed an approach to input/output reduction problem that typically occurs in organizations with a centralized decision-making environment. They discussed how a DEA-based model can be used to determine an optimal input/output reduction plan.

Toloo, et al. [13] formulated and solved two mixed integer programming models, based on individual and aggregate efficiency scores, for developing the idea of selective measures. In fact, the authors modified the standard constant returns to scale model of DEA to obtain a model that could select the performance measures so that the number of DMUs, inputs and outputs meet the mentioned rule of thumb. Toloo and Tichý [14] extended Toloo, et al's [13] approach to propose multiplier and envelopment forms of DEA models for selecting inputs/outputs under variable returns to scale assumption. It is proved that multiplier form leads to the maximum efficiency scores while the maximum discrimination between efficient units is achieved by applying the envelopment form.

All the mentioned approaches use two-stage analysis for selecting input/output measures in DEA, where selective inputs/outputs are chosen in the first stage and then DMUs are evaluated in the second stage based on the selected measures. The current study suggests a common weight DEA model in order to select inputs/outputs and evaluate all DMUs, simultaneously, in a single-stage.

The rest of the paper is organized as follows: In Section 2, a variable returns to scale DEA model based on the common weights approach is presented. A single stage selecting approach is introduced in Section 3. Section 4 illustrates the applicability of the proposed approach by using a real data set of the banking industry in Iran. The paper concludes in Section 5.

2 Common set of weights model

Consider a set of n DMUs, each consuming various amounts of m inputs to produce s outputs. Let $\mathbf{x}_j = (x_{1j}, \dots, x_{mj}) \in \mathbb{R}^m$ and $\mathbf{y}_j = (y_{1j}, \dots, y_{sj}) \in \mathbb{R}^s$ represent the input and output semi-positive vectors for $DMU_j (j = 1, \dots, n)$, respectively. Charnes, et al.[4] proposed a linear programming model, which is referred to as CCR (**C**harnes, **C**ooper and **R**hodes), for evaluating the relative efficiencies among DMUs with multiple inputs and multiple outputs under constant returns to scale (CRS) technology. Banker, et al. [3] suggested the following model, which is known as BCC (**B**anker, **C**harnes and **C**ooper), to handle the variable returns to scale (VRS) situation:

$$\begin{array}{lll}
\max & \theta_o = \sum_{r=1}^{s} u_r y_{ro} + u_o & \\
\text{s.t.} & \sum_{i=1}^{m} v_i x_{io} = 1 & \\
& \sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} \le 0 & j = 1, \dots, n \\
& v_i \ge \varepsilon \quad i = 1,2, \dots, m & \\
& u_r \ge \varepsilon \quad r = 1,2, \dots, s &
\end{array} \tag{2}$$

where u_r and v_i are the set of output and input weights, respectively, u_o is a variable with free in sign; and $\varepsilon > 0$ is the non-Archimedean infinitesimal, which is employed to hinder the weights to become zero [1]. Let the optimal solution of model (2) be $(\theta_o^*, \boldsymbol{u}^*, \boldsymbol{v}^*, u_o^*)$. The value of θ_o^* is named *BCC-efficiency score* of DMU_o. If $\theta_o^* = 1$, then DMU_o is *BCC-efficient* and otherwise is *BCC-inefficient*.

Model (2) finds input/output weights that are DMU-specific, and it is acceptable for individual circumstances of operation of DMUs, in practice, it can fail to discriminate on the performance of DMUs. In fact, most units attain the maximum or near maximum efficiency score. On the other hand, there are situations in which the different DMUs experience similar circumstances and therefore, using input/output weights that differ substantially across DMUs may not be warranted. When that is the case, both the inputs and the outputs should be aggregated by using weights that are common to all the DMUs [10]. There are different models have been proposed to evaluate and rank DMUs based on the common weights (CW) approach, one of the most appropriate of them is the following model which is named the minimax deviation CW model.

$$\begin{array}{lll}
\min & d_{max} & \\
\text{s.t.} & \sum_{i=1}^{m} v_i x_{ij} \le 1 & j = 1, \dots, n \\
& \sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0 & j = 1, \dots, n \\
& d_{max} - d_j \ge 0 & j = 1, \dots, n \\
& v_i \ge \varepsilon & i = 1,2, \dots, m \\
& u_r \ge \varepsilon & r = 1,2, \dots, s
\end{array} \tag{3}$$

In this model, DMU_j is *CW-efficient* if and only if $d_j^* = 0$ or, equivalently, $\rho_j^* = \frac{\sum_{r=1}^{s} u_r^* y_{rj} + u_0^*}{\sum_{i=1}^{m} v_i^* x_{ij}} = 1$, and otherwise it is *CW-inefficient*. Based on this notation, the value of ρ_j^* is called *CW-efficiency score* of DMU_j. To see the role of the non-Archimedean epsilon in CW-DEA models see Toloo [12].

It is worth noticing that model (3) is an aggregated model and generally, CW-efficiency score of a DMU is less than or equal to its BCC-efficiency score and therefore the discriminating power of model (3) is higher than model (2). Nevertheless, it is possible that there exists some cases which the number of inputs/outputs and DMUs contravene the rule of thumb and discriminating power of CW model (3) is also reduced. In the next section, we propose a selecting DEA model that can handle the situations, in which the manager has to select some of the performance measures among suggested inputs and outputs.

3 Selecting inputs/outputs in DEA

Consider n DMUs, with m inputs and s outputs, where $n < \max\{3(m+s), m \times s\}$. In this situation, one approach for having a sharper discrimination among DMUs is decreasing the number of inputs/outputs $(m+s)$, such that the rule of thumb (1) is met. Let s_1 and s_2 denote subsets of outputs corresponding to fixed-output and selective-output measures, respectively. Similarly, assume that m_1 and m_2 are the parallel subsets of inputs. Toloo and Tichý [14] presented the following mixed integer programming model to reduce the number of inputs/outputs, as an individual BCC-based model.

$$
\begin{array}{ll}
\max \sum_{r=1}^{s} u_r y_{ro} + u_0 & \\
\text{s.t.} & \\
\sum_{i=1}^{m} v_i x_{io} = 1 & \\
\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} \leq 0 & j = 1, \dots, n \\
\sum_{r \in s_2} b_r^y + \sum_{i \in m_2} b_i^x \leq \min\left\{\left[\frac{n}{3}\right], 2\sqrt{n}\right\} - (|m_1| + |s_1|) & \\
\varepsilon b_i^x \leq v_i \leq M b_i^x & i \in m_2 \\
\varepsilon b_r^y \leq u_r \leq M b_r^y & r \in s_2 \\
b_i^x, b_r^y \in \{0,1\} & i \in m_2, r \in s_2 \\
v_i, u_r \geq \varepsilon & i \in m_1, r \in s_1
\end{array}
\tag{4}
$$

where $\varepsilon > 0$ is a small number and M is a large positive number. Binary variables b_i^x and b_r^y are associated with selective input $i \in m_2$ and selective output $r \in s_2$, respectively. It should be noted that b_i^x is equal to 1 if its associated input is selected, and the same result can be achieved for b_r^y. Toloo and Tichý [14] proved that the selecting model (4) holds the rule of thumb.

In the Toloo and Tichý's method, since the model (4) uses an individual approach, selected variables for an individual DMU is different from the selected variables of other DMUs, and so there is no stability for this method. To tackle this issue they proposed an aggregate two-stage approach, which selective measures are chosen in the first stage and then the DMUs are evaluated in the second stage based on the selected measures. Here, we propose the following CW-DEA model which selects inputs/outputs and finds efficiency of DMUs, simultaneously.

$$
\begin{array}{ll}
\min d_{\max} & \\
\text{s.t.} & \\
\sum_{i=1}^{m} v_i \left(\sum_{j=1}^{n} x_{ij}\right) = 1 & \\
\sum_{r=1}^{s} u_r y_{rj} + u_0 - \sum_{i=1}^{m} v_i x_{ij} + d_j = 0 & j = 1, \dots, n \\
\sum_{r \in s_2} b_r^y + \sum_{i \in m_2} b_i^x \leq \min\left\{\left[\frac{n}{3}\right], 2\sqrt{n}\right\} - (|m_1| + |s_1|) & \\
\sum_{i \in m_2} b_i^x \geq 1 & \\
\sum_{r \in s_2} b_r^y \geq 1 & \\
d_{\max} - d_j \geq 0 & j = 1,,, n \\
\varepsilon b_i^x \leq v_i \leq M b_i^x & i \in m_2 \\
\varepsilon b_r^y \leq u_r \leq M b_r^y & r \in s_2 \\
d_j \geq 0 & j = 1, \dots, n \\
b_i^x, b_r^y \in \{0,1\} & i \in m_2, r \in s_2 \\
v_i, u_r \geq \varepsilon & i \in m_1, r \in s_1
\end{array}
\tag{5}
$$

It should be mentioned here that the n normalization constraints $\sum_{i=1}^{m} v_i x_{ij} \leq 1$ $(j = 1,2,..n)$ are replaced with the aggregate normalization constraint $\sum_{i=1}^{m} v_i \left(\sum_{j=1}^{n} x_{ij}\right) = 1$ in order to reduce the number of constraints as much as possible. Constraints $\sum_{i \in m_2} b_i^x \geq 1$ and $\sum_{r \in s_2} b_r^y \geq 1$ force that at least on of selective inputs and one of selective outputs are selected.

Theorem 1. The presented model (5) will meet the rule of thumb.

Proof. Let $(\mathbf{u}^*, u_0^*, \mathbf{v}^*, \mathbf{b}^{y*}, \mathbf{b}^{x*})$ be the optimal solution of selecting model (5). As inspection makes clear, if $b_p^{x^*} = 1$, then the selective input measure p is selected. In a similar manner, the selective output q is selected when $b_q^{y^*} = 1$. Hence, the total number of involved inputs and output, including fixed and selective measures is equal to $\left(|m_1| + \sum_{i \in m_2} b_i^{x^*}\right) + \left(|s_1| + \sum_{r \in s_2} b_r^{y^*}\right)$. Now, taking the constraint $\sum_{r \in s_2} b_r^y + \sum_{i \in m_2} b_i^x \leq \min\{[n/3], 2\sqrt{n}\} - (|m_1| + |s_1|)$ into consideration, two cases may arise:

(i) $[n/3] = \min\{[n/3], 2\sqrt{n}\}$ which implies $n \geq 3\left(\left(|m_1| + \sum_{i \in m_2} b_i^{x^*}\right) + \left(|s_1| + \sum_{r \in s_2} b_r^{y^*}\right)\right)$

(ii) $2\sqrt{n} = min\{[n/3], 2\sqrt{n}\}$ which considering the constraint $\sum_{r \in s_2} b_r^y + \sum_{i \in m_2} b_i^x \leq 2\sqrt{n} - (|m_1| + |s_1|)$ leads to $n \geq \left(|m_1| + \sum_{i \in m_2} b_i^{x^*}\right) \times \left(|s_1| + \sum_{r \in s_2} b_r^{y^*}\right)$.

As a result, from the optimal solution of the selecting model (5) we obtain $n \geq \max\left\{3\left(\left(|m_1| + \sum_{i \in m_2} b_i^{x^*}\right) + \left(|s_1| + \sum_{r \in s_2} b_r^{y^*}\right)\right), \left(\left(|m_1| + \sum_{i \in m_2} b_i^{x^*}\right)\left(|s_1| + \sum_{r \in s_2} b_r^{y^*}\right)\right)\right\}$ which completes the proof. ■

In the next section, we utilize a real data set of the banking industry in order to show the applicability of the proposed CW selecting approach.

4 Application

This section illustrates the proposed approach through a real data set obtained from annual reports of 12 branches of a private bank in Iran, the inputs and outputs measures are suggested as follows:

Inputs: Employees, Number of accounts, Assets, Space, Costs, Expenses,

Outputs: Number of transactions, Deposits, Loans, Check card, Credit card, OTP.

Table (1) exhibits the data set and, BCC-efficiency and CW-efficiency scores. The last two columns in the table show that all DMUs are BCC-efficient and about 60% of banks (i.e. 7 out of 12) are CW-efficient. Indeed BCC model fails to discriminate DMUs and discrimination power of CW model is not strong enough because of an inadequate number of performance measures ($n = 12 < 36 = \max\{3(m + s), m \times s\}$). To get an acceptable result, we have to decrease the number of inputs/outputs such that the rule of thumb is held. To this end, we apply presented selecting model (5) to the data set in Table (1). Referencing Toloo and Tichý [14], in more than 80% of the bank studies 'Employees' is considered as an input and subsequently it is reasonable to consider it as a fixed input. Hence, we consider the number of employees as the fixed input and the other inputs and outputs as selective measures. Optimal solution shows that 'Costs' is the selected input and 'Loans' and 'Check card' are the selected outputs. Table (2) summarizes all fixed and selected data and also acceptable CW-efficiency scores which is obtained by solving model (5). As can be extracted from the table, there are 2 CW-efficient banks out of 12 banks which shows the proposed model is succeed in decreasing the percentage of efficient DMUs from 60% to 16% with two CW-efficient DMUs. Also, the number of involved input/output measures and the number of DMUs meet the rule of thumb.

DMU	Inputs						Outputs						efficiency	
	Employ-ees	No. of accounts	Assets	Space	Costs	Ex-penses	No. of transactions	Deposits	Loans	Check card	Credit card	OTP	BCC	CW
1	33	1250	1753	97	10020	3137	5214	72149	57537	5105	4839	25	1.00	0.87
2	21	3217	1155	61	8427	2180	5145	42654	52485	2797	2697	5	1.00	0.99
3	15	1475	829	44	5283	2887	3913	45732	14237	3795	3500	32	1.00	1.00
4	18	1689	1023	52	5856	2606	4559	53323	37418	1858	1746	8	1.00	1.00
5	27	2669	1536	79	7326	1989	5031	49153	47139	4811	4578	31	1.00	1.00
6	24	7175	1367	70	8326	3727	5053	92365	55543	6840	6588	45	1.00	1.00
7	21	2120	1193	61	6525	3473	4762	64235	22347	5382	5188	22	1.00	0.98
8	21	1464	1111	61	11135	1524	4307	42012	73925	3187	2984	22	1.00	1.00
9	21	8924	1182	68	6920	3573	5331	69360	27246	3743	3524	24	1.00	0.85
10	21	2388	1069	61	5864	2523	4004	51438	26531	4360	4140	17	1.00	1.00
11	18	4714	992	52	5039	2398	2342	39948	20223	2688	2574	36	1.00	0.79
12	21	1866	1180	62	8378	3165	4238	154284	43928	4182	4008	18	1.00	1.00

Table 1 Bank data (Source: Authors’ calculation, 2015)

DMU	Inputs		Output		Single Stage CW-efficiency
	Employees	Costs	Loans	Check card	
1	33	10020	57537	5105	0.82
2	21	8427	52485	2797	0.88
3	15	5283	14237	3795	0.74
4	18	5856	37418	1858	1.00
5	27	7326	47139	4811	0.98
6	24	8326	55543	6840	1.00
7	21	6525	22347	5382	0.75
8	21	11135	73925	3187	0.86
9	21	6920	27246	3743	0.74
10	21	5864	26531	4360	0.88
11	18	5039	20223	2688	0.86
12	21	8378	43928	4182	0.82

Table 2 Selected inputs/outputs and CW-efficiencies (Source: Authors’ calculation).

5 Conclusion

The lack of discrimination among efficient DMUs is an important challenge in DEA. There are some different reasons causing weak discrimination among DMUs, one of them is the high ratio of inputs/outputs number to the number of DMUs, in such case a large percentage of DMUs will be identified as efficient. One way to tackle this issue is selecting some of inputs/outputs, based on the rule of thumb, in a manner that the discrimination of DMUs is improved. In this paper, we dealt with selecting inputs/outputs and modified the common weight BCC model to obtain a selecting model with the aim of selecting a set of acceptable performance measures. Moreover, the proposed approach can take the fixed performance measures into consideration. A real application of banking industry in Iran is used to show the applicability of the introduced model. It was shown that the approach leads to the maximum discrimination between efficient DMUs. As a future research, an epsilon-free approach [11] can be extended in order to exclude the non-Archimedean epsilon from the suggested model.

Acknowledgements

The research was supported by the European Social Fund within the project CZ.1.07/2.3.00/20.0296 and the Czech Science Foundation through project No. 16-17810S. The study was also supported by SGS project of VŠB-TUO SP2016/116 and the Office of Vice Chancellor for Research of Islamic Azad University, Sirjan Branch.

References

[1] Amin, G. R. and M. Toloo: A polynomial-time algorithm for finding ε in DEA models. *Computers & Operations Research*, **31** (2004), 803-805.

[2] Amirteimoori, A. and A. Emrouznejad: Optimal input/output reduction in production processes. *Decision Support Systems*, **52** (2012), 742-747.

[3] Banker, R. D., A. Charnes, and W. W. Cooper: Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, **30** (1984), 1078-1092.

[4] Charnes, A., W. W. Cooper, and E. Rhodes: Measuring the efficiency of decision making units. *European Journal of Operational Research*, **2** (1978), 429-444.

[5] Cooper, W. W., L. M. Seiford, and T. K., *Data Envelopment Analysis: Comprehensive Text with Models, Applications, References and DEA-Solver Software*, 2nd ed. Springer, Berlin, 2007.

[6] Dyson, R. G., R. Allen, A. S. Camanho, V. V. Podinovski, C. S. Sarrico, and E. A. Shale: Pitfalls and protocols in DEA. *European Journal of Operational Research*, **132** (2001), 245-259.

[7] Jenkins, L. and M. Anderson: A multivariate statistical approach to reducing the number of variables in data envelopment analysis. *European Journal of Operational Research*, **147** (2003), 51-61.

[8] Morita, H. and N. K. Avkiran: Selecting inputs and outputs in data envelopment analysis by designing statistical experiments. *Journal of the Operations Research Society of Japan*, **52** (2009), 163-173.

[9] Nataraja, N. R. and A. L. Johnson: Guidelines for using variable selection techniques in data envelopment analysis. *European Journal of Operational Research*, **215** (2011), 662-669.

[10] Ramón, N., J. L. Ruiz, and I. Sirvent: Common sets of weights as summaries of DEA profiles of weights: With an application to the ranking of professional tennis players. *Expert Systems with Applications*, **39** (2012), 4882-4889.

[11] Toloo, M.: An epsilon-free approach for finding the most efficient unit in DEA. *Applied Mathematical Modelling*, **38** (2014), 3182-3192.

[12] Toloo, M.: The role of non-Archimedean epsilon in finding the most efficient unit: With an application of professional tennis players. *Applied Mathematical Modelling*, **38** (2014), 5334-5346.

[13] Toloo, M., M. Barat, and A. Masoumzadeh: Selective measures in data envelopment analysis. *Annals of Operations Research*, **226** (2014), 623-642.

[14] Toloo, M. and T. Tichý: Two alternative approaches for selecting performance measures in data envelopment analysis. *Measurement*, **65** (2015), 29-40.

Employee Selection – Case Study Applying Analytic Hierarchy Process

Zuzana Kiszová[1]

Abstract. The employee selection is one of the key activities of human resources management in organizations. The result of this process influences performance and success of work teams, education and training costs. An interview, personality and performance tests and assessment centres are methods often used in selection process in organizations. Some supplementary methods can be applied.

The case study demonstrates one part of director assistant selection. The selection process in this case is divided in three parts. As the first step, unsuitable candidates not satisfying necessary conditions are eliminated. As the second step, candidates are completing performance test. As the third step, three best candidates are interviewed by personnel manager and director. The case study focuses on the second part – the task completing as a performance test.

The usual way of determining results applied in the selected company is total sum of points. Modification of this procedure by analytic hierarchy process (AHP) is suggested in this paper. Particular tasks are pairwise compared and their priorities are derived. Results gained by both methods are compared.

Keywords: employee selection, analytic hierarchy process, priorities.

JEL Classification: C44, O15
AMS Classification: 90B50, 90C29, 91B06

1 Introduction

Effectively invested disposable resources are one of company competitiveness conditions. Enterprises use material, financial and human resources to reach their goals. All these resources are of strategic significance, they are connected to each other and one cannot be used without the others.

Human resources management includes many personnel activities. Employee selection is one of them. It influences performance and success of work teams or education and training costs. This activity is concerned in selection of a candidate satisfying some general and particular characteristics and conditions. Various methods are used for this purpose. Analytic hierarchy process can be applied in the selection process, too; for example as in [6] for criteria selection, in [2] for human performance improvement, in [7] for competency models or in the fuzzy modification as in [1], [5] or [10].

This paper consists of 3 chapters. The first one presents some employee selection methods briefly. The second chapter is focused on analytic hierarchy process and the third one occupies with the case study of employee selection using the analytic hierarchy process. Then the conclusion follows.

2 Employee selection

Recruitment and selection go hand in hand with each other. Recruitment includes defining requirements and attracting candidates. Selecting candidates means sifting applications, interviewing, testing, assessing candidates, assessment centres, offering employment, obtaining references [3]. To the most popular selecting methods belong interview, assessment centres and tests.

Some types of *interviews* may be distinguished:

- *individual interview* – there is one person on each side, i.e. one candidate and one interviewer; it gives opportunity to become more relaxed and feel comfortable;
- *interviewing panel* and *selection board* – one candidate and two or more interviewers attend the interview, which enables to all participants to bring up their questions, but the candidate may become nervous and uncomfortable.

[1] Silesian University in Opava, School of Business Administration in Karviná, Department of Informatics and Mathematics, Karviná, Univerzitní náměstí 1934/3, kiszova@opf.slu.cz.

Assessment centre is comprehensive method of selection and/or evaluation. It enables to observe candidates' behavior in different situations: team work, conflict management, communication with other people. It is focused especially on social skills.

Many types of *tests* may be used in the selection, for example personality tests, intelligence tests or skills and knowledge tests [8]. Some of them are of psychological character, some of them are more qualification-focused. These tests may be evaluated by internal employee or outsourced to a specialized company.

3 Analytic hierarchy process

Analytic hierarchy process is multicriteria decision-making method. The problem is structured in a hierarchy of three (or more) levels. The goal of the problem represents the highest level, the second one belongs to criteria, i. e. substantial factors influencing the decision (or evaluation), and alternatives to be assessed are on the last level of hierarchy. The criteria may be quantitative and qualitative, too. Quantitative criteria are of minimizing or maximizing character.

The pairwise comparisons method is used to derive unknown/undetermined weights (priorities) of objects on each hierarchy level. All objects are compared to each other by couples. If there are numerical characteristics of object, these are pair-compared. If the characteristics of objects are qualitative, the nine-point scale is applied to express the difference of preferences in couple of objects. Number one means equality, number nine represents extreme difference between objects, see [11]. See table 1.

Intensity of importance	Definition
1	Equal importance
2	Weak
3	Moderate Importance
4	Moderate plus
5	Strong Importance
6	Strong plus
7	Very strong Importance
8	Very, very strong
9	Extreme importance

Table 1 The nine-point scale

Values of the pairwise comparisons represent estimation of weight ratio of two compared elements of the same hierarchic level:

$$a_{ij} = \frac{w_i}{w_j}, \tag{1}$$

where a_{ij} is value of pairwise comparison between the i-th and j-th object, w_i is weight of the i-th object, w_j is weight of the j-th object. The i-th object is equal to itself, corresponding value is 1.

There is multiplicative reciprocity between pair-compared objects:

$$a_{ji} = \frac{1}{a_{ij}} \text{ or} \qquad a_{ij} \cdot a_{ji} = 1. \tag{2}$$

Values of pairwise comparisons are inserted in the pairwise comparison matrix $\mathbf{A}$. Maximal eigenvalue $\lambda_{\max}$ and corresponding eigenvector $\mathbf{w}$ are to be calculated according to the characteristic equation:

$$\mathbf{Aw} = \lambda_{\max}\mathbf{w}. \tag{3}$$

Some special attributes of this matrix ensure relatively simple calculation of its maximal eigenvalue $\lambda_{\max}$ and corresponding eigenvector $\mathbf{w}$. When normalized, i.e. $\sum_{i=1}^{n} w_i = 1$, element w_i of vector $\mathbf{w}$ represents the relative importance of the i -th object.

The pairwise comparison matrix is *square*. All n objects of given hierarchical level are compared to each other and the $n \times n$ matrix is created. It is enough to execute $(n^2 - n)/2$ pairwise comparisons with respect to the reciprocity.

The matrix is *nonnegative*, too. If pairwise comparisons are expressed by the nine-point scale, the possible values are $\{1/9; 1/8; ...; 1/2; 1; 2; ...; 8; 9\}$. If pairwise comparisons are expressed by real number ratio, the value may be negative. Sufficiently large positive number has to be added to all pair-compared entry values to get nonnegative matrix.

The pairwise comparison matrix is *irreducible*. That means it is not possible to rearrange the columns and rows to get zero submatrix. This attribute is ensured when expressing pairwise comparisons by the nine-point scale. If the pairwise comparison value got by the real number ratio is zero, it is necessary to add sufficiently large positive number to all entry values.

The Perron-Frobenius theorem ensures existence of the maximal eigenvalue and corresponding eigenvector including positive components for such matrix (see [9], page 673). The Wieland theorem is applied to derive the eigenvector, see e.g. [4]:

$$\mathbf{w} = \lim_{k \to \infty} \frac{\mathbf{A}^k \mathbf{e}}{\mathbf{e}^{\mathrm{T}} \mathbf{A}^k \mathbf{e}} \tag{4}$$

where $\mathbf{A}^k$ is the k -th power of matrix $\mathbf{A}$, $\mathbf{e}$ is vector of ones, i.e. $\mathbf{e}^{\mathrm{T}} = (1; 1; 1; ...; 1)$.

Some inconsistency may appear in pairwise comparisons. It means the following consistency condition is not satisfied:

$$a_{ij} \cdot a_{jk} = a_{ik} \text{ for all } i, j, k = 1,2,...,n. \tag{5}$$

Inconsistency is measured by inconsistency index I_c. It is calculated for $n \times n$ matrix as follows:

$$I_c = \frac{\lambda_{\max} - n}{n - 1}. \tag{6}$$

The inconsistency index must not exceed the threshold of 10 %. In such a case the matrix is considered to be sufficiently consistent. Otherwise the pairwise comparisons have to be reassessed.

Weighted sum is calculated when weights of all criteria and weights of all alternatives according to all criteria are derived. The result is overall weights of alternatives with regard to the goal. This result gives final ranking of alternatives.

4 Case study

The case study concerns with selection of an employee in a medium enterprise. This high-tech enterprise is focused on automated measuring and testing systems, electricity quality monitoring systems, industrial testers and camera systems. The enterprise looks for a director assistant.

Some essential requirements of education and skills and some information about the job are announced: secondary education, typing skills, English communication skills, Microsoft Office knowledge and driving skills. Applications and curricula vitae are sent to the personnel department. Some unsuitable candidates not satisfying necessary conditions are eliminated. Completing performance test is the first selection round. Candidates have to complete six written tasks by means of computer and the Internet: transport order, accommodation reservation, costing, translation from Czech to English, public contract administration and personality presentation. These tasks are usual part of the director assistant work. Three best candidates are chosen and they are invited to second selection round – the interview with personnel manager and director. Here the notions of salary, work tasks and working hours are discussed.

The case study focuses on the first round – the performance test comprising six tasks:

- *transport order* (TO) – some information about the transporter, loading and unloading place and the shipment parameters are given and candidates write a transport order in Czech and find insurance options information; formal arrangement, content and insurance options are evaluated, 2 + 2 + 1 points may be given;
- *accommodation reservation* (AR) – some information about the hotel, date and room equipment are given and candidates shall write an accommodation reservation in English; formal arrangement and content are evaluated, 2 + 2 points may be given;
- *costing* (C) – some information about cars parameters, average fuel consumption, driven distance in some months are given; candidates shall create tables of fuel consumption, driven distance and average fuel costs in Microsoft Office Excel (they shall find out a price figure source, too); procedure and results, formulas used in MSO Excel and price figure source finding out are evaluated, 2 + 2 + 1 points may be given;
- *translation from Czech to English* (T) – a professional text in English is given and candidates shall read it out in English and translate it in Czech using the Internet; pronunciation, translation and text understanding are evaluated, 1 + 1 + 2 point may be given;
- *public contract administration* (PCA) – some materials about a public contract are given, candidates shall describe necessary documentation and their procuring; list of documents and way of their procuring are evaluated, 2 + 2 points may be given;
- *personality presentation* (PP) – candidates shall create short presentation about themselves using Microsoft Office PowerPoint; creativity and applied effects as well as content are evaluated, 2 + 2 points may be given.

Five suitable candidates are in the first selection round. Let's call them candidate A, candidate B, candidate C, candidate D and candidate E. Their performances in each task are executed by the personnel manager, results are in table 2.

Candidate	**TO**	**AR**	**C**	**T**	**PCA**	**PP**	**total**
Candidate A	1 + 2 + 1	2 + 1	2 + 1 + 1	0,5 + 0,5 + 2	0,5 + 0,5	1,5 + 1	17,5
Candidate B	1,5 + 1 + 0	1 + 0,5	2 + 0,5 + 1	1 + 0,5 + 1	0,5 + 0	2 + 2	14,5
Candidate C	2 + 1,5 + 1	2 + 1,5	2 + 1,5 + 1	0,5 + 0,5 + 1	1 + 0,5	1,5 + 1,5	19
Candidate D	1 + 1 + 0	0,5 + 0,5	0,5 + 0,5 +0	0,5 + 0,5 +0	0,5 + 0,5	1 + 1	8
Candidate E	1,5 + 1 + 0	0,5 + 0,5	2 + 1 + 1	0 + 0 + 0	1 + 0,5	1 + 1	11

Table 2 Candidates evaluation

Original evaluation procedure consists in total sum of points (last column in table 2). The personnel manager wishes to differentiate significance of tasks because some activities prevail in the director assistant job. These tasks should have stronger influence on the decision making. Personnel manager is not able to determine task weights directly, she compares all tasks in pairs according to their significance in the director assistant job and creates pairwise comparison matrix of tasks (figure 1).

	TO	AR	C	T	PCA	PP
TO	1	3	3	5	1/4	7
AR	1/3	1	3	4	1/2	6
C	1/3	1/3	1	2	1/5	4
T	1/5	1/4	1/2	1	1/6	2
PCA	4	2	5	6	1	7
PP	1/7	1/6	1/4	1/2	1/7	1

Figure 1 Pairwise comparison matrix of tasks

Final rankings of candidates derived by original and weighted procedure are given in table 3.

	Candidate A	**Candidate B**	**Candidate C**	**Candidate D**	**Candidate E**
Original	2	3	1	5	4
Weighted	2	4	1	5	3

Table 3 Candidates ranking in original and weighted procedure

Let's remain that the best three candidates go to the second selection round. As we can see, candidate C is the best one in both cases and candidate A is the second best according to both types of selection procedure. Candidate D is always on the last position. Candidate B would go to the second selection round according to original procedure, but he is positioned as the fourth one in the weighted approach. Candidate's E situation is opposite, he is ranked fourth by the original procedure and third according to weighted approach.

As we can see, there is one difference in ranking of candidates. But this variance changes the triple composition of candidates going to the second selection round and it may cause different result of the whole employee selection process. Ranking alternations are brought about by deriving priorities of decisive factors. If candidate gets high score in highly valued factor, his total score is rising faster than if he gets high score in less valued factor. If one candidate gets high score in highly valued factor and the other one is evaluated only little points in the same factor, the gap between them becomes wider. This is the reason of changes is candidates ranking derived by original and modified (weighted) procedure.

5 Conclusion

Analytic hierarchy process and the eigenvector method can be used successfully in human resources management including employee selection. Sometimes people are not sure how to set considered factors priorities and that may be reason of getting unsatisfactory results in evaluating processes. The possibility of deriving still unknown/undetermined weights of object is considerable advantage of analytic hierarchy process. These weights reflect evaluators' preferences, the assessment and selection becomes more realistic and results are more acceptable for the evaluator.

In this paper, eigenvector method was applied in the case study of employee selection process. The personnel manager provided additional indispensable information, i.e. she performed pairwise comparisons of tasks. Unknown weights of decisive factors were derived by the eigenvector method and the original procedure of determining candidate ranking was modified from total sum of points to total weighted sum of points. This adaptation brought minor change of candidates ranking and different triple of candidates going to the second selection round was gained. Priorities inclusion may cause major, minor or no changes in candidates ranking – it depends on individual setting of the selection process.

Acknowledgements

This paper is supported by GA CR project no. 14-02424S.

References

[1] Ablhamid, R. K., Santoso, B., and Muslim, M. A.: Decision Making and Evaluation System for Employee Recruitment Using Fuzzy Analytic Hierarchy Process. *International Refereed Journal of Engineering and Science* **2** (2013), 24-31.

[2] Albayrak, E.: Using analytic hierarchy process (AHP) to improve human performance: An application of multiple criteria decision making problem. *Journal of Intelligent Manufacturing* **15** (2004), 491-503.

[3] Armstrong, M.: *A Handbook of Human Resource Management Practice*. Kogan Page, London, 2006.

[4] Gavalec, M., Ramík, J., and Zimmermann, K.: *Decision Making and Optimization: Special Matrices and Their Applications in Economics and Management*. Springer International Publishing, Berlin, 2015.

[5] Gungor, Z., Serhadlıoglu, S., and Kesen, S. E.: A fuzzy AHP approach to personnel selection problem. *Applied Soft Computing* **9** (2009), 641–646.

[6] Hsiao, W.-H., Chang T.-S., Huang, M.-S., and Chen, Y.-C.: Selection criteria of recruitment for information systems employees: Using the analytic hierarchy process (AHP) method. *African Journal of Business Management* **5** (2011), 6201-6209.

[7] Kashi, K., and Friedrich, V.: Manager's Core Competencies: Applying the Analytic Hierarchy Process Method in Human Resources. In: Proceedings of the 9th Eureopean Conference on Management Leadership and Governance (SemmelrockPicej, MT; Novak, A, eds.). ACAD CONFERENCES LTD., Klagenfurt, 2013, 384-393.

[8] Koubek, J.: *Řízení lidských zdrojů*. Management Press, Praha, 2006.
[9] Meyer, C. D.: *Matrix Analysis and Applied Linear Algebra: Book and Solutions Manual.* SIAM, Philadelphia, 2000.
[10] Rouyendegh, B. D., and Erkan, T. E.: Selection of academic staff using the fuzzy analytic hierarchy process (FAHP): A pilot study. *Technical Gazette* **19** (2012), 923-929.
[11] Saaty, T. L.: *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy process*. RWS Publications, Pittsburgh, 1994.

Three dimensional Bin Packing Problem in batch scheduling.

František Koblasa[1], Miroslav Vavroušek[2],František Manlig[3]

Abstract. Nowadays planning is aided with Advanced Planning and Scheduling Systems that are calculating with constrained capacities. However there is used common scheduling approach of assigning jobs to resources that complete the work in specific time. Processing tasks, which are composed from segments of several preceding jobs, are not scheduled in this way, but long term planning approach is used.

This paper is following recent trends in manufacturing systems and operations research focusing on scheduling special technologies as rapid prototyping, heat treatment or surface finishing technologies. These technologies require two stage scheduling approach of work space partitioning and classical batch scheduling.

This article is proposing early research done by reviewing three dimensional Bin Packing Problem and by developing and testing constructive algorithm inspired by MinfMaxr approach with goal to discuss further approach of special technologies batch scheduling. Batch scheduling scheme and its possible implementation in to Advanced Planning and Scheduling system is than discussed.

Key words: Space Partitioning, Bin Packing Problem, Scheduling, Constructive Algorithm

JEL Classification: L23
AMS Classification:90C27

1 Introduction

Fulfilment of the fourth industrial revolution i.e. Industry 4.0 is widely debated with ease and scepticism but it is sure that will contains basic principles of automation with aid of methods of operation research as artificial intelligence [21] , multi-criteria decision making [9], ending with transportation optimization[23].

Production planning and scheduling using information systems usually consist of hierarchic module data sharing. The first hierarchic level focuses on material availability (Material Requirements Planning) by planning long term utilization of material given by forecast of customer demand. There is also limited job planning, which is based on production lead times. However, this approach provide limited accuracy of setting due date of each jobs because it is not considering machine utilization.

The second level is than focused on calculating resources, usually machines, utilization (Manufacturing Resource Planning). Second level provide corrections of first level provided due dates by moving term planned tasks along planning horizon by forward and backward scheduling while checking 100% machine utilization. Levelling machine utilization below desired level can be done automatically by moving jobs further in time or, and more likely, by information system user interventions. This kind of schedule is reflecting utilization and production lead times on one side and not taking in account job conflicts on machines (more jobs are planned on one resource at same time) on other side.

The third level known as advanced planning and scheduling is than removing job conflict by so called scheduling in to constrained capacities. Constrained capacities usually means machines and in the case of considering workers or setup jigs we are talking about multi constrained manufacturing planning (models).

The fourth level is used to day to day operative planning which purpose is to limit random events in the system (defects, machine breakdowns, missing workers etc.).

This paper is focussing on the third level of production planning and mainly on cases where there is possible to schedule several jobs in one tasks and resource is constrained by working space of machines. This problem is addressed as scheduling of P-batch where batch processing time is equal to the maximum processing time among

[1]Department of Manufacturing Systems and Automation, Faculty of Mechanical Engineering,
Technical University of Liberec, Studentská 2, Liberec 1, Czech Republic, e-mail: frantisek.koblasa@tul.cz.
[2]Department of Manufacturing Systems and Automation, Faculty of Mechanical Engineering,
Technical University of Liberec, Studentská 2, Liberec 1, Czech Republic, e-mail: miroslav.vavrousek@tul.cz.
[3] Department of Manufacturing Systems and Automation, Faculty of Mechanical Engineering,
Technical University of Liberec, Studentská 2, Liberec 1, Czech Republic, e-mail: frantisek.manlig@tul.cz..

all tasks. This problem has wide variety of application in manufacturing systems as mechanical engineering (heat treating furnaces [20], during blasting [14] or burnishing [19]), electrical engineering (chip testing operations [13], long term heating or baking-out [12], wafer fabrication [1], manufacturing of ceramic semiconductors [17]) and services (operation Medical processing of sterilization as shown in [18] and [22]) as:
P-batch scheduling is in this paper is taking in account placing incompatible job families with product different dimensions in to the working space. This problem is addressed as Bin Packing Problem. Following chapters are dealing with three dimensional Bin Packing Problem (3DBPP). There are reviewed basic classification and solution methods of mentioned problem together with preliminary testing of designed Maximal Box algorithm on benchmark problems. Batch scheduling problem together with special technology constraints to 3DBPP and usability of designed algorithm is than discussed.

2 Three dimensional Bin Packing Problem

Our problem which contains batch scheduling of various types of product focus on placing object in special technologies like heating treatment or 3D printing (ie. Rapid Prototyping- RP) devices with goal to minimize overall process time so number of batches i.e. bins. This problem can be modeled or described as bin packing problem. We are focusing on bin packing problem, where complex shape of inserted object in to the bin is simplified to three-dimensional object - box.

The Three-Dimensional Bin Packing Problem (3D-BPP) consists of orthogonally packing all the items into the minimum number of bins. 3DBPP is strongly NP-hard as it is a generalization of the well-known one dimensional Bin Packing Problem (1D-BPP), in which a set of n positive values w_j has to be partitioned into the minimum number of subsets so that the total value in each subset does not exceed a given bin capacity W.

3D BPP has lot of applications which can be divided in to two basic classes:

- Three-dimensional optimal packing - space optimization and volume utilization with rectangular shaped boxes as container loading optimization, pallet building and truck loading, air cargo load planning, warehouse management systems etc.
- Three-dimensional Limited Resource Scheduling - multi-dimensional containment heuristic approaches as multi-dimensional limited resource scheduling and constraint based resource planning, job-shop scheduling, finite capacity scheduling optimization, logistics planning, time tabling optimization algorithms etc.

There are various methods how to solve BPP begging with classical methods like branch and bound [16], mix linear integer programming [8], local search [24] evolution base [6] or hybrid approach [26]. Solving method is usually dependent on additional constraints i.e. problems. Most models deal with packing connected problems as warehouse management [25], distribution logistics [3] but also with manufacturing [7].

We are solving 3D BPP to implement in three-dimensional Limited Resource Scheduling, so combining BPP with job shop scheduling, which is solved by hybrid genetic algorithm [10].

In our case, there are given a set of n rectangular-shaped items, each characterized by width w_j , height h_j and depth d_j ($j \in J = \{1;\dots; n\}$), and an unlimited number of identical three-dimensional containers (bins) having width W, height H and depth D. We assume that the items may be rotated within all axes (x,y,z). Following chapters describes and tests constructive based algorithm (Maximum box algorithm), to solve before described problem.

2.1 Maximum box algorithm

Maximum box algorithm is inspired by Maximal rectangle algorithms presented in [16] and [11]. The set of placed boxes is arranged according to the volume from the biggest to the smallest. The first bin of given proportions and a box of free space covering the entire space of the bin are created at the preparation phase of the algorithm.
The placed boxes are consecutively chosen from the arranged set. Task has no solution if the first block has bigger volume than the space of the bin. The first block is placed in the corner of the free space. The free space is divided into three new free spaces which covers the space of the cut out boxes. Following boxes are placed in to newly created spaces. Boxes are rotated in all axes if they do not fit at first try. New bin is created if there is no other option of placement. Following blocks (third and following ones) are placed in to newly design free spaces as it is shown at Figure 1. Free spaces which are only subspaces of the bigger ones are eliminated.

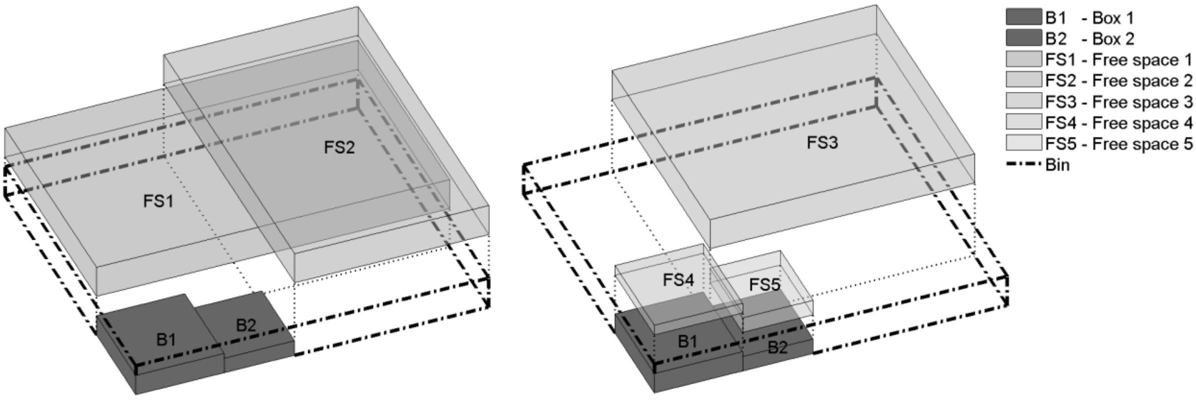


Figure 1 Free space fragmentation

Algorithm is than placing and rotating boxes, by procedure described below, until there are no boxes available:

Data structure

Bins contain set of Bin
Bin contain set of FreeSpace
FreeSpace = {*w*; *d*; *h*; centerX; centerY; centerZ}
Boxes contain set of Box
Box = {w; d; h; centerX; centerY; centerZ}

Initialize:

```
create new Bin B in the Bins
add new FreeSpace F = (W; D; H) to B
```

Pack:

```
sort Box in the Boxes by size of volume from the biggest
foreach Box X in the Boxes
    Bool placed = False

    foreach Bin B in the bins
        foreach FreeSpace F in the Bin B
            for rot = 1:6
                B = rotate X by rot

                if F.w>=X.w and F.d>=X.d and F.h>=X.h
                    placed = True
                    compute F - X and subdivide the result into at most six new FreeSpaces G1;..;G6
                    add G1;...;G6 to B
                    break
        if placed = True
            break
    if placed = False
        create new Bin B in the Bins = (W; D; H)
        add new FreeSpace F = (W; D; H) to B
        compute F - X and subdivide the result into at most six new FreeSpaces G1;...;G6
        add G1;...;G6 to B
    optimize FreeSpaces in the Bin B (Erase FreeSpace, which is only subspace)
    sort FreeSpaces in the Bin B by size of the volume from the smallest
```

2.2 Test problems and Results

We considered three types of instances published in [16]. All test instances are available at [4]. The first class of instances are generalizations of the instances considered by Martello and Vigo in [15].

The first class instance has bin size is $W = H = D = 100$ and five types of items are considered (u.r. stands for „Uniformly random") with:

- Type 1: w_j u.r. in [1;1/2W], h_j u.r. in [2/3H; H], d_j u.r. in [2/3D; D].
- Type 2: w_j u.r. in [2/3W; W], h_j u.r. in [1;1/2H], d_j u.r. in [2/3D; D].
- Type 3: w_j u.r. in [2/3W; W], h_j u.r. in [2/3H; H], d_j u.r. in [1;1/2D].

- Type 4: w_j u.r. in [1/2W; W], h_j u.r. in [1/2H; H], d_j u.r. in [1/2D; D].
- Type 5: w_j u.r. in [1;1/2W], h_j u.r. in [1;1/2H], d_j u.r. in [1;1/2D].

The second class is a generalization of the instances presented by Berkey and Wang [2]:

- Type 6: bin size W = H = D = 10; items with w_j ; h_j ; d_j are u.r. in [1; 10].
- Type 7: bin size W = H = D = 40; items with w_j ; h_j ; d_j are u.r. in [1; 35].
- Type 8: bin size W = H = D = 100; items with w_j ; h_j ; d_j are u.r. in [1; 100].

The third class consists of “all-fill” problems [20]:

- Type 9: instances have a known solution with three bins, since the items are generated by cutting the bins into smaller parts. For a problem with n items, bins 1 and 2 are cut into $n/3$ items each, while bin 3 is cut into n-2(n/3) items. The cutting is made using procedure presented in [16].

Table 1 presents computation time results (*t* [*s*]), where (0.00) time represent result where computation time was below one hundred of a second and (–) represents situation, where ONEBIN algorithm was not able to find solution within 1000 s which satisfied lower bound definition in [16].

n	Method	Type																	
		1		2		3		4		5		6		7		8		9	
		t [s]	σ	*t* [s]	σ	*t* [s]	σ	*t* [s]	σ	*t* [s]	σ	*t* [s]	σ	*t* [s]	σ	*t* [s]	σ	*t* [s]	σ
10	MaxBox	0.02	0.28	0.02	0.50	0.02	0.50	0.02	0.78	0.02	0.47	0.01	0.25	0.01	0.33	0.02	0.35	0.01	0.17
	ONEBIN	0.00	0.32	0.00	0.59	0.00	0.64	0.00	0.83	0.01	0.67	0.00	0.45	0.01	0.67	0.01	0.40	0.00	0.00
15	MaxBox	0.05	0.34	0.04	0.32	0.04	0.44	0.06	0.70	0.03	0.44	0.02	0.23	0.03	0.38	0.04	0.30	0.01	0.37
	ONEBIN	0.00	0.47	0.00	0.45	0.00	0.50	0.00	0.74	0.01	0.61	0.01	0.47	0.01	0.43	0.01	0.37	0.00	0.00
20	MaxBox	0.09	0.36	0.09	0.30	0.08	0.33	0.11	0.80	0.07	0.33	0.03	0.26	0.06	0.40	0.08	0.33	0.02	0.30
	ONEBIN	0.01	0.43	0.01	0.53	0.00	0.43	0.00	0.86	0.02	0.44	0.01	0.38	0.02	0.52	0.01	0.48	0.01	0.00
25	MaxBox	0.14	0.28	0.14	0.37	0.13	0.43	0.19	0.82	0.11	0.34	0.05	0.20	0.10	0.19	0.13	0.29	0.04	0.33
	ONEBIN	0.05	0.37	0.01	0.43	0.08	0.51	0.00	0.86	0.39	0.44	0.03	0.31	–	0.42	0.04	0.45	0.01	0.00
30	MaxBox	0.21	0.33	0.21	0.39	0.21	0.36	0.28	0.78	0.17	0.39	0.07	0.17	0.14	0.34	0.19	0.26	0.06	0.33
	ONEBIN	3.40	0.43	0.40	0.43	0.48	0.41	0.01	0.81	0.05	0.42	0.79	0.26	–	0.45	0.07	0.40	1.96	0.00
35	MaxBox	0.32	0.33	0.31	0.32	0.31	0.36	0.43	0.83	0.26	0.36	0.10	0.16	0.20	0.32	0.28	0.36	0.09	0.40
	ONEBIN	12.02	0.35	10.02	0.38	2.51	0.37	0.01	0.87	–	0.52	0.38	0.22	–	0.43	–	0.44	15.97	0.00
40	MaxBox	0.45	0.34	0.45	0.35	0.43	0.41	0.62	0.88	0.35	0.36	0.12	0.17	0.28	0.29	0.37	0.30	0.13	0.37
	ONEBIN	35.15	0.33	–	0.39	–	0.42	0.01	0.90	–	0.49	1.73	0.24	–	0.62	28.93	0.35	–	0.07
45	MaxBox	0.60	0.33	0.59	0.35	0.55	0.39	0.83	0.91	0.47	0.34	0.16	0.18	0.35	0.34	0.50	0.27	0.16	0.37
	ONEBIN	190.98	0.39	–	0.42	–	0.40	0.01	0.92	–	0.42	2.42	0.22	–	0.61	–	0.33	–	0.47
50	MaxBox	0.76	0.36	0.76	0.36	0.74	0.42	1.04	0.85	0.62	0.34	0.18	0.14	0.47	0.27	0.64	0.24	0.24	0.33
	ONEBIN	–	0.41	–	0.39	–	0.39	1.67	0.86	–	0.49	1.25	0.17	–	0.47	–	0.41	–	0.63
60	MaxBox	1.17	0.31	1.17	0.34	1.15	0.32	1.75	0.94	0.95	0.32	0.27	0.11	0.72	0.25	1.06	0.26	0.32	0.33
	ONEBIN	–	0.37	–	0.40	–	0.38	–	0.94	–	0.50	34.35	0.17	–	0.52	–	0.35	–	0.67
70	MaxBox	1.74	0.31	1.72	0.28	1.75	0.33	2.48	0.91	1.39	0.29	0.36	0.12	1.07	0.27	1.55	0.28	0.54	0.33
	ONEBIN	–	0.39	–	0.35	–	0.40	–	0.93	–	0.46	–	0.14	–	0.55	–	0.40	–	0.80
80	MaxBox	2.45	0.27	2.45	0.30	2.30	0.29	3.63	0.92	1.91	0.27	0.50	0.11	1.35	0.32	2.19	0.24	0.69	0.43
	ONEBIN	–	0.35	–	0.40	–	0.42	–	0.94	–	0.53	–	0.19	–	0.60	–	0.41	–	1.00
90	MaxBox	3.35	0.30	3.41	0.34	3.27	0.30	5.11	0.93	2.74	0.30	0.58	0.13	1.97	0.23	2.85	0.21	0.96	0.33
	ONEBIN	–	0.39	–	0.41	–	0.38	–	0.93	–	0.50	–	0.18	–	0.54	–	0.35	–	1.07

Table 1 Maxbox and ONEBIN algorithm comparison

For each type (1-9) there where generated 10 tests with n parts $n \in \langle 10; 15; 20;\ldots; 45\rangle \wedge \langle 50; 60; 70;\ldots; 90\rangle$.
All tests are comparing results of designed algorithm with ONEBIN approach presented by Martello and Vigo [15]. MaxBox and ONEBIN algorithm where compared with lower bound represented by total bin volume which is necessary to pack (1)

$$V_j = W_j \times H_j \times D_j \tag{1}$$

and deviation σ (2), as there is no possibility to get lower number of bins than that is given by volume can be achieved, is than calculated as (V_j – part volume, V_{bin} –bin volume R_i – number of bins given by method):

$$\sigma = \frac{\sum_{i=1}^{10} R_i - \left\lceil \frac{\sum_{j=1}^{n} V_j}{V_{bin}} \right\rceil}{10} \qquad (2)$$

Comparing MaxBox and ONEBIN scheme we can see, that in the models with small number of parts (below 25) branching scheme gets results faster than MaxBox, however its time prolongs considerably as solution space widens (in some cases more than 1000s). The values of deviance to lower bound shows, that in most cases MaxBox algorithm performs same or better than ONEBIN, however it was not able to reach optimums (e.g. lower bounds). This is expected result as MaxBox is constructive algorithm which only optimization technique uses sorting by volume and trying to fit objects to free space by rotation only if position was not suitable in the first pack try.

3 Technology based constraints of batch scheduling

The purpose of this research is to schedule manufacturing resources (e.g. constraints) with incompatible job families, non-identical job sizes and dynamic job arrivals with objective function of maximizing the utilization of the batch processors or to minimize the total weighted tardiness. General approach is dividing scheduling steps in to loading available batches (including part properties – to indicate job families of manufacturing process), sorting parts by families (i.e. by technology, material, mass, surface), design manufacturing batches from families (solved by bin packing problem) and create space decomposition schedule for operative planning. Than is necessary to assign respective bins to a suitable machine and to determine the completion time.

There is necessary to take in mind constraints based on the technology. Focussing on the special technology of RP we had considered basic constraints as in [5]. Laser melting or polyjet (in our laboratories) are not suitable for 3D BPP as there is not possible to pile up objects in the working space. However technologies like FDM can use MaxBox with restrictions to rotate object. Constraints to rotate parts vertically are based on required surface, and flexibility strength properties. Processing time used for scheduling is than based on sum of processing parts with respect of building support layers.

4 Conclusion an further research

Designed MaxBox algorithm shoved promising result and could be used operation connected to technologies as Rapid prototyping or heat treatment as tool to design manufacturing batches. However, it is necessary to make in-deep analysis of particular technology process at defined machine to schedule whole system in the holistic point of view. Further research will focus on development of meta-heuristic based on evolution algorithm with hybridization by searching procedure with use of knowledge presented in this article and on definition of the multi-objective function, which could cover most of the enterprise requirements as machine utilization, meeting due dates etc.

Acknowledgements

This publication was written at the Technical University of Liberec as part of the project " Project 21130 - Research and development in the field of 3D technology, manufacturing systems and automation" with the support of the Specific University Research Grant, as provided by the Ministry of Education, Youth and Sports of the Czech Republic in the year 2016.

References

[1] Bang, J. Y., Kim, Y. D. and Choi, S. W.: Multiproduct Lot Merging–Splitting Algorithms for Semiconductor Wafer Fabrication. *Semiconductor Manufacturing*, IEEE Transactions, **25**, 2, (2012), 200-210.

[2] Berkey, J. O., and Pearl Y. W.:Two-dimensional finite bin-packing algorithms. *Journal of the operational research society* (1987), 423-429.

[3] Bortfeldt, A., and Homberger, J.: Packing first, routing second—a heuristic for the vehicle routing and loading problem. *Computers & Operations Research* (2013), 873-885.

[4] David Pisinger's optimization codes: General three-dimensional Bin-packing Problem [online]. [cit. 2016-05-02]. available at http://www.diku.dk/~pisinger/codes.html

[5] Freens, J. P., Adan, I. J., Pogromsky, A. Y., and Ploegmakers, H.: Automating the production planning of a 3D printing factory. In: *Proceedings of the 2015 Winter Simulation Conference* (2136-2147). IEEE Press. 2015.

[6] Gonçalves, J. F., and Resende, M. G.: A biased random key genetic algorithm for 2D and 3D bin packing problems. *International Journal of Production Economics*, **145**, 2, (2013) 500-510.

[7] Gutin, G., Tommy J. and Anders Y.: Batched bin packing. *Discrete Optimization* 2.1 (2005): 71-82.

[8] Hifi, M., et al.: A linear programming approach for the three-dimensional bin-packing problem. *Electronic Notes in Discrete Mathematics* 36 (2010), 993-1000.

[9] Chyna, V.; Kuncova, M. and Seknickova, J.: Estimation of weights in multi-criteria decision-making optimization models. In: *Proceedings of 30th International Conference on Mathematical Methods in Economics*, 2012, 355-360

[10] Koblasa, F. and Manlig, F.: Application of Adaptive Evolution Algorithm on real-world Flexible Job Shop Scheduling Problems. In: *Proceedings of 32nd International Conference Mathematical Methods in Economics* . First edition. Olomouc, 2014, 425-430. ISBN 978-80-244-4209-9

[11] Koblasa, F., Vavroušek, M. and Manlig, F.: Two-dimensional Bin Packing Problem in batch scheduling. In: *Proceedings of 33rd International Conference Mathematical Methods in Economics*. 1. vyd. University of West Bohemia, Plzeň: University of West Bohemia, 2015. 354 – 359. ISBN 978-80-261-0539-8.

[12] Koh, S.-G., Koo, P.-H., Ha, J.-W., Lee, W.-S. Scheduling parallel batch processing machines with arbitrary job sizes and incompatible job families. *International Journal of Production Research*. **42**, 19, 4091-4107. (2004), ISSN 0020-7543,.

[13] Koh, S.G., Koo, P.H., Kim, D.C., and Hur, W.S.: *Scheduling a single batch processing machine witharbitrary job sizes and incompatible job families*. *International Journal of Production Economics*. **98**, 81-96. (2005), ISSN 0925-5273.

[14] Kumar, V.Kumar, S.Chan, F.T. and S.Tiwari, M.K.: Auction-based approach to resolve the scheduling problem in the steel making process. *International Journal of Production Research*, **44**, 8. 1503-1522. (2006) ISSN 1366-588,

[15] Martello S. and Vigo D.: Exact solution of the two-dimensional finite bin packing problem. *Management Science*. (1998)

[16] Martello, S., Pisinger, D., and Vigo, D.: The three-dimensional bin packing problem. *Operations Research*, **48**, 2, (2000), 256-267.

[17] Mathirajan, M. and A.I. Sivakumar,: A literature review, classification and simple meta-analysis on scheduling of batch processors in semiconductor. *J. Adv. Manuf. Technol*., **29**:(2006), 990-1001.

[18] Ozturk, O. Espinouse, M-L.Mascolo, M.D,and Gouin, A.: Makespan minimisation on parallel batch processing machines with non-identical job sizes and release dates. *International Journal of Production Research*, Taylor & Francis: STM, Behavioural Science and Public Health Titles, **50**, 20, (2012), 6022-6035,

[19] Przybylski, W., and Ściborski, B.: Studies on the influence of technological variants of finishing machining on flow of parts in flexible manufacturing. *Advances in Manufacturing Science and Technology*, **34**,1, (2010), 21-32.

[20] Ramasubramanian, M, Mathirajan, M and Ramachandran V.: Minimizing makespan on a Single Heat-Treatment Furnace in Steel Casting Industry. *International Journal of Services and Operations Management* (Inderscience Publication), **7**, (2010), 112-142.

[21] Raska, P.and Ulrych, Z.: Comparison of optimisation methods tested on testing functions and discrete event simulation models, *International Journal of Simulation and Process Modelling*, **10**, 3, (2015), 279-293.

[22] Rossi, A., Puppato, A. and Lanzetta, M.: Heuristics for scheduling a two-stage hybrid flow shop with parallel batching machines: application at a hospital sterilisation plant. *International Journal of Production Research*, 51, **8**, (2013), 2363-2376,

[23] Šlaichová, E. Štichhauerová, E. and Zbránková, M.: Use of linear programming method to constructing a model for reduction of emission in a selected company. In: *Proceedings of the 33nd International Conference Mathematical Methods in Economics*. 1. vyd. Cheb: ZČU, 2015, 805 – 811. ISBN 978-80-261-0539-8.

[24] Viegas, J. L., Vieira, S. M., Henriques, E. M., and Sousa, J. M.: A Tabu Search Algorithm for the 3D Bin Packing Problem in the Steel Industry. InCONTROLO'2014: In: *Proceedings of the 11th Portuguese Conference on Automatic Control* , 2015, 355-364.

[25] Wu, Y., et al.: Three-dimensional bin packing problem with variable bin height. *European journal of operational research* (2010), 347-355.

[26] Xueping L., Zhaoxia Z., Kaike Z.,:A genetic algorithm for the three-dimensional bin packing problem with heterogeneous bins In: *Proceedings of the 2014 Industrial and Systems Engineering Research Conference*., 2014

Modelling of Regional Price Levels in the Districts of the Czech Republic

Aleš Kocourek[1], Jana Šimanová[2], Jiří Šmída[3]

Abstract. The aim of this article is to suggest and apply methods for estimation of the regional price levels in Czech districts. Its purpose is to provide an instrument for more precise and more realistic comparison of regional standard of living of households across the regions of the Czech Republic. The article contributes to solution of the often discussed problem of nominal income indicators as benchmark of social-economic disparities. Nominal indicators provide distorted information about social and economic position of inhabitants of a region because they do not reflect the regional differences in the costs of living. Authors use basic set of regional price levels in 36 districts (LAU 1) processed by original authors´ certified methodology. This set of the basic results – regional price levels - has been further extended to whole Czech Republic by using econometric modelling methods. The results reflect regional price level differences in twelve CZ-COICOP Headings - market prices of goods, services, as well as housing and rentals. The findings underpin the need for a more accurate specification of economic and social disparities on a lower regional level.

Keywords: regional price levels, real incomes, LAU 1, standard of living.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

The regional policies of the European Union (EU) are targeted among others at sustainable development of regions and improving the citizen's quality of life. The regional convergence has been one of the major issues of economic analyses, while almost a third of the EU budget is set aside for the cohesion policy [25]. The primary indicator for assessment of regional economic performance is the regional gross domestic product compared on the European level in so-called purchasing parity standard (PPS). The PPS is calculated by the Eurostat within the Eurostat-OECD International Comparison Program on the national level and as such it does not take into account the differences in price levels across the regions. [3] Although the regional price levels may constitute an important factor when assessing the economic development of a region, this issue has until recently not received much attention either in the world, in the EU, or in the Czech Republic. [4]

The first attempts to measure the regional price levels in the Czech Republic have been carried out by Musil et al. [17] and Čadil et al. [4] The aim of this paper is to update and rectify their results using slightly more advanced methods of calculation and data processing.

The purpose of this paper is to suggest and apply methods for estimation of the regional price levels in Czech districts (LAU 1) as an instrument for estimating the real standard of living in the Czech regions. Authors use basic set of regional price levels in 36 districts processed by original authors´ certified methodology. This set of basic results – regional price levels - has been further extended to whole Czech Republic by using econometric modelling methods. The results reflect regional differences in market prices of goods, services, as well as housing and rentals. The findings underpin the need for a more accurate specification of economic and social disparities on a lower regional level.

2 Importance and topicality

The need to measure regional price levels originated in the new concept of regional policies which should be generally directed more at the people living in the region than at the area of the region. [7] The problem is, the nominal income indicators provide distorted information about social and economic position of inhabitants of a region because they do not reflect the regional differences in the costs of living. After all, even Kahoun [9] and

[1] Technical University of Liberec/Faculty of Economics, Department of Economics, Studentská 2, 461 17, Liberec, Czech Republic, ales.kocourek@tul.cz
[2] Technical University of Liberec/Faculty of Economics, Department of Economics, Studentská 2, 461 17, Liberec, Czech Republic, jana.simanova@tul.cz
[3] Technical University of Liberec/Faculty of Science, Humanities and Education, Department of Applied Mathematics, Studentská 2, 461 17, Liberec, Czech Republic, jiri.smida@tul.cz

Viturka [26] admit the price levels can vary locally and regionally, especially due to different prices of services and real estate.

In the last ten years, the issue of regional price levels has been addressed by several authors, whose works are often based on regionalization of national price indexes. In the European countries, the attempts to regionalize the price indexes are usually hindered by insufficient or random investigation of prices in the respective regions. At present, the regional price levels are systematically measured and published in the USA, in the UK, and in Australia.

In Germany, the published estimates of regional price levels are based on price survey carried out in 50 German cities in 1994. The first German author, who exploited the price investigation from the viewpoint of regional price levels, was Ströhl. [24] His followers, Schultze [23], Kosfeld et al. [14]), Kosfeld and Eckey [13], and Roos [21,22] look for possible ways of price level estimation in the regions where they have explanatory data at their disposal. They frequently apply econometric modelling and complement the calculation of regional price levels with a real estate price index [13] [14] [20]. Other, often one-off efforts of regional price levels calculations have been carried out in Italy in Pittau et al. [18], China in Brandt and Holz [2] or Gong and Meng [8], Austria in Matzka and Nachbagauer [16] or also in Slovakia in Radvanský and Fuchs [19]. In the Czech Republic, the regional price levels were estimated by Musil et al. [17] on a common consumer basket and by Čadil et al. [4] on a set of regional consumer baskets. They applied the Eurostat-OECD International Comparison Program methods with a certain simplification. They used a national concept (rather than domestic) and calculated the regional price levels for the Czech regions (NUTS 3) based on the historical data from 2007. [17] [4]

3 Methods and data sources

The process of *RPI* construction for 36 Czech districts (using original data from the extensive price surveys in 36 districts carried out by Czech statistical office) was certified by the Ministry of Regional Development of the Czech Republic in December 2015. [10] It is based on the Eurostat-OECD International Comparison Program methods. Results for 36 districts have been published in database of the research project “Regional price index as indicator of real social and economic disparities”. [27]

The aim of further research is to design and apply econometric model that enable to extend the results to whole Czech Republic. We followed a procedure similar to Roos [21], but estimated the partial regional price levels for each of the twelve CZ-COICOP Headings, where CZ-COICOP 01 represents Food and non-alcoholic beverages, 02 - Alcoholic beverages, tobacco and narcotics, 03 - Clothing and footwear, 04 - Housing, water, electricity, gas and other fuels, 05 - Furnishings, household equipment and routine household maintenance, 06 - Health, 07 - Transport, 08 - Communication, 09 - Recreation and culture, 10 - Education, 11 - Restaurants and hotels, 12 - Miscellaneous goods and services.

We tested nearly fifty indicators available for the period 2011–2013 for all 78 districts (LAU 1) of the Czech Republic. However, neither average wage, nor net disposable household incomes were available at the time of estimation on the LAU 1 level. Data on average income after taxation were provided by the General Financial Directorate of the Czech Republic. All the data used for our estimates were recalculated so that they express the average share of a certain district when bilaterally compared to all other districts in the Czech Republic.

The outcomes of our estimations are summed up in the following set of equations (1) – (12). All the parameters were proved significant at the 95% confidence level. All the models passed the Durbin-Watson test on residuals autocorrelation.

$$RPI_{COI01} = 0.991 - 0.020 dens + 0.048 income - 0.018 BU_A \tag{1}$$

$$RPI_{COI02} = -0.184\, pop_{dis} + 0.037 income + 1.091 BU_{ind} + 0.055 BU_G \tag{2}$$

$$RPI_{COI03} = -2.727 - 0.344\, pop_{dis} + 0.094 income + 4.105 BU_{ind} - 0.127 BU_C \tag{3}$$

$$RPI_{COI04} = 0.721 + 0.292 house - 0.023 BU_L \tag{4}$$

$$RPI_{COI05} = 0.961 + 0.148\, pop_{dis} - 0.107 BU_G \tag{5}$$

$$RPI_{COI06} = 1.952 - 0.977\, pop_{15-64} + 0.013\, pop_{>20K} + 0.055 income - 0.062\, phys \tag{6}$$

$$RPI_{COI07} = 0.906 - 0.040 road_{1st} + 0.135 BU_H \tag{7}$$

$$RPI_{COI08} = 1.035 - 0.011 dens - 0.022\, pop_{<5K} \tag{8}$$

$$RPI_{COI09} = 0.984 - 0.038 dens - 0.013 accom + 0.066 BU_{corp} \tag{9}$$

$$RPI_{COI10} = 0.527 + 0.268\, pop_{dis} + 0.193 income \tag{10}$$

$$RPI_{COI11} = 0.864 - 0.031 accom + 0.157 BU_R \tag{11}$$

$$RPI_{COI12} = 0.936 - 0.025 road_{2nd} + 0.083 BU_G \tag{12}$$

where RPI_{COI} are the partial regional price level indexes for CZ-COICOP Headings, predictors are explained in the table below:

$pop_{15\text{-}60}$	share of population in the age from 15 to 60 years
$pop_{<5K}$	share of population living in cities of less than 5,000 inhabitants
$pop_{>20K}$	share of population living in cities of more than 20,000 inhabitants
pop_{dis}	share of population living in the district city
dens	specific population density
income	share of average income of economically active person in the district to an average income in the Czech Republic
accom	share of accommodation capacity (number of beds) to population
phys	count of physicians per 100,000 inhabitants
house	average market price of a dwelling
$road_{1st}$	number of kilometres of 1st class roads per 10,000 inhabitants
$road_{2nd}$	number of kilometres of 2nd class roads per 10,000 inhabitants
BU_{corp}	number of corporations based in the district per 1,000 inhabitants
BU_{ind}	number of individual business units based in the district per 1,000 inhabitants
BU_A	number of business units operating in agriculture, forestry, and fishery per 1,000 inhabitants
BU_C	number of business units operating in the field of manufacturing per 1,000 inhabitants
BU_G	number of business units operating in wholesale and retail trade per 1,000 inhabitants
BU_H	number of business operating in transportation and storage per 1,000 inhabitants
BU_L	number of business units operating in the field of real estate activities per 1,000 inhabitants
BU_R	number of business units operating in the field of arts, entertainment, and recreation per 1.000 inhabitants

Achieved values of adjusted coefficients of determination ($R^2_{adj.}$) and of standard errors of estimates (SEE) are summed up in the Table 1. Their values indicate varying, but still acceptable qualities of the twelve regression models. The aggregation of the twelve fractional regional price-level indexes for each CZ-COICOP Heading (RPI_{COI}) to the overall value of regional price-level index followed a procedure analogical to aggregation of the *RPI* itself. [10]

	RPI	***COI*01**	***COI*02**	***COI*03**	***COI*04**	***COI*05**	***COI*06**	***COI*07**	***COI*08**	***COI*09**	***COI*10**	***COI*11**	***COI*12**
$R^2_{adj.}$	**0.883**	0.468	0.298	0.341	0.743	0.232	0.247	0.598	0.292	0.496	0.766	0.334	0.458
SEE	**0.012**	0.014	0.016	0.038	0.032	0.021	0.034	0.029	0.009	0.023	0.058	0.047	0.037

Table 1 Adjusted Coefficients of Determination ($R^2_{adj.}$) and Standard Errors of Estimates (SEE)

Source: authors' calculations based on (CZSO [5])

4 Results

The results of our calculations are summed up in the Table 2 below. It is apparent that the differences in the regional price levels are to the highest extent influenced by the CZ-COICOP Heading 04 (Housing, Water, Gas, Electricity, and Other Fuels), Heading 10 (Education), and Heading 11 (Restaurants and Hotels) – i.e. immobile commodities.

Code	**District**	***RPI***	***COI*01**	***COI*02**	***COI*03**	***COI*04**	***COI*05**	***COI*06**	***COI*07**	***COI*08**	***COI*09**	***COI*10**	***COI*11**	***COI*12**
CZ0100	Praha	**1.172**	1.012	1.007	1.059	1.424	1.007	1.047	1.158	1.009	1.105	1.480	1.117	1.133
CZ0201	Benešov *	**1.025**	1.003	1.012	1.057	1.063	1.014	0.997	1.039	0.996	1.008	1.035	1.007	0.976
CZ0202	Beroun *	**1.051**	1.016	1.008	1.026	1.116	1.000	0.993	1.103	0.993	1.022	1.074	1.029	1.036
CZ0203	Kladno	**1.046**	1.004	0.995	0.984	1.108	0.988	1.055	1.030	0.986	1.010	1.220	1.044	1.084
CZ0204	Kolín	**1.040**	1.037	1.019	1.062	1.062	0.976	1.029	1.008	1.005	1.066	1.096	1.019	1.044
CZ0205	Kutná Hora *	**1.013**	1.009	1.004	1.010	1.057	0.999	1.023	0.997	1.001	1.008	0.968	0.980	0.948
CZ0206	Mělník *	**1.044**	1.007	1.001	1.016	1.114	1.001	0.981	1.067	1.004	1.014	1.056	1.002	1.066
CZ0207	Mladá Boleslav *	**1.026**	1.022	1.009	1.007	1.091	0.985	0.997	0.971	1.003	1.023	1.081	1.015	0.983
CZ0208	Nymburk	**1.023**	1.022	1.015	1.028	1.096	1.011	0.942	0.984	1.012	0.991	1.054	0.931	0.960
CZ0209	Praha-východ *	**1.110**	1.059	1.049	1.147	1.244	1.000	1.059	1.131	0.989	1.031	1.140	1.052	1.105
CZ020A	Praha-západ *	**1.129**	1.056	1.042	1.137	1.308	1.008	1.072	1.137	0.979	1.048	1.204	1.072	1.116
CZ020B	Příbram	**1.028**	1.010	0.989	1.056	1.037	1.005	1.074	1.029	0.998	1.040	1.027	1.026	1.043
CZ020C	Rakovník *	**1.005**	0.999	0.991	0.975	1.024	0.992	0.978	1.000	0.999	1.025	0.979	0.994	0.984
CZ0311	České Budějovice	**1.027**	1.033	0.987	1.045	1.035	1.026	0.979	1.020	1.001	1.059	1.051	1.067	0.987
CZ0312	Český Krumlov *	**0.972**	0.982	0.980	1.014	0.954	1.037	0.957	0.966	0.988	0.960	0.995	0.940	0.954
CZ0313	Jindřichův Hradec *	**0.968**	0.992	0.989	1.004	0.966	1.012	1.005	0.905	0.999	0.984	0.974	0.940	0.919
CZ0314	Písek *	**0.990**	1.002	0.997	1.016	0.984	1.017	1.042	0.932	1.009	0.999	1.020	1.015	0.969
CZ0315	Prachatice *	**0.969**	0.969	0.986	1.014	0.946	1.027	0.967	0.956	0.993	0.975	0.940	0.973	0.949
CZ0316	Strakonice	**0.978**	1.032	0.979	0.941	0.955	1.022	0.970	0.917	1.011	0.990	0.928	0.982	0.978
CZ0317	Tábor	**1.001**	1.002	0.994	1.071	0.977	0.990	0.976	1.001	0.995	0.992	0.964	1.076	1.030
CZ0321	Domažlice *	**0.980**	0.995	0.993	0.959	0.946	0.991	0.985	0.986	0.990	1.019	0.978	1.023	0.961
CZ0322	Klatovy	**0.960**	0.977	0.989	0.962	0.914	1.018	1.006	0.945	1.005	0.999	0.993	1.029	0.946

CZ0323	Plzeň-město	**1.038**	1.013	1.005	0.994	1.071	0.986	0.975	1.039	1.001	0.999	1.268	1.073	1.080
CZ0324	Plzeň-jih *	**0.993**	1.003	0.999	0.989	0.991	0.998	0.984	0.980	0.988	1.013	0.981	0.987	0.971
CZ0325	Plzeň-sever *	**1.000**	1.001	1.006	1.014	1.017	0.998	1.001	0.990	0.978	0.998	1.011	0.966	0.978
CZ0326	Rokycany *	**1.016**	1.007	1.007	1.001	1.016	0.996	0.993	1.085	0.990	1.011	1.020	1.003	1.004
CZ0327	Tachov *	**0.977**	0.982	0.983	0.946	0.932	0.992	0.942	1.044	1.005	1.036	0.967	0.971	0.960
CZ0411	Cheb	**0.971**	0.999	1.012	0.930	0.907	0.990	0.980	0.997	1.006	1.024	0.980	1.020	0.999
CZ0412	Karlovy Vary	**0.995**	0.993	1.014	1.133	0.947	1.029	0.971	1.052	1.002	1.021	1.124	0.921	1.022
CZ0413	Sokolov *	**0.961**	0.979	0.992	0.975	0.896	0.986	0.967	0.999	0.996	0.973	0.963	0.991	0.982
CZ0421	Děčín	**0.993**	1.007	0.961	0.991	0.915	1.019	1.013	1.048	1.016	1.014	0.918	1.017	1.108
CZ0422	Chomutov *	**0.949**	0.992	0.989	0.971	0.848	0.984	0.999	0.981	1.015	0.988	0.967	1.001	0.987
CZ0423	Litoměřice *	**0.994**	0.998	0.996	1.008	0.975	1.003	0.997	1.012	1.000	1.008	0.991	0.998	0.983
CZ0424	Louny *	**0.972**	1.001	0.995	0.983	0.929	0.990	0.978	0.960	1.009	1.014	0.899	0.997	0.956
CZ0425	Most *	**0.946**	0.981	0.996	0.975	0.836	0.970	1.031	0.986	1.004	0.972	1.049	1.009	1.015
CZ0426	Teplice	**1.000**	1.010	1.000	0.970	0.969	0.975	1.015	1.074	1.001	0.989	0.942	1.011	1.027
CZ0427	Ústí nad Labem	**0.973**	0.976	0.994	0.941	0.938	0.909	0.984	1.038	1.017	0.988	1.032	0.943	1.026
CZ0511	Česká Lípa *	**0.981**	0.994	1.001	0.987	0.953	0.998	0.991	0.979	1.008	0.977	0.960	0.971	0.999
CZ0512	Jablonec n. Nisou *	**0.991**	0.991	1.006	0.973	0.984	1.011	1.038	0.988	1.000	0.962	1.053	1.001	0.995
CZ0513	Liberec	**1.043**	0.994	1.007	1.069	1.076	1.030	1.049	1.042	1.010	1.033	1.089	1.052	1.063
CZ0514	Semily *	**1.001**	0.997	1.008	1.029	1.036	1.017	0.987	0.982	0.997	0.949	1.037	0.960	0.980
CZ0521	Hradec Králové	**1.056**	1.016	1.022	1.003	1.164	0.977	1.051	1.040	1.028	0.993	1.033	1.064	0.982
CZ0522	Jičín *	**1.004**	1.005	1.005	0.985	1.024	0.998	0.983	0.992	1.002	1.007	1.009	0.992	0.968
CZ0523	Náchod	**0.983**	1.001	1.005	0.990	0.985	0.977	1.064	0.977	0.995	0.984	0.948	0.950	0.944
CZ0524	Rychnov n. Kněž. *	**0.996**	1.002	1.005	1.009	1.000	0.997	0.998	0.983	0.998	0.994	0.989	0.985	0.979
CZ0525	Trutnov *	**0.983**	0.994	1.000	1.029	0.993	1.006	1.009	0.957	1.002	0.938	1.026	0.904	0.991
CZ0531	Chrudim	**0.977**	1.017	1.001	0.953	0.957	0.973	0.985	0.966	0.994	1.028	0.988	1.027	0.915
CZ0532	Pardubice	**1.046**	1.016	1.026	1.072	1.055	1.012	1.047	1.049	1.008	1.045	1.154	1.100	1.069
CZ0533	Svitavy *	**0.974**	0.997	0.994	0.973	0.965	0.994	0.965	0.931	0.997	0.993	0.907	0.973	0.948
CZ0534	Ústí nad Orlicí *	**0.983**	1.000	1.006	0.990	0.973	0.986	0.986	0.948	1.001	0.988	0.950	1.000	0.974
CZ0631	Havlíčkův Brod *	**0.973**	0.999	0.987	0.965	0.952	1.011	0.997	0.941	1.002	1.002	0.968	0.989	0.933
CZ0632	Jihlava	**0.986**	0.997	1.007	1.010	0.948	0.999	1.076	1.039	1.009	0.989	0.901	0.953	1.002
CZ0633	Pelhřimov *	**0.978**	0.997	0.993	0.989	0.959	1.011	0.996	0.969	0.997	1.003	0.997	0.968	0.933
CZ0634	Třebíč *	**0.973**	0.996	0.982	0.945	0.944	1.011	0.995	0.971	0.997	1.002	0.902	0.987	0.949
CZ0635	Žďár nad Sázavou	**0.967**	1.000	0.991	1.000	0.943	0.993	0.939	0.969	0.994	0.981	0.978	0.900	0.963
CZ0641	Blansko *	**0.996**	0.997	0.989	0.952	1.024	1.001	0.994	0.977	0.988	0.987	0.908	0.992	0.977
CZ0642	Brno-město	**1.091**	1.021	1.013	0.991	1.221	1.015	1.016	0.991	0.999	1.041	1.171	1.164	1.121
CZ0643	Brno-venkov *	**1.026**	1.010	1.001	0.986	1.074	1.006	1.018	1.037	0.979	1.005	0.970	0.998	1.026
CZ0644	Břeclav *	**0.993**	0.992	0.988	0.964	0.990	1.002	0.975	1.020	0.992	0.994	0.888	0.948	1.027
CZ0645	Hodonín	**0.993**	1.001	1.004	0.985	0.986	1.018	0.990	0.972	1.000	0.986	0.958	0.997	1.010
CZ0646	Vyškov *	**1.007**	0.999	0.993	0.982	1.023	1.011	0.991	1.002	0.998	1.006	0.925	1.006	1.012
CZ0647	Znojmo	**0.982**	1.009	1.000	1.008	0.940	1.030	1.005	0.947	0.991	0.986	0.833	1.000	1.020
CZ0711	Jeseník *	**0.972**	0.976	0.991	1.037	0.918	1.034	0.964	0.982	0.989	0.964	0.957	0.948	1.063
CZ0712	Olomouc	**1.009**	0.986	0.994	1.005	1.017	1.000	0.960	1.042	1.000	1.002	0.959	1.084	0.997
CZ0713	Prostějov *	**0.992**	1.001	0.999	0.968	0.973	0.993	1.018	1.003	0.993	0.999	0.926	1.041	0.978
CZ0714	Přerov	**0.990**	0.992	1.012	0.971	0.996	0.993	0.973	0.961	1.004	0.964	0.992	1.063	0.985
CZ0715	Šumperk	**0.970**	0.971	1.007	1.020	0.962	1.028	1.010	0.946	1.002	0.967	0.805	1.022	0.925
CZ0721	Kroměříž *	**0.994**	1.002	0.991	0.969	0.979	1.006	0.986	1.006	1.005	1.001	0.937	1.014	1.000
CZ0722	Uherské Hradiště	**1.015**	1.002	1.016	0.969	1.037	1.015	1.009	0.980	0.978	1.016	0.928	1.020	1.052
CZ0723	Vsetín	**1.002**	0.991	0.999	1.006	1.038	1.010	1.024	0.989	0.998	0.950	0.937	0.907	1.026
CZ0724	Zlín	**1.038**	1.007	1.002	0.989	1.112	0.983	0.972	1.030	0.998	0.995	1.115	0.989	1.044
CZ0801	Bruntál	**0.938**	0.938	0.991	0.988	0.901	0.992	0.990	0.915	1.015	0.944	0.879	0.945	0.950
CZ0802	Frýdek-Místek *	**0.995**	0.997	1.000	0.989	1.002	0.988	1.001	0.985	0.997	0.977	0.979	0.975	1.016
CZ0803	Karviná	**0.976**	0.990	0.994	1.013	0.959	0.987	1.026	0.962	1.001	0.949	0.926	0.954	0.995
CZ0804	Nový Jičín	**0.979**	0.957	0.984	0.982	0.948	1.001	1.031	0.966	1.024	1.021	1.056	0.997	1.038
CZ0805	Opava	**1.009**	0.984	1.003	0.907	1.061	0.970	0.926	1.045	1.002	0.970	0.962	1.053	0.987
CZ0806	Ostrava-město	**1.007**	0.992	1.005	1.008	1.015	0.978	1.039	1.043	1.020	1.021	1.079	0.956	0.986

Table 2 Regional Price-Level Index (*RPI*) at LAU 1 Level and Its Breakdown to CZ-COICOP Headings

Note: results for districts with asterisks * are based on estimates
Source: authors' calculations based on (CZSO [5])

Regional price-level results also reflect themselves well in the structurally affected and economically weak regions (lower price levels in Teplice, Karviná, Nový Jičín, and in Hodonín, Znojmo, Přerov, Šumperk, Bruntál). Ostrava and Opava remain very close to the mean value. The cartogram in Figure 1 indicates the regional price levels for LAU 1 regions.

5 Conclusion

The purpose of Regional Price-level Index is to enable an assessment of spatial differences in the costs of living of an average household. In terms of spatial comparison, the index needs to include all relevant expenditures which can indicate interregional differences and which are purchased by households. These are mainly goods and services which cannot be provided supra-regionally (common food, local services) and market prices of rentals and real estate. The immobile commodities (housing, education, accommodation, catering) represent the main source of regional price-level differences.

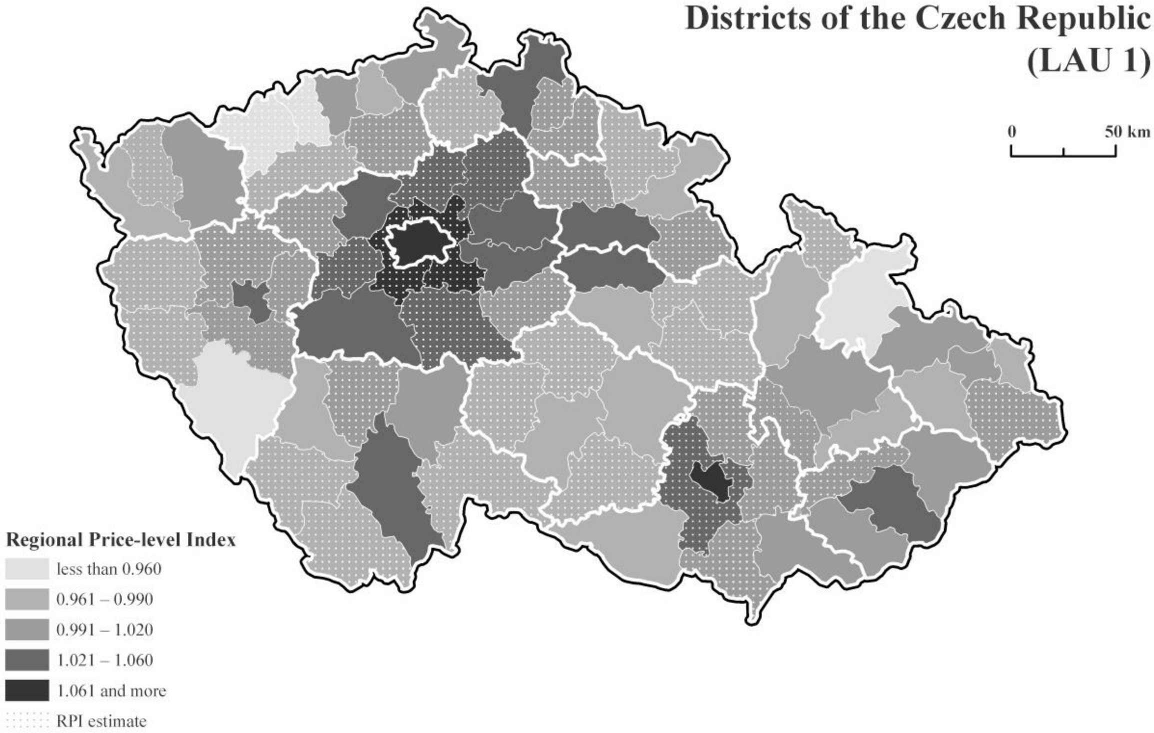


Figure 1 Regional Price-Level Index (*RPI*) at LAU 1 level

Source: authors' own calculations and processing based on (ArcData [1], CZSO [5])

The purpose of the *RPI*, however, is also a source of its shortcoming. It should be used and applied carefully, because it is clear, the average household is not a household of unemployed, or of pensioners. The social status is usually connected with a consumer behavior, differing significantly from the consumer behavior of an average household. Therefore, it should be strictly used together with or applied to average income indicators (average wage in a certain region, average net disposable household income, etc.)

The real income indicator would make the state and development of social and economic disparities on the regional and sub-regional levels more precise. [26] [15] [9]

According to the preliminary results of Kocourek and Šimanová [12] and Kocourek et al. [11], the real regional disparities in the income of households in the Czech Republic are smaller than so far published nominal ones, which is consistent with findings of Čadil et al. [4] Therefore, it seems very useful (if not necessary) to measure or at least estimate the price levels on the most detailed scale available. Significant differences in cost of living can be identified even within the former districts in the Czech Republic (LAU 1), a price level homogeneity on the level of NUTS 3 or NUTS 2 is therefore another very strong and hardly justifiable precondition.

Although on the lower levels of territorial division (LAU 1 and smaller) the income indicators are also very difficult to measure or reliably estimate, even the regional price-level index alone seems to provide a very important information.

Acknowledgements

This article presents results of the applied research project TD020047 “Regional Price Index as the Indicator of the Real Social and Economic Disparities” supported by the Technology Agency of the Czech Republic, the Omega Programme. Special thanks belong to the staff of the Price Statistics Department of the Czech Statistical Office (esp. RNDr. Jiří Mrázek and Ing. Pavla Šedivá) for their willingness to provide data and consultation related to data collection and processing.

References

[1] ArcData: *ArcČR 500: Digital geographic database 1:500 000* [online]. Praha: ArcData, 2014. [cit. 20.4.2015]. <http://www.arcdata.cz/produkty-a-sluzby/geogra ficka-data/arccr-500/>.

[2] Brandt, L. and Holz, C. A.: Spatial Price Differences in China: Estimates and Implications. *Economic Development and Cultural Change* **55** (2006), 43–86.

[3] Čadil, J., and Mazouch, P.: PPS and EU Regional Price Level Problem. *The Open Political Science Journal* **4** (2011), 1–5.

[4] Čadil, J., Mazouch, P., Musil, P. and Kramulová, J.: True Regional Purchasing Power: Evidence from the Czech Republic. *Post-Communist Economies* 26 (2014), 241–256.

[5] CZSO: Data from Internal Database of the Czech Statistical Office. Prague: Czech Statistical Office, 2014.

[6] Gibbons, S., Overman, H. G. and Pelkonen, P.: Wage Disparities in Britain: People or Place? *Spatial Economics Research Centre Discussion Paper* **60** (2010), 1–48.

[7] Gibbons, S., Overman, H. G. and Resende, G.: Real Earnings Disparities in Britain. *Spatial Economics Research Centre Discussion Paper* **65** (2011), 1–44.

[8] Gong, C. H. and Meng, X.: *Regional Price Differences in Urban China 1986 – 2001: Estimation and Implication.* Germany: Institute for the Study of Labor, 2008, IZA Discussion Papers 3621.

[9] Kahoun, J.: Měření regionálního HDP: důchodový a produkční přístup. *Ekonomické listy* **2** (2011), 3–13.

[10] Kocourek, A., Šimanová, J., Rozkovec, J., and Kraft, J.: Metodika stanovení regionálních spotřebních košů a cenovýc parit pro kalkulaci regionálních cenových hladin (regionálního cenového indexu) [online]. Liberec: Technická univerzita v Liberci, 2015. [cit. 15. 4. 2016]. <http://vyzkum.ef.tul.cz/td020047/metodika/rch.pdf>

[11] Kocourek, A., Šimanová, J., and Kraft, J.: Regionalization of the Consumer Price Index in the Czech Republic. In: *The 8th International Days of Statistics and Economics Conference Proceedings*. Prague: University of Economics, 2014, 1487–1496.

[12] Kocourek, A., and Šimanová, J.: Nominal vs. Real Regional Income Disparities in Selected Cities of the Czech Republic. In: *Proceedings of the 12th International Conference Liberec Economic Forum 2015.* Liberec: Technická univerzita v Liberci, 2015, 192–200.

[13] Kosfeld, R., and Eckey, H. F.: Market Access, Regional Price Level and Wage Disparities: The German Case. *Jahrbuch für Regionalwissenschaft* **30** (2010), 105–128.

[14] Kosfeld, R., Eckey, H. F., and Laudrisen, J.: Disparities in Prices and Income across German NUTS 3 Regions. *Applied Economics Quarterly* 54 (2008), 123–141.

[15] Martinčík, D.: Ekonomicko-sociální úroveň krajů – komplexní srovnávací analýza. *E+M Ekonomie a management* **11** (2008), 14–25.

[16] Matzka, C., and Nachbagauer, A.: Reale Kaufkraft 2008: Einkommen unter Berücksichtigung des Regionalen Preisniveaus [online]. *ÖGM Studie.* Vienna, Austria: Österreichische Gesellschaft für Marketing, 2009. [cit. 2015-05-17]. <http://www.ogm.at/inhalt/2012/04/RealeKaufkraft11.pdf>

[17] Musil, P., Kramulová, J., Čadil, J., and Mazouch, P.: Application of Regional Price Levels on Estimation of Regional Macro-Aggregates Per Capita in PPS. *Statistics and Economy Journal* **49** (2012), 3–13.

[18] Pittau, M. G., Zelli, R., and Massari, R.: Do Spatial Price Indices Reshuffle the Italian Income Distribution? *Modern Economy* **2** (2011), 259–265.

[19] Radvanský, M., and Fuchs, L.: *Computing Real Income at NUTS 3 Regions* [online]. Brusel, Belgie: EcoMod, 2009. [cit. 2014-05-13]. <http://ecomod.net/system/files/Computing real income at NUTS 3 regions Radvansky Fuchs.pdf>

[20] Roos, M. W. M.: *Regional Price Levels in Germany* [online]. Dortmund: Universität Mannheim, 2003. [cit. 15.12.2014]. <http://www.vwl.uni-mannheim.de/brownbag/roos.2003.pdf>.

[21] Roos, M. W. M.: Earnings Disparities in Unified Germany: Nominal versus Real. *Jahrbuch für Regionalwissenschaft* **26** (2006), 171–189.

[22] Roos, M. W. M.: Regional Price Levels in Germany. *Applied Economics* **38** (2006), 1553–1566.

[23] Schultze, C. L.: The Consumer Price Index: Conceptual Issues and Practical Suggestions. *Journal of Economic Perspectives* **17** (2003), 3–22.

[24] Ströhl, G.: Zwischenörtlicher Vergleich des Verbraucherpreisniveaus in 50 Städten. *Wirtschaft und Statistik*, 1994, iss. 6, pp. 415–434.

[25] Terem, P., Čajka, P., and Rysová, L. Europe 2020 Strategy: Evaluation, Implementation, and Prognoses for the Slovak Republic. *Economics & Sociology* **8** (2015), 154–171.

[26] Viturka, M.: Konkurenceschopnost regionů a možnosti jejího hodnocení. *Politická ekonomie* **55** (2007), 638–658.

[27] Žižka, T., Kocourek, A., and Šimanová, J.: Strukturovaná výpočtová databáze regionálních cenových indexů pro měření regionálních cenových hladin v členění dle CZ-COICOP [online]. Liberec: Technická univerzita v Liberci, 2015. [cit. 2015-05-17]. <http://vyzkum.ef.tul.cz/td020047/index.php?content=databaze>.

Discrete Dynamic Endogenous Growth Model: Derivation, Calibration and Simulation

Jan Kodera[1], Quang Van Tran[1] , Miloslav Vošvrda[2]

Abstract. Endogenous economic growth model were developed to improve traditional growth models with exogenous technological changes. There are several approaches how to incorporate technological progress into a growth model. Romer was the first author who has introduced it by expanding the variety of intermediate goods. Overall, the growth models are often continuous. In our paper we formulate a discrete version of Romer's model with endogenous technological change based on expanding variety of intermediates, both in the final good sector and in the research-development sector, where the target is to maximize present value of the returns from discovering of intermediate goods which should prevail introducing costs. These discrete version then will be calibrated by a numerical example. Our aim is to find the solution and analyse the development of economic variables with respect to external changes.

Keywords: growth model, endogenous technological progress, Romer's model, discrete optimization problem, impulse response analysis

JEL classification: C51, E32
AMS classification: 62P20, 91B62

1 Introduction

In our paper we study the features of a specific model of economic growth based on the structural approach. In growth models, technological progress is a very important determinant for modelling of economic growth. In Sollow-Swan's model technological progress is treated exogenously and independently of factors of production [1]. However, the role of technological progress as a process resulting from internal causes is more natural and more relevant as we can often observe. Taking this fact into account, endogenous growth model links technological progress with growth models with optimal consumer behaviour first proposed by Ramsey [6]. There are several ways how to introduce endogenous technological progress. One of them is to introduce it as an expanding variety of intermediate products suggested by Romer [7] and developed in [3], [4] and [5]. In this approach the incorporation of technological progress into the model differs from the Schumpeterian quality ladders approach [2]. Both approaches can be used to analyze the behavior of firms, their production, markup dynamics and the implications for endogenous fluctuations and growth. We stick to the traditional Romer's approach, but we derive a discrete time version of an endogenous growth model with expanding variety of intermediates as an alternative to the continuous time approach. The reason for this approach is that continuous time models are resistent to calibrations, simulations and verifications. On the other hand, a discrete time version of an endogenous growth model allows us to apply all DSGE modeling technique to investigate its behavior which is a DSGE endogenous growth model symbiosis. Then the calibration and impulse response analysis will be carried out on the derived model to examine and quantify the effect of exogenous factors on the development of endogenous variables in the model.

2 The discrete model with expanding variety of intermediates

In order to derive the discrete version of the endogenous growth model we choose the traditional procedure similar to the one for the continuous version. There are three sectors in the model. The first one is the

[1]University of Economics in Prague, Department of Statistics and Probability, nám. W. Churchilla 4, Praha 3 - 130 67, {kodera,tran}@vse.cz
[2]Institute of Information Theory and Automation, Department of Econometrics, Pod Vodárenskou věží 4, Praha 8, 182 08, vosvrda@utia.cas.cz

final production sector represented by one aggregated firm. The second sector is a sector of Research and Development firms which consists of a continuum of entities. The sector of households is filled with unique household. All three sectors interact to construe a market equilibrium.

2.1 Modelling of the Final Product Producers

Let us assume that final goods are produced by one aggregate firm. Its production depends on constant amount of labour L and on intermediate products produced by N_t different Research and Development firms

$$Y_t = A_t L^{1-\alpha} \sum_{j=1}^{N_t} X_t^{\alpha}(j). \tag{1}$$

where Y_t is the output of final goods. The variable $X_t(j)$ denotes the amount of j-th intermediate product employed for final goods production. Technological progress takes the form of expansion of N_t which is a number of specialised intermediate products. To better understand the effect of the extension of N for the increase of final production let us assume that for the production of final goods Y_t the same amount of $X_t(j) = X_t$ is used. Using this assumption in expression (1) we get

$$Y_t = A_t L^{1-\alpha} N_t X_t^{\alpha} = A_t L^{1-\alpha} (N_t X_t)^{\alpha} N_t^{1-\alpha} \tag{2}$$

The production function (2) exhibits constant returns to scale with respect to $N_t X_t$ and L. Notice that if X_t increases the marginal product of intermediate goods decreases. If N_t increases the marginal product of intermediate goods is constant. It can be observed that the marginal product of intermediate goods depends on which component of $N_t X_t$ changes. The most intensive factor of growth therefore is the number of intermediate goods.

Expression (1) requires N_t to be an integer, but N_t is a quantity which denotes technological complexity what positive number describes more properly. From that reason we re-formulate expression (1) to

$$Y_t = A_t L^{1-\alpha} \int_0^{N_t} X_t^{\alpha}(j)\mathrm{d}j \tag{3}$$

and we will use it in the following text. Firms producing final goods maximize their profit. The profit for final goods produced by final product firm at time t is given by

$$Y_t - w_t L - \int_0^{N_t} P_t(j) X_t(j)\mathrm{d}j = A_t L^{1-\alpha} \int_0^{N_t} X_t^{\alpha}(j)\mathrm{d}j - W_t L - \int_0^{N_t} P_t(j) X_t(j)\mathrm{d}j \tag{4}$$

The firm producing final goods Y_t maximizes its profit with respect to intermediates and labour. Derived necessary conditions are used for the derivation of demand function for intermediates. Necessary conditions for the maximization with respect to intermediates are the Euler equations for degenerate problem of variation calculus

$$A_t \alpha L^{1-\alpha} X_t^{\alpha-1}(j) - P(j) = 0 \tag{5}$$

The demand for $X_t(j)$ depending on $P_t(j)$ is given by

$$X_t(j) = L \left(\frac{A_t \alpha}{P_t(j)} \right)^{1/(1-\alpha)} \tag{6}$$

The necessary condition for profit maximization with respect to labour is

$$(1-\alpha) A_t L^{-\alpha} \int_0^{N_t} X_t^{\alpha}(j)\mathrm{d}j - W_t = (1-\alpha)\frac{Y_t}{L} - W_t = 0 \tag{7}$$

Hence

$$(1-\alpha)\frac{Y_t}{L} = W_t \tag{8}$$

2.2 Research firms

Research firms transform one unit of final product to one unit of intermediate product of type j. The profit of a research firm at time τ is given by

$$\pi_\tau(j) = (P_\tau(j) - 1)X_\tau(j) \tag{9}$$

Let's choose $t \geq 0$ fixed. Then the present value of profits of Research and Development firms at time t is given by

$$V_t(j) = \sum_{\upsilon=0}^{\infty} \pi_{t+\upsilon}(j) Q_{t,\upsilon} \tag{10}$$

where

$$\begin{array}{rcl} Q_{t,0} & = & 1 \\ Q_{t,1} & = & \frac{1}{(1+r_t)} \\ Q_{t,\upsilon} & = & \frac{1}{(1+r_t)\times\cdots\times(1+r_{t+\upsilon-1})},\ \upsilon = 2, 3, \ldots \end{array} \tag{11}$$

The research and development firm maximizes its present value subject to demand functions (6). To solve present value maximization problem we substitute equation (9) into equation (10) and using expression (6) we get for the profit of j-th firm

$$\pi_\tau(j) = (P_\tau(j) - 1)L \left(\frac{A_t \alpha}{P_\tau(j)} \right)^{1/(1-\alpha)} \tag{12}$$

Substituting the equation displayed above into (10) we get

$$V_t(j) = \sum_{\upsilon=0}^{\infty} (P_{t+\upsilon}(j) - 1)L \left(\frac{A_t \alpha}{P_{t+\upsilon}(j)} \right)^{1/(1-\alpha)} Q_{t,\upsilon} \tag{13}$$

We obtain the necessary condition for the maximum of present value by taking the derivative of elements of infinite sum with respect to $P_{t+\upsilon}(j)$ which we will put equal to zero:

$$\frac{\mathrm{d}\pi_t(j)}{\mathrm{d}P_t(j)} = L \left[\left(\frac{A_t \alpha}{P_t(j)} \right)^{1/1-\alpha} + \frac{1}{1-\alpha} [1 - P_t(j)] \left(\frac{A_t \alpha}{P_t(j)} \right)^{1/1-\alpha} \frac{1}{P_t(j)} \right] = 0 \tag{14}$$

Having solved it, we get $P_t(j) = 1/\alpha$. Each Research and Development firm quotes the same optimum monopolistic price which is constant over time. The production of the j-th firm providing that the price $P_t(j) = 1/\alpha$ is given by

$$X_t(j) = L(A_t \alpha^2)^{1/1-\alpha} \tag{15}$$

which means that intermediate firms produce the same quantity of product. Then aggregate production of intermediates is

$$X_t = \int_0^{N_t} X_t(j)\mathrm{d}j = \int_0^{N_t} L(A_t \alpha^2)^{1/1-\alpha} \mathrm{d}j = L N_t (A_t \alpha^2)^{1/1-\alpha} \tag{16}$$

Substituting (16) into (3) $X_t(j)$ we have:

$$Y_t = A_t L^{1-\alpha} N_t L^\alpha (A_t \alpha^2)^{\alpha/(1-\alpha)} = A_t^{\frac{1}{1-\alpha}} \alpha^{2\alpha/(1-\alpha)} L N_t \tag{17}$$

Finally we will express present value of Research and Development firm replacing $1/\alpha$ for $P_t(j)$ into (13).

$$V_t = \frac{1-\alpha}{\alpha} L(A_t \alpha^2)^{1/1-\alpha} \sum_{\upsilon=0}^{\infty} Q_{t,\upsilon} \tag{18}$$

It is clear that that present value doesn't depend on j, so we omit it.

Let us assume that there is free entry into sector of Research and Development firms. Under this condition $V_t = \eta$, where η denotes constant costs of starting business in the Research and Development sector. If $V_t < \eta$ no one starts activity in Research and Development. If $V_t > \eta$ anybody can enter the sector and the price of intermediary and profits declines. Decreasing profits will give decline to present values. The process stops as soon as the equilibrium between present value and starting cost is restored.

In the equilibrium the present value of the Research and Development firm is given by (18). To express the present value of the firm at time $t+1$, we write

$$V_{t+1} = \frac{1-\alpha}{\alpha} L(A_t \alpha^2)^{1/1-\alpha} \sum_{\upsilon=0}^{\infty} Q_{t+1,\upsilon} \tag{19}$$

Notice that $V_t = V_{t+1} = \eta$ in equilibrium, so we have

$$V_t - \frac{1}{1+r_t} V_{t+1} = \frac{r_t}{1+r_t}\eta = \frac{1-\alpha}{\alpha} L(A\alpha^2)^{1/1-\alpha} \frac{1}{1+r_t} \tag{20}$$

After a small rearrangement we get

$$r_t = \frac{1-\alpha}{\eta\alpha} L(A_t\alpha^2)^{\frac{1}{1-\alpha}} \tag{21}$$

Using equation (17), then equation (21) can be written as

$$r_t = \alpha(1-\alpha)\frac{1}{\eta}\frac{Y_t}{N_t} \tag{22}$$

2.3 Households

Households maximize utility in infinite time horizon and receive wage rate W_t for constant amount of labour L supplied to economy. Households' utility functional is given by

$$U = \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\theta} - 1}{1-\theta} \tag{23}$$

where C_t denotes consumption of the household, β subjective discount factor and $1-\theta$ is elasticity of the utility function. Budget equations of households generally describe the dynamics of assets owned by household. The households are the owners of Research and Development firms whose assets are evaluated as ηN_t.

$$\eta(N_{t+1} - N_t) = W_t L + r_t \eta N_t - C_t \tag{24}$$

Rearranging it, we have

$$N_{t+1} = \frac{1}{\eta}(W_t L - C_t) + (1+r_t)N_t \tag{25}$$

To obtain necessary conditions for maximum of utility functional (23) we use Lagrange functional

$$\mathfrak{L} = \sum_{t=0}^{\infty} \beta^t \left\{ \frac{C_t^{1-\theta} - 1}{1-\theta} + \mu_t[\eta(N_{t+1} - N_t) - W_t L + \eta r_t N_t - C_t] \right\}$$

Taking a derivative of Lagrange functional with respect to C_t and N_{t+1} and put them equal zero. After excluding μ_t we get so called Euler equation (we leave out the expectation operator)

$$C_{t+1} = [(1+r_t)\beta]^{(1/\theta)} C_t \tag{26}$$

which is the first necessary condition of the consumer maximization problem. The second necessary condition is budget equation (25).

2.4 Market Equilibrium of the Model

Market equilibrium is given by the following equations. The first equation is the necessary condition for final product firm. It was derived in section 2 as equation (8), for better readability we display it again.

$$W_t = (1-\alpha)\frac{Y_t}{L} \tag{27}$$

The second equation is derived in the Section 3 as equation (22) and it is derived from equilibrium condition for Research and Development firms. Let's display it again.

$$r_t = (1-\alpha)\alpha \frac{1}{\eta} \frac{Y_t}{N_t} \tag{28}$$

The expression for Y_t is derived from production function in the Section 2 in the form

$$Y_t = A_t L^{1-\alpha} X_t^{\alpha} N^{1-\alpha} = A_t^{1/(1-\alpha)} \alpha^{2/(1-\alpha)} L N_t \tag{29}$$

The model hence consists of equations (25)-(29).

3 Calibration and Simulation

First we log-linearize the system of equations (25)-(29) describing the economy behaviour. As the model like the AK model exhibits no actual steady state, but only a quasi-steady state, in which variables C, Y, W and N growth at the same rate denoted g. Let $\hat{C}, \hat{Y}, \hat{W}, \hat{N}$ be the values of C_t, Y_t, W_t and N_t respectively at the beginning of the steady state. Let define $z_t = \ln Z_t - \ln \hat{Z}$ as the deviation of an endogenous variable from its corresponding steady state value. Then c_t, y_t, w_t and n_t the deviations of C_t, Y_t, W_t and N_t from $\hat{C}, \hat{Y}, \hat{W}, \hat{N}$. Let $\bar{a} = \ln A, \bar{r}$ be the values of a_t and r_t at the steady state and these values are unchanged along the balanced growth trajectory. We keep the labor force constant and normalize it to one. The log-linearized system is as follows

$$\begin{aligned}
&(1+g)^{\theta}\theta(c_{t+1} - c_t) = \bar{r}\beta r_t \\
&(1+g)\hat{N} n_{t+1} = \tfrac{1}{\eta}(L\hat{W} w_t - \hat{C} c_t) + \bar{r} r_t + (1 + \bar{r} r_t)\hat{N} n_t \\
&y_t = \tfrac{1}{1-\alpha} a_t + n_t \\
&r_t = y_t - n_t \\
&x_t = y_t \\
&w_t = y_t \\
&a_t = \rho a_{t-1} + \epsilon_t.
\end{aligned} \tag{30}$$

The system has only one exogenous variable which is the productivity level a_t. We set $\alpha = 0.5$, $\beta = 0.91$, $\theta = 0.5$, $\rho = 0.9$, $\bar{a} = 0$, $g = 0.01$. The size of the productivity shock is set at size of 0.1. The system is solved in Dynare together with impulse response analysis. The results of impulse response analysis are shown in Figure 1. The results of impulse response analysis show that a positive productivity shock leads to a jump of values of production and wage whose size is twice of the size of the shock due to terms $\frac{1}{1-\alpha}$. The size of consumption growth is even higher while r_t and n_t grow slower than y_t. The reason is that as r_t rises, the value of monopoly profit decreases which demotivate the research firms at the beginning. When interest rate goes down, the number of intermediate goods grows again. The economy as the whole then forms a new higher steady state level then it was before the shock. This implication of the impulse response analysis is consistent with common economic wisdom indicating the applicability of our model.

4 Conclusion

In this paper we have derived a simple endogenous growth discrete time model with expanding variety of intermediate goods in the discrete time fashion. The problem of growth models is that steady state does not exist. In our model it is replaced by steady state growth. Then the gap variables in the model are defined as the deviation of model variables from their corresponding steady state growth variables. The model in this form can be easily solved and the follow-up simulation can be performed on it. We have conducted this task with the help of Dynare and the results we have obtained using our model are quite consistent with economic theory. Our model can be extended in many aspects and make it more complicated to study the effect of various factors on numerous economic variables.

Acknowledgements

The support from the Czech Science Foundation under Grant P402/12/G097 is gratefully acknowledged.

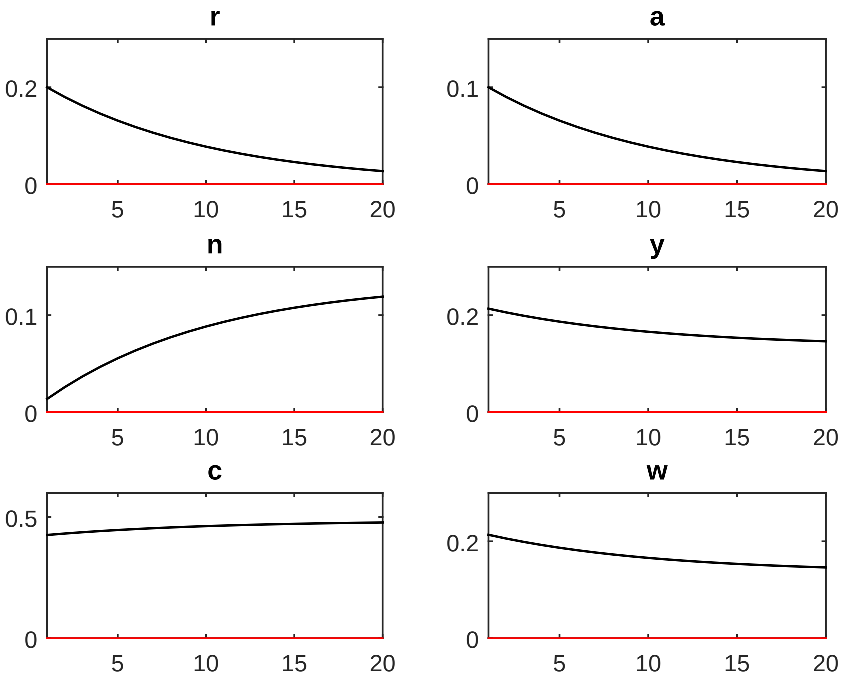


Figure 1 The responses of endogenous variables to productivity shock

References

[1] Barro, R. J., and Sala-i-Martin, X.: *Economic Growth*. MIT Press Cambridge, 2004.

[2] Barrau, N. G.: Schumpeterian foundations of real business cycles. *MPRA Paper* 9430, University Library of Munich, 2008.

[3] Holden, T.: Products, patents and productivity persistence: A DSGE model of endogenous growth. *Dynare Working Papers Series*, WP No. **4**, May 2011.

[4] Jaimovich, N.: Firm dynamics and markup variations: Implications for sunspot equilibria and endogenous economic fluctuations. *Journal of Economic Theory* **137**, 1 (2007), 300-325.

[5] Jones, C. I.: R&D-Based Models of Economic Growth. *Journal of Political Economy* **103**, August (1995), 759–784.

[6] Ramsey, F. P.: A mathematical theory of saving. *The Economic Journal* **38**, 152 (1928), 543—559.

[7] Romer, P. M.: Endogenous technological change. *The Journal of Political Economy* **98**, 5 (1990), 71–102.

Robust design of tariff zones in integrated transport systems

Michal Kohání[1]

Abstract. Designing of tariff system in integrated transport system (ITS) is an important issue. There are several approaches how to design a tariff system, for example a division the region into smaller areas - tariff zones can be used. One of the fundamental input data in the zone tariff planning is origin-destination (OD) matrix as the demand matrix which can have a major influence on the system design, but it can usually contain some uncertainties. To deal with this problem and to test robustness of the design several approaches can be used. In this paper we present two integer programming based approaches, one of them is based on the theory of fuzzy sets and second one on the model reformulation with parameters modeling the impact of fare changes on the demand for traveling. We formulate mathematical model of the tariff zones design problem and present both approaches. Designed approaches are verified on the real data set from the Žilina municipality, where we have an OD matrix calculated using passengers' smart card transactions for a period of one week (approx. 110 000 transaction records). To solve the problem we use a universal optimization tool Xpress.

Keywords: Tariff planning, tariff zones design, IP solver, fuzzy approach, demand, fares.

JEL Classification: C44, C61, D81, R42
AMS Classification: 90C08, 90C10, 90C035, 90C70

1 Introduction

Sustainable transportation and transport integration are connected with a large numbers of challenges, especially in public transportation within the cities and suburban regions. One of the way how motivate citizens to use public transportation is effective integrated transport system (ITS), which integrates several transportation means (buses, trams, trains, etc.). Designing of effective public transport systems includes solving of various optimization problems such as line network planning, coordination of connections in transport nodes, optimization of connection supply, minimization of time losses related to the changing of travel connection etc. Designing of tariff system in ITS is also an important issue, tariff system should be understandable and motivating for passengers, but, on the other hand, effective for transport operators and municipalities. [2] [8]

When transport authorities plan the regional public transportation, one of the problems they deal with is the problem of the tariff and the ticket prices. As was mentioned in [2] and [8], one of possible tariff system is the counting zone tariff system. In this system the region is divided into smaller sub-regions - tariff zones and the price for travelling depends on the origin and the destination zone and on the number of travelled zones on the trip. For all trips hold that the price for trips passing the same number of zones must be equal. With zone system is strongly connected the problem of tariff zones design and new fares. In various publications, e.g. [2] [7] [8], are proposed approaches for the zone design. One of frequently used approach is solving of the counting zones tariff system where the goal is to design the zones so that the new and the old price for most of the trips are as close as possible.

The results of solved problem can be affected by various factors. Resulting tariff zone partitioning design should be able to resist possible imperfections and the resulting solution should be robust. To deal with the robustness of the design several approaches can be used. In this paper we present two integer programming based approaches. One of them is based on the theory of fuzzy sets where possible uncertain values in OD matrix can be modeled as triangle fuzzy numbers and the objective function can be transformed to the fuzzy inequality with the level of satisfaction and than it can be solved using Tanaka-Asai's approach [9] to maximize the level of satisfaction and find appropriate solution. Second approach is based the model reformulation with parameters modeling possible change in demand for traveling caused by the fare changes in the zone tariff system.

This paper will be organized as follows. Firstly, we present the model of the tariff zones design problem for counting zones tariff system and the way of linearization of the model. Secondly, we describe both approaches to

[1] University of Žilina, Department of Mathematical Methods and Operations Research, Univerzitná 8215/1, 01026 Žilina, Slovakia, michal.kohani@fri.uniza.sk.

study the robust design of the proposed model. Thirdly, proposed approaches are verified on the real data set from the Žilina municipality, where we have the OD matrix calculated using passengers' smart card transactions for a period of one week (approx. 110 000 transaction records). To solve the problem we use a universal optimization tool Xpress.

2 Mathematical model of the tariff zones design problem

Construction of the zone tariff problem was inspired by the model of the *p*-median problem. All stops in the public transport network constitute the set of nodes I. If there is direct connection by public transport line between two stops i and j from the set I, these stops are connected by the edge $(i,j) \in E$,. Symbol E denotes the set of edges. For each pair of stops i and j is value c_{ij} the current price of travelling between these two stops. The number of passengers between stops i and j is b_{ij} (OD matrix). We expect to create at most p tariff zones.

To describe decisions in the model, we introduce binary variables y_i, z_{ij} and w_{rs}. Variables y_i represent a decision about "fictional" centre of the zone. Variable y_i is equal to 1 if there is a centre of the zone in node i and 0 otherwise. For each pair of stations i and j binary variable z_{ij} is equal to 1 if the station j is assigned to the zone with centre in the node i and 0 otherwise. To be able to calculate new price of the trip between nodes i and j, we need to calculate the number of zones crossed on this trip. Accordingly to the technique proposed in [2] and [3], this calculation can be easily replaced by the calculation of crossed zone borders. We assume that the stop can be assigned only to one zone and then the border between zones is on the edge. We will introduce the binary variable w_{rs} for each existing edge $(r, s) \in E$, which is equal to 1 if stops r and s are in different zones and is equal to 0 otherwise. We need to determine the path used for travelling between stops i and j for the calculation of number of crossed borders. To be able to do so, we can calculate parameter a_{ij}^{rs}, where the used paths will be observed. a_{ij}^{rs} is equal to 1 if the edge (r,s) is used for travelling from stop i to stop j and 0 otherwise.

Another issue related to this problem is a new price setting for travelling in zone tariff system. Based on previous research we use approach with two different unit prices that was described in [3] – parameter f_1 is a price for travelling in the first zone and parameter f_2 is a unit price for travelling in each additional zone. New price n_{ij}, determined by the number of crossed zones will be calculated according to the definition (1) as follows:

$$n_{ij} = f_1 + \sum_{(r,s)\in E} f_2 a_{ij}^{rs} w_{rs} \tag{1}$$

When determining the optimization criteria we can find various approaches to objective function in the literature. Minimisation of the maximal deviation between the current price (or fair fare [7][8]) and new price determined by the number of crossed minimisation of the average deviation between current and new price for all passengers were described in [2] and another approach based on maximising the revenue of the transportation companies was proposed in [8]. According to the advices of experts in [8] and previous research in [3], we use the average deviation approach. Objective function of the model can be written as (2):

$$\text{Minimize } dev_{avg} = \frac{\sum_{i\in I}\sum_{j\in I} \left|c_{ij} - n_{ij}\right| b_{ij}}{\sum_{i\in I}\sum_{j\in I} b_{ij}} \tag{2}$$

Nevertheless, designed objective function (2) is not a linear function. To be able to solve the problem using an IP solver, we have to reformulate the objective function. Reformulation of the objective function was described in [3]. We introduce new non-negative real variables u_{ij} and v_{ij}. Variables u_{ij} represent the calculated price deviations for travelling in case that current price c_{ij} is higher than new price n_{ij} and variables v_{ij} represent the calculated price deviations for travelling in opposite case. Reformulated objective function can be written in the form (3) and to describe relation between original and new variables binding constrains (9) have to be added into the model. The mathematical model of the tariff zones partitioning problem can be written in the following form:

$$\text{Minimize } dev_{avg} = \frac{\sum_{i\in I}\sum_{j\in I} u_{ij} b_{ij} + \sum_{i\in I}\sum_{j\in I} v_{ij} b_{ij}}{\sum_{i\in I}\sum_{j\in I} b_{ij}} \tag{3}$$

$$\text{subject to } \sum_{i\in I} z_{ij} = 1, \text{ for } j \in I \tag{4}$$

$$z_{ij} \le y_i, \, for \, i, j \in I \tag{5}$$

$$n_{ij} = f_1 + \sum_{(r,s) \in E} f_2 a_{ij}^{rs} w_{rs}, \, for \, i, j \in I \tag{6}$$

$$z_{ij} - z_{ik} \le w_{jk}, \, for \, i \in I, (j,k) \in E \tag{7}$$

$$\sum_{i \in I} y_i \le p \tag{8}$$

$$c_{ij} - n_{ij} = u_{ij} - v_{ij}, \, for \, i, j \in I \tag{9}$$

$$z_{ij} \in \{0,1\}, \, for \, i, j \in I \tag{10}$$

$$y_i \in \{0,1\}, \, for \, i \in I \tag{11}$$

$$w_{ij} \in \{0,1\}, \, for \, (i, j) \in E \tag{12}$$

$$n_{ij} \ge 0, \, for \, i, j \in I \tag{13}$$

$$u_{ij} \ge 0, \, for \, i, j \in I \tag{14}$$

$$v_{ij} \ge 0, \, for \, i, j \in I \tag{15}$$

Conditions (4) ensure that each stop will be assigned to only one zone exactly. Conditions (5) ensure that the stop j will be assigned only to existing centre of the zone. Conditions (7) are coupling constraints between variables for allocation of the stop to zone and variables for determining the zone border on the edge (j,k). Condition (8) ensures that we will create at most p tariff zones.

3 Robust design of tariff zones

The results of solved problem can be affected by various factors. Resulting tariff zone partitioning design should be able to resist possible imperfections and the resulting solution should be robust. First presented approach to deal with the robustness is based on the theory of fuzzy sets and second approach is based the model reformulation with parameters modeling possible change in demand for traveling.

3.1 Fuzzy sets approach

In this first approach is based on the technique described in [9] and [10] and detailed description of this approach was mentioned in [5]. Uncertain demand can be modeled as a triangle fuzzy number $\underline{b_{ij}} = <b_{ij}^1, b_{ij}^2, b_{ij}^3>$ with its associated normalized membership function $\mu(b_{ij}) \in <0,1>$. Minimization process of optimized objective function can be replaced by the intention that the objective function value of the resulting solution should belong to the fuzzy set of sufficiently small objective function values at as high level of satisfaction h as possible. Result of the objective function transformation to the form of fuzzy inequality is described by (17). Mathematical model of finding the highest level of satisfaction h can be described as follows:

$$Maximize \;\; h \tag{16}$$

$$subject \, to \; (1-h)\left(\frac{\sum_{i \in I}\sum_{j \in I} u_{ij} b_{ij}^1 + \sum_{i \in I}\sum_{j \in I} v_{ij} b_{ij}^1}{\sum_{i \in I}\sum_{j \in I} b_{ij}^1} \right) + h\left(\frac{\sum_{i \in I}\sum_{j \in I} u_{ij} b_{ij}^2 + \sum_{i \in I}\sum_{j \in I} v_{ij} b_{ij}^2}{\sum_{i \in I}\sum_{j \in I} b_{ij}^2} \right) \le hF_1 + (1-h)F_2 \tag{17}$$

$$subject \, to \; (4)-(15), h \ge 0$$

This model can be solved using Tanaka-Asai's approach [9] which maximizes the level of satisfaction and finds appropriate solution of the counting zone tariff system. We gradually increase the value of h and for given value of h we solve linear problem (16), (4) – (15), (17) with original decision variables and original objective function (3). The process terminates, when no feasible solution exists. Optimal solution obtained for the last (highest) value of h is the resulting solution of the problem.

3.2 Model reformulation approach with impact parameters

Prices for traveling have major influence on demand in public transport. As was mentioned earlier, the goal is to design the zones in that the new and the old price for most of the trips are as close as possible. The increase in prices can cause reducing the number of transported passengers. On the contrary, a price reduction may affect the attractiveness of transport and hence the slight increase in the demand and the number of transported passengers.

In the second approach we reformulate original objective function (3) of the problem. This approach was discussed in [4].We introduce nonnegative coefficients d and e. Coefficient d represents the increase rate in number of passengers in the case of lower new prices, coefficient e represents the decrease rate in number of passengers in the case of higher new prices. Then we can reformulate mathematical model to the form:

$$Minimize\ dev_{\max} = \frac{\sum_{i\in I}\sum_{j\in J}(1+d).u_{ij}b_{ij} + \sum_{i\in I}\sum_{j\in J}(1-e).v_{ij}b_{ij}}{\sum_{i\in I}\sum_{j\in J}b_{ij}} \tag{18}$$

$$subject\ to\ (4)-(15)$$

4 Computational study

To test proposed approaches to study robustness of obtained solutions and zone partitioning, we use the test network of public transport in Žilina Municipality with approx. 85 thousand inhabitants that consists of 120 stops. Schematic map of the network is shown in the Figure 1. Circles represents bus stops and the diameter of circles represent approximate number of passengers using the stop for starting or ending their journeys.

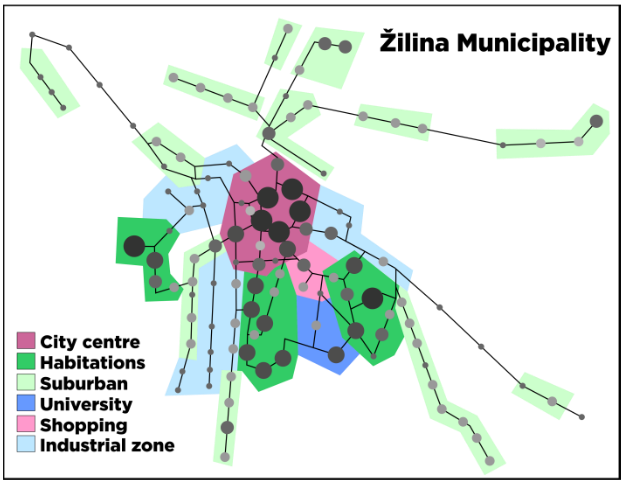


Figure 1: Zilina Municipality

4.1 OD matrix calculation

In cases where passengers in transportation use smart cards, we can obtain more accurate data about the passengers' journeys even in cases where these data are incomplete. In our case we are dealing with the municipal public transportation, where we usually know just information about the boarding stop and the destination stop must be estimated. Since all passengers with smart cards are obliged to validate their card when boarding the vehicle, we can get following information from smart card transactions:

- serial number of passenger's smart card,
- name and ID number of boarding stop, date and time of boarding,
- bus line and connection (trip) number and route variant.

To obtain OD matrix based on these data, we use a trip-chaining algorithm similar to the algorithms mentioned in [1] and [6]. The basic idea of the algorithm assumes that the boarding stop of passenger's immediately following journey is the destination stop of current journey. To be able to follow the idea of chaining passengers' trips, we need to sort all transaction data by unique serial card number, day and time.

Trip-chaining algorithm involves also assumptions that some transaction records can represent not a return trip, but can be a record of an interchange between lines. To be able to handle this, algorithm includes conditions for evaluating transfers (time limits between consecutive smart card records, distance between possible transfer stops in the case of walking between different stops). Estimation of destination stop can also handle the possibility of different boarding/destination stops on the return journey.

To calculate OD matrix we use data set provided by Žilina public transport operator DPMŽ which consist from 111293 records of 11300 smart cards in the time period between October 6 and October 12, 2014. We calculate OD matrix b_{ij} and track passengers' journey to be able to calculate values in the matrix a_{ij}^{rs}. We evaluate various setting for transfers in trip-chaining algorithm to get OD matrix with the highest percentage of the successfully processed records from the data set. We use values 30 and 60 minutes as the value of parameter maximum transfer time and values 300, 600 and 900 meters as the value of maximum distance between stops. We calculated OD matrix with approximately 80 % of successfully calculated destination stops.

4.2 Numerical experiments

With calculated OD matrix we are able to test proposed approaches to study robustness of obtained solutions and zone partitioning. As the value of the parameter c_{ij} we use current prices of travelling which depend partly on the distance. For all journeys up to 5 stops without transfer passenger pays 0.55 €, for all journeys for more than 5 stops without transfer 0.65 € and for a journey with transfer 0.80 €.

In the computational study we use as parameter f_1 value 0.55 € and as parameter f_2 value 0.1 € which were obtained in previous research [6] as best values. Values of parameter $p = \{3, 4, 5, 6, 7\}$.

In the first approach based on the fuzzy sets, we use various triangle fuzzy numbers $<b_{ij}^1, b_{ij}^2, b_{ij}^3>$ to describe OD matrix. First setting is described by triangle fuzzy number $<0.8*b_{ij}, b_{ij}, 1.2*b_{ij}>$, where $0.8*b_{ij}$ means 80% from original values in OD matrix etc. Second setting is described by triangle fuzzy number $<0.9*b_{ij}, b_{ij}, 1.2*b_{ij}>$ and setting 3 is described by triangle fuzzy number $<0.8*b_{ij}, b_{ij}, 1.1*b_{ij}>$. In the second approach based on model reformulation with impact parameters were $d = \{0, 0.01, 0.02, 0.03\}$ what represent increase in demand up to 3 % and $e = \{0, 0.02, 0.04, 0.06\}$ what represent decrease in demand up to 6 %.

Results of numerical experiments are shown in the Table 1. Column denoted as *p* represent values of parameter *p* used in the set of calculations. In both approaches we evaluate the best obtained value of the solution calculates according to the (3) for all possible parameter setting (column *ObjBest*), number of differently assigned stops (assignment to different zones with different parameter setting of the solution method) is in the column denoted as *Diff* and number of trips affected by different assignment of stops is in the column denoted as *Trips*. All numerical experiments were performed in Xpress [11] on a computer equipped with Intel Core 2 Duo E6850 with parameters 3 GHz and 3.5 GB RAM.

	Fuzzy aproach			**Model reformulation approach**		
p	***ObjBest***	***Diff.***	***Trips***	***ObjBest***	***Diff.***	***Trips***
3	5261.05	6	856	5260.30	7	726
4	5235.20	4	956	5223.85	4	956
5	5270.05	6	1230	5267.15	6	1230
6	5295.00	6	1057	5299.05	5	997
7	5315.65	5	997	5312.75	5	997

Table 1 Results of numerical experiments

5 Conclusion

In this paper we presented two integer programming based approaches to test robustness of the zone tariff design problem. One of approaches is based on the theory of fuzzy sets and second one on the model reformulation with parameters modeling the impact of fare changes on the demand for traveling. Designed approaches were verified on the real data set from the Žilina municipality, where we have an OD matrix calculated using passengers' smart card transactions for a period of one week (approx. 110 000 transaction records). To perform all numerical experiments we used IP-solver Xpress.

By comparing of obtained results with various settings of parameter p we can conclude that the changes in the demand do not have big influence on the zone partitioning and differences between both used methods are relatively small. In the worst case only 7 out of 120 stops are assigned to the different zones and the number of affected trips (customers) in comparing with total number of trips is relatively small. Differently assigned stops are mainly stops with small demand. This allows us to conclude that the proposed solution methods are enough robust to be used in the case of changing the tariff system.

Newly created zones are mostly in suburban parts of the network with smaller demand for travelling and higher distances to the centre, as it is common in most transportation system. The major drawback of this process was the time complexity of the problem. Computation time varies from few second up to 2 hours for one particular parameter setting.

Acknowledgements

This work was supported by the research grant VEGA 1/0463/16 "*Economically efficient charging infrastructure deployment for electric vehicles in smart cities and communities*" and VEGA 1/0518/15 "*Resilient rescue systems with uncertain accessibility of service*". We would like to thank to "*Centre of excellence for systems and services of intelligent transport*" (ITMS 26220120050) for built-up infrastructure, which was used.

References

[1] Bagchi, M., and White, P. R.: The Potential of Public Transport Smart Card Data. *Transport Policy* **12,** 5 (2005), 464–474.

[2] Hamacher, H. W., and Schöbel. A.: Design of Zone Tariff Systems in Public Transportation. *Operations Research* **52** (2004), 897–908.

[3] Kohání, M.: Zone partitioning problem with given prices and number of zones in counting zones tariff system. In: *SOR ´13: proceedings of the 12th International Symposium on Operational Research: Dolenjske Toplice, Slovenia* (2013), 75–80.

[4] Kohání, M.: Impact of changes in fares and demand to the tariff zones design. In: *Mathematical methods in economics: proceedings of the 31st International Conference*, College of Polytechnics Jihlava (2013), 417–422.

[5] Kohání, M.: Modeling and handling of uncertain demand in counting zones tariff system. In: *Mathematical methods in economics: proceedings of the 33rd International Conference*, University of West Bohemia, Cheb (2015), 371–376.

[6] Kohání, M.: Tariff zones design in integrated transport systems: a case study for the Žilina municipality. In: *Mechanics, energy, environment: proceedings of the 7th international conference on Urban rehabilitation and sustainability (URES '15),* Rome, Italy, WSEAS Press (2015), 91–97.

[7] Palúch, S.: On a fair fare rating on a bus line. In: *Communications: scientific letters of the University of Žilina* **15** (2013), 25–28.

[8] Schöbel, A.: *Optimization in Public Transportation: Stop Location. Delay Management and Tariff Zone Design in a Public Transportation Network*. Springer, 2006.

[9] Tanaka, H., and Asai, K.: Fuzzy linear programming problems with fuzzy numbers. *Fuzzy Sets and Systems* **13** (1984), 1–10.

[10] Teodorović, D., and Vukadinović. K.: *Traffic Control and Transport Planning: A Fuzzy Sets and Neural Networks Approach*. Kluver Academic Publishers, Boston, Dordrecht, London, 1998.

[11] *XPRESS-Mosel User guide*. Fair Isaac Corporation, Birmingham, 2012.

China and EU as trade competitors: different methodological approaches

Veronika Končiková[1], Jakub Buček[2], Robert Barca[3]

Abstract. Gravity models have been widely used in the international economy. Numerous papers have used gravity models as their main methodology to estimate the impacts and determinants of international flows of goods, capital or labor. However, the specifications of gravity models vary a lot. Numerous researches have focused on how to correctly specify the gravity models in order to get the best results. In our paper we study the impact of Chinese exports on the exports of EU15 countries in the period 2010-2014. However, the literature about gravity models is wide and different authors recommend different estimation techniques in their papers. The aim of this paper is to offer different specifications of trade gravity models and compare the results for our research topic. We find out those different specifications indeed deliver different results; however, the differences are not substantial. We can conclude that the use of different methodological approaches does not differ the answer for our research question. In both cases the models show that we can observe negative, but marginal, impact of Chinese high-tech exports on the exports of selected European countries.

Keywords: gravity model, trade, China, exports, EU, high-tech

JEL Classification: C40, C49
AMS Classification: 91B60

1 Introduction

Gravity models are widely used in the international trade literature. They are used to evaluate flows in different fields of international economics including trade, finances as well as migration. In the trade economics they address various research questions such as the impact of exchange rates, historical and linguistic heritage or increased competition from emerging countries on the trade flows. However, we observe not only the variety of topics which are answered with the help of the gravity models, but we also observe an important variety in the estimation techniques. This paper tries to compare a use of two different estimation techniques in order to assess whether the results are strongly affected by the choice of the estimation technique.

In our model we try to answer a research question: “Are Chinese exports crowding out high-tech exports of the EU-15 into the OECD countries in the period 2010-2014?” However, the main goal of this paper is to compare two different approaches to estimate gravity models and try to evaluate if the estimation techniques have substantial impact on the results. In our paper we first present the motivation and rational behind our research question by addressing the rapid increase of Chinese exports in the last decade and we discuss the shift in the Chinese export structure by analyzing China as the main high-tech product exporter. In the following chapter we briefly present the gravity models and their use in trade economics. This chapter is followed by the model specification where our methodology is revealed and the variables and data used are presented in detail. The next chapter evaluates the results of both our techniques and we compare the outcomes. At the end the conclusion is offered.

2 China and trade

China is mostly known as a country which exports mainly garments and textiles. However, with a short glance at the data showing the structure of Chinese exports, we might find out that these are no longer the main export articles. In this chapter we first talk about the impressive rise of Chinese exports in both intensive and extensive

[1]Masaryk University, Faculty of Economics and Administration, Department of Economics, Lipová 41a, 602 00 Brno, 207077@mail.muni.cz.
[2]Masaryk University, Faculty of Economics and Administration, Department of Economics, Lipová 41a, 602 00 Brno, jakubbucek@mail.muni.cz.
[3]Masaryk University, Faculty of Economics and Administration, Department of Economics, Lipová 41a, 602 00 Brno, 393670@mail.muni.cz.

margin. In the second part we discuss why the high-tech exports are particularly interesting for economic researchers and how they support the motivation of our research question.

2.1 The rise of Chinese exports

Ever since the economic reforms introduced by Deng Xiaoping, China has undergone an important economic transformation which included its opening to the world trade. While in the early sixties Chinese external trade was virtually nonexistent or insignificant given the size of Chinese economy, in the year 2006, when the share of exports on Chinese GDP reached its peak level, the share of exports increased to 35,65% of GDP as shown in [16]. In only few decades China has become the world´s biggest exporter of goods. According to [17] the share of Chinese exports of the goods and services on the world exports in approaching 10%. This rapid rise in Chinese exports has enhanced economic discussion on what impact the Chinese development might have on the developing countries. Probably the most important work is [4] which is using gravity models to estimate the impact of the Chinese rise on other Asian economies.

2.2 The export of high-tech goods

Another outstanding development in Chinese exports is the evolution of its structure. While most people still see China mainly as the exporter of cheap garments, the data demonstrate that China has moved towards the production of more sophisticated products. Rodrik in [14] argues that the sophistication of Chinese production is even higher than its income per capita would suggest. Rodrik's paper stirred a debate about sophistication of Chinese exports. Most opponents[4] suggest that given the global value chains observed in contemporary trade, Chinese exports are not as sophisticated as the trade data suggest. They argue that the export of sophisticated products from China is accompanied by import of highly sophisticated intermediate goods from developed countries and therefore Chinese role is only in assembling operations which have very low value added.

However, this knowledge is not going to affect our research. The data confirm that the amount of high-tech goods from China has increased massively. According to [12] electronics creates 35% of total Chinese exports. Moreover, the discussion on sophistication of Chinese products does not refute that we observe the rise in high-tech exports from China. It only highlights that while China is an important exporter of high-tech products, it does not necessary mean that its value-added is high as well. Our paper uses the data of final goods exports to assess the competition of final goods on the third markets. Therefore it does not include the analysis concerning the global value chains. But the lessons learnt from the debate on the sophistication of Chinese exports need to be taken into consideration when evaluating the results of our estimations.

The importance of structural change in Chinese exports lies in the shift from literature evaluating the impact of Chinese rise on developing countries to the literature which assesses the impact of China on the developed countries. [6] is one of the first papers to study impact of Chinese exports on developed countries. Our research follows this literature which focuses on the impact Chinese rise in exports has on the developed economies, more precisely on their exports.

3 Gravity models in trade economics

Gravity models have long history of usage in trade economics. They were first used in the studies of Tinbergen [15] and Pöyhönen [13]. But ever since this time gravity models became much more sophisticated, both – in its theoretical foundations as well as in its estimation techniques. At first, gravity models were accused of being atheoretical. However, already in [3] Bergstrand proposed a microeconomic foundation for gravity models. The theoretical reasoning behind the gravity models was further developed over the time. Among the most important milestones are papers [7], [5], [2] and [8]. All these studies try to improve the theoretical basis as well as the specification of the gravity models.

4 Model specification

Our paper derives from the previous literature on gravity models. The model is specified as follows:

$$\ln(\mathrm{EX}_{i,j,t}) = \beta_0 + \beta_1 ln(\mathrm{ChEx}_{j,t}) + \beta' \mathrm{X}_{i.j.t} + \vartheta_{i,j,t}$$

$\mathrm{EX}_{i,j,t}$ are exports of all European countries "i" into all destination countries "j" in the time "t". For the purpose of our research the most important variable is $\mathrm{ChEx}_{j,t}$ which in our model represents Chinese export competition. $\mathrm{ChEx}_{j,t}$ therefore represents Chinese exports to country "j" in the time "t". ϑ represents the error term.

[4]See for example [1], [11] and [18]

X is a matrix of gravity type variables. Given the previous researches on gravity models we add the following variables into our model:

- GDP of the exporting European country “i” in the year “t”;
- GDP of the importing OECD country “j” in the year “t”;
- Population of the exporting European country “i” in the year “t”;
- Population of the importing OECD country “j” in the year “t”;
- Distance between the exporting EU country “i” and importing OECD country “j” (measured as distances between their capital cities);
- Dummy for the borders between the trade partners;
- Dummy for the language similarity between the trade partners.

There are lots of approaches towards gravity models. The following methods are mentioned by [9] as estimation techniques used in the existing gravity model literature:

- Truncated OLS;
- OLS ($1+T_{ij}$);
- Tobit (censored regression)
- Panel fixed effects
- Panel random effects
- Heckman two-step
- Poisson Pseudo Maximum Likelihood
- Nonlinear Least Squares
- Feasible Generalized Least Squares
- Gamma Pseudo Maximum Likelihood

Given that the study [10] proves that different estimation techniques might lead to different results, we decided to compare two different estimation techniques in our model.

First estimation will be based on panel data model. Given that the pooled OLS model would give biased estimations and the test has shown that the random effects are more suitable then the fixed effects, we opted to estimate our model with the help of panel data random effects. The second estimation is based on the Poisson model which is one of the alternative techniques used for gravity models estimations. In the next chapter we will estimate our model using the programs gretl and R.

5 Results

The results for using the panel random effects are available in the Table 1:

Variable	estimate	std. error	t-value	pr(>|t|)	
(intercept)	-3.6448e+01	2.4864e+00	-14.6593	< 2.2e-16	***
ChExj,t (China exports)	-1.7149e-01	4.2757e-02	-4.0108	6.240e-05	***
GDP it (EU)	1.1598e+00	9.9102e-02	11.7028	< 2.2e-16	***
GDP jt (OECD)	1.3209e+00	1.1227e-01	11.7655	< 2.2e-16	***
Population it (EU)	-2.3264e-01	9.0485e-02	-2.5710	0.0102005	*
Population jt (OECD)	-2.3991e-01	9.0502e-02	-2.6508	0.0080827	**
Language dummy	9.5903e-01	2.4780e-01	3.8702	0.0001117	***
Distance	-1.1545e-04	1.4124e-05	-8.1738	4.814e-16	***
Border dummy	1.0308e+00	2.2674e-01	4.5459	5.745e-06	***

Table 1 Panel random effects

The results for using Poisson model are available in the Table 2:

Variable	estimate	std. error	t-value	pr(>|t|)	
(intercept)	-4.774e+01	1.466e+00	-32.567	< 2e-16	***
ChExj,t (China exports)	-8.946e-02	5.006e-03	-17.871	< 2e-16	***
GDP it (EU)	1.712e-01	9.411e-03	18.190	< 2e-16	***
GDP jt (OECD)	2.928e+00	3.788e-02	77.306	< 2e-16	***

Population it (EU)	-7.699e-01	5.398e-02	-14.263	< 2e-16	***
Population jt (OECD)	-7.805e-01	5.402e-02	-14.448	< 2e-16	***
Language dummy	-1.909e-01	4.004e-01	-0.477	0.6336	
Distance	-1.087e-04	1.715e-05	-6.339	2.32e-10	***
Border dummy	8.369e-01	3.523e-01	2.375	0.0175	*
Sigma	2.340e-01	1.240e-02	18.872	< 2e-16	***

Table 2 Poisson estimation

When we try to compare the results we can see that indeed we observe differences between the two estimation techniques used in our paper. However, the differences are not substantial. When it comes to the most basic gravity models variable, including GDP of exporting as well as GDP of importing country, they are exactly as the theory behind the gravity models predicts. In both cases the GDP of exporting as well as importing countries have positive impact on the mutual trade between the EU and OECD countries in the high-tech exports. The only difference we find is that the Poisson model gives more importance to the impact GDP of importing countries has on mutual trade comparing to the GDP of exporting countries.

Another typical variable for gravity type models is distance between the exporting and importing country. This variable is a proxy to estimate transaction and transportation costs between countries. In the gravity models we assume that higher the distance between the two countries, bigger are the transactions and transportation costs. We can see that the distance variable gives the predicted results and the impact on the trade is negative. Nevertheless, we can see that the impact is marginal. This can be caused by the fact that our dataset is limited to a specific set of countries which are close to each other hence the transportation costs might be minimal. In addition to this, all the countries concerned are among developed countries and therefore the trade infrastructure might be on high-level and transaction costs therefore might be lower.

The variable which is crucial for our research is the variable for Chinese exports to the OECD countries. We can see that this variable has negative impact on EU-15 exports to OECD countries according to both our estimations. Therefore we can assume that the Chinese high-tech products between the years 2010-2014 has become to tread out the European exports from OECD countries. However, it is important to notice that the effect is very low.

The most significant difference is in the results for language dummy variable. The dummy is 1 if at least 9% of the population shares the same language and it is 0 if they do not share a common language. The common language should smooth the trade and therefore in the gravity models we expect positive impact on the bilateral trade. We can see that in the case of panel data estimation we can truly find a positive impact of the language dummy. However, in the Poisson model we find a negative impact of language dummy on the trade between EU-15 and OECD countries. But we can see that in Poisson model the language dummy is not significant. This problem might have emerged from the fact that we use a limited dataset and therefore the information contained in our data is insufficient for proving the language similarity importance for the trade flows. In order to prove that our model is consistent we did check our model for robustness by excluding the insignificant variable "language dummy". The table 3 proves that by removing this regressor our model leads to the same conclusions.

The rest of the variables behave as the theory behind gravity models suggest. The border dummy shows that when countries have mutual borders they are more likely to trade together. And the negative impact of population is logical given that in our model we included variables of GDP and not the variable GDP per capita.

When it comes to the accuracy measures we can provide coefficient of determination for the random effects model and McFadden´s pseudo-coefficient of determination for the Poisson estimation. For the random effects model the coefficient of determination equals to 0.646488. For the Poisson model, the McFadden's R^2 equals 0.64432869. However, we cannot compare these two accuracy measures. We need to keep in mind that these two models use different principles and therefore the dependant variables in the two models differ. While in the random effect models we use the logarithm of the dependant variable $EX_{i,j,t}$, in the Poisson model, the variable we explain is $EX_{i,j,t}/1000000$.

Variable	estimate	std. error	t-value	pr(>|t|)	
(intercept)	-4.768e+01	1.459e+00	-32.683	< 2e-16	***
ChExj,t (China exports)	-8.941e-02	5.005e-03	-17.865	< 2e-16	***
GDP it (EU)	1.713e-01	9.408e-03	18.204	< 2e-16	***
GDP jt (OECD)	2.928e+00	3.788e-02	77.294	< 2e-16	***

Population it (EU)	-7.716e-01	5.382e-02	-14.337	< 2e-16	***
Population jt (OECD)	-7.822e-01	5.386e-02	-14.523	< 2e-16	***
Distance	-1.110e-04	1.637e-05	-6.783	1.18e-11	***
Border dummy	7.796e-01	3.279e-01	2.377	0.0174	*
Sigma	2.340e-01	1.240e-02	18.870	< 2e-16	***

Table 3 Poisson estimation without language variable

6 Conclusion

The main aim of this paper was to compare different estimation techniques of gravity models. For this purpose we chose panel random effects and Poisson model. We found out that the results show slight differences, however no significant divergence is observed. And the use of different estimation method does not impact the main conclusions of the model.

The research question for this model was to estimate whether Chinese growth in the high-tech exports impacts exports of EU15 countries on the OECD markets in the period 2010-2014. Both of out models suggest that there is a competition between Chinese and European high-tech exports on OECD markets. However, at least for the years investigated in our paper this impact seems to be marginal.

Acknowledgements

This work was supported by the specific research at Masaryk University MUNI/A/0915/2015.

References

[1] Amiti, M., and Freund, C.: *An Anatomy of China's Export Growth*. In Fenstra, W. – Wei, S. China's Growing Role in the World Trade. The University of Chicago Press, Chicago, 2010.

[2] Anderson, J. E., and Van Wincoop, E.: Gravity with Gravitas: A solution to the Border Puzzle. *American Economic Review* **93** (2003), 1268-1290.

[3] Bergstrand, J. H.: The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. *The Review of Economics and Statistics* **67** (1985), 474-481.

[4] Eechengreen, B., Rhee, Y., and Tong, H.: China and the Exports of Other Asian Countries. *Review of World Economics* **143** (2007), 201-226.

[5] Feenstra, R. C.: Border Effect and the Gravity Equation: Consistent Methods for Estimation. *Scottish Journal of Political Economy* **49** (2002), 491-506.

[6] Giovannetti, G., Sanfilippo, M., and Velucchi, M.: *The "China effect" on EU Exports to OECD markets - A focus on Italy*. Working Paper, 2011.

[7] Helpman, E.: Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries. *Journal of the Japanese and International Economics* **1** (1987), 62-91.

[8] Helpman, E., Melitz, M., and Rubenstein, Y.: Estimating Trade Flows: Trading Partners aadn Trading Volumes. *Quarterly Journal of Economics* **123** (2008), 441-487.

[9] Gómez, E. H.: Comparing alternative methods to estimate gravity models of bilateral trade. *Empirical Economics* **44** (2013), 1084-1111.

[10] Gómez, E. H., and Milgram, J. B.: *Are estimation techniques neutral to estimate gravity equations? An application to the impact of EMU on third countries´ exports*. Working Paper, 2009.

[11] Nataraj, G., and Tandon, A.: China's changing Export Structure: A factor based analysis. *Economic and Political Weekly* **66** (2011), 130-136.

[12] *OEC*. 2015. Accessed 10.12.2015 on the web page: <http://atlas.media.mit.edu/en/>

[13] Pöyhönen, P. A.: A Tentative Model for the Volume of Trade between Countries. *Weltwirtschaftliches-Archiv* **71** (1963), 179-182.

[14] Rodrik, D.: What's so special about China's exports? *China &World Economy* **14** (2006), 1-19.

[15] Tinbergen, J.: *Shaping the World Economy: Suggestions for an International Economic Policy*. The Twentieth Century Fund, New York, 1962.

[16] *World databank*. 2015a. Accessed 10.12.2015 on the web page: <http://databank.worldbank.org/data/reports.aspx?source=2&series=NE.EXP.GNFS.ZS>

[17] *World databank*. 2015b. Accessed 10.12.2015 on the web page: <http://databank.worldbank.org/data/reports.aspx?source=2&series=BX.GSR.MRCH.CD>

[18] Xing, Y.: *The People's Republic of China's High-Tech Exports: Myth and Reality*. ADBI Working Paper Series, 2012.

Portfolio selection problem with the third-order stochastic dominance constraints

Miloš Kopa[1]

Abstract. The paper deals with portfolio selection problems which maximize mean portfolio return under constraints that the random return outperform a random benchmark. The outperformance can be understood in several different ways. In this paper, we focus on the mean maximization under third-order stochastic dominance constraints. The third-order stochastic dominance constraints are approximated by so called "super-convex" third-order stochastic dominance constraints which compare the semivariance functions in various grid points. First, we compute the optimal solution of the problem when an ultra-fine grid is used, i.e. super-convex third-order stochastic dominance is a very good approximation of the third-order stochastic dominance. Then, we decrease the number of grid (partition) points (and consequently increase the speed of computations) and we compare the optimal solutions and optimal objective values for various numbers of partition points between each other. Finally, we use the second-order stochastic dominance constraints instead of the third-order ones and we again analyze the changes in the optimal solution and the optimal objective value.

Keywords: Portfolio selection problem, the third-order stochastic dominance constraints, computational complexity

JEL classification: D81, G11
AMS classification: 91B16, 91B30

1 Introduction

Stochastic dominance is, in general, a relation between two random variables. This relation allows for comparison of random returns, losses, or other random outcomes of the investments. The most commonly used stochastic dominance relations are so called the first-order, the second-order and the third-order stochastic dominance. We say that a random variable X dominates a random variable Y with respect to the N-th order stochastic dominance, $N = 1, 2, ...$ if

$$\mathbb{E}u(X) \geq \mathbb{E}u(Y) \quad \text{for all} \quad u \in U_N$$

where the set of admissible utility functions for the N-th order stochastic dominance is defined as follows:

$$U_N = \{u(x) \in D^N : (-1)^k u^{(k)} \leq 0, \quad \forall x, \quad k = 1, ..., N\}.$$

In particular, the first-order stochastic dominance (FSD) is generated by the set of all non-decreasing utility functions (all non-satiated decision makers are considered), while the second order stochastic dominance (SSD) focuses only on the risk averters, that is, only concave and non-decreasing functions are admissible. The basics go back to 1960th, see [20], [8], [9], [22] or [23]. In these papers, several sufficient and necessary conditions for the N-th order stochastic dominance, $N = 1, 2, 3$, were derived. For example, a random variable X dominates a random variable Y with respect to the N-th order stochastic dominance, $N = 1, 2$ if and only if

$$F_X^{(N)}(z) \leq F_Y^{(N)}(z) \quad \forall z \in \mathbb{R}$$

where $F_X^{(1)}(z)$ is a cumulative distribution function of X and

$$F_X^{(2)}(z) = \int_{-\infty}^{z} F_X^{(1)}(t)dt.$$

[1]Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, 186 75 Prague 8, Czech Republic, kopa@karlin.mff.cuni.cz

See [13] for more details. In last years, a substantial development of the stochastic dominance applications in finance was observed; for example in the following directions:

- necessary and sufficient conditions for portfolio efficiency with respect to stochastic dominance criteria, see e.g. [18], [7], [12], [17]
- portfolio enhancement using stochastic dominance rules, see e.g. [21], [10], [19]
- DEA equivalent models in the sense that a portfolio is classified efficient with respect to a stochastic dominance relation if and only if it is efficient with respect to a particular Data Envelopment Analysis model, see [1], [2], [3] for more details.
- more robust version of stochastic dominance relations and stochastic dominance efficiency, see e.g. [4], [11], [5] or [6].
- more general stochastic dominance (ordering) rules, see e.g. [15], [14] and references therein, or [16].

In this paper we focus on the portfolio selection problems which maximize expected return under the third order stochastic dominance constraints. We follow [19] because they argue in favor of the third order stochastic dominance (TSD) as follows: "TSD is less restrictive than SSD, because it requires a preference ordering only for 'skewness lovers', or those risk averters who exhibit decreasing risk aversion. This assumption is accepted by financial economists based on compelling theoretical and empirical arguments. The relaxation of the dominance restriction improves the feasible combinations of return and risk. In particular, TSD is well suited for constructing enhanced portfolios with less downside risk and more upside potential than the benchmark. The SSD criterion ignores these solutions if they are sub-optimal for some skewness haters." Since TSD condition requires in general infinitely many constraints, [19] proposed very tight approximation (sufficient condition) called super-convex TSD (SCTSD). The goal of this paper is to analyze the quality of the approximation for various numbers of partition points. Moreover, SSD is another sufficient condition for TSD, therefore we observe what happens if SSD constraints are used instead of TSD constraints. Contrary to [19], we provide the analysis using monthly data (returns) of 25 Fama French representative portfolios. The analysis uses historical data 1995 - 2014 from the Kenneth French library.

The remainder of this paper is structured as follows. Section 2 presents a notation and basic properties of the SSD relations as well as of the TSD relations. It is followed by a summary of portfolio selection models with SSD constraints based on [12] and with TSD constraints following [19] in Section 3. These models assume an empirical distribution of the returns. Section 4 shows the results of the empirical part. The paper is summarized and concluded in Section 5.

2 Stochastic dominance relations

We consider a random vector $\mathbf{r} = (r_1, r_2, ..., r_N)'$ of returns of N assets in T equiprobable scenarios. The returns of the assets for the various scenarios are given by

$$X = \begin{pmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^T \end{pmatrix}$$

where $\mathbf{x}^t = (x_1^t, x_2^t, \ldots, x_N^t)$ is the t-th row of matrix X. We will use $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_N)'$ for a vector of portfolio weights and the portfolio possibilities are given by

$$\Lambda = \{\boldsymbol{\lambda} \in R^N | \mathbf{1}'\boldsymbol{\lambda} = 1, \ \ \lambda_n \geq 0, \ \ n = 1, 2, \ldots, N\}.$$

Let $F^{(1)}_{\mathbf{r}'\boldsymbol{\lambda}}(x)$ denote the cumulative probability distribution function of returns of portfolio $\boldsymbol{\lambda}$. The *twice cumulative probability distribution function* of returns of portfolio $\boldsymbol{\lambda}$ is given by:

$$F^{(2)}_{\mathbf{r}'\boldsymbol{\lambda}}(t) = \int_{-\infty}^{t} F^{(1)}_{\mathbf{r}'\boldsymbol{\lambda}}(x)dx \tag{1}$$

and the *three times cumulative probability distribution function* of returns of portfolio $\boldsymbol{\lambda}$ is similarly defined as:

$$F^{(3)}_{\mathbf{r}'\boldsymbol{\lambda}}(y) \quad = \quad \int_{-\infty}^{y} F^{(2)}_{\mathbf{r}'\boldsymbol{\lambda}}(t)dt. \tag{2}$$

Definition 1. Portfolio $\boldsymbol{\lambda} \in \Lambda$ dominates portfolio $\boldsymbol{\tau} \in \Lambda$ by the second order stochastic dominance ($\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$) if

$$F^{(2)}_{\mathbf{r}'\boldsymbol{\lambda}}(t) \leq F^{(2)}_{\mathbf{r}'\boldsymbol{\tau}}(t) \qquad \forall t \in \mathbb{R}.$$

Portfolio $\boldsymbol{\lambda} \in \Lambda$ dominates portfolio $\boldsymbol{\tau} \in \Lambda$ by the third order stochastic dominance ($\mathbf{r}'\boldsymbol{\lambda} \succ_{TSD} \mathbf{r}'\boldsymbol{\tau}$) if

$$F^{(3)}_{\mathbf{r}'\boldsymbol{\lambda}}(t) \leq F^{(3)}_{\mathbf{r}'\boldsymbol{\tau}}(t) \qquad \forall t \in \mathbb{R}$$

and $\mathbb{E}(\mathbf{r}'\boldsymbol{\lambda}) \geq \mathbb{E}(\mathbf{r}'\boldsymbol{\tau})$.

This relation is sometimes called a weak SSD and the equivalent definition, presented in e.g. [9], [13] or [12] is based on comparison of expected utility of portfolio returns:

$$\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau} \Longleftrightarrow \mathbb{E}u(\mathbf{r}'\boldsymbol{\lambda}) \geq \mathbb{E}u(\mathbf{r}'\boldsymbol{\tau})$$

for all concave utility functions u. Similarly:

$$\mathbf{r}'\boldsymbol{\lambda} \succ_{TSD} \mathbf{r}'\boldsymbol{\tau} \Longleftrightarrow \mathbb{E}u(\mathbf{r}'\boldsymbol{\lambda}) \geq \mathbb{E}u(\mathbf{r}'\boldsymbol{\tau})$$

for all concave utility functions u having a nonnegative third derivative everywhere.

3 Portfolio selection models

In this paper, we are dealing with portfolio selection models with stochastic dominance constraints in the form:

$$\begin{aligned} & \max \mathbb{E}(\mathbf{r}'\boldsymbol{\lambda}) \\ \text{s.t.} \quad & \mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau} \\ & \boldsymbol{\lambda} \in \Lambda \end{aligned} \tag{3}$$

and

$$\begin{aligned} & \max \mathbb{E}(\mathbf{r}'\boldsymbol{\lambda}) \\ \text{s.t.} \quad & \mathbf{r}'\boldsymbol{\lambda} \succ_{TSD} \mathbf{r}'\boldsymbol{\tau} \\ & \boldsymbol{\lambda} \in \Lambda \end{aligned} \tag{4}$$

where $\boldsymbol{\tau}$ is a given benchmark (reference) portfolio. Since we consider a discrete distribution of returns with equiprobable scenarios (atoms) we can (3) formulate as a linear programming problem, following [12]:

$$\max_{d_s, v_{t,s}, \boldsymbol{\lambda}} \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}^t \boldsymbol{\lambda} \tag{5}$$

$$\begin{aligned} \text{s.t.} \quad -T^{-1}\mathbf{x}^t\boldsymbol{\lambda} + \tfrac{1}{s} d_s - v_{t,s} + \tfrac{1}{s}\textstyle\sum_{k=1}^{T} v_{k,s} \quad &\leq \quad -\tfrac{1}{Ts}\textstyle\sum_{k=1}^{s} \mathbf{x}^k\boldsymbol{\tau} \\ & \quad\quad\; t, s = 1, 2, ..., T \\ v_{t,s} \quad &\geq \quad 0, \quad t, s = 1, 2, ..., T \\ d_s \quad &\geq \quad 0, \quad s = 1, 2, ..., T. \\ \boldsymbol{\lambda} \quad &\in \quad \Lambda \end{aligned}$$

Assuming a partition $(z_1, ..., z_K)$ such that $z_1 \leq \mathbf{x}^t\boldsymbol{\tau}$ and $z_K \geq \mathbf{x}^t\boldsymbol{\tau}$ for all $t = 1, 2, ..., T$, [19] proposed a SCTSD model with a tight sufficient condition for $\mathbf{r}'\boldsymbol{\lambda} \succ_{TSD} \mathbf{r}'\boldsymbol{\tau}$:

$$\max_{\boldsymbol{\lambda}} \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}^t\boldsymbol{\lambda} \tag{6}$$

$$\begin{array}{rcl}
\text{s.t.}\quad \dfrac{1+e_k}{2T}\displaystyle\sum_{t=1}^{T}\theta_{k,t}^2 & \leq & F^{(3)}_{\mathbf{r}'\boldsymbol{\tau}}(z_k), \quad k = 1, ..., K \\
\theta_{k,t} & \geq & z_k - \mathbf{x}^t\boldsymbol{\lambda} \quad k = 1, ..., K, t = 1, ..., T \\
\dfrac{1}{T}\displaystyle\sum_{t=1}^{T}\mathbf{x}^t\boldsymbol{\lambda} & \geq & \dfrac{1}{T}\displaystyle\sum_{t=1}^{T}\mathbf{x}^t\boldsymbol{\tau} \\
\boldsymbol{\lambda} & \in & \Lambda \\
\theta_{k,t} & \geq & 0, \quad k = 1, ..., K, t = 1, ..., T
\end{array}$$

where $e_1 = e_2 = 0$ and

$$e_k = \frac{F^{(3)}_{\mathbf{r}'\boldsymbol{\tau}}(z_k)}{F^{(3)}_{\mathbf{r}'\boldsymbol{\tau}}(z_{k-1}) + F^{(2)}_{\mathbf{r}'\boldsymbol{\tau}}(z_{k-1})(z_k - z_{k-1})} - 1, \quad k = 3, ..., K.$$

Since the solution of problem (6) depends on the partition $(z_1, ..., z_K)$ we compare the optimal solutions and optimal objective values of (6) for various partitions among each other, see the next section. The higher number of partition points K is, the higher optimal objective value is reached, but the longer computation time is needed.

4 Empirical study

In this section we apply problems (5) and (6) to monthly returns of 25 Fama-French portfolios (base assets) formed on size and book-to-market, such that the portfolios are the intersections of 5 portfolios formed on size (market equity, ME) and 5 portfolios formed on the ratio of book equity to market equity (BE/ME). Moreover we included returns of one-year US tbill. As the reference (benchmark) portfolio $\boldsymbol{\tau}$, we consider the market US portfolio, proxied by the CRSP index. We took data from period 1995-2014 (240 scenarios of returns).

The goal of these problem is to find a portfolio with the maximal mean return which outperforms the US market portfolio in sense of the second order stochastic dominance (5) or third order stochastic dominance (6). The optimal portfolios dominates the benchmark portfolio by SSD (TSD). Problem (6) uses a sufficient condition for TSD and we will empirically explore how tight the condition is in dependence on the number of partition points K. To simplify the analysis, we consider only equidistant partitions, i.e. $z_k - z_{k-1}$ is constant (not depending on k). Finally we replace TSD constraints by the SSD constrains, because SSD is another sufficient condition for TSD. The following table summarizes the results in terms of optimal solutions, optimal objective values and computer times. From all 25 assets, only 2 of them appears in the optimal solution portfolios as well as the t-bill. The first asset (A1) is a value weighted average of stocks having the smallest size (ME=1 of 5) and almost the highest BE/ME (BE/ME = 4 of 5). The second asset (A2) is a value weighted average of stocks having medium size (ME=3 of 5) and the highest BE/ME (BE/ME = 5 of 5).

5 Conclusions

In this paper we analyzed the accuracy of the optimal solutions and optimal objective values of the mean maximizing portfolio selection problem with the third-order stochastic dominance constrains. These constraints guarantee that the solution portfolio is preferred to the US market portfolio by all investors having increasing, concave utility functions with nonnegative third derivatives. First, we compute the problem for the ultra-fine partition, i.e. with 1000 equidistant partition points. We found that it has

	Optimal portfolios			Optimal	Relative	Computer
	A1	A2	T-bill	objective	error	time(s)
K=1000	0.7304	0.1442	0.1254	1.22746	0%	475
K=800	0.7303	0.1443	0.1254	1.22745	0.001%	420
K=600	0.7303	0.1443	0.1254	1.22744	0.002%	280
K=400	0.7304	0.1441	0.1255	1.22741	0.004%	182
K=200	0.7299	0.1444	0.1256	1.22722	0.020%	102
K=100	0.7294	0.1444	0.1263	1.22648	0.080%	63
K=50	0.7271	0.1440	0.1289	1.22349	0.323%	28
K=25	0.7192	0.1430	0.1378	1.21316	1.165%	14
SSD	0.6527	0.1870	0.1603	1.18892	3.140%	85

Table 1 Numerical results using GAMS software with solver IPOPT

no sense to use more-fine partition because it would have no effect on the optimal solution. Secondly, we decrease the number of partition points to 800, 600, 400, 200, 100, 50, 25. Table 1 shows that even if we use 90% partition reduction (i.e. 100 instead of 1000 partition points) that relative error of the optimal objective value is less than 0.1%. And, moreover, the computer time decreases by 85%. Finally, we compare these results with results of the problem with SSD constraints. We find that if SSD is used as the approximation (sufficient condition) of TSD then the relative error of the optimal objective value is approximately 3% and it takes more computer time than the above TSD case with K=100 partition points.

For future research, this study can be improved in various ways. For example, longer historical data can be used or different frequency data can be considered (mainly with daily returns). In addition, one can consider the stochastic dominance relations in a more robust way as it was done in [11] or [5], [6] for the first and the second-order stochastic dominance, using contamination techniques and the worst-case approach. Alternatively, one can compare the results also with the case when the first-order stochastic dominance is used. Unfortunately, all these improvements would lead to much more computationally demanding optimization problems.

Acknowledgements

The paper was supported by the grant No. 15-02938S of the Czech Science Foundation.

References

[1] Branda, M., and Kopa, M.: DEA-risk Efficiency and Stochastic Dominance Efficiency of Stock Indices. *Czech Journal of Economics and Finance* **62** (2012), 106–124.

[2] Branda, M., and Kopa, M.: On relations between DEA-risk models and stochastic dominance efficiency tests. *Central European Journal of Operations Research* **22** (2014), 13–35.

[3] Branda, M., and Kopa, M.: DEA models equivalent to general N-th order stochastic dominance efficiency tests. *Operations Research Letters* **44** (2016), 285–289.

[4] Dentcheva, D., and Ruszczyński, A.: Robust stochastic dominance and its application to risk-averse optimization. *Mathematical Programming, Series B* **123** (2010), 85–100.

[5] Dupačová, J., and Kopa, M.: Robustness in stochastic programs with risk constraints. *Annals of Operations Research* **200** (2012), 55–74.

[6] Dupačová, J., and Kopa, M.: Robustness of optimal portfolios under risk and stochastic dominance constraints. *European Journal of Operational Research* **234** (2014), 434–441.

[7] Grechuk, B.: A simple SSD-efficiency test. *Optimization Letters* **8** (2014), 2135–2143.

[8] Hadar, J., and Russell, W. R.: Rules for Ordering Uncertain Prospects. *American Economic Review* **59** (1969), 25–34.

[9] Hanoch, G., and Levy, H.: The efficient analysis of choices involving risk. *Review of Economic Studies* **36** (1969), 335–346.

[10] Hodder, E., Jackwerth, J.C., and Kolokolova, O.: Improved Portfolio Choice Using Second-Order Stochastic Dominance. *Review of Finance* **19** (2015), 1623–1647.

[11] Kopa, M.: Measuring of second-order stochastic dominance portfolio efficiency. *Kybernetika* **46** (2010), 488–500.

[12] Kopa, M., and Post, T.: A general test for SSD portfolio efficiency. *OR Spectrum* **37** (2015), 703–734.

[13] Levy, H.: *Stochastic dominance: Investment decision making under uncertainty.* Second edition, Springer, New York, 2006.

[14] Ortobelli, S., Lando, T., Petronio, F., and Tichy, T.: Asymptotic stochastic dominance rules for sums of i.i.d. random variables. *Journal of computational and applied mathematics* **300** (2016), 432–448.

[15] Ortobelli, S., Shalit, H., and Fabozzi, F.: Portfolio selection problems consistent with given preference orderings. *International Journal of Theoretical and Applied Finance* **16** (2013), 38p.

[16] Post, T.: Standard Stochastic Dominance. *European Journal of Operational Research* **248** (2016), 1009–1020.

[17] Post, T., Fang, Y., and Kopa, M.: Linear Tests for DARA Stochastic Dominance. *Management Science* **61** (2015), 1615–1629.

[18] Post, T., and Kopa, M.: General Linear Formulations of Stochastic Dominance Criteria. *European Journal of Operational Research* **230** (2013), 321–332.

[19] Post, T., and Kopa, M.: Portfolio Choice based on Third-degree Stochastic Dominance. Forthcoming in *Management Science.*

[20] Quirk, J. P., and Saposnik, R.: Admissibility and measurable utility functions. *Review of Economic Studies* **29** (1962), 140-146.

[21] Roman, D., Mitra, G., and Zverovich, V.: Enhanced indexation based on second-order stochastic dominance. *European Journal of Operational Research* **228** (2013), 273–281.

[22] Rothschild, M., and Stiglitz, J.E.: Increasing risk: I. A definition. *Journal of Economic Theory* **2** (1970), 225–243.

[23] Whitmore, G.A.: Third-degree Stochastic Dominance. *American Economic Review* **60** (1970), 457–459.

Income Inequality in V4+ Countries at NUTS 2 Level

Eva Kotlánová[1], Kamila Turečková[2]

Abstract. Income inequality could be a social problem because it causes poverty, generally negatively affects society and it could create a space for social and economic discrimination. Nowadays it is not a problem to find data about it in a form of Gini coefficient, which is one of the best known indicators of measuring income inequality, at national level, but relevant data at regional level does not exist. The NUTS classification is a hierarchical system for dividing up the economic territory of the European Union and among others is used for the purpose of socio-economic analyses of the regions. Regions eligible for support from cohesion policy have been defined at NUTS 2 level. The aim of this paper is to measure and compare income inequality in V4+ countries at NUTS 2 level in a period of 2005 – 2013. Because of the regional income distribution data unavailability income inequality will be analysed through the standard deviation and coefficient of variation. The analysis of disposable incomes at NUTS 2 level in selected countries will be made on empirical data from the regional Statistics on disposable income of households published by Eurostat.

Keywords: Income Inequality, V4+ Countries, NUTS 2, Standard Deviation, Coefficient of Variation.

JEL Classification: C13, D32, I32, R13
AMS Classification: 62P20

1 Introduction

Income inequality has become one of an important issue in countries all over the world. This term is closely associated with poverty, affects negatively society and it could create a space for social and economic discrimination. Some studies [5], [6]found the relationship between income inequality and economic growth. Indicators of income inequality are source of information for EU's main investment policy – cohesion (regional) policy. The bulk of Cohesion Policy funding is concentrated on less developed European countries and regions in order to help them to catch up and to reduce the economic, social and territorial disparities that still exist in the EU. It targets all regions and cities in the European Union in order to support job creation, business competitiveness, economic growth, sustainable development, and to improve citizens' quality of life. All these factors could conduce to improve income inequality in the EU regions.

The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the EU for the purpose of collection, development and harmonisation of European regional statistics, for socio-economic analyses of the regions (NUTS 1, NUTS 2, NUTS 3) and for framing of EU regional policies.The current NUTS 2013 classification is valid from 1 January 2015 and lists 98 regions at NUTS 1, 276 regions at NUTS 2 and 1342 regions at NUTS 3 level.Regions eligible for support from cohesion policy have been defined at NUTS 2 level.

One of the best known and used measures of income inequality is Gini coefficient and its graphical representation – Lorenz curve [8]. Eurostat [3] publish databases of Gini coefficient for each member country but relevant data of distribution of income by quintiles (which are fundamental for Gini coefficient's calculation) at regional level does not exist. The aim of this paper is to measure and compare income inequality in V4+ countries at NUTS 2 level in a period of 2005 – 2013 through the method of coefficient of variation.

The text of the article below this Section 1, Introduction, will be organized as follow: Section 2 provides some theoretical introduction to measurement of income inequality, in detail the method of coefficient of variation is characterized here. Section 3 is oriented to methodology and characterization of data that was used. Section 4 contains the empirical analysis of the income inequality in six selected European countries using the method of coefficient of variation. There is also indicated the development of income inequality during analyzed

[1]Silesian University in Opava, School of Business Administration in Karvina, Department of Economics and Public Administration, Univerzitní nám. 1934/3, Karviná 733 40, kotlanova@opf.slu.cz.
[2]Silesian University in Opava, School of Business Administration in Karvina, Department of Economics and Public Administration, Univerzitní nám. 1934/3, Karviná 733 40, tureckova@opf.slu.cz.

period with Gini coefficient. Finally the conclusion, Section 5, highlights some major conclusions of detailed analysis of income inequality in V4+ countries at NUTS 2 level.

2 How to measure income inequality at regional level NUTS 2

Income inequality is a natural part of every human society. It could reflect a measure of poverty and redistribution of income. Among methods how to measure income inequality are mostly included Gini coefficient and its graphical representation – Lorenz curve., Robin Hood Index, indicator S80/S20 Income Quintile Share Ratio or method of non-weighted average deviation (for more see [1], [7], [9], [10]). All these indicators are based on distribution of income by quintiles (deciles), so it is necessary to know income allocation of households. This data is not available at regional level, so we have to find another way how to measure income inequality. Coefficient of variation is one of the possibility that appears to be appropriate for it.

The standard formulation of a coefficient of variation (CV), the ratio of the standard deviation to the mean, applies in the single variable setting. The CV is often presented as the given ratio multiplied by 100.

$$CV = \frac{SSD}{\bar{x}} * 100[\%] \tag{1}$$

Where: SSD is sample standard deviation,

($\bar{x}$) is the arithmetic mean.

The CV for a single variable aims to describe the dispersion of the variable in a way that does not depend on the variable's measurement unit. The higher the CV, the greater the dispersion in the variable. The CV for a variable can easily be calculated using the information from a typical variable summary (and sometimes the CV will be returned by default in the variable summary).

The standard deviation (SD)is a measure that is used to quantify the amount of variation or dispersion of a set of data values.A low standard deviation indicates that the data points tend to be close to the mean (also called the expected value) of the set, while a high standard deviation indicates that the data points are spread out over a wider range of values [2]. In our case we are interested in sample standard deviation:

$$SSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} x_i^2 - N\bar{x}^2} \tag{2}$$

Where: N present number of values that we have available,

x_i presents the i-th indicator,

($\bar{x}$) is the arithmetic mean.

A large standard deviation indicates that the data points can spread far from the mean and a small standard deviation indicates that they are clustered closely around the mean.

The arithmetic mean of a set of values is the quantity commonly called "the" mean or the average:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{3}$$

3 Methodology and data

From a methodological perspective, the work is based on data gained by Eurostat, concretely from the database of Disposable income of private households by NUTS 2 region for the Czech Republic, Slovakia, Poland, Hungary, Austria and Slovenia (in Slovenia data for 2013 is missing). The covered period includes years 2005-2013 because of missing credible data which is not available for a longer period. The disposable income of private households is the balance of primary income (operating surplus/mixed income plus compensation of employees plus property income received minus property income paid) and the redistribution of income in cash. These transactions comprise of social contributions paid, social benefits received in cash, current paid taxes on income and wealth paid, as well as other current transfers. Disposable income does not include social transfers in kind coming from public administrations or non-profit institutions serving households [4].

Calculations of coefficient of variation and value of standard deviation (SSD) are based on calculations using formulas (1), (2) and (3). This type of measurement of income inequality were described in the text above. The ware used was MS Excel. All calculations and graphical analysis are authors own.

4 Results of Empirical Analysis

First of all Figure 1 shows the development of values of Gini coefficient in countries V4+. The value of Gini can range from 0 (the country has an even distribution of income in relation to households) to 1 (all income in a society is concentrated with one household or individual). In some databases the scale from 0 to 100 is used [3].

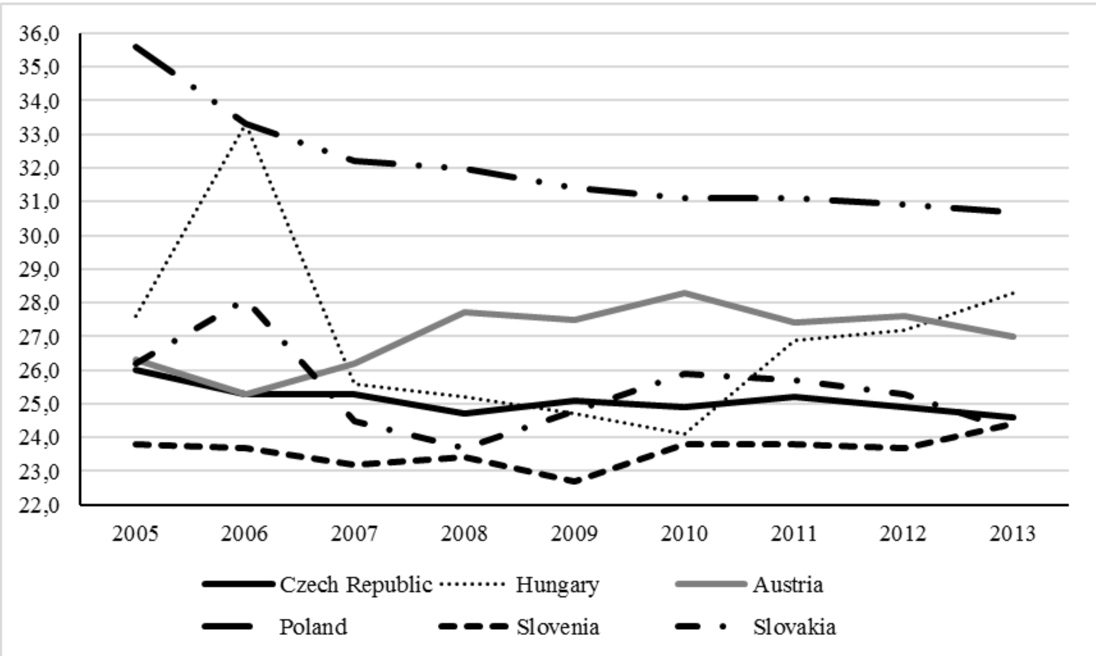


Figure 1 Values of Gini coefficient in V4+ countries (6), 2005 - 2013.

As we can see in Figure 1 the best results are in Slovenia, where the Gini coefficient is the lowest and its development is stable with moderate rising in last two years. The highest income inequality was at the beginning of covered period in Poland and although there is a downtrend, it is still the highest value of Gini in 2013. The largest variety can be seen in Hungary, where the Gini coefficient increased sharply in 2006 and in 2007 declined very fast. From 2010 it has been in progress. In Hungary there are many changes in fiscal policy, so this could be the consequences.

Empirical analysis was made on basis of the net disposal income of households (Euro) by NUTS 2 region in V4+ countries. From the total number of 276 NUTS 2 regions in the EU, there are 8 NUTS 2 regions in the Czech Republic, 7 in Hungary, 16 in Poland, 4 in Slovakia, 9 in Austria and 2 NUTS 2 in Slovenia.

Table 1 displays the calculated values of sample standard deviation and coefficient of variation. The arithmetic mean is the chosen given point and the highest values of sample standard deviation and coefficient of variation the highest the income inequality is.

Country	Czech Republic		Hungary		Poland	
Year/Variable	CV	SSD	CV	SSD	CV	SSD
2005	15.15	810.64	27.69	1,234.04	10.77	418.15
2006	15.76	949.81	22.82	1,007.35	10.24	435.81
2007	15.18	984.80	20.03	972.72	10.06	475.73
2008	14.44	1,108.09	11.32	587.16	9.56	527.69
2009	13.99	1,037.08	8.30	389.14	10.62	513.71
2010	15.60	1,212.73	6.88	335.23	10.34	563.41
2011	14.73	1,181.71	15.09	748.01	10.36	577.58
2012	14.06	1,110.98	11.31	565.69	10.89	623.56
2013	13.85	1,059.65	7.60	390.97	11.38	657.14
Country	Slovakia		Austria		Slovenia	
2005	32.44	1,443.38	2.75	504.98	10.10	848.53
2006	29.39	1,454.88	2.71	519.62	7.19	636.40
2007	28.79	1,763.28	2.65	528.63	7.29	707.11
2008	27.62	1,988.30	2.73	560.01	9.43	989.95
2009	29.82	2,281.08	2.88	589.49	8.24	848.53
2010	26.03	2,036.95	2.58	531.51	7.52	777.82
2011	28.04	2,250.00	2.85	602.08	7.37	777.82
2012	28.77	2,344.50	2.51	552.27	6.27	636.40
2013	28.29	2,376.27	2.78	609.87	NA	NA

Table 1 Values of coefficient of variation and sample standard deviation at NUTS 2 level in V4+ countries (6), 2005-2013.

Table 1 shows different results to estimates of Gini coefficient in selected countries. We can find the lowest values of coefficient of variation in Austria. There are not very large variances of net disposable incomes through the NUTS 2 region. In 2013 there was a highest disposable income in Wien and Vorarlberg (22,800 euro) and the lowest was 21,100 euro in Kärnten region. Results in Slovenia are biased by the fact that this small country is divided just into two NUTS 2 regions, but results in Table 1 correspond to the results of Gini Coefficient.

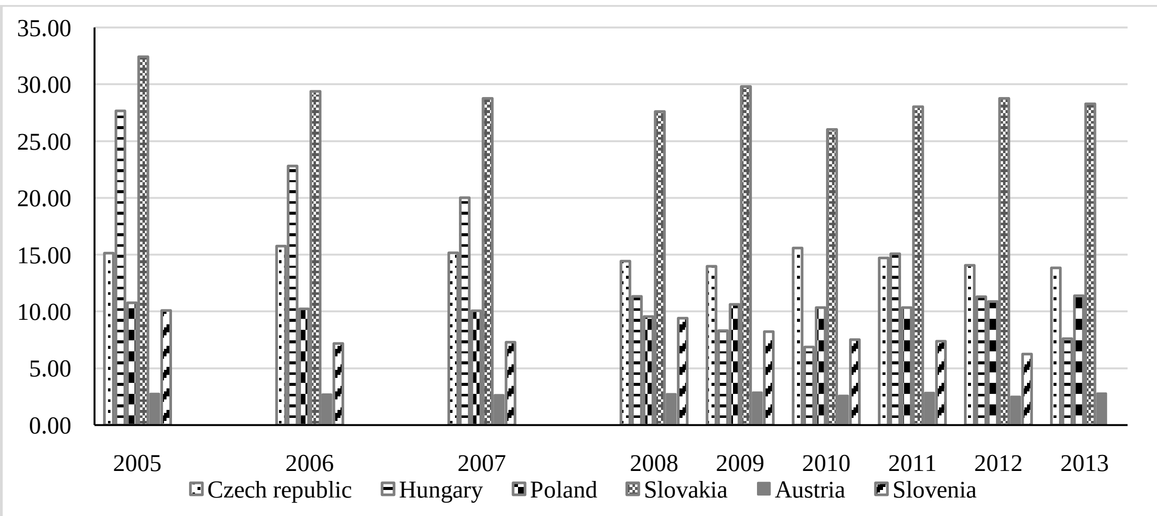


Figure 2 Values of coefficient of variationin countries V4+ countries (6), 2005 – 2013.

Without any doubt the largest country is Poland which is divided into 16 NUTS 2 regions. In such a case, it could be expected, that the variances between the centre and outlying regions should be large. But when we take a look at coefficients of variation, this premise is false. The highest variability in results we can see in Hungary. To a certain extent (delays) it corresponds to results of Gini coefficient. We suppose that this variability relates to Hungarian political atmosphere, unstable government and fiscal policy.

As we can see in Figure 2 the largest variances across the NUTS 2 regions are in Slovakia and the Czech Republic. The reason we can find in a fact that in Slovakia the difference between disposable incomes in capital city and the periphery (East Slovakia) was bigger than 2,000 euro in 2013, when net disposable income in the periphery is twice lower than in the capital city. In the Czech Republic there are two problematic regions (Severozápad and Moravskoslezsko) where the net disposable incomes are the lowest. These areas have problems with unemployment and efflux of labour and young, well-educated people.

5 Conclusion

Despite of the unavailable data for Gini coefficient calculation at regional level NUTS 2, we tried to estimate income inequality in the selected countries of V4+. We chose the method of coefficient of variation of net disposable income which we found as the most suitable. Data used in this paper comes from the Eurostat database of disposable income of private households by NUTS 2 and the covered period was 2005-2013.

Using the method of coefficient of variation the results show that the largest income inequality at NUTS 2 level is in Slovakia and in the Czech Republic. These relatively small economies evince the highest value of coefficient of variation. As we expected the lowest level of coefficient of variation was proved in Austria. Where do we see the causes of this income inequality? Despite the fact that the selected indicator does not include social transfers in kind coming from public administrations, the fiscal policy is very important. For example institute of minimal wage, system of current taxes on income or system of retirement play the role. In the long term the Czech Republic and Slovakia have problems with reforms in area of pension, health service and education system. All these factors could affect disposable incomes directly or indirectly. To improve them, it is essential to have a stable, powerful and effective government.

Acknowledgements

This paper was supported by the project SGS/13/2015 "Influence of Selected Macroeconomic and Microeconomic Determinants on the Competitiveness of Regions and Firms in Countries of the Visegrad Group Plus".

References

[1] Atkinson, A. B.: On the Measurement of Inequality. *Journal of Economic Theory* **2** (1970), 244–263.
[2] Bland, J. M., and Altman, D. G.: Statistics notes: measurement error. *BMJ* **313** (1996), 41– 42.
[3] Eurostat. Statistics: Gini coefficient of equivalised disposable income [online]. [2016-01-29]. Available from: http://ec.europa.eu/eurostat/web/products-datasets/-/ilc_di12.
[4] Eurostat. Statistics: Income of households by NUTS 2 regions [online]. [2016-01-30]. Available from:http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.
[5] Kotlánová, E.: Income Inequality and Economic Growth. In: *Proceedings of the 13th International Conference Economic Policy in the European Union Member Countries* (Machová, Z., and Tichá, M., eds.). VŠB-TU EKF Ostrava, Ostrava, 2015, 288–294.
[6] Kuznets, I.: Economic Growth and Income Inequality. *American Economic Review* **65,** 1 (1955), 1–28.
[7] Schutz, R.: On the Measurement of Income Inequality. *American Economic Review* **41,** 1 (1951), 107–122.
[8] Turečková, K.: Příjmové nerovnosti a jejich měření. *Acta Academica Karviniensia* **1** (2007), 191–198.
[9] Turečková, K.: Income Inequality by Methodof Non-weighted Average Absolute Deviation: Case Study of Central and Eastern European Countries. *Equilibrium* **10,** 4 (2015), 99–110.
[10] Turečková, K., and Kotlánová, E.: Method of Non-weighted Average Absolute Deviation in Context of Measuring Income Inequality by Gini Coefficient: case study ofVisegrad group countries. In: *Conference-Proceedings: 33rd International Conference Mathematical Methods 2015* (Martinčík, D., Ircingová, J., and Janeček, P., eds). University of West Bohemia, Plzeň, 2015, 835–840.

Optimal Portfolio Selection with Different Approximated Returns

Noureddine Kouaissah[1]

Abstract. In this paper, we investigate the impact of the approximation methods on the large-scale portfolio selection problems. In particular, we compare conditional expectation estimation from both parametric and nonparametric regression models. In this context, we use a general nonparametric multivariate regression framework to cope with several problems arising with linear regression approximation. To this end, we firstly reduce the dimensionality of the large-scale portfolio by performing a principal component analysis on the stock returns. Then, within k-fund separation model, we use different methodologies to approximate the portfolio returns. Finally, through an empirical analysis, we show the impact of regression estimation on the ex-post sample paths of several portfolio strategies. The proposed empirical analysis confirms that it is better using nonparametric regression rather than the classical parametric regression.

Keywords: conditional expectation estimators, large-scale portfolio selection, dimensionality reduction, rewards measure.

JEL Classification: G11, C44, G10
AMS Classification: 28A25

1 Introduction

The portfolio selection problem is mainly based on the assumption that investors allocate their wealth across a number of selected assets in order to maximize their expected utility. The first rigorous approximating model to the portfolio selection problem was introduced by Markowitz [8], where the return and risk are modeled in terms of portfolio mean and variance. Generally, the mean-variance approach works well with Gaussian distribution, which is a very restrictive assumption. Indeed, the Gaussian distributional assumption of financial return series is mostly rejected, see for instance Rachev and Mittnik [15] and the references therein. It follows that several alternative approaches to portfolio selection has been proposed, see among others Rachev *et al.* [16], Farinelli *et al.* [4] and the references therein. According to many researchers, see among others Papp *et al*. [13] and Kondor *et al.* [7] the portfolio selection problem is extremely related to the estimation of inputs, statistical parameters, which describe the dependence structure of the returns. The contribution of this paper lies in this context. In particular, we investigate the impact of different approximation methods on large-scale portfolio selection problems.

The first contribution of this paper is to assess the approximation effects, using conditional expectation estimators, on large-scale portfolio selection problems. Using k-fund separation model (see Ross [18]) and principal components analysis, we examine the impact of the regression analysis on portfolio theory. In particular, we use a general nonparametric multivariate regression framework to cope with several problems arising with linear regression approximation. The ordinary least squares (OLS) is commonly used as method to estimate the coefficient of *k*-fund separation model, see among others Ortobelli and Tichý [12]. Generally, OLS procedure works well with Gaussian distribution and when there exist substantial linearity of the data set. Unfortunately, as shown by Nolan *et al*. [10], errors with a heavier tailed distribution can extremely affect the estimated OLS regression coefficients. In addition, we typically observe the nonlinearity in the data set used to estimate the returns. These are the reasons why we decided to use the nonparametric regression analysis as a framework for approximating the returns.

The second contribution of this paper is an ex-post empirical analysis on the use of different conditional expectation estimators in the portfolio theory. In particular, we compare conditional expectation estimation from both parametric and nonparametric regression models. Thus, we show their effects on the ex-post sample paths of several portfolio strategies. For this problem we perform an empirical analysis utilizing all the active stocks of the S&P 500 index.

The rest of the paper is organized as follows. In section 2, we set up the portfolio dimensionality problem and we briefly review some performance ratios. In section 3, we summarize the approximation with parametric and nonparametric regression analysis. In section 4, we provide empirical comparison among different portfolio strategies. Our conclusions are summarized in section 5.

[1] University of Bergamo, Italy & VSB-TU of Ostrava , Czech Republic, noureddine.kouaissah@unibg.it

2 Large-scale portfolio selection problems

Optimal portfolio selection concerns careful decision making about the portfolio composition. Essentially, the problem of choosing a portfolio is a problem of choice under uncertainty. In this context, a fundamental theory of asset choice under uncertainty is the expected utility. There are two different approaches to the problem of portfolio selection under uncertainty stemming from this theory. One of them is the stochastic dominance approach, while the second is reward–risk analysis, according to which, the portfolio choice is made with respect to two criteria – the expected portfolio return and portfolio risk. In particular, a portfolio is preferred to another one if it has higher expected return and lower risk. The relationship between the two approaches is an ongoing research topic; see among others Stoyanov *et al.* [22], and Ortobelli *et al.* [11].

In this section, we focus on the estimation issues within large-scale portfolio selection problems. According to several studies, see among others Rachev *et al.* [14], the portfolio dimensionality problems is enormously related to the estimation of inputs, statistical parameters, which describe the dependence structure of the returns. Thus, in order to have realistic approximation of the portfolio risk-reward measures Papp *et al*. [13] and Kondor *et al.* [7] have shown that the number of observations should increase with the number of assets. Therefore, it is important to find the right trade-off between a statistical approximation of the historical observations depending only on a few parameters and the number of historical series.

Theoretically, there exist different methods to reduce the dimensionality of large-scale portfolio. In this study, we combine two well-known methodologies. Firstly, we preselect the 'best' assets (with respect to performance, optimizing either the Sharpe Ratio or the Rachev Ratio). Secondly, we apply a Principal Component Analysis (PCA) to the correlation matrix of the preselected returns to determine those few principal components that account for most of the portfolio's variability. In practice, we first preselect 100 assets, from S&P 500 components, that have the highest Rachev ratio. Then, we reduce the dimensionality of the portfolio problem as we approximate the returns series through a multifactor model that depends on an appropriate number of parameters, see Ross [18]. Therefore, we approximate the returns by regressing them on those few principal components in order to improve the robustness of the approximations of the performance measures in the portfolio selection model (see, among others, Biglova et al. [1]). Essentially, we replace the original n (n = 100) correlated time series z_i with the n uncorrelated time series R_i assuming that each z_i is a linear combination of R_i. Following this, we call *portfolios factors* f_i the s portfolios R_i with a significant variance, while the remaining $n – s$ portfolios with very small variances are summarized by an error ε. Accordingly, each series z_i is a linear combination of the factors plus a small-uncorrelated noise.

$$z_i = b_{i,0} + \sum_{j=1}^{s} b_{i,j} f_j + \sum_{j=s+1}^{n} b_{i,j} R_j = b_{i,0} + \sum_{j=1}^{s} b_{i,j} f_j + \varepsilon_i, \quad i = 1,...,n \tag{1}$$

where, z_i is the gross return for asset i, $b_{i,0}$ is the fixed intercept for asset i, $b_{i,j}$ is the coefficient for the factor f_j, s is the number of factors, ε_i is the error term for asset i and n is the number of assets.

PCA can be applied either to the Pearson correlation matrix or to any linear correlation measure, for further discussion see Ortobelli and Tichý [12]. Applying PCA to the classical Pearson correlation matrix, we choose the first factors that represent the highest variability of the data set. Hence, we choose the first factors that explain in average more than 75% of the total variance. Therefore, each series z $z_i\left(i = 1,...,100\right)$ can be represented as a linear combination of selected first factors plus a small-uncorrelated noise

$$z_i = b_{i,0} + \sum_{j=1}^{s} b_{i,j} f_j + \varepsilon_i \tag{2}$$

Once we approximate the portfolio return and risk measures, we apply portfolio selection optimization problems to the approximated portfolio returns:

$$x' z \cong x' \hat{b}_0 + \sum_{j=1}^{s} x' \hat{b}_j f_j \tag{3}$$

where, $\hat{b}_j = \left[\hat{b}_{1,j},....,\hat{b}_{n,j}\right]'$ is the vector of estimated coefficients $\hat{b}_{i,j}\left(j = 0,1,...,s\right)$.

To approximate the portfolio returns consistently, and in order to cope with several problems arising with parametric regression, a promising methodology based on nonparametric regression is used and hence serves for the comparison to the classical OLS estimator in the empirical analysis.

The portfolio selection problem is typically examined in a reward–risk framework, according to which, the portfolio choice is made with respect to two criteria – the expected portfolio return and portfolio risk (see, among others, Ortobelli *et al.* [11], Rachev *et al*. [16] and the reference therein). We review two performance measures used in the empirical analysis section: the Sharpe Ratio and the Rachev Ratio.

Sharpe ratio [21]. The Sharpe ratio is used to characterize how well the return of an asset compensates the investor for the risk taken. In particular, it suggests that investors should maximize the Sharpe ratio (SR), given by

$$SR(x'\mathrm{z}) = \frac{E(x'\mathrm{z}) - z_0}{\sigma_{x'\mathrm{z}}}, \tag{4}$$

where, $E(x'z)$ is the portfolio expected returns, z_0 is the risk-free return and $\sigma_{x'z}$ is the portfolio standard deviation.

Rachev ratio. The Rachev ratio, see Biglova *et al.* [1], is the ratio between the average of earnings and the mean of losses. i.e.:

$$RR\left(x'z,\alpha,\beta\right) = \frac{CVaR_\beta\left(z_b - x'z\right)}{CVaR_\alpha\left(x'z - z_b\right)} \tag{5}$$

where Conditional Value-at-Risk (*CVaR*) is a coherent risk measure (see for instance Rockafellar and Uryasev [17]) defined as:

$$CVaR_\alpha\left(X\right) = \frac{1}{\alpha}\int_0^\alpha VaR_q\left(X\right)dq,$$

and

$$VaR_q\left(X\right) = -F_X^{-1}\left(q\right) = -inf\left\{x \setminus P\left(X \le x\right) > q\right\},$$

is the Value-at-Risk (*VaR*) of the random return *X*. If we assume a continuous distribution for the probability law of *X*, then $CVaR_\alpha\left(X\right) = -E\left(X \mid X \le -VaR_\alpha\left(X\right)\right)$, thus *CVaR* can be understood as the average loss beyond *VaR*. For further performance measures that have been proposed in the literature, see among others Farinelli *et al.* [4] and the references therein. Finally, to conclude this section, we refer the readers to some recent studies (see Rachev *et al*. [16]; Stoyanov *et al*. [22]) that classify the computational complexity of reward-risk portfolio selection problems.

3 Approximation with parametric and nonparametric methods

Regression analysis is certainly one of the most suitable and largely used statistical method. In essence, it explores the dependency of the so-called dependent variable on one (or more) explanatory or independent variables.

$$Y = E(Y \mid X = \boldsymbol{x}) + \varepsilon = m(\boldsymbol{x}) + \varepsilon\ . \tag{6}$$

Fundamentally, if we know the form of the function $m(\boldsymbol{x}) = E(Y \mid X = \boldsymbol{x})$, (e.g. polynomial, exponential, etc.), then we can estimate the unknown parameters of $m(\boldsymbol{x})$ with several methods (e.g. least squares). On the other hand, if we do not know the general form of $m(\boldsymbol{x})$, except that it is a continuous and smooth function, then we can approximate it with a non-parametric technique, as proposed by E. A. Nadaraya [9] and G. S. Watson [23]. Using factor model the ordinary least squares (OLS) is widely used as procedure to estimate the coefficient of *k*-fund separation model, see among others Ortobelli and Tichý [12]. Generally, OLS procedure works well with Gaussian distribution and when there exist substantial linearity of the data set. Unfortunately, as shown by Nolan *et al*. [10], errors with a heavier tailed distribution can extremely affect the estimated OLS regression coefficients. Furthermore, we commonly observe the nonlinearity in the data set used to estimate the returns. For this reason, we aim to remedy this deficiency and propose to use the nonparametric regression analysis, which allows data search for an appropriate function that represent well the available data, without assuming any specific form of the function.

3.1 Multivariate kernel methodology

The multivariate nonparametric regression model considered here is as follows

$$y_i = m\left(\boldsymbol{x}_i\right) + \varepsilon_i, \quad \text{for } i = 1,2,...,n \tag{7}$$

where, the y_i 's are observed random variables, $\mathbf{x}_\mathrm{i} \in \mathbb{R}^\mathrm{d}$ and ε_i are assumed to be independent and identically distributed (i.i.d) with mean zero and variance σ_m^2. The most popular method for estimating $m(\cdot)$ is multivariate version of the Nadaraya-Watson kernel estimator, for more details about nonparametric regression, see Härdle [5] or Härdle and Müller [6]. Unfortunately, the Nadaraya-Watson estimator suffers from certain disadvantages. In particular, it corresponds to the local constant fit and presents bias at the boundaries. To overcome these shortcomings, substantial attention focuses on a larger class of kernel estimators given by the locally weighted least squares, see among others Ruppert and Wand [19]. In particular, an estimate of regression function $m(\boldsymbol{x})$ which is $\hat{\alpha}$ given by minimization of the following criterion

$$\text{Minimize } \sum_{i=1}^{n} \{y_i - \alpha - b^T(\boldsymbol{x}_i - \boldsymbol{x})\}^2 K_{\boldsymbol{H}}(\boldsymbol{x}_i - \boldsymbol{x}), \quad (8)$$

where $\boldsymbol{H}$ is $d \times d$ symmetric positive definite matrix that depends on n. For further discussion on the properties of multivariate locally weighted least squares regression, see Ruppert and Wand [19]. Using weighted least squares matrix theory, they derive the leading bias and variance terms for general multivariate kernel weights.

Several scholars revealed that the choice of kernel is not crucial, whereas the performance of the nonparametric regression is more a question of the bandwidth selection. Fan and Gijbels [3] present a survey on bandwidth choice for univariate local polynomial smoothing technique, which includes the Nadaraya-Watson estimator as a special case. However, the literature provides little guidance for those embarking on multivariate setting, which is definitely a significant problem in empirical analysis. Future research attempts to overcome this gap and focuses on optimal bandwidth selection method. The well-known bandwidth selection methods are the rule-of-thumb and the plug-in bandwidth choice. Especially, the former method is the normal reference rule proposed in Bowman and Azzalini [2]. For general multivariate kernel estimators, Scott [20] suggests the following rule:

$$\text{Scott's rule in } \Re^d\text{:} \quad \hat{h}_i = \hat{\sigma}_i \, n^{-1/(d+4)}, \quad (9)$$

where, $\hat{\sigma}_i$ is the estimate of the standard deviation of each $\boldsymbol{x}_i$, and n is the sample size. This technique of bandwidth choice has the property that minimizes the so-called mean integrated squared error (MISE) of the estimate. To summary, in the next section, we use a normal kernel function and as a rule for bandwidth selection the one given by Scott [20].

4 An empirical ex-post analysis

In this section, we examine the ex-post impact of two different estimators considering two portfolio problems: portfolio dimensionality reduction problems and portfolio performance ones. We use all active stocks on S&P 500 index from March 17, 2003 to February 24, 2015 using the previous 3000 daily observations. The data set is taken from Thomson-Reuters DataStream. In particular, starting from September 12, 2003 we preselect the 100 stocks with the highest Rachev performance ratio (5). Then, using the preselected stock, we have to reduce the dimensionality of the portfolio problem. Thus, as suggested by Ortobelli and Tichý [12], we perform a PCA on the return of the selected stocks in order to identify few numbers of factors that represent the highest return variability. We apply PCA on Person correlation matrix and then we regress the series on these factors so that we are able to approximate the returns, as suggested in section 2.

The aim of this section is to investigate the impact of different approximation methods on large-scale portfolio selection problems. Therefore, we first compare the classical OLS and Ruppert and Wand (hereinafter RW) estimator presented in (9) in terms of the Sum of the MSE (Sum-MSE) computed at each recalibration time of the portfolio. Then, we show the impact of two estimators on sample paths of two different portfolio strategies. In particular, on the preselected gross returns, we maximize both Sharpe and Rachev ratios such that the vector of weights x belongs to the simplex:

$$S = \left\{ x \in R^n \setminus \sum_{i=1}^{n} x_i = 1; x_i \geq 0; x_i \leq 0.1 \right\}.$$

This means that short sale is not allowed and we invest not more than 10% in each asset. In this empirical analysis, we ignore the transaction costs, we assume risk free rate zero and α and β in Rachev ratio are set to 5%.

We use a moving average window of 125 working days for the computation of each optimal portfolio and we recalibrate the portfolio every 20 days. First, we compare some average statistics of the Sum-MSE calculated from each methods. The results of this analysis are reported in Table 1.

	Mean(Sum-MSE)	St dev.(Sum-MSE)	Total(Sum-MSE)
RW	0.0148	0.0194	2.1162
OLS	0.0583	0.0769	8.3358

Table 1: Average of some statistics of the Sum-MSE obtained by different estimators

Table 1 reports some statistics of the sum-MSE obtained with two methods, namely the classical OLS and the nonparametric technique based on RW estimator. As expected, we observe that RW estimator is much better than the classical OLS. This means that the nonparametric estimators are performing much better than the parametric estimators are. Indeed, the RW estimator outperforms the classical OLS in terms of the mean, standard deviation and total Sum of MSE. Since the true function of the conditional expectation is not known a priori, a study based on simulation analysis confirms that the RW estimator performs well.

In order to evaluate the impact of the OLS and RW estimators on portfolio theory we go further with the following empirical analysis. Starting with an initial wealth $W_0 = 1$ that we invest on September 12, 2003, we evaluate the ex-post wealth sample paths obtained from two performance ratios (4) and (5). Thus, at the k-th recalibration time, the following steps are performed for Sharpe and Rachev strategies:
Step 1 Preselect the 100 stocks with the highest Rachev performance ratio.
Step 2 Apply a PCA as suggested in section 2 and choose the first factors that guarantee more than 75% of the global variance, then we approximate the gross returns by applying the OLS and Kernel method.
Step 3 Determine the market portfolio $x_M^{(k)}$ that optimizes the portfolio problem applied to the approximated returns.
Step 4 Since we recalibrate the portfolio every 20 days, we calculate the ex-post final wealth as follows:

$$W_{t_{k+1}} = W_{t_k}\left(\left(x_M^{(k)}\right)' z_{(t_{k+1})}^{(expost)}\right), \tag{10}$$

where $z_{(t_{k+1})}^{(expost)}$ is the vector of observed gross returns in the period between t_k and t_{k+1}, such that $t_{k+1} = t_k + 20$.

We apply the four steps until the observations are available. The results of this analysis are reported in Figure 1.

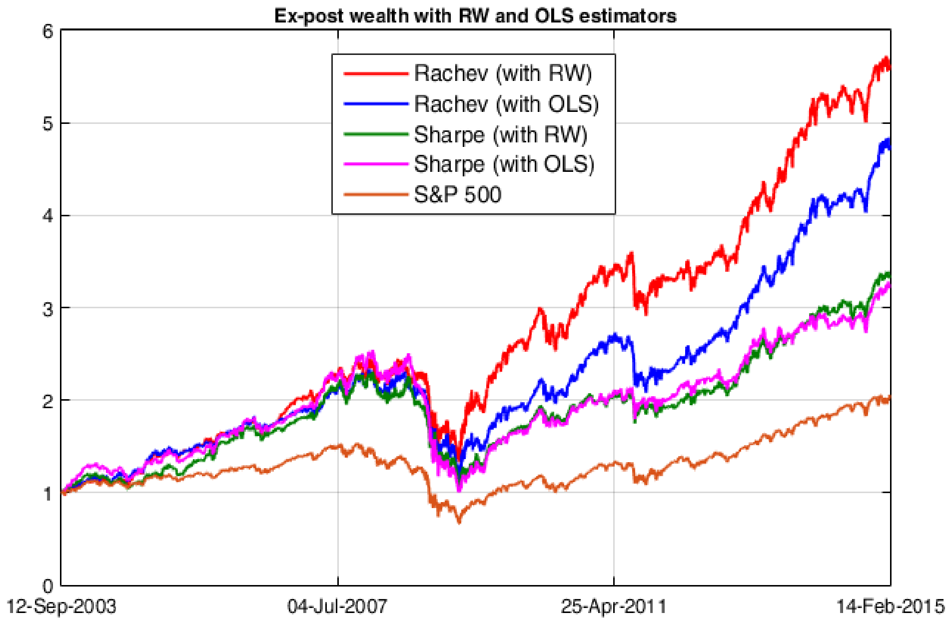


Figure 1: Ex-post wealth obtained with Sharpe and Rachev performance measures when RW and OLS estimators approximate returns, compared with S&P 500 benchmark.

Figure 1 reports the ex-post wealth evolution obtained with two performance ratios, Sharpe and Rachev, when the portfolio returns are approximated by either OLS or RW estimators. On one hand, we observe that Rachev strategy with RW estimator outperforms better than the other strategies. On the other hand, both strategies (Rachev and Sharpe) based on RW estimator are performing better than the same strategies based on OLS estimator. Moreover, all presented strategies are higher than the S&P 500 benchmark. To sum up, the performance strategies that give the best results are those based on the returns approximated with RW estimator.

Overall, from this analysis, we deduce that the RW estimator is very suitable to the large-scale portfolio within k-fund separation model. Therefore, we believe that nonparametric techniques are much more effective, when used as procedure to approximate the portfolio returns, than the classical OLS.

5 Conclusion

In this paper, we consider the impact of two alternative approximation methods on the large-scale portfolio selection problems. In particular, we reduce the dimensionality of the large-scale portfolio by performing a PCA on the stock returns. Then, within k-fund separation model, we use two distinct methodologies to approximate the portfolio returns, namely the classical OLS and RW estimator. Thus, we examine the impact of the two procedures on the U.S. stock market. First, a comparison among two techniques is performed in terms of the Sum-MSE. Secondly, we compare the ex post wealth obtained with two different portfolio performances when the returns are approximated either by the OLS or kernel method. From the comparison among two different strategies, we deduce that the best results are those based on the returns approximated with kernel methodology (RW). Therefore, the nonparametric regression techniques are much more effective and performing, when used to approximate the portfolio returns, than the classical OLS.

Acknowledgements

This paper has been supported by the Italian funds ex MURST 60% 2015 and 2016 and MIUR PRIN MISURA Project, 2013–2016. The research was also supported through the Czech Science Foundation (GACR) under project 16-23699S, by VSB-TU Ostrava under the SGS project SP2016/11, and the European Social Fund in the framework of CZ.1.07/2.3.00/20.0296.

References

[1] Biglova, A., Ortobelli, S., Rachev, S., Stoyanov, S. (2004). Different approaches to risk estimation in portfolio theory. *Journal of Portfolio Management*, **31**(1), 103-112.

[2] Bowman, A.W., Azzalini, A. (1997). Applied Smoothing Techniques for Data Analysis. Oxford University Press, London.

[3] Fan. J., Gijbels, I. (1996). Local polynomial modelling and its applications. Chapman and Hall, London.

[4] Farinelli, S., Ferreira, M., Rossello, D. (2008). Beyond Sharpe ratio: Optimal asset allocation using different performance ratios. *Journal of Banking & Finance*, **32**(10), 2057-2063.

[5] Härdle, W. (1990). Applied Nonparametric Regression. Cambridge University Press, Cambridge.

[6] Härdle, W., Müller, M. (2000). Multivariate and semiparametric kernel regression. In: Schimek, M.G. (Ed.), Smoothing and Regression: Approaches, Computation, and Application. John Wiley & Sons, New York, 357-392.

[7] Kondor, I., Pafka, S., Nagy, G. (2007). Noise sensitivity of portfolio selection under various risk measures. *Journal of Banking and Finance*, **31**, 1545-1573.

[8] Markowitz, H. M. (1952). The utility of wealth. *Journal of Political Economy*, **60**, 151-158.

[9] Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and its Applications*, **9**(1), 141-142.

[10] Nolan, J.P., Ojeda-Revah, D. (2013). Linear and nonlinear regression with stable errors. *Journal of Econometrics*, **172**(2), 186–194.

[11] Ortobelli, S., Petronio, F., Lando, T. (2015). A portfolio return definition coherent with the investors preferences. *IMA Journal of Management Mathematics*, 1-13.

[12] Ortobelli, S., Tichý, T. (2015). On the impact of semidefinite positive correlation measures in portfolio theory. *Annals of Operations Research*, **235**(1), 625-652.

[13] Papp, G., Pafka, S., Nowak, M.A., Kondor, I. (2005). Random matrix filtering in portfolio optimization. *ACTA Physica Polonica B*, **36**, 2757-2765.

[14] Rachev, S.T., Menn, C., Fabozzi, F.J. (2005). Fat-tailed and Skewed Asset Return Distributions: Implications for Risk Management, Portfolio Selection, and Option Pricing. John Wiley & Sons, New York.

[15] Rachev, S.T., Mittnik, S. (2000). Stable Paretian Models in Finance. John Wiley & Sons, Chichester.

[16] Rachev, S., Ortobelli, S., Stoyanov, S., Fabozzi, F. Biglova, A. (2008). Desirable properties of an ideal risk measure in portfolio theory. *International Journal of Theoretical and Applied Finance*, **11**(1), 19-54.

[17] Rockafellar, R.T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, **26**(7), 1443-1471.

[18] Ross, S. (1978). Mutual fund separation in financial theory-the separating distributions. *Journal of Economic Theory*, **17**, 254-286.

[19] Ruppert, D., Wand, M.P. (1994). Multivariate locally weighted last squares regression. *The annals of Statistics*, **22**(3), 1346-1370.

[20] Scott, D.W. (2015). Multivariate Density Estimation: Theory, Practice, and Visualization. John Wiley & Sons, New York.

[21] Sharpe, W.F. (1994). The Sharpe ratio. *Journal of Portfolio Management*. Fall, 45-58.

[22] Stoyanov, S., Rachev, S., Fabozzi, F. (2007). Optimal financial portfolios. *Applied Mathematical Finance*, **14**(5), 401-436.

[23] Watson, G. S. (1964). Smooth regression analysis. *Sankhya, Series A*, **26**(4), 359-372.

Simulation model of Czech small farming business

Igor Krejčí[1], Gabriela Koláčková[2], Ivana Tichá[3]

Abstract. The paper deals with the dynamic simulation of small farming business. The model parameters are stated according to the official statistics and qualitative research that is used mainly in order to define and specify structure of the system. For the purpose of the paper we had to define the "small farming business" precisely to be able to quantify the model parameters on the basis of various data sources using different methodology and therefore with significantly different values of indicators. Missing values for the model variables are estimated on the basis of Powell optimisation.

System dynamics simulation model shows possible development scenarios. Whether the scenarios are pessimistic or optimistic, the parameters are based on real historical conditions. It is typical for a small farmer that prices of both input and output of the transformation process are out of his reach and non-negotiable. The simulation results depict a dependence of the small farming business on the environment constrains and the weak market position. In this paper, we focus mainly on the scenarios of animal production of the defined small farming business.

Keywords: Small farming business, system dynamics, computer simulation, scenarios.

JEL Classification: C63, Q12
AMS Classification: 91B55, 93C15

1 Introduction

The diversification of the economic production of small farms has been growing in recent years [3]. In [14] we introduce the system dynamics model that depicts rather weak competitiveness of small farms due to the impossibility to influence the prices of the production and inputs. The simulation model is based on official data and qualitative research [23] which describes the farmers' motivation to remain in agricultural sector. Commonly, it is not the profit maximisation what keeps them running their business. Very often they do not leave the market despite adverse conditions but stay because of tradition or "hobby" reasons. The care of animals without a consequent market production is a common hobby reason and its analysis is the aim of the paper. In this paper we show the impact of those non-economic reason to stay in business on the economic situation of the farmer.

Systems' behaviour grows from the systems structure [17]. However, the attributes of dynamic complexity such as non-linearity, delays, feedback structure or the self-organisation [30] together with the bounded rationality [28, 29], mental modelling [8] and cognitive limitations of human brain [1, 19] leads to the counterintuitivity of the systems' behaviour and responses. Therefore, many authors stress the computer simulation as a tool necessary for understanding the structure, behaviour, sources of policy resistance and indication of leverage points [10, 15, 18, 30]. System dynamics stress the difference between desired and actual states of variables, the equilibrium is not a precondition but could arise from the structure and interaction of elements in the system [11, 30].

Application of system dynamics in agriculture is not new. [31] analysed cooperatives in Swiss agribusiness with the impact on reorganisation project. [27] identifies the factors of eco-agriculture promotion in county in China. Also [16] analyse the eco-agriculture in province in China, the analysis shows the limits for further development. [25, 26] show that the main impulse for the conversion to organic farming in Slovenia are subsidies.

In our paper, we focus on Czech average small farm. For these purposes we define the small farmer according to main statistical classifications. The modelled farmer is "Agricultural entrepreneur – natural person" [3], in institutional sub-sector 142 "Recipients of Property Income and Transfers", with the production category in group

[1] Czech University of Life Sciences Prague, Department of Systems of Engineering department, Kamýcka 129, Prague 6 - Suchdol, Czech Republic, krejcii@pef.czu.cz.
[2] Czech University of Life Sciences Prague, Department of Management, Kamýcka 129, Prague 6 - Suchdol, Czech Republic, kolackovag@pef.czu.cz.
[3] Czech University of Life Sciences Prague, Department of Management, Kamýcka 129, Prague 6 - Suchdol, Czech Republic, ticha@pef.czu.cz.

01.11 “Growing of cereals (except rice), leguminous crops and oil seeds” according to both classification of economic activities CZ-NACE rev. 2 and classification of products CZ-CPA [5].

2 Material and Methods

Beside the well-known characteristics of agriculture production as seasonal dependency and dependency on subsidies [12, 32], the model reflects the attributes typical for small farms [23]:

- Owner’s satisfaction is superordinate to profitability;
- The criterion that we call “hobby” reasons is quantified as 5 heads of livestock
- Family members often retake a role of the employed staff;

Figure 1 shows the simplified stock and flow diagram where the crucial stock variables (accumulations) are denoted in boxes and flow variables indicating inflows and outflows of the stocks are denoted by pipes. Clouds denote the boundaries of the examined system.

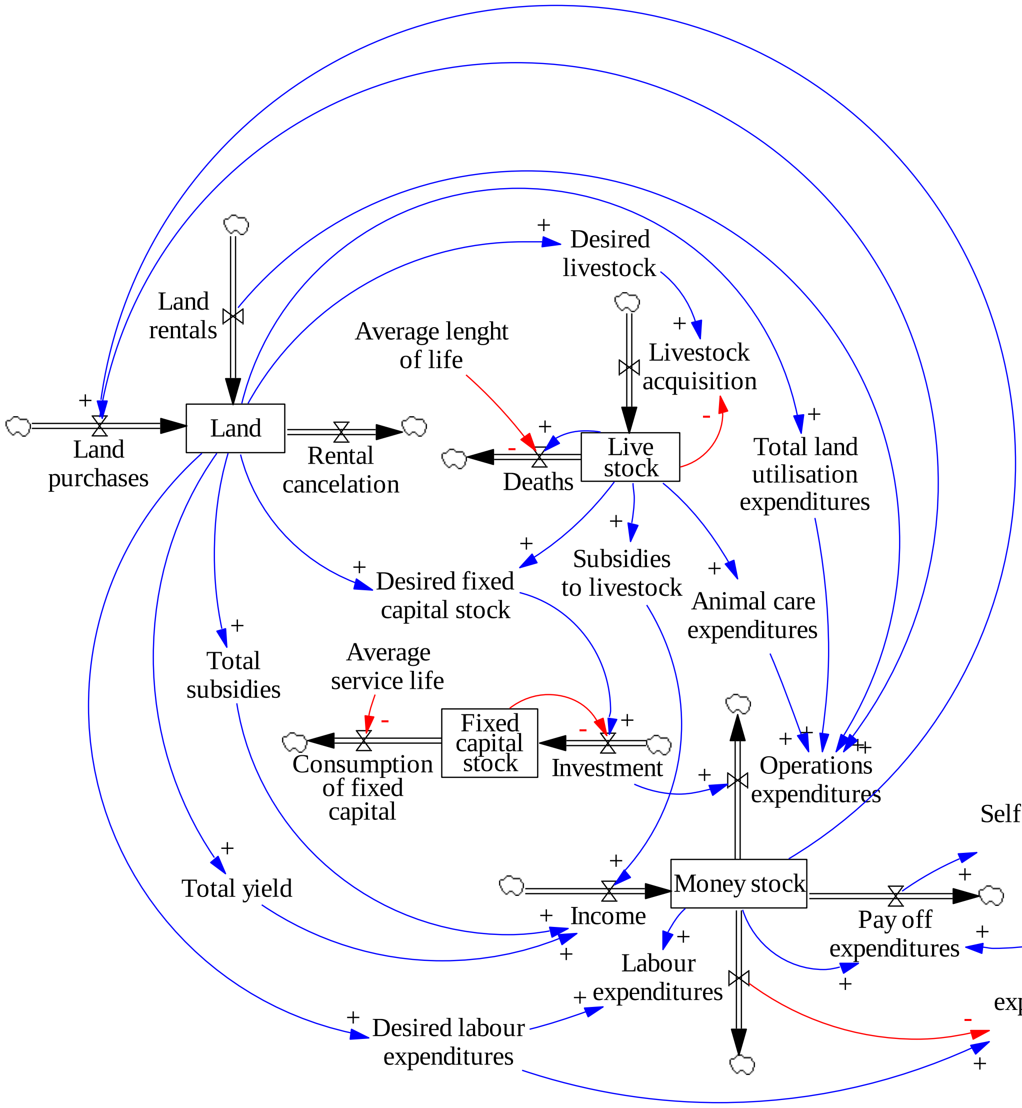


Figure 1 Small farmer – simplified stock and flow diagram

The polarity of causal links has a mathematical interpretation. For positive polarity, *y* increases (decreases) above (below) a value that it would have gain, if *x* increased (decreased) [30]:

$$\frac{\delta y}{\delta x} > 0\,. \tag{1}$$

For negative polarity, *y* decreases (increases) below (above) a value that it would have gain, if *x* increased (decreased) [30]:

$$\frac{\delta y}{\delta x} < 0 \,. \tag{2}$$

Box variable represents the definite integral [30]:

$$s = \int_{T_0}^{T} (i - o)\mathrm{d}t + s_{T_0} \,, \tag{3}$$

where s is stock variable, i indicates inflows, o indicates outflows, T_0 is initial time, T is current time and t is any time between T and T_0.

Model parameters are quantified mainly on the basis of the agriculture accounts [3], fixed capital is quantified according to national accounts statistics [2, 4] and international standards for measuring fixed capital [21]. To define the initial arable land area of the small farm we preferred the data from Czech Statistical Office. Data from green book [20] are based on different thresholds that could lead to significantly different estimates of the area of arable land.

Average utilised arable land in 2013 was 23.2 ha according to green book but 36.8 ha according to structure survey done by Statistical Office. This is due to lower thresholds for arable land, livestock etc. used for purposes of the green book (e.g. minimum 1 ha of arable land in green book and 5 ha in structural survey). The difference between these thresholds results in the very different sample size (n = 16,523 in structural survey and 26,076 in green book). As the model reflects the farm that have the main income from the crop production we consider the higher threshold used by Czech Statistical Office as more relevant. For these purposes, we asked for more detailed data that are not published regularly – data on farmers that really utilise the land (i.e. without sole animal producers that do not utilise any land). This resulted in another increase of the average area of arable land to 50.7 ha in 2013. Our simulation assumes a production of three main crops in the Czech agriculture– wheat, barley and rape.

Czech Statistical Office also provided us more detailed data on Labour force in annual work units (AWU) as these indicators are not regularly published on the level of Agricultural entrepreneur – natural person. Data on wages and labour force are based on official wage statistics [6].

Data on subsidies area from [9] and necessary unpublished parameters are from Institute of Agricultural Economics and Information. Average expenditures on crop production (fertilisation, seeds, crop protection) and livestock (such as consumption of feed, medication and veterinary care) are based on data from [13]

To depict the farmers' point of view we introduce a variable denoted as *self-debt*. That variable accumulates the difference between desired and real labour expenditures. In case the possible income of the farmer (and other workers/family members) is below the average in the agriculture it is perceived as a bad period but still, the farmers assume better days to come. The *self-debt* is paid off in periods of surplus.

Table 1 describes the model boundary. Model boundary contains the most significant endogenous, exogenous and excluded variables [22]. Therefore, the table depicts the simplification of the real system and also the real situation of the small farmer as the pure receiver of both inputs' and outputs' prices.

Endogenous variables	Exogenous variables	Excluded variables
Consumption of fixed capital	Average animal care costs	Bank loan
Fixed capital stock	Average gross wage	Calf breeding
Investment	Average land rent	Interest rate
Labour expenditures	Average land rentals	"Other" crops
Labour force	Average land utilisation costs	"Other" kinds of animals
Land purchases	Average price of cattle	Bankruptcy threshold
Livestock	Average subsidies to land	
Livestock acquisition	Average subsidies to livestock	
Money stock	Land price	
Own land	Price per ton of production	
Rented land	Yield per ha	
Self-debt		

Table 1 Small farm model boundary table

The missing values of parameters (mainly delay times and decision making thresholds) were estimated by Powel optimisation [7, 24] with a goal to minimise the difference between simulated and real average land area. The secondary criterion was to minimise the difference between the desired and real labour expenditures, the last goal was to reach desired level of livestock before 2012.

For the simulation we use Vensim DSS simulation software. For the calculations we used Euler integration with time step set to $dt = 0.03125$. The complete simulation model has nearly 150 variables and parameters and its detailed description including the description of main feedback loops is in [14].

3 Results and discussion

Figure 2 shows the farmer‘s cash flow in standard scenario. The standard scenario represents the settings when the farmer decides when to buy a new land and the quality of the production is average. In this case, the farmer buys the livestock around 2006 and the cash flow with and without livestock diverse from that time.

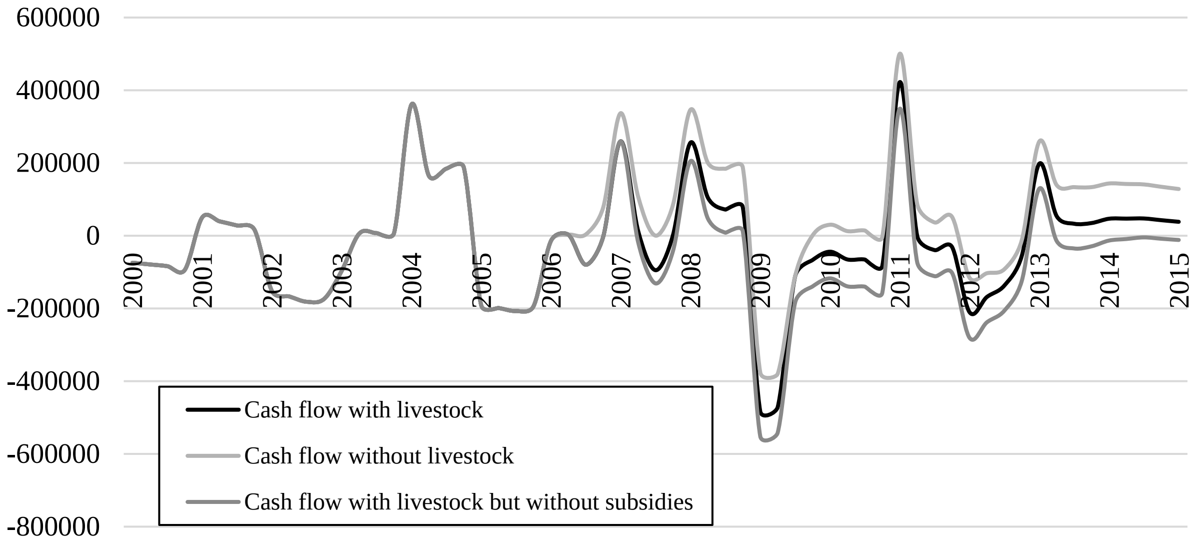


Figure 2 Cash flow – standard scenario (CZK, prices of 2010)

Similarly to [14] we tested different scenarios with various levels of optimism and pessimism. The optimistic scenario is represented by a forecast that expects the bread quality of crop and yield and prices on the level of the three years (2012-2014) maximum. On the other hand, the pessimistic scenario assumes feed quality of production and prices on the level of the three years minimum (i.e. prices of 2014 for wheat and rape and prices of 2012 for barley). Pessimistic scenario II is the most pessimistic from the selected set. The settings assume feed quality and prices of year 2010 for all products.

Moreover, each scenario has variation whether the farmer obtains subsidies on livestock or not. Figure 3 shows the share of livestock expenditures on the farmer's income. Highest share at the beginning is caused mainly by the initial acquisition of the livestock. The decreasing trend is also caused by increasing land area and therefore the increasing income.

The average difference between the optimistic and the most pessimistic scenario is 2.10%. Hobby expenditures that do not exceed 7% of income do not seem to be determining the overall unfavourable situation of small farms. For small farmers the livestock is often the reason to stay in the agribusiness. Therefore, it is also a common source of subject diversity in the agriculture and it preserves the continuity of the rural tradition. On the other hand, despite the low share of livestock expenditures, the figure 2 points out the importance of the subsidies to keep the cash flow positive.

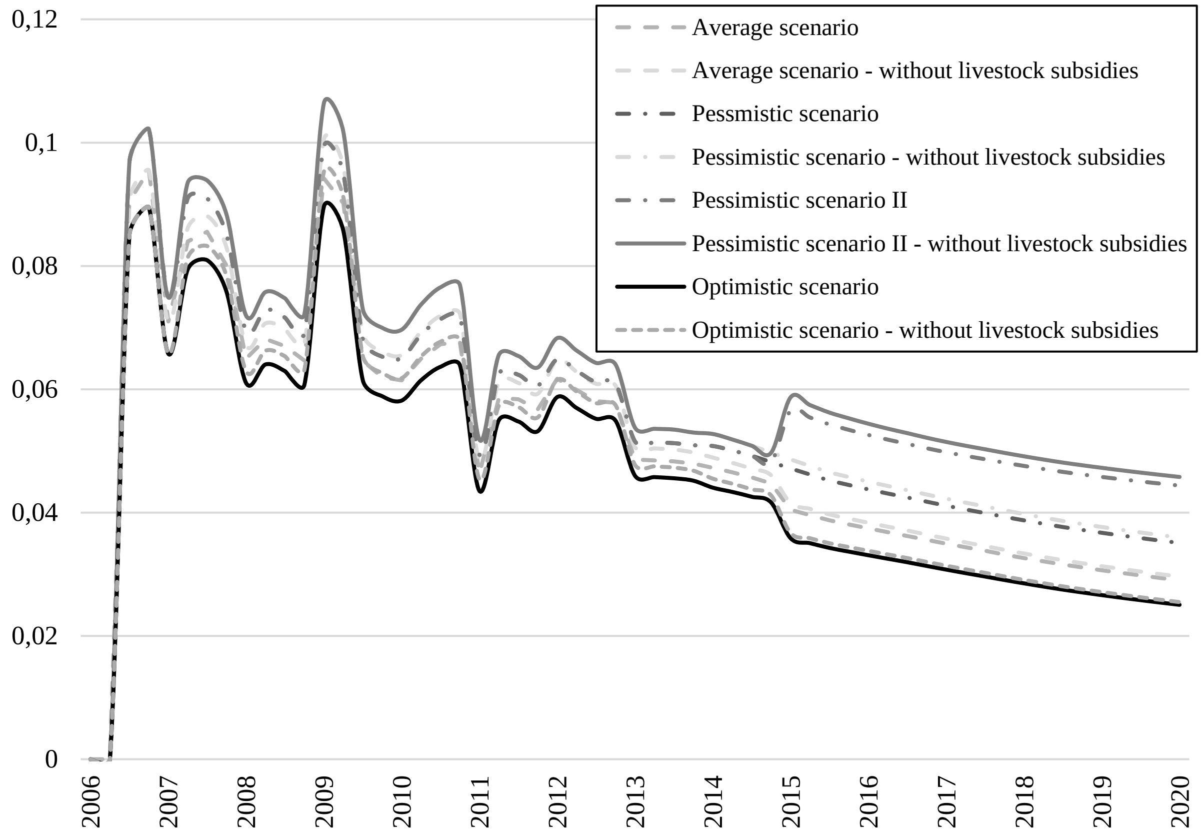


Figure 3 Share of livestock expenditures on total income

4 Conclusion

In [14] we show that 2011 was an important turning point where small farms' situation was getting better and self-debt was continuously paid off. On the other hand, the simulation clearly shows the weak market position of small farm and supports the necessity of business diversification as the source of farmers' economic stability. The non-economic reasons to stay in agribusiness cannot prevail if the farmer is not profitable or the profit is too low for the long period of time.

In this paper we show that the important non-economic factor – the livestock – does not affect the significant part of the farmers' income. On the other hand, despite the share on the income is low, it is necessary to mention that even the small part can result in the negative cash flow. The livestock subsidies are crucial in this case. Therefore, it shows how economic instruments support the non-economic reasons of farmers to stay in agribusiness. Our future research will focus on transformation of livestock from the hobby reason to fully developed economic activity with breeding and increasing capacity above the hobby level.

Acknowledgements

The article is supported by the grant project of the Internal Grant Agency of the Czech University of Life Sciences Prague "Performance implications of business models adopted by the Czech agribusiness SME´s", No. 20131005.

References

[1] Cowan N.: The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* **24** (2001) 87-114.

[2] Czech Statistical Office: *Gross National Income Inventory*, [online]. Czech Statistical Office, Prague. 2002 Available: http://apl.czso.cz/nufile/GNI_CZ_en.pdf [1 November 2015).

[3] Czech Statistical Office: *Agriculture*, [online]. Czech Statistical Office, Prague. 2016. Available: www.czso.cz/csu/czso/agriculture_ekon [15 February 2016].

[4] Czech Statistical Office: *Annual National Accounts*, [online]. Czech Statistical Office, Prague. 2016 Available: apl.czso.cz/pll/rocenka/rocenka.indexnu_en [15 February 2016].

[5] Czech Statistical Office: *Classifications*, [online]. Czech Statistical Office, Prague. 2016. Available: www.czso.cz/eng/redakce.nsf/i/classifications [15 February 2016].

[6] Czech Statistical Office: *Wages – time series*, [online]. Czech Statistical Office, Prague. 2016 Available: www.czso.cz/csu/czso/pmz_ts [15 February 2016].
[7] Dangerfield B., Roberts C.: Optimisation as a statistical estimation tool: An example in testing the AIDS treatment-free incubation period distribution. *System Dynamics Review* **15** (1999) 273-291.
[8] Doyle J. K. and Ford D. N.: Mental models concepts for system dynamics research. *System Dynamics Review* **14** (1998) 3-29.
[9] Foltýn I., Zedníčková I., Kopeček P., Vávra V., Humpál J.: *Predikce rentability zemědělských komodit do roku 2014 (certifikovaná metodika)* (Prediction of agriculture comodities rentability till 2014 (certified methodology)), [online]. Institute of Agricultural Economics and Information, Prague. 2010. Available: http://www.uzei.cz/data/usr_001_cz_soubory/metodika_rentability.pdf [3 March 2015].
[10] Forrester J. W.: *Idustrial dynamics*. Pegasus Communications Waltham, 1961.
[11] Forrester N. B.: The role of econometric techniques in dynamic modeling: systematic bias in the estimation of stock adjustment models. *System Dynamics Review* **3** (1987) 45-67.
[12] Hlavsa T., Aulová R.: Analysis of the Effect of Legal Form and Size Group on the Capital Structure of Agricultural Businesses of Legal Entities. *AGRIS on-line Papers in Economics and Informatics* **4** (2013) 91-104.
[13] Institute of Agricultural Economics and Information: *Costs of Agricultural Products*, [online]. Institute of Agricultural Economics and Information, Prague. 2014. Available: www.iaei.cz/costs-of-agricultural-products/ [5 January 2015].
[14] Koláčková, G, Krejčí, I., Tichá, I. Dynamics of small farmers' behaviour - scenario simulations. In press. (accepted for publishing in Agricultural Economics), 2016.
[15] Krejčí, I., Kvasnička, R., Dömeová, L. Introducing System Dynamics at CULS Prague. *Journal on Efficiency and Responsibility in Education and Science* **4** (2011), 187-196.
[16] Li F. J., Dong S. Ch., Li F. (2012): A system dynamics model for analysing the eco-agriculture system with policy recommendations. *Ecological Modelling*, **227**: 34-45.
[17] Meadows D. H.: *Thinking in Systems. A Primer.* Wright D. (ed.) Chelsea Green Publishing Company, White River Junction. 2008.
[18] Mildeová S., Dalihod M., Exnarová A.: Mental Shift Towards Systems Thinking Skills in Computer Science. *Journal on Efficiency and Responsibility in Education and Science* **5** (2012) 25-35.
[19] Miller G. A.: Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review* **63** (1956), 81-97.
[20] Ministry of Agriculture of the Czech Republic: *Zelené zprávy (Zemědělství, eAgri)* (Green book, Agriculture) [online]. Ministry of Agriculture of the Czech Republic, Prague. 2015 Available: http://eagri.cz/public/web/mze/zemedelstvi/publikace-a-dokumenty/zelene-zpravy/ [3 March 2015].
[21] OECD: *Measuring Capital - OECD Manual 2009: Second edition*. OECD Publishing, Paris, 2009.
[22] Pierson, K. Sterman, J. D.: Cyclical dynamics of airline industry earnings. *System Dynamics Review* **29** (2013), 129–156.
[23] Poláková J., Koláčková G., Tichá I.: Business model for Czech agri-business. *Scientia Agriculturae Bohemica* **46** (2015), 128-136
[24] Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P.: *Numerical Recipes in C: The Art of Scientific Computation. 2nd ed.* Cambridge University Press, New York, 1992.
[25] Rozman Č., Pažek K., Prišenk J., Škraba A., Kljajić M.: System Dynamics Model for Policy Scenarios of Organic Farming Development. *Organizacija* **45** (2012), 212-218.
[26] Rozman Č., Pažek K., Škraba A., Turk J., Kljajić M. (2012b): Determination of Effective Policies for Ecological Agriculture Development with System Dynamics – Case Study in Slovenia. In: *Proceedings of the 30th International Conference of the System Dynamics Society*, (Huseman E. and Lane D., eds), System Dynamics Society, St. Gallen, 2012.
[27] Shi T., Gill R.: Developing effective policies for the sustainable development of ecological agriculture in China: the case study of Jinshan County with a systems dynamics model. *Ecological Economics*, **53** (2005), 223-246.
[28] Simon H. A.: Rational choice and the structure of environment. *Psychological Review*, **63** (1956), 129-138.
[29] Simon H. A.: Rational Decision Making in Business Organizations. American Economic Review, **69** (1979), 493-513.
[30] Sterman J. D.: *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, Boston, 2000.
[31] Weber M., Schwaninger M.: Transforming an agricultural trade organization: a system-dynamics-based intervention. *System Dynamics Review*, **18** (2002), 381-401.
[32] Žídková D., Řezbová H., Rosochatecká E.: Analysis of Development of Investments in the Agricultural Sector of the Czech Republic. *Agris on-line Papers in Economics and Informatics* **3** (2011), 33-43.

Evaluation of market risk models: changing the weights in the DEA approach

Aleš Kresta[1], Tomáš Tichý[2], Mehdi Toloo[3]

Abstract. Value at Risk (VaR) models are important part of financial risk management and thus concern especially banks and insurance companies. While improper amount of capital would have negative impact on entity performance and might even lead to bankruptcy, application of an improper model could have the same implications. In our previous research, we have suggested to use an alternative approach for the integrated evaluation of market risk model quality, based on the data envelopment analysis. Such approach allows us to examine the efficiency of available models in a comprehensive way. In this contribution we reconsider the DEA models by changing the weights restriction, ie. we consider both strictly positive as well as non-negative weights. The results obtained on real data of stock market index show some interesting findings and relations among selected models. Overall results show that NIG model with estimation horizon of 1 to 2 years should probably be preferred.

Keywords: Value at Risk, backtesting, efficiency, weights, Data Envelopment Analysis.

JEL classification: C44, G13
AMS classification: 35, 90C15

1 Introduction

Market risk constitutes important part of risk profile of financial institutions active at financial markets and especially those, that are active internationally. Since 1996, also market risk creates capital requirements – the Basel Accord suggests to measure the market risk to which the portfolio of financial institution is exposed to by internal models based on Value at Risk (VaR), ie. to estimate the left quantile of the probability distribution of future portfolio returns.

Notwithstanding, a financial institution can use its internal model only if it fulfils prespecified qualitative and quantitative criteria. It includes, besides others, comparison of recorded failures of the model (ie. observed loss to the portfolio is higher than estimated VaR) with its assumed number. The procedure of such comparison (see eg. [9]) is called backtesting since it is performed on a series of past data in a such way that on a given day t previous n observations are used to estimate the parameters of the model to obtain VaR for the future. Next, at time $t+1$, we record 0 if VaR is higher than the true loss and 1 otherwise. This is repeated over m days to obtain sequence of 1's and 0's. There are available several tests to asses the quality of the model, see e.g. [3, 2, 7], including references therein for a brief review. Unfortunately, it is not clear how to rank models with different levels of complexity and thus time costs.

Recently, with Basel III proposals an alternative measure (CVaR), see eg. [10], which takes into account the conditional expected loss, has been considered to replace VaR. However, estimation as well as backtesting of CVaR is much more complicated, since it crucially depends on returns behaviour in the tails. Moreover, backtesting procedure should be evaluated on longer series of data rather than on just over one year period as was assumed in Basel I and II.

The aim of our research is to evaluate various implementations of selected market risk models via

[1]VŠB – Technical University of Ostrava, Department of Finance, Sokolská 33, 701 21 Ostrava, Czech Republic, ales.kresta@tul.cz
[2]VŠB – Technical University of Ostrava, Department of Finance, Sokolská 33, 701 21 Ostrava, Czech Republic, tomas.tichy@vsb.cz
[3]VŠB – Technical University of Ostrava, Department of Systems Engineering, Sokolská 33, 701 21 Ostrava, Czech Republic, mehdi.toloo@vsb.cz

DEA approach. It has three steps. First, we apply the DEA methodology (a method that seeks a frontier to envelop data with data acting in a critical role in the process) to assess which estimation period can be regarded as efficient for a given model. Next, we take into account the mutual characteristics in order to compare particular models among the others. Finally, we try to change the weight restriction (strictly positive instead of non-negative) so that the method is more capable to identify efficient model specifications with higher sensitivity. We consider four model types HS (Historical Simulation), GI (stochastic simulation with Gaussian Innovations), NIG (stochastic simulation with Normal Inverse Gaussian innovations), and AGG (stochastic simulation with AR(1)-GARCH(1,1) model for volatility and Gaussian innovations). Since CVaR is assumed in Basel III, we consider a whole set of probability levels of VaR for the left tail of the probability distribution of returns so that as a result we get more information about the whole probability distribution. All models are evaluated for daily data of US equity index S&P 500 over previous 16 years plus another 4 years of data used for initial model estimation only.

The paper proceeds as follows. The next section is focused on market risk modelling, including specification of selected models, while Section 3 shows the DEA model definition. Crucial part of the paper is Section 4 in which we show illustrative example using three periods of US stock market index.

2 Market risk modelling and measuring

Market risk is the risk coming from the fluctuation of market prices. Its efficient modelling, measuring and managing is a crucial step for running the business of financial institutions, such as banks, insurance companies or pension funds, smoothly, without negative impacts of financial distress or even default.

Market risk evaluation

For a long time, a key risk measure used within the risk management of financial institutions has been the Value at Risk (VaR). Assuming a random variable X following a given distribution, VaR over a time of length Δt at confidence level α (ie. on a probability level $p = 1 - \alpha$) can be obtained as follows:

$$VaR_X(\Delta t, \alpha) = -F_X^{-1}(1-\alpha). \tag{1}$$

Here, $-F_X^{-1}(p)$ denotes the inverse function to the distribution function of random variable X for p. Assuming, for example, the Gaussian distribution, it can be further decomposed into the mean (the expected value) of random variable X over Δt, $\mu_X(\Delta t)$, and the product of its standard deviation, $\sigma_X(\Delta t)$, and $F_{\mathcal{N}}^{-1}(p)$ – p-th percentile of standard normal distribution (Gaussian distribution with zero mean and unit variance).

The efficiency of the estimation of VaR measure can be easily evaluated using past data. The incurred losses are compared with the risk measure (VaR) estimated in the past for each available observation (mostly closing prices on a given day from selected horizon). Such procedure is called backtesting. If the loss is higher than the risk measure, we call such observation an exception. Depending on the length of the horizon and selected probability level, we get an assumed number of exceptions – for example, if there is 100 observations and the risk is measured at 1% probability level, we should assume just 1 exception. Obviously, if the observed number of exceptions is not statistically different from the assumed number, we should regard the model as acceptable for risk measuring.

Recently, however, partly as an answer to the subprime crises of 2008 and subsequent financial crises, the financial institutions are recommended to switch to more complex risk measures than just quantile based VaR. One of such measures, which is moreover coherent, is called CVaR (ie., conditional Value at Risk):

$$cVaR_X(\Delta t, p) = -\mathbb{E}[x|x < -VaR(\Delta t, p)]. \tag{2}$$

It measures the expected value of the losses exceeding a given level (VaR here).

A considerable trouble, related to the application of CVaR measure for risk management of financial institutions, is how to implement the backtesting procedure. While backtesting the VaR is quite easy, CVaR is effectively based on continuum of VaR measures and optimal way of its testing is therefore not clear. It probably leaded the regulators to a little controversial recommendation – calculate the capital requirements using CVaR, but evaluate the quality of the model using VaR, each at different probability level.

As a response we therefore suggest here to implement VaR backtesting at several levels of probability, ie. calculate the quality of the probability distribution estimates at discrete points. The efficiency of these estimates is than mutually evaluated using Data Envelopment Approach (DEA).

Market risk estimation

We focus on the application of three different, but basics kind of models, which we denote as follows: HS (Historical Simulation), GI (stochastic simulation with Gaussian Innovations), NIG (stochastic simulation with Normal Inverse Gaussian innovations).

HS. Historical simulation has been reported as the most common model for VaR estimation in large commercial banks by several studies, see eg. [2]. Its basic form, which we apply here, estimates VaR at a given probability level as related quantile calculated from previous n observations. Obviously, in case that the quantile is not available directly (far tail and short time series), we apply extrapolation procedure using the two nearest values.

GI. The next two models are based on stochastic (Monte Carlo) simulation. First, standard approach assuming that the log-returns are normally distributed (Gaussianity) is applied. That is, we assume as a process driving the innovations a Wiener process with mean value and variance based on standard normal distribution $\mathcal{N}(0,1)$ with probability density function $f_{\mathcal{N}}(x)$ defined as follows:

$$f_{\mathcal{N}}(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}. \tag{3}$$

NIG. A quite popular model, which allows to fit also skewness and kurtosis, is NIG model (*normal inverse Gaussian model*). NIG model can be defined by the following characteristic function ($\alpha > 0$, $-\alpha < \beta < \alpha$, $\delta > 0$):

$$\phi_{\mathcal{NIG}}(x,t;\alpha,\beta,\delta) = \exp\left[-t\delta\left(\sqrt{\alpha^2-(\beta+\imath x)^2}-\sqrt{\alpha^2-\beta^2}\right)\right]. \tag{4}$$

Then, the density function is given as follows:

$$f_{\mathcal{NIG}}(x,t;\alpha,\beta,\delta) = \frac{\alpha\delta}{\pi}\exp\left(\delta\sqrt{\alpha^2-\beta^2}+\beta x\right)\frac{K_1\left(\alpha\sqrt{\delta^2+x^2}\right)}{\sqrt{\delta^2+x^2}}, \tag{5}$$

where $K_\lambda(x)$ is modified Bessel function:

$$K_\lambda(x) = \frac{1}{2}\int_0^{-\infty} y^{\lambda-1}\exp\left(-\frac{1}{2}x(y+y^{-1})\right)dy. \tag{6}$$

Alternatively, following the definition of the Brownian motion driven by inverse Gaussian (IG) process, ie. process $\mathcal{I}(t;\nu)$ with drift ν, which at time $\mathcal{I} \sim \mathcal{IG}[t;\nu]$ reaches level t, as follows:

$$\mathcal{NIG}(\mathcal{I}(t;\nu);\theta,\vartheta) = \theta\mathcal{I}_t + \vartheta\mathcal{Z}(\mathcal{I}_t) = \theta\mathcal{I}_t + \vartheta\sqrt{\mathcal{I}_t}\varepsilon. \tag{7}$$

In this case we can formulate the characteristic function as follows:

$$\phi_{\mathcal{NIG}}(x;\nu,\theta,\vartheta) = \exp\left[\frac{1}{\nu}-\frac{1}{\nu}\left(\sqrt{1+x^2\vartheta^2\nu-2\theta\nu\imath}\right)\right], \tag{8}$$

which results into: $\theta = \delta\beta/\sqrt{\alpha^2-\beta^2}$, $\vartheta = \frac{\sqrt{\delta\sqrt{\alpha^2-\beta^2}}}{\sqrt{\alpha-\beta}\sqrt{\alpha+\beta}}$ and $\nu = \left(\delta\sqrt{\alpha^2-\beta^2}\right)^{-1}$. Density function $f_{\mathcal{IG}}(x;\delta,\alpha,\beta)$ can be rewritten as:

$$\frac{\delta}{\sqrt{2\pi}}\exp(\delta\sqrt{\alpha^2-\beta^2})x^{-3/2}\exp\left(-\frac{1}{2}\left(\delta^2x^{-1}+(\alpha^2-\beta^2)x\right)\right). \tag{9}$$

The parameters of NIG model can be estimated either via maximum likelihood approach or by method of moments. We apply the latter since it seems to be more efficient from the point of view of time costs. Notwithstanding, we should keep in mind that especially for short time series it can happen that estimated kurtosis is too low or even lower than 3 – since for such cases NIG model has no solution, we need to increase it artificially.

3 DEA approach with weight constraints

Data Envelopment Analysis (DEA) is a mathematical approach to measure a set of homogeneous Decision Making Unit (DMU) with multiple inputs and multiple outputs. This method, originated by [4], maximizes the ratio of the weighted sum of outputs to the weighted sum of inputs for a DMU, subject to the condition that the same ratio for all DMUs must be less than or equal to one. Mathematically, suppose there are n DMUs ($DMU_j,\ j = 1, \dots, n$) with m inputs, $\mathbf{x}_j = (x_{1j}, \dots, x_{mj})$, and s outputs, $\mathbf{y}_j = (y_{1j}, \dots, y_{sj})$. Note that inputs and outputs in DEA are quantitative data which are available for each DMU. The following fractional programing model which is called CCR (adapted from Charnes, Cooper and Rhodes) measures the relative efficiency score of under evaluation DMU, i.e. DMU_o, $o \in 1, \dots, n$,

$$\max \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}}$$

subject to

$$\begin{aligned} \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} &\leq 1 \quad j = 1, \dots, n \\ u_r &\geq \epsilon \quad \forall r \\ v_i &\geq \epsilon \quad \forall i \end{aligned} \tag{10}$$

where $\mathbf{v} = (v_1, \dots, v_m)$ and $\mathbf{u} = (u_1, \dots, u_s)$ are the unknown input and output weights and ϵ is the non-Archimedean infinitesimal which is added to prevent input and output weights from zero (for more details see [1]). The fractional programming model (6) can be converted to the following linear programming model, which can be solved straightforwardly:

$$\max \theta = \sum_{r=1}^{s} u_r y_{ro}$$

subject to

$$\begin{aligned} \sum_{i=1}^{m} v_i x_{io} &= 1 \\ \sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} &\leq 0 \quad j = 1, \dots, n \\ u_r &\geq \epsilon \quad \forall r \\ v_i &\geq \epsilon \quad \forall i \end{aligned} \tag{11}$$

In our preliminary application, see [8], we have observed that the basic specification of CCR model cannot distinguish among models with relatively good performance so that most of them were evaluated as efficient. We therefore try to evaluate the model under an alternative specification of DEA by replacing the non-negativity constraint of weights by a strictly positive one (see [11] for a comprehensive discussion of various DEA models).

4 Experimental study

In order to evaluate the performance of particular VaR models and their various implementations due to different estimation periods, we first calculate the efficiency score via CCR model. We calculate the efficiency first independently for each of the model and after that mutually to obtain integrated efficiency. For the calculation we use the daily closing prices of S&P 500 index over last 16 years. This period is also split into two sub-periods of 8 years to examine the impact of the subprime crises of 2008.

First, we estimate VaR for each of the models considered (ie. GI, NIG, HS) at 16 probability levels, ie. for 0.15, 0.14,. . . , 0.01, and 0.005, respectively, assuming one estimation window, and compare it with observed losses over whole horizon so that we get the number of observed exceptions. Than we change the window and calculate the exceptions again, so that we get the differences between the assumed and observed exceptions for each model assuming various estimation windows, running from about 1 month

(21 days) up to approximately 4/8 years. Note that we penalize both positive and negative differences in the number of model failures (exceptions) since none of them can be regarded as efficient in terms of capital (lower number of failures leads to too high capital, which means poor economic performance, while higher number of failures leads to too low capital and increases the risk of default). We proceed in the same way also for both subperiods.

The results are apparent from Figure 1. We can observe that (1) GI model is mostly useless since it is quantified as efficient only occasionally – the reason is that it cannot capture the behaviour of real returns due to lack of parameters; (2) the whole period supports the usage of HS model, while the subperiods show that we should give preference to NIG model – it seems that in the long horizon NIG model sometimes (specific estimation window) provide incorrect estimate and therefore is beaten by HS model, which, though generally not so accurate, does not show such failure; (3) as indicated in the previous point, the very long evaluation horizon makes the results to be averaged, ie., the efficiency index is generally higher for the long period than for the subperiods; (4) also in the long horizon, replacing the non-negativity of weights constraint by the strictly positive weights has almost no effect, while it significantly decreases the number of units evaluated as efficient within the subperiods, especially with the shortest and long estimation horizons. On this basis, the overall recommendation is to use rather the NIG model with estimation horizon between little less than one year and two years.

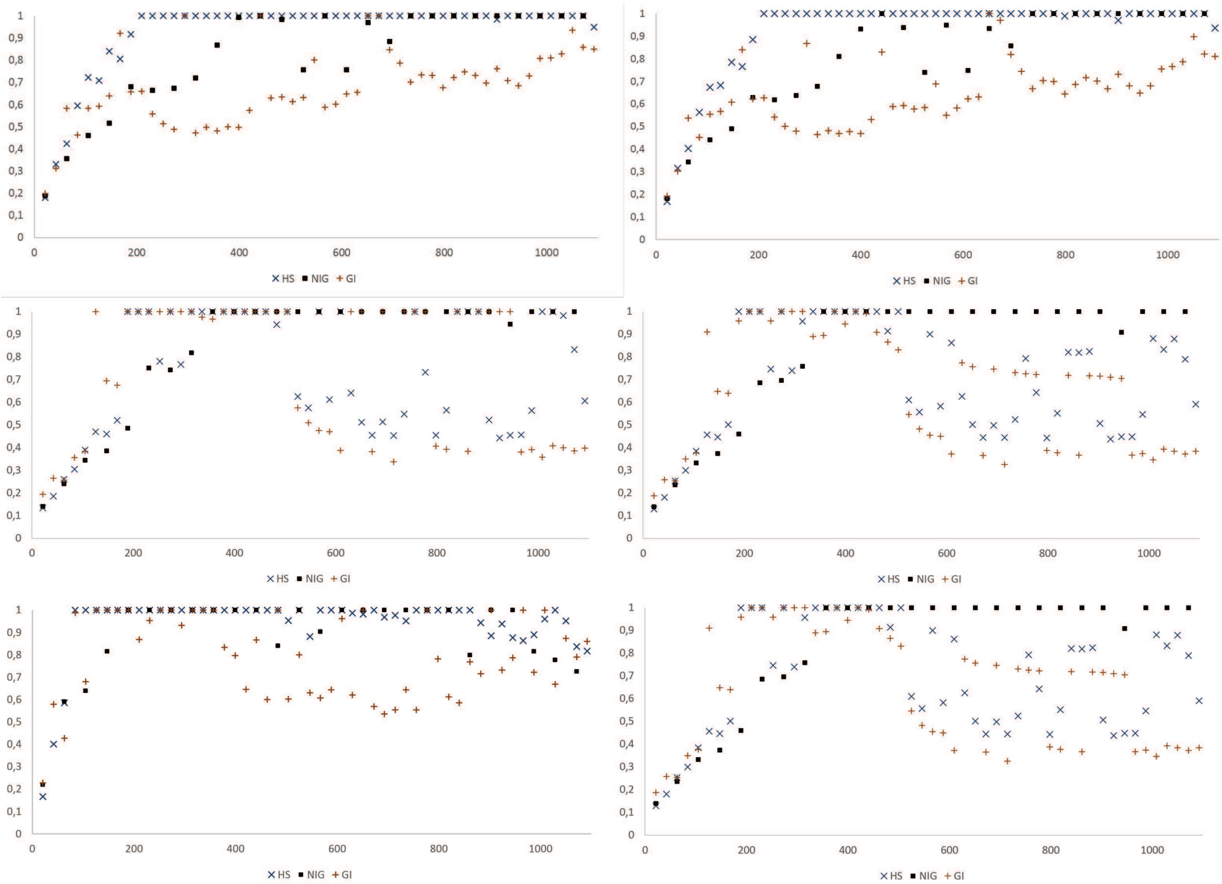


Figure 1 Efficiency of models over 2000–2015 (top), 2000–2008 (middle), 2008–2016 (bottom)

5 Conclusion

In this short contribution we have focused on the comprehensive evaluation of market risk models, including selection of efficient estimation horizons. Since the classic non-negativity of weights constraint often leads to classification of a really large number of models as efficient, ie. the method cannot distinguish among efficient and almost efficient models in comparison to surely inefficient models, we suggested to use the strictly-positive weights constraint only. Overall results show that NIG model with estimation horizon of 1 to 2 years should probably be preferred – this recommendation is not strong since the model does not behave well on the longest time period we have considered and thus is beaten by HS approach. The further directions of research might be focused on evaluation of portfolio models, autoregressive models or consideration of some more DEA alternatives to further decrease the number of efficient models. One might be interested to study the behaviour of the model in far tails, ie. with very low p, though these results are generally very case sensitive.

Acknowledgements

The research was supported by an SGS project of VSB-TU Ostrava under No. SP2016/11 and the Czech Science Foundation through project No. 13-13142S. The support is greatly acknowledged.

References

[1] Amin, G.R., Toloo, M.: A polynomial-time algorithm for finding in DEA models. *Computers and Operations Research* **31** (2004), 803–805.

[2] Berkowitz, J., O'Brien, J.: How accurate are value-at-risk models at commercial banks? *Journal of Finance* **57** (2002), 1093–1111.

[3] Berkowitz, J., Christoffersen, P.F., Pelletier, D.: Evaluating Value-at-Risk Models with Desk-Level Data. *Management Science* **57** (2011), 2213–2227.

[4] Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. *European Journal of Operational Research* **2** (1978), 429–464.

[5] Cooper, W.W., Ruiz, J., Sirvent, I.: Choosing weights from alternative optimal solutions of dual multiplier models in DEA. *European Journal of Operational Research* **180** (2007), 443–458.

[6] Cooper, W.W., Seiford, L.M., Tone, K.: *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, second ed. Kluwer Academic Publishers, 2007.

[7] Kresta, A., Tichý, T.: International Equity Portfolio Risk Modeling: The case of NIG model and ordinary copula functions. *Finance a úvěr – Czech Journal of Economics and Finance* **61** (2012), 141–151.

[8] Kresta, A., Tichý, T., Toloo, M.: Examination of market risk estimation models via DEA approach. *9th International Iranian Operation Research conference*, Shiraz University of Technology, 241–255, 2016.

[9] Resti, A., Sironi, A.: *Risk Management and Shareholders' Value in Banking: From Risk Measurement Models to Capital Allocation Policies.* Wiley, 2007.

[10] Rockafeller, R.T., Uraysev, S.: Conditional Value at Risk for general loss distributions. *Journal of Banking and Finance* **26** (2002), 1443–1471

[11] Toloo, M.: *Data Envelopement Analysis with Selected Models and Applications.* Series of Advanced Economic Issues. VSB-TU, Ostrava, 2014.

[12] Toloo, M., Masoumzadeh, A., Barat, M.: Finding an initial basic feasible solution for DEA models with an application on bank industry, *Computational Economics*, 2004.

Capital market efficiency in the Ising model environment: Local and global effects

Ladislav Kristoufek[1], Miloslav Vosvrda[2]

Abstract. Financial Ising model is one of the simplest agent-based models (building on a parallel between capital markets and the Ising model of ferromagnetism) mimicking the most important stylized facts of financial returns such as no serial correlation, fat tails, volatility clustering and volatility persistence on the verge of non-stationarity. We present results of Monte Carlo simulation study investigating the relationship between parameters of the model (related to herding and minority game behaviors) and crucial characteristics of capital market efficiency (with respect to the efficient market hypothesis). We find a strongly non-linear relationship between these which opens possibilities for further research. Specifically, the existence of both herding and minority game behavior of market participants are necessary for attaining the efficient market in the sense of the efficient market hypothesis.

Keywords: Ising model, efficient market hypothesis, Monte Carlo simulation.

JEL classification: G02, G14, G17
AMS classification: 91G60, 91G70

1 Introduction

Agent-based models (ABM) have attracted much attention in economics and finance in recent years [10, 12, 21] as they describe the reality better than simplified models of traditional economics and finance. The crucial innovation lies in assuming a boundedly rational economic agent [20, 18] instead of a perfectly rational representative agent with homogeneous expectations [16, 14]. In these models, agents make decision without utility maximization but usually using simple heuristics. The resulting systems are majorly driven endogenously, i.e. without exogenous shocks forcing the dynamics.

In finance, the founding contributions were laid by Brock and Hommes models [3, 4] characteristic by strategy-switching agents and possible bifurcation dynamics. Here, we focus on one of the simplest ABMs built on a parallel between ferromagnetism and market dynamics, i.e. the Ising model adjusted for financial economics. In the model, economic agents participating in the market are spins of a magnet. In the same way as the spins, the agents are influenced by (make their decisions based on) their neighbors, or agents with similar beliefs, but also by the overall market sentiment and activity. Such model has been shown to mimic the basic financial stylized facts successfully [2]. We focus on the model parameters and how they influence price and returns dynamics in the optics of the efficient market hypothesis. The attention is given to finding a combination of parameters which yields an efficient market or dynamics close to it. We show that the effects of parameters are more complicated than one might expect and their influence is apparently non-linear. This opens further research options which are shortly discussed as well.

[1]Institute of Information Theory and Automation, Czech Academy of Sciences, Pod Vodarenskou vezi 4, Prague 8, CZ-182 08, Czech Republic; Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Opletalova 26, Prague 1, CZ-110 00, Czech Republic, kristouf@utia.cas.cz

[2]Institute of Information Theory and Automation, Czech Academy of Sciences, Pod Vodarenskou vezi 4, Prague 8, CZ-182 08, Czech Republic; Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Opletalova 26, Prague 1, CZ-110 00, Czech Republic, vosvrda@utia.cas.cz

2 Methodology

In this section, we provide a brief introduction to the Ising model adjusted for financial markets and we shortly discuss the essence of the efficient market hypothesis.

2.1 Ising model

As a representative of the agent-based models applied to finance and financial economics, we opt for a simple Ising model adjusted for financial markets as proposed by Bornholdt [2]. The model builds on a combination of the standard Ising model of ferromagnetism (with local field interactions) [11] and a minority game behavior of market agents [1, 5]. Financial market is represented by a square lattice (usually with torus-like neighborhoods) with a side of N, i.e. with N^2 elements representing market agents. These elements are referred to as spins due to their magnetization of either $+1$ or -1. This spin orientation is translated into a financial market as either a buy or a sell signal (decision), respectively. The spin orientation of agent i for a time period t is labelled as $S_i(t)$. For each agent i, the local field $h_i(t)$ for a time period t is defined as

$$h_i(t) = \sum_{j=1}^{N} J_{ij} S_j(t) - \alpha C_i(t) \frac{1}{N} \sum_{j=1}^{N} S_j(t). \tag{1}$$

The first term is defined as a local Ising Hamiltonian with neighbor interactions J_{ij}. This is the reference to the standard Ising model. The second term represents the minority game dynamics as it depends on the total magnetization of the system $M(t) \equiv \frac{1}{N}\sum_{j=1}^{N} S_j(t)$ at time t with sensitivity α. $C_i(t)$ gives the strategy of spin i. Orientation of spin i at time $t+1$ is given as

$$\begin{aligned} S_i(t+1) &= +1 \quad \text{with} \quad p = [1 + \exp(-2\beta h_i(t))]^{-1}, \\ S_i(t+1) &= -1 \quad \text{with} \quad 1 - p, \end{aligned}$$

which is directly connected to Eq. 1 with an additional sensitivity β, which is parallel to the inverse temperature of the original Ising model.

The strategy term $C_i(t)$ is given as a general term in Eq. 1 which can be further specified. A popular choice is to highlight the minority game behavior of the spin by allowing the strategy to change with respect to the total magnetization and the spin's own orientation. This specification also allows for more strategy types. Bornholdt [2] proposes the following dynamics:

$$C_i(t+1) = -C_i(t) \text{ if } \alpha S_i(t) C_i(t) \sum_{j=1}^{N} S_j(t) < 0 \tag{2}$$

A simple alternative is to keep the strategy spin update immediately, which reduces the local field equation to

$$h_i(t) = \sum_{j=1}^{N} J_{ij} S_j(t) - \alpha S_i(t) \left| \frac{1}{N} \sum_{j=1}^{N} S_j(t) \right|, \tag{3}$$

i.e. it does not depend on the strategy of any spin at all.

The price dynamics of the system is extracted directly from the magnetization dynamics so that

$$\log P(t) = M(t) \equiv \frac{1}{N} \sum_{j=1}^{N} S_j(t). \tag{4}$$

2.2 Efficient market hypothesis

Efficient market hypothesis (EMH) has been a cornerstone of modern financial economics for decades. Even though its validity has been challenged on many fronts, it still remains the firm theoretical basis of the financial economics theory [6, 15]. In the fundamental paper, Fama [8] summarizes the empirical validations of the theoretical papers of himself [7] and Samuelson [17]. The theory is revised and made clearer in Fama's 1991 paper [9] where the market efficiency is split into three forms based on availability of information. From mathematical standpoint, the historical papers [7, 17] are more important as they provide specific model forms of an efficient market. Specifically, Fama [7] connects the (logarithmic) price process of an efficient market to a random walk and Samuelson [17] specifies it as a martingale. Implications for the statistical properties of the returns process of the efficient market are straightforward. For the former, the returns are expected to be serially uncorrelated and follow the Gaussian (normal) distribution, which implies independence. For the latter, only the serial uncorrelatedness is implied. We thus have two straightforward implications of the market efficiency – normally distributed (for the random walk definition) and serially uncorrelated (for both random walk and martingale definition) returns – which we use in the simulations presented in the next section.

3 Results and Discussion

3.1 Simulation setting

We are interested in the ability of the Ising model defined between Eqs. 1-4 to meet the criteria attributed to the efficient capital market, i.e. normality and serial uncorrelatedness of returns. To test these, we use the Shapiro-Wilk test [19] and Ljung-Box test [13], respectively.

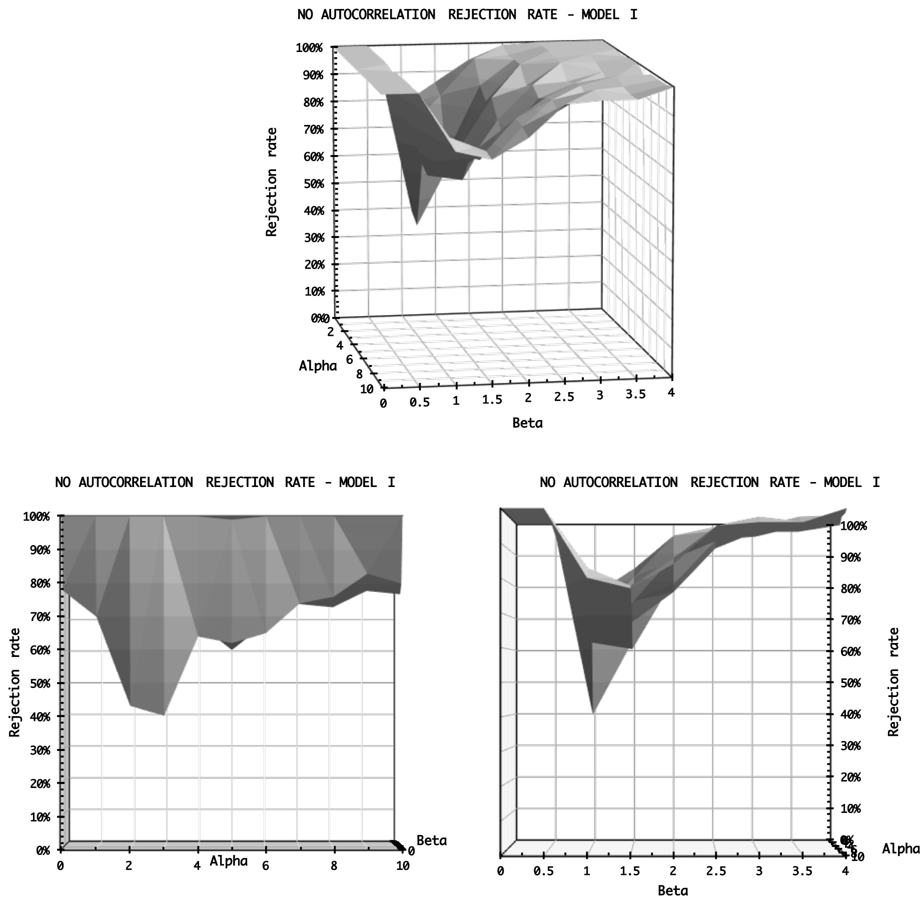


Figure 1: Rejection rates of no serial correlation hypothesis for Model I according to Eq. 1. Parameter α varies between 0 and 10 with a step of 1, and parameter β between 0 and 4 with a step of 0.5. Other parameters are set at $T = 1000$ and $N = 25$, neighborhood interactions J_{ij} are set to the nearest neighbors and the spin itself with a weight of 1, and 0 otherwise. We provide a 3D view as well as focusing on parameters separately.

There are two crucial parameters in the model – α and β – which can influence the prices and returns dynamics emerging from the model. We vary these two parameters and study how it influences the rejection rate of normality and uncorrelatedness with respective tests. In other words, we are interested in a proportion of times these tests reject (with a significance level of 0.10) market efficiency of series generated by the financial Ising model with specified parameters. Based on findings of previous research [2], we manipulate α between 0 and 10 with a step of 1 and β between 0 and 4 with a step of 0.5. We fix the time series length $T = 1000$ and the number of agents in the market to $N^2 = 25^2 = 625$. The neighborhood influence J_{ij} is set equal to 1 for the nearest neighbors and the spin's own position (five spins in total), and 0 otherwise. For each setting, we perform 100 simulations. Two specifications are studied – Model I given by Eq. 3, i.e. with fixed strategy spins, and Model II given by Eq. 1, i.e. with variable strategy spins. The code in R is available upon request.

3.2 Main findings

The findings are summarized in Figs. 1 and 2 for Model I and Model II, respectively. The 3D charts summarize the results (rejection rates) for simulations described in the previous section. Before turning to these, it needs to be noted that for the normality testing, only the case when $\beta = 0$ gives the rejection rates around 10% whereas for $\beta > 0$, normality is rejected practically always. This is true for both specifications of the model and regardless the values of parameter α. These are thus not represented graphically.

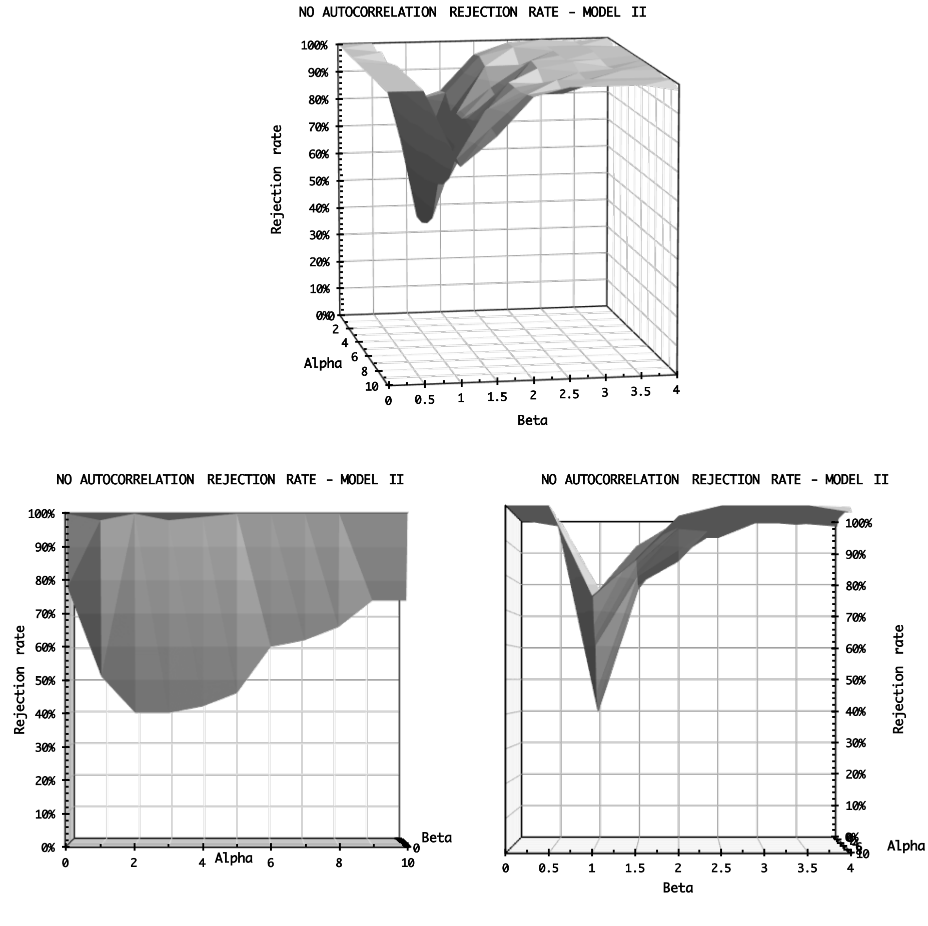


Figure 2: Rejection rates of no serial correlation hypothesis for Model II according to Eq. 3. Parameter α varies between 0 and 10 with a step of 1, and parameter β between 0 and 4 with a step of 0.5. Other parameters are set at $T = 1000$ and $N = 25$, neighborhood interactions J_{ij} are set to the nearest neighbors and the spin itself with a weight of 1, and 0 otherwise. We provide a 3D view as well as focusing on parameters separately.

We now turn to the tests of uncorrelatedness. For both models, we find a strongly non-linear dependence between rejection of no serial correlation hypothesis and the model parameters. For the sensitivity to the global magnetization (parameter α), we find a minimal rejection rate of approximately 40% at $\alpha = 3$ for Model I and at $\alpha = 2$ for Model II. For $\alpha = 0$, the rejection rate is around 80% for both models, and the same is true for the other boundary of $\alpha = 10$. The serial correlation dynamics thus emerges both for no reaction to the total magnetization, i.e. avoiding the influence of the overall market situation, and for a strong minority game behavior. There is thus no simple outcome such that a minority game behavior induces a serial correlation structure or the other way around. Such structure emerges for both extremes and market gets closer to efficiency for a setting in between.

Qualitatively similar results are found for the β parameter, i.e. the sensitivity to the local field. The minimal rejection rate is found at $\beta = 1$ for both models. For $\beta < 1$, the no serial correlation hypothesis is rejected practically always. The relationship between β and the rejection rate is smoother for $\beta > 1$ but still the rejection rate gets very close to 100% for $\beta > 3$ for Model I and $\beta > 2$ for Model II.

These preliminary results suggest the following. First, there is no simple linear relationship between market efficiency and model parameters. This poses a problem for policy makers potentially trying to get the market closer to efficiency as there is no simple answer to this endeavor. Second, which is tightly connected to the first, more detailed (smoother) simulations need to be undertaken to find a more precise efficient setting. And third, inclusion of the strategy spin plays no important role for this task.

Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. FP7-SSH-612955 (FinMaP) and the Czech Science Foundation project No. P402/12/G097 "DYME – Dynamic Models in Economics".

References

[1] Arthur, W.B.: Inductive reasoning and bounded rationality. *American Economic Review* **84** (1994), 406–411.

[2] Bornholdt, S.: Expectation bubbles in a spin model of markets: Intermittency from frustration across scales. *International Journal of Modern Physics C* **12**, 5 (2001), 667–674.

[3] Brock, W. A. and Hommes, C. H.: A rational route to randomness. *Econometrica* **65**, 5 (1997), 1059–1095.

[4] Brock, W. A. and Hommes, C. H.: Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics & Control* **22** (1998), 1235–1274.

[5] Challet, D. and Zhang, Y.-C.: Emergence of cooperation and organization in an evolutionary game. *Physica A* **246** (1997), 407–418.

[6] Cont, R.: Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* **1**, 2 (2001), 223–236.

[7] Fama, E.: The behavior of stock market prices. *Journal of Business* **38** (1965), 34–105.

[8] Fama, E.: Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance* **25** (1970), 383–417.

[9] Fama, E.: Efficient Capital Markets: II. *Journal of Finance* **46**, 5 (1991), 1575–1617.

[10] Hommes, C. H.: Chapter 23: Heterogeneous agent models in economics and finance. In L. Tesfatsion and K.L. Judd, editors, *Handbook of Computational Economics*, Elsevier (2006), 1109–1186.

[11] Ising, E.: Beitrag zur theorie des ferromagnetismus. *Zeitschrift fur Physik* **31**, 1 (1925), 253–258.

[12] LeBaron, B. and Tesfatsion, L.: Modeling macroeconomies as open-ended dynamic systems of interacting agents. *The American Economic Review* **98**, 2 (2008), 246–250.

[13] Ljung, G. M. and Box, G. E. P.: On a measure of a lack of fit in time series models. *Biometrika* **65**, 2 (1978), 297–303.

[14] Lucas Jr., R. E.: Expectations and the neutrality of money. *Journal of Economic Theory* **4**, 2 (1972), 103–124.

[15] Malkiel, B.: The efficient market hypothesis and its critics. *Journal of Economic Perspectives* **17**, 1 (2003), 59–82.

[16] Muth, J. F.: Rational expectations and the theory of price movements. *Econometrica* **29**, 3 (1961), 315–335.

[17] Samuelson, P.: Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* **6** (1965), 41–49.

[18] Sargent. T. J.: *Bounded Rationality in Macroeconomics.* Oxford: Clarendon Press, 1993.

[19] Shapiro, S. S. and Wilk, M. B.: An analysis of variance test for normality (complete samples). *Biometrika* **52**, 3-4 (1965), 591–611.

[20] Simon, H. A.: A behavioral model of rational choice. *The Quarterly Journal of Economics* **69**, 1 (1955), 99–118.

[21] Stiglitz, J. E. and Gallegati, M.: Heterogeneous interacting agent models for understanding monetary economies. *Eastern Economic Journal* **37** (2011), 6–12.

The Role of Banking and Shadow Banking Sector in the Euro Area: Relationship Assessment via Dynamic Correlation

Zuzana Kučerová[1], Jitka Poměnková[2]

Abstract. Financial sector comprises many types of institutions, among others shadow banking institutions. Shadow banking activities can be associated with the potential financial instability because they are realised outside the regulated banking sector. The aim of the paper is to assess the relationship between the banking and shadow banking activities and credit lending standards imposed by banking institutions in Euro Area countries. For such aim we use correlation analysis. In deep view into dependence structure is done via dynamic correlation. We use quarterly data over the period 2003-2015. However, an attention is paid to the period during and after the financial crisis where the particular role of shadow bank institutions seems to be important. We conclude that banking institutions of the analysed Euro Area countries originated fewer loans in reaction to tighter lending standards (imposed by banking institutions on their loans). However, the expected positive impact of tighter lending standards, higher interest rates and higher capital requirements on loans provided by shadow banking institutions is confirmed only in some periods and frequency ranges.

Keywords: dynamic correlation, segmentation correlation, banking sector, shadow banking sector, banking sector

JEL Classification: E42, E44, E51
AMS Classification: 62P20, 62M10, 91B84

1 Introduction

Shadow banking institutions became a part of the financial system; they take part in the securitisation process and use sophisticated securitisation techniques to intermediate credit and produce structured financial products. As Bakk-Simon et al. [1] state, shadow banking *"...refers to activities related to credit intermediation, liquidity and maturity transformation that take place outside the regulated banking system*". Commonly used definition of the shadow banking sector is that of the FSB [9]; it describes the shadow banking system "*...as the system of credit intermediation that involves entities and activities fully or partially outside the regular banking system, or non-bank credit intermediation in short*". The shadow banking activities could be a source of potential financial instability and could increase the probability of a severe financial contagion. Therefore, these activities should be monitored and the interconnectedness between banking (regulated) and shadow banking (unregulated) sector should be measured. The Financial Stability Board (FSB) recommends national authorities to enhance their monitoring framework to uncover potential risks hidden in the shadow banking sector by means of the application of a stylised monitoring process (see FSB [9]). In our paper, we use the general definition of FSB [9], i.e. we focus on such activities realised fully or partially outside the traditional or regulated banking system.

The relation between economic indicators would be a task for a long time. In the front of interest, there is the research of the mutual relations and its measuring. Traditional analyses were performed in the time domain and were based on the correlation and similar methods (see work of Engle and Granger [4]), or with the idea of co-integration, common features (Engle and Kozicki [5]), common cycles and co-dependence (Vahid and Engle [16], [17]). Consequent question resulted into an analysis of common features, into questions how to quantify the degree of synchronisation and how to analyse the evolution of such a synchronisation in time. However, the basic approach still uses correlation analysis and its modification such as moving correlation.

In the past several decades, the following methodological approaches proceeded to use the spectral and cross spectral analyses which allow a detailed study of time series structure (Iacobucci and Noullez [10], Poměnková and Maršálek [15]). Thus, for the evaluation of relation, the method of coherency, squared coherency, dynamic correlation and phase shift can be used. Croux et al. [2] provide a theoretical background of dynamic correlation and phase shift methods and of coherency or squared coherency (with a practical application on business cycles

[1] VSB-Technical University of Ostrava, Faculty of Economics, Department of National Economy, Sokolská třída 33, Ostrava, Czech Republic, e-mail: zuzana.kucerova@vsb.cz.
[2] Brno University of Technology, Faculty of Electrical Engineering and Communication, Technická 12, 616 00 Brno, Czech Republic, e-mail: pomenkaj@feec.vutbr.cz.

in Europe and the USA). The use of the dynamic correlation approach can be also seen in Fidrmuc et al. [8] who estimate the determinants of output comovement among OECD countries. Poměnková et al. [14] focus on changes in dynamic correlation of economic cycles during the financial crisis among V4 countries, Germany and euro area; results show different responses to a symmetric shock in V4 countries. Kučerová and Poměnková [12] assess the relationship between financial and trade integration in the new EU member countries using more complex view via the classical, moving and finally dynamic correlation.

The aim of the paper is to measure the relationship between the banking and shadow banking sector activities and credit lending standards imposed by banking institutions in Euro Area countries in period 2003-2015. The analysis is conducted only for Euro Area countries: Austria, Belgium, Cyprus, Estonia, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Slovakia, Slovenia, Spain, and Portugal. For such aim we use correlation analysis. In deep view into dependence structure is done via dynamic correlation. The structure of the paper is as follows. The second section describes data and indicators and section three presents methods used in our paper. In the fourth section, we present the results of our analysis. Section five brings conclusion.

2 Data and indicators

The database of the European flow-of-funds data does not provide sufficient data to analyse the shadow banking in the Euro Area. However, it is possible to combine data from two ECB statistical online databases – *monetary statistics* and *euro area accounts* (EEA) – and find a proxy which enables us to roughly measure shadow banking activities (Bakk-Simon et al. [1]). In the EEA database, we focus on financial corporations which comprise monetary financial institutions (MFIs), other financial intermediaries (OFIs) and insurance corporations and pension funds (ICPFs) (see ECB [6]). MFIs represent the regulated banking sector and include central banks, credit institutions and money markets funds. The OFIs can be regarded as a part of the sector of the shadow banking activities. In our paper, we use these institutions as representatives of the shadow banking sector. We also use data from the ECB *Bank Lending Survey* (BLS); these data are available since 2003:Q1 (see ECB [7]).

Several categories of indicators can be used as a proxy of the size of the shadow banking sector.[3] In our paper, we measure the activities of the banking and shadow banking sector using *the difference of the cumulative level of long-term loans provided by the relevant sector*. Differences are used to focus on the dynamics (i.e. net loans provided) and not on the overall level of the outstanding amount of loans. As a measure of lending standards, we use a measure defined by ECB [7] – the diffusion index; it is based on the quarterly responses to questions concerning lending standards (i.e. information from the supply side) and is measured as the weighted difference between the share of banks reporting that lending standards have been tightened and the share of banks reporting that they have been eased (banks who have answered “considerably” are given score 1 and banks that answered “somewhat” are given score 0.5). Positive values of the index indicate that a larger proportion of banks have tightened lending standards and *vice versa*. The representations of time series are depicted in Figure 1.

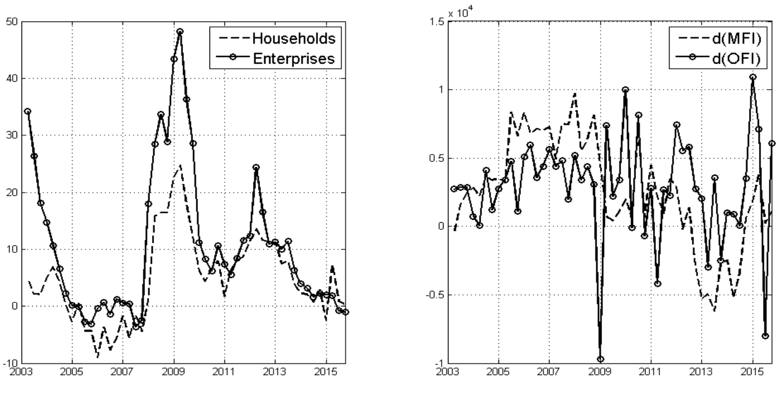


Figure 1 Time series representation, Source: ECB [6], ECB [7]

3 Methods

For analysis of dependence between time series we use classical approach represented by classic correlation coefficient as well as dynamic correlation. In case of classical correlation and regression analysis we use as input time series data described in previous part. In such way, we are going to correlate Household, i.e. Enterprises,

[3] See e.g. Bakk-Simon *et al*. [1] or Kučerová [11].

and MFI, i.e. OFI. Because correlation method is well known, we skip its description here. In case of dynamic correlation, we follow Croux et al. [2] who measured co-movement between two time series via dynamic form. Let us have two time series *y* and *z*. Thus, we can measure their similarity according to the following form:

$$\rho_{yz}(\omega_1,\omega_2)=\frac{\int_{\omega_1}^{\omega_2} C_{yz}(\omega)d\omega}{\sqrt{\int_{\omega_1}^{\omega_2} S_z(\omega)d\omega \int_{\omega_1}^{\omega_2} S_y(\omega)d\omega}}, \tag{1}$$

where *Cyz* is a co-spectrum (the real part of the cross-spectrum) and *Sy*, *Sz* are the individual spectra of time series *y* and *z* for frequencies ω. Integrating the eq. (1) in the frequency band from ω_1 to ω_2 evaluates the common behaviour of two time series in the given band of frequencies. For $\omega_1=0$, $\omega_2=\pi$ the integration is done over the whole defined frequency range and thus the dynamic correlation coefficient corresponds to the classical correlation coefficient (Fidrmuc et al. [8], Kučerová and Poměnková [12]). If our attention is turn into different frequency range (shorter one), we can change the edges of integrals ω_1 and ω_2 to evaluate specific correlation of the frequency range of our interest.

4 Results

The analysis of the relationship between the credit lending standards imposed by banks on long-term loans provided to households (denoted as Households) and to enterprises (denoted as Enterprises) and the volume of new long-term loans (expressed using the first difference) provided by both MFIs (variable d(MFI)) and OFIs (variable d(OFI)) follows several steps. Firstly, we verify results from the study Kučerová and Poměnková [13] on an updated data sample (longer time series and data sample containing 18 countries). However, we extend the correlation analysis assuming lag order of one quarter in case of the variable d(OFI) (not in case of the variable d(MFI)). The reason is that we intend to test whether economic agents (enterprises and households) need some time to move from MFIs to OFIs in case MFIs refuse to provide the loan (particularly in times of tighter lending standards). Moreover, we work with two time periods, i.e. 2003-2015 and 2008-2015, in order to identify the changes after the world financial crisis. Consequently, we proceed with in deep view into dependence structure via dynamic correlation analysis on an updated new data set.

In the first step, we perform a classical correlation analysis for the 2003-2015 and 2008-2015 period, both with lag 0 (i.e. no lag) and 1 (i.e. the lag of one quarter) only in case of the variable OFI; the results are presented in Table 1. It is apparent that there is a negative correlation between the lending standards of banks and loans provided by banks in the 2003-2015 periods i.e. tighter lending standards are connected with lower level of new loans provided by banks. The same holds for non-banks which is not consistent with our hypothesis. However, the results are different in the 2008-2015 period; the results confirm a positive relationship between lending standards and loans provided by banks. It could be interpreted as a result of the financial crisis when banks became hesitant in the policy of providing loans. In this context, it is possible to discuss a possibility of credit crunch: even though the lending standards of banks were eased the banks did not originate new loans (probably as a result if higher perceived risk). All the results are similar in case of lag 0 and lag 1.

	2003-2015				2008-2015			
Lag	0	0	1	1	0	0	1	1
	Households	Enterprises	Households	Enterprises	Households	Enterprises	Households	Enterprises
Households	1	0,8356***	1	0,8345***	1		1	0,9378***
Enterprises	0,8356***	1	0,8345***	1	0,9405***	1	0,9378***	1
d(MFI)	-0,3035**	-0,1167	-0,3096**	-0,1251	0,3072*	0,3291*	0,3183*	0,3387*
d(OFI)	-0,1285	-0,0784	-0,1505	-0,0251	-0,0147	-0,0144	-0,0204	0,0933

Table 1 Classical correlation coefficients

Note: statistically significant at: *** 1%, ** 5%, * 10%

However, the classical (static) correlation gives us information just about a linear positive or negative relation between indicators. In our case, the negative classical correlation between lending standards and loans provided by both banks and non-banks in the 2003-2015 and the positive classical correlation between lending standards and loans provided by banks is confirmed in the 2008-2015 period. Unfortunately, results given by classical correlation can be insufficient and does not provide in deep view into dependence structure. Therefore, we proceed with dynamic correlation (correlation in frequency domain).

Focusing on dynamic correlation for time period 2003-2015 without lag (Figure 2a) we can find several frequency ranges with the high level of correlation. The first range is (0-0.2) representing long cycle such a trend or a cycle of the length of 10 quarters with negative correlation values. The second range is (0.4-0.7) representing short cycles of the length of 3-5 quarters with both positive and negative correlation values. The third range is (0.9-1) representing rapid changing component (cycles of the length of 2 quarters) with negative correlation values. Imposing one lag order for time period 2003-2015 (Figure 2b) we can see generally same tendency of dynamic curve trend in the frequency range (0-0.7). The difference (between figures 2a and 2b) occurs in the frequency range (0.7-1) for the dependence of loans by non-banks and lending standards for enterprises taking opposite trend-direction (from negative to positive level) with higher dynamic correlation values. The trend-direction in case of loans by non-banks and lending standards for households is also opposite with lower dynamic correlation values.

Focusing on the period 2008-2015 (Figure 2c-2d), the results allow almost the same partition into frequency ranges (0-0.2), (0.4-0.7) and (0.7-1) such in time period 2003-2015 (Figure 2a, 2b). For the period 2008-2015 (Figure 2c-2d) in the range (0-0.2), we can identify significantly higher level of correlation for bank loans and lending standards for households, resp. enterprises. In this range, correlation of non-bank loans and households, resp. enterprises, takes an insignificant level. In the second range (0.4-0.7) for the period 2008-2015, the dynamic correlation curves have the same shape for all indicators and lags 0. In case of lag 1, there is a similarity between the comovement of lending standards and bank loans both for households and enterprises as well as non-banks loans both for households and enterprises. We can also say that the general tendency (increase/decrease) for lag 1 is the same. This result is contrasts to the results for the whole period (2003-2015). In the frequency range (0.7-1), the trend-direction of dependence of lending standards for households takes positive values for both lags. The most dramatic change of dynamic correlation curve in the 2003-2015 and after shortcut 2008-2015 in the frequency range (0-0.3) covers business cycle frequencies (6-32 quarters, i.e. the frequency range (0.06-0.33)).

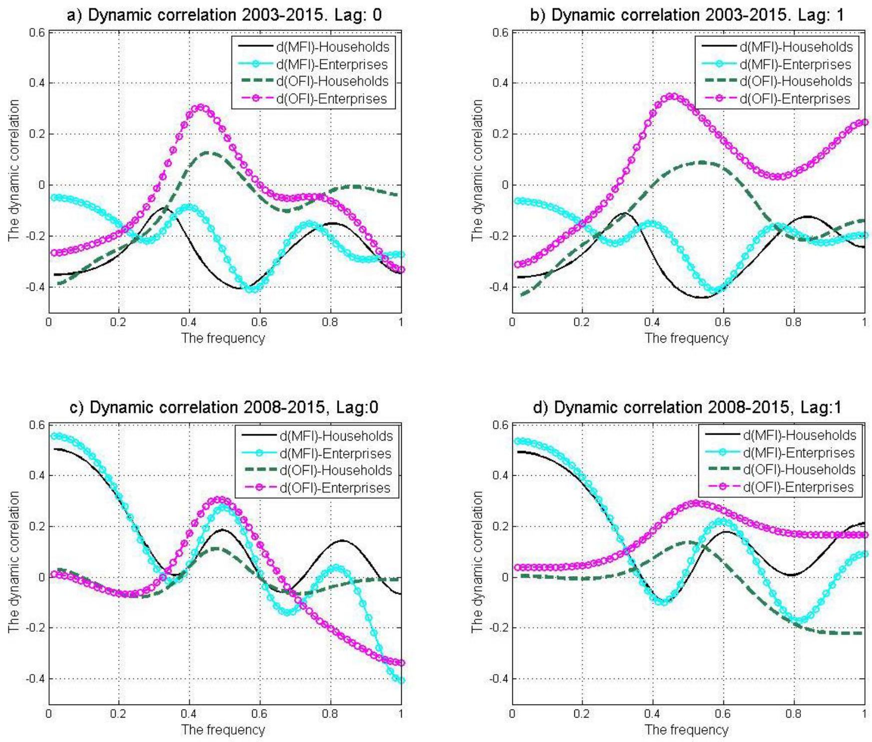


Figure 2 Dynamic correlation

Now, we proceed into the calculation of the correlation level on pre-define frequency ranges. Therefore, according to the formula (1), we calculate three segmentation correlation coefficients corresponding to the frequency ranges (0-0.3), (0.3-0.7) and (0.7-1). Denote the dynamic correlation done over the whole defined frequency range corresponds to the classical correlation coefficient. Results are presented in the Table 2.

In the period 2003-2015, for the lag 0 and 1, we can identify the expected negative correlation of lending standards (imposed by banks on loans for households and enterprises) and loans provided by banks. Additionally, in the frequency range (0.3-0.7) we can identify the expected positive correlation of both types of lending standards and loans provided by non-banks. In the period 2008-2015 we identified following results. Firstly, we identified the positive correlation of lending standards and bank loans in almost all three frequency ranges. Secondly, we identified the positive correlation of lending standards and non-banks loans in the frequency range (0.3-0.7) for lagged and non-lagged data. Thirdly, we identified the positive correlation in the frequency range (0-0.3) and (0.7-1) for lagged data and lending standards for enterprises. Therefore, it is evident that there is a stronger positive relationship after 2008 between new long-term loans provided by non-banks and lending standards (imposed by banks) on loans for enterprises than in case of lending standards on loans for households.

However, the significance of these results is slightly limited. Focusing on the dependence between lending standards for households and long-term loans provided by banks we can see the same level of correlation for frequency range (0-0.3) and (0.3-0.7) for lagged as well as non-lagged data. Such result was not confirmed for the shortened period 2008-2015. Here, the significant level of correlation was achieved only for the frequency range (0-0.3) representing long and business cycles. In case of lending standards for household and long-term loans provided by non-banks, we can see a significant correlation only for the frequency range (0-0.3) for 2003-2015 and lag 0 and 1. In case of the shortened period, the dependence vanished. Focusing on the dependence between enterprises and d(MFI) we can find only one significant dependence which is for time period 2003-2015 for lag 1 in the frequency range (0.3-0.7).

2003-2015	0-0.3, i.e. 25Q-6.7Q		0.3-0.7, i.e. 6.7Q-2.8Q		0.7-1, i.e. 2.8Q-2Q	
Lag 0	Households	Enterprises	Households	Enterprises	Households	Enterprises
Households	1	0,8763***	1	0,6928***	1	0,4795**
Enterprises	0,8763***	1	0,6928***	1	0,4795**	1
d(MFI)	-0,3162**	-0,0924	-0,2754**	-0,2139	-0,2211	-0,2347
d(OFI)	-0,2701*	-0,1880	0,0047	0,1225	-0,0326	-0,1442
2003-2015	0-0.3, i.e. 25Q-6.7Q		0.3-0.7, i.e. 6.7Q-2.8Q		0.7-1, i.e. 2.8Q-2Q	
Lag 1	Households	Enterprises	Households	Enterprises	Households	Enterprises
Households	1	0,8767***	1	0,6756***	1	0,5035***
Enterprises	0,8767***	1	0,6756***	1	0,5035***	1
d(MFI)	-0,3209**	-0,1000	-0,3176**	-0,2503*	-0,1844	-0,1956
d(OFI)	-0,2771**	-0,1803	0,0156	0,2073	-0,1758	0,1118
2008-2015	0-0.3, i.e. 25Q-6.7Q		0.3-0.7, i.e. 6.7Q-2.8Q		0.7-1, i.e. 2.8Q-2Q	
Lag 0	Households	Enterprises	Households	Enterprises	Households	Enterprises
Households	1	0,9711***	1	0,8711***	1	0,6724***
Enterprises	0,9711***	1	0,8711***	1	0,6724***	1
d(MFI)	0,3922***	0,4257***	0,0523	0,0647	0,0531	-0,1139
d(OFI)	-0,0296	-0,0316	0,0159	0,1378	-0,0309	-0,2349
2008-2015	0-0.3, i.e. 25Q-6.7Q		0.3-0.7, i.e. 6.7Q-2.8Q		0.7-1, i.e. 2.8Q-2Q	
Lag 1	Households	Enterprises	Households	Enterprises	Households	Enterprises
Households	1	0,9687***	1	0,8766***	1	0,7271***
Enterprises	0,9687***	1	0,8776***	1	0,7271***	1
d(MFI)	0,4133***	0,4471***	0,0435	0,0511	0,0934	-0,0533
d(OFI)	-0,0002	0,0463	0,0573	0,2082	-0,1906	0,1724

Table 2 Dynamic correlation coefficients

Note: statistically significant at: *** 1%, ** 5%, * 10%

5 Conclusion

The paper was focused on the problem of interconnectedness of the banking sector and the shadow banking sector with a view to credit standards and the supply of loans. The aim of the paper was to measure the relationship between the banking and shadow banking sector activities and credit lending standards imposed by banking institutions in Euro Area countries in period 2003-2015. Using the classical correlation, it is clear that there is a negative correlation between the lending standards of banks and loans provided by both banks and non-banks in the 2003-2015 period, i.e. tighter lending standards are connected with lower level of new bank loans. In the 2008-2015 period the results confirm a positive relationship between lending standards and bank loans probably as a result of the financial crisis. It can indicate a possible credit crunch: even though the lending standards of banks were eased the banks did not originate new loans (as a result if higher perceived risk). Using the dynamic correlation, we identified the positive correlation of lending standards and bank loans in almost all three frequency ranges. We conclude there is a stronger positive relationship (after 2008) between new long-term loans provided by non-banks and lending standards (imposed by banks) on loans for enterprises than in case of lending standards on loans for households. Focusing on the dependence between lending standards for households and long-term bank loans we can see the same level of correlation for frequency range (0-0.3) and (0.3-0.7) for lagged as well as non-lagged data. Such result was not confirmed for the shortened period 2008-2015. Here, the significant level of correlation was achieved only for the frequency range (0-0.3) representing long and business cycles. Banking institutions originated less loans in reaction to tighter lending standards imposed by these banking institutions on their loans which is consistent with expected results; this finding is also consistent with findings of Demiroglu *et al*. [3]. However, the expected positive impact of tighter lending standards (concerning bank loans) on loans provided by shadow banking institutions (i.e. OFIs) has not been confirmed in all cases.

Acknowledgements

This paper was supported by the VSB-Technical University of Ostrava, Faculty of Economics under SGS Grant [no. SP2016/101]. Research described in this paper was financed by Czech Ministry of Education in frame of National Sustainability Program under grant LO1401. For research, infrastructure of the SIX Center was used.

References

[1] Bakk-Simon, K., Borgioli, S., Giron, C., Hempell, H., Maddaloni, A., Recine, F., Rosati, S.: Shadow banking in the Euro Area: An Overview [online]. *European Central Bank Occasional Paper* **133** (2012), [cit. 2014-07-12]. Available at: http://www.ecb.europa.eu/pub/pdf/scpops/ecbocp133.pdf.

[2] Croux, Ch., Forni, M., Reichlin, L.: A Measure of Comovement for Economic Variables: Theory And Empirics. *The Review of Economics and Statistics* **83** (2001), 232-241.

[3] Demiroglu, C., James, Ch., Kizilaslan, A.: Bank lending standards and access to lines of credit. *Journal of Money, Credit and Banking* **44** (2012), 1063-1089.

[4] Engle, R. F., Granger, C. W.: Cointegration and error correction: representation, estimation and testing. *Econometrica* **55** (1987), 251–276.

[5] Engle, R. F., Kozicki, S.: Testing for common features. *Journal of Business & Economic Statistics* **11** (1993), 369-380.

[6] European Central Bank: *Statistical Data Warehouse: Euro Area accounts* [online database]. (2016a), [cit. 2016-03-12]. Available at: http://sdw.ecb.europa.eu/reports.do?node=1000002340.

[7] European Central Bank: *Statistical Data Warehouse: Bank Lending Survey* [online database]. (2016b), [cit. 2016-03-10]. Available at: https://sdw.ecb.europa.eu/browse.do?node=9484572.

[8] Fidrmuc, J., Ikeda, T., Iwatsubo, K.: International transmission of business cycles: Evidence from dynamic correlations. *Economic Letters* **114** (2012), 252-255.

[9] Financial Stability Board: *Global Shadow Banking Monitoring Report 2013* [online]. Financial Stability Board, Basel, 2013, [cit. 2014-07-12]. Available at: http://www.financialstabilityboard.org/publications/r_131114.htm.

[10] Iacobucci A., Noullez, A.: A frequency selective filter for short-length time series. *Computational Economics* **25** (2005), 75–102.

[11] Kučerová, Z.: Monitoring the Shadow Banking Sector in the Euro Area. In: *Proceedings of the 12th International Scientific Conference Economic Policy in the European Union Member Countries: 16.-18. 9. 2014* (Tvrdoň, M., Majerová, I., eds.). Opava: Silesian University in Opava, 2014, 462-472.

[12] Kučerová, Z., Poměnková, J.: Financial and Trade Integration of Selected EU Regions: Dynamic Correlation and Wavelet Approach. *Ekonomický časopis* **63** (2015a), 686-704.

[13] Kučerová, Z., Poměnková, J.: The Activity of the Banking and Shadow Banking Sector vs. Lending Standards in Selected Euro Area Countries: A Wavelet Approach. In: *Proceedings of the 33rd International Conference on Mathematical Methods in Economics: 9.-11. 9. 2015* (Martinčík, D., Ircingová, J., Janeček, P., eds.). Cheb: University of West Bohemia in Plzeň, 2015.

[14] Poměnková, J., Kapounek, S., Maršálek, R.: Variability of Dynamic Correlation – the Evidence of Sectoral Specialization in V4 Countries. *Prague Economic Papers* **3** (2014), 371-384.

[15] Poměnková, J., Maršálek, R.: Time and frequency domain in the business cycle structure. *Agricultural Economics* **58** (2012), 332-346.

[16] Vahid, F., Engle, R. F.: Codependent cycles. *Journal of Econometrics* **80** (1997), 199-221.

[17] Vahid, F., Engle, R. F.: Common trends and common cycles. *Journal of Applied Econometrics* **8** (1993), 341–60.

Estimation of Alpha Stable Distribution without Numerical Difficulties

Jaromír Kukal[1], Quang Van Tran[1]

Abstract. Alpha stable distribution is frequently used for modeling of financial asset returns. This heavy-tail distribution has four parameters which need to be identified. Identification via numerical integration tends to be time consuming. Therefore, it asks for some alternative method. Our novel approach is based on Nolan parametrization of standardized characteristic function of alpha stable distribution which is expanded using Taylor expansion for alpha close to one. The standardized pdf is obtained by fast Fourier transform for fixed parameters alpha and beta in the outer optimization loop. Log likelihood function is realized as look up table with linear interpolation. Therefore, the remaining parameters can be easily determined in the inner optimization loop. This two-phase likelihood maximization methodology is then applied to artificially simulated data for verification and after that to actual returns series of several stock market indices.

Keywords: alpha stable distribution, heavy tails, parameter estimation, fast Fourier transform, maximum likelihood method, financial market indices

JEL classification: C61
AMS classification: 62G07

1 Introduction

Financial asset returns dispose a leptokurtic characteristic of their distribution which normal distribution often is unable to capture. However, this heavy tail feature of asset returns can be modeled by stable distributions. This family of distributions were first introduced by Lévy (1924, [5]) when studying the sums of i.i.d. random variables. The stable distributions is characterized by four parameters, but due to the lack of closed forms of densities and distribution functions, they are difficult to be estimated. There are several estimation methods of which MLE is most often used and accurate one, but it is very time consuming due to high computational complexity. Therefore, a fast and reliable estimation algorithm is in quest. To meet this demand, we propose an algorithm for estimation of parameters of stable distributions via Fast Fourier Transform. In this paper, we verify the workability of our method on an artificially generated data and then it is applied on real financial data to estimate parameters of returns distribution. The data are Prague stock market index PX and German stock market index DAX from the last two years.

2 Alpha Stable Distribution

Generally, besides a few exceptions the closed forms for alpha stable distributions do not exist. The only way to describe α stable law is the characteristic function. However, there are many different parametrizations for it. The most often used form of the characteristic function of standard alpha stable distributions [9] is the parametrization according to Zolotarev [11], who modifies Samorodnitsky and Taqqu's parametrization [10] and it has the logarithmic form as follows

$$\ln \psi(t) = \begin{cases} - \mid t \mid^{\alpha} \{1 + i\beta \mathrm{sign}(t)(\mid t \mid^{1-\alpha} - 1) \tan \frac{\alpha\pi}{2}\}, \text{ for } \alpha \neq 1, \\ - \mid t \mid \{1 + i\beta \mathrm{sign}(t) \frac{2}{\pi} \ln \mid t \mid\}, \text{ for } \alpha = 1. \end{cases} \tag{1}$$

Needless to say that in the formula shown above only two parameters of the distribution can be observed and the remaining two are hidden due to standardization.

[1]Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Trojanova 13, Praha 2, Czech Republic, {kukal,tranvqua}@fjfi.cvut.cz

Traditional approach according to Nolan (1997, [8]) evaluates the probability density function (PDF) $f(x;\alpha,\beta)$ directly from $\psi(t)$. Setting $\zeta = -\beta \tan \frac{\alpha\pi}{2}$, the density calculation of a standard stable random variable $f(x;\alpha,\beta)$ depends on α and x as follows:

- if $\alpha \neq 1$ and $x \neq \zeta$
$$f(x;\alpha,\beta) = \frac{\alpha(x-\zeta)^{\frac{1}{1-\alpha}}}{\pi \mid \alpha - 1 \mid} \int_{-\theta_0}^{\frac{\pi}{2}} V(\theta;\alpha,\beta) \exp\{-(x-\zeta)^{\frac{\alpha}{\alpha-1}} V(\theta;\alpha,\beta)\} \mathrm{d}\theta, \tag{2}$$
for $x > \zeta$ and $f(x;\alpha,\beta) = f(-x;\alpha,-\beta)$ for $x < \zeta$.
- if $\alpha \neq 1$ and $x = \zeta$
$$f(x;\alpha,\beta) = \frac{\Gamma\left(1+\frac{1}{\alpha}\right)\cos(\xi)}{\pi(1+\zeta^2)^{\frac{1}{2\alpha}}}, \tag{3}$$
- if $\alpha = 1$
$$f(x;1,\beta) = \begin{cases} \frac{1}{2|\beta|} \exp(\frac{x\pi}{2\beta}) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V(\theta;1,\beta) \exp\{-\exp(\frac{x\pi}{2\beta}) V(\theta;1,\beta)\} \mathrm{d}\theta, & \beta \neq 0 \\ \frac{1}{\pi(1+x^2)}, & \beta = 0 \end{cases} \tag{4}$$
where
$$\xi = \begin{cases} \frac{1}{\alpha} \arctan(-\zeta), & \alpha \neq 1 \\ \frac{\pi}{2}, & \alpha = 1 \end{cases}$$
and
$$V(\theta;\alpha,\beta) = \begin{cases} (\cos(\alpha\xi))^{\frac{1}{\alpha-1}} \left(\frac{\cos(\theta)}{\sin(\alpha(\xi+\theta))}\right)^{\frac{\alpha}{\alpha-1}} \frac{\cos[\alpha\xi+(\alpha-1)\theta]}{\cos(\theta)} \\ \frac{2}{\pi}\left(\frac{0.5\pi+\beta\theta}{\cos(\theta)}\right) \exp\{\frac{1}{\beta}(0.5\pi+\beta\theta)\tan(\theta)\} \end{cases}$$

The calculation of cumulative distribution function (CDF) $F(x;\alpha,\beta)$ also depends on parameters as follows:

- if $\alpha \neq 1$ and $x \neq \zeta$
$$F(x;\alpha,\beta) = c_1(\alpha,\beta) + \frac{\mathrm{sign}(1-\alpha)}{\pi} \int_{-\zeta}^{\frac{\pi}{2}} \exp\{-(x-\zeta)^{\frac{\alpha}{\alpha-1}} V(\theta;\alpha,\beta)\} \mathrm{d}\theta \tag{5}$$
if $x > \zeta$ and $F(x;\alpha,\beta) = 1 - F(-x;\alpha,-\beta)$ if $x < \zeta$, where
$$c_1(\alpha,\beta) = \begin{cases} \frac{1}{\pi}\left(\frac{\pi}{2} - \xi\right) & \text{if } \alpha < 1 \\ 1 & \text{otherwise} \end{cases}$$
- if $\alpha \neq 1$ and $x = \zeta$
$$F(x;\alpha,\beta) = \frac{1}{\pi}\left(\frac{\pi}{2} - \xi\right), \tag{6}$$
- if $\alpha = 1$
$$F(x;1,\beta) = \begin{cases} \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp\{-\exp(-\frac{x\pi}{2\beta}) V(\theta;1,\beta)\} \mathrm{d}\theta & \text{if } \beta > 0 \\ \frac{1}{2} + \frac{1}{\pi}\arctan(x) & \text{if } \beta = 0 \\ 1 - F(x;1,-\beta) & \text{if } \beta < 0 \end{cases} \tag{7}$$

More information on PDF and CDF of alpha stable distribution can be found in [3].

Numerical integration for every point x is time consuming operation which is the first disadvantage of this traditional approach. Both PDF and CDF are continuous functions in the neighborhood of $\alpha = 1$ or $x = \zeta$. Unfortunately, the numeric evaluation for $\alpha \approx 1$ or $x \approx \zeta$ generates discontinuities due to truncation error of ill-posed formulas as the second disadvantage. The objectives of our research is to eliminate these problems in the case of alpha stable PDF calculations as useful for identification of models with this noise type.

3 Expanded Nolan Formula

Discontinuities in PDF calculations result from the fact that Nolan formula in (1) switches in the neighborhood of $\alpha = 1$. To eliminate this effect, we reformulate Nolan formula for $\alpha \neq 1$ as

$$\ln \psi(t) = - \mid t \mid^{\alpha} \cdot (1 + i\beta \text{sign}(t) \cdot \varphi(\alpha, t))) \tag{8}$$

where

$$\varphi(\alpha, t) = \tan \frac{\alpha\pi}{2} (\mid t \mid^{1-\alpha} - 1). \tag{9}$$

After the substitutions $Q = 2\pi^{-1} \ln \mid t \mid, y = \pi(1 - \alpha)/2$ one obtains

$$\varphi(\alpha, t) = \phi(Q, y) = \frac{\exp(Qy) - 1}{\tan y} \tag{10}$$

which is not defined for $y = 0$ of course. To avoid numerical difficulties, we expand $\phi(Q, y)$ according to powers of y for $y \to 0$. Using only finite Taylor expansion, we use approximation formula

$$\phi(Q, y) = \sum_{k=0}^{n} \frac{\rho_k(Q) \cdot y^k}{k!} \tag{11}$$

instead of (9) for $\mid y \mid < \varepsilon$ where the first seven terms are

$$\begin{aligned}
\rho_0(Q) &= Q, \\
\rho_1(Q) &= Q^2/2, \\
\rho_2(Q) &= Q^3/3 - 2Q/3, \\
\rho_3(Q) &= Q^4/4 - Q^2, \\
\rho_4(Q) &= Q^5/5 - 4Q^2/3 - 8Q/15, \\
\rho_5(Q) &= 5Q^6/6 - 25Q^4/6 - 20Q^2/3, \\
\rho_6(Q) &= Q^7/7 - 2Q^5 - 3Q^3/3 - 32Q/21.
\end{aligned} \tag{12}$$

When $t = 0$ we obtain $\psi(0) = 1$ from the definition of the characteristic function. Using numeric representation with minimum positive number $realmin > 0$ depending on software implementation, the critical value of Q is

$$Q_{\text{crit}} = 2\pi^{-1} \ln \mid realmin \mid . \tag{13}$$

The fast convergence of series (11) is guaranteed for $\mid Q_{\text{crit}} \cdot y \mid < 1$ and therefore we have to set

$$\varepsilon < \mid Q_{\text{crit}} \mid^{-1} . \tag{14}$$

4 Numeric Calculation of PDF by Fast Fourier Transform

The probability density function (PDF) of any continuous distribution can be derived from the characteristic function by Fourier transform [1] as

$$\text{f}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \psi(t) \exp(-\text{i}tx) \, \text{d}t. \tag{15}$$

Fast Fourier Transform (FFT) [4] is used for numeric realization of look up table. The algorithm operates as follows:

- The characteristic function $\psi(t)$ is evaluated in 2^N equidistant points within range $\mid t \mid \leq t_{\max}$.
- Resulting finite series is proceeded by FFT and only real part of its result is used as dependent variable in the look up table.
- The independent variable in this table is also calculated from equidistant condition.

The final look up table consists of pairs (x_k, f_k) for $k = 1, \dots, 2^N$. The value of PDF i.e. f(x) is easily obtained by linear interpolation in the table. This procedure is suitable for parameter determination of any model with stable distribution and fixed values of parameters α, β using maximum likelihood method. In this case, the one-off evaluation of look up table is less time consuming in comparison with traditional approach.
In the case of unknown parameters $\alpha \in (0, 2], \beta \in [-1, +1]$ we can repeatedly perform previous algorithm on 2D grid to map the likelihood and construct confidence domain of parameters α, β. The second possibility is to maximize the likelihood in outer optimization loop.

5 Implementation and results

The methodology we have proposed in the previous sections now is applied to three sets of data. The first one is generated artificially with similar characteristics as the real datasets with $\alpha = 1.5$ and $\beta = 0.4$. The remaining one are stock market indices from Prague and German stock exchanges. These indices are transformed into their corresponding logarithmic returns. The descriptive statistics of the data used for our analysis are shown in Table 1.

Consequently, four parameters of α-stable distribution of these series are estimated. For this purpose, the range is set at $R = 10^5$ and the FFT order $N = 20$. All required calculations are performed in Matlab. The estimation results are displayed in Table 2. The obtained results clearly indicate that our estimation approach can reliably recover the values of parameters of α-stable distribution from simulated data. Regarding the distribution of real financial data, the estimated values of α in both cases are less than 2 suggesting that normal distribution is not the appropriate one for modeling asset returns, especially in the case of returns of index PX. Further, for both shape parameters α and β we calculate the so called critical region based on their loglikelihood ratio. According to [6], the boundary of 100(1 - α)% significance region can be determined as follows

$$\ln L(\mathbf{b}) - \ln L(\beta) = \frac{\chi^2_{1-\alpha}}{2}. \qquad (16)$$

Characteristic	Simulated data	Index PX	Index DAX
Mean	-0.001050	0.000198	2.83E-05
Median	-0.002970	-0.000256	-0.000783
Maximum	0.327567	0.047141	0.048165
Minimum	-0.303741	-0.044719	-0.048521
Std. Dev.	0.030427	0.009796	0.013876
Skewness	1.707822	0.262182	0.148451
Kurtosis	55.27038	5.491211	3.537781

Table 1 Descriptive statistics of analyzed data

Series	α	β	δ	γ
Simulated	1.5552	0.4592	-0.0046	0.0103
PX	1.8121	0.3963	$-3.40.10^{-4}$	0.0060
DAX	1.9382	0.6112	$-4.40.10^{-4}$	0.0095

Table 2 Estimation results of α-stable distribution

The left panels in Figures 1, 2, and 3 show the estimated confidence regions of α and β of stable distribution of our data. In these panels, the point estimates of α and β are marked as a cross and the confidence regions are delimited by a bold line. The results clearly display non-zero character of β of our simulated data while the possibility of zero β of real returns are not excluded as well as the value of α of distribution of DAX returns is not statistically different from two. In the right-hand side panels of these Figures the theoretical distributions of data which are calculated using estimated values of their parameters by our methodology are compared with their corresponding empirical distributions in the form of their histograms. Here the cohesion of these two distributions are clearly observed.

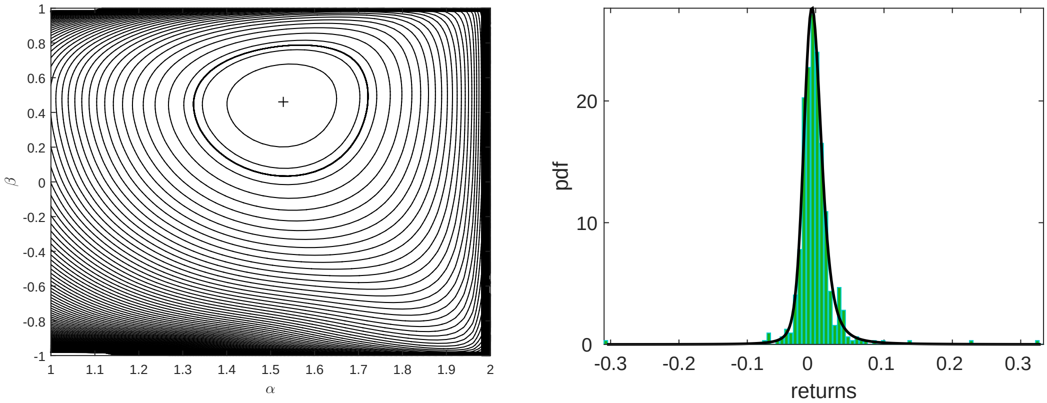


Figure 1 The critical region of α and β and the estimated pdf of simulated data

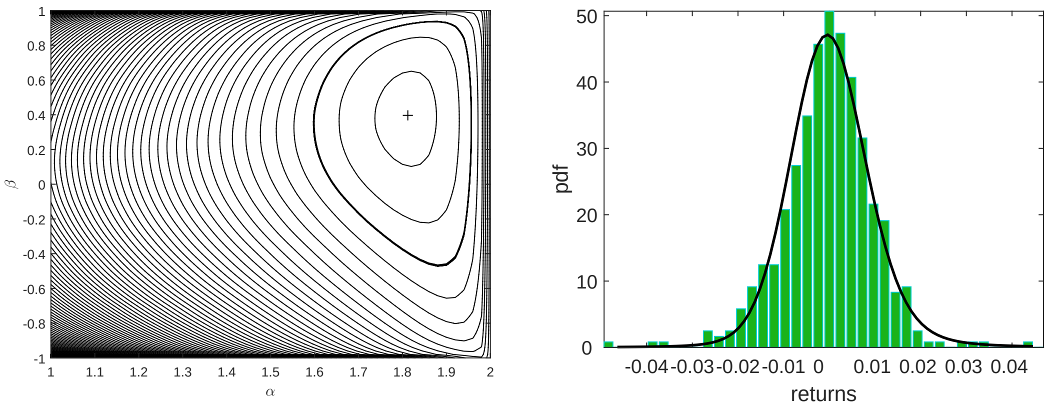


Figure 2 The critical region of α and β and the estimated pdf of PX returns

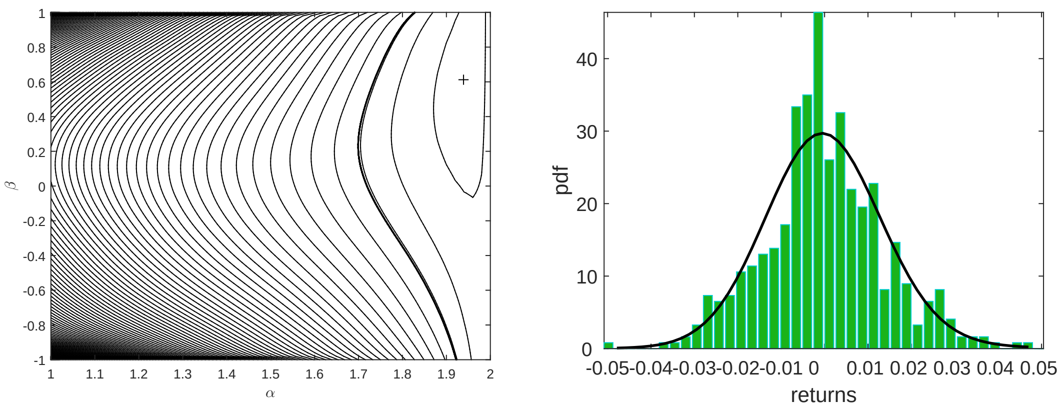


Figure 3 The critical region of α and β and the estimated pdf of DAX returns

6 Conclusion

It has been known for some time that it is possible to use characteristic function of a distribution to estimate its corresponding pdf. In this paper, we have used this technique to estimate the α-stable distribution. This approach is faster than using traditional MLE approach. Unlike [2] and [7], our methodology is used to estimate all four parameters of the distribution in a two-stage optimization procedure. To verify the functionality of this procedure, we test it both on artificially generated and real data. In the simulated data case, our method is capable of extracting the value of the distribution. In the case with real financial data, our method has also succeeded. Its results show that parameters of α-stable distribution estimated by this method provide a goodfit when we compare the theoretical distribution using the estimates with the empirical one. These results indicate that the method we propose is a viable alternative to the traditional estimation approach.

Acknowledgements

The authors acknowledge the financial support of CTU in Prague under Grant SGS14/208/OHK4/3T/14.

References

[1] Bracewell, R. N.: *The Fourier Transform and Its Applications* 3e. McGraw-Hil, Boston, 2000.

[2] Belov, A.: On the Computation of the Probability Density Function of Stable Distribution. In: *Mathematical Modelling and Analysis, Proceedings of the 10th International Conference MMA, Trakai*, 2005, 333–341.

[3] Borak, S., Hardle, W., and Weron, R.: Stable Distributions. *SFB 649 Discussion Paper* 2005-008, Center for Applied Statistics and Economics, Humbolt-Universitat zu Berlin, 2005.

[4] Brigham, E. O.: *Fast Fourier Transform and Its Applications*, Pearson, New Jersey, 1988.

[5] Lévy, P.: Théorie des erreurs la loi de Gauss et les lois exceptionelles. *Bulletin de la Société de France* **52** (1924), 49 - 85.

[6] Meloun, M., and Militký, J.: *Statistické zpracování experimentálních dat.* Edice PLUS, Pardubice, 1994.

[7] Mittnik, S., Doganoglu, T., and Chenyao, D.: Computing the probability density function of the stable Paretian distribution, *Mathematical and Computer Modelling* **29** (1999), 235 - 240.

[8] Nolan, J. P.: Numerical calculation of stable densities and distribution functions, *Communications in Statistics - Stochastic Models* **13** (1997), 759 - 774.

[9] Nolan, J. P.: Maximum likelihood estimation and diagnostics for stable distributions. In: *Lévy Processes* (O. E. Barndorff-Nielsen, T. Mikosch, S. Resnick, eds.), Brikhauser, Boston, 2001.

[10] Samorodnitsky, G., and Taqqu, M. S.: *Stable Non-Gaussian Random Processes.* Chapman and Hall, New York, 1994.

[11] Zolotarev, V. M.: *One-Dimensional Stable Distributions*, American Mathematical Society, 1986.

Cost Minimization and Electricity Consumption Growth in the Czech Households in Case of the New Conditions

Martina Kuncová[1]

Abstract. Situation in the Czech Republic concerning the electricity market has been changing every year since the liberalization of the market in the year 2006. The electricity consumption was rising till the year 2008 then it felt down (because of the financial crisis) and the next growth was not so big. In the last years the whole consumption in the country is falling. This might be the reason why the Energy Regulatory Office suggests completely new conditions for the households connected with the electricity price. The current annual cost of the electricity are influenced mainly by the height of the annual consumption. But the new suggestion gives more importance to the circuit breaker. This article tries to answer the questions for what consumption the annual cost will be lower in the new conditions (when D25d tariff rate is used) and for how many MWh the annual consumption of household can rise to have the same cost as in the new conditions case.

Keywords: Electricity prices, Annual Consumption, Suppliers, Optimization model.

JEL Classification: C44, C63, O13
AMS Classification: 90C15

1 Introduction

Electricity belongs to the essential commodities as it is strongly related with the modern age and huge development of its usage. The transformation of the electricity market in the Czech Republic into the liberalized one started in the end of the first liberalization period in Europe (for companies in 2002). The second period of the liberalization process was aimed at the households (in the Czech Republic it started in 2006). Since 2007 the liberalization of the retail market in EU should be finished and the households can choose the electricity supplier or switch to another one [2]. In majority of European Union countries the switching rate was rising during the last years but still a lot of households have not participate actively in the market by exercising choice among available suppliers [1]. In many countries including the Czech Republic the liberalization of the market led to the increasing number of suppliers and their products. Table 1 describes selected EU countries and the number of offers and suppliers available to household consumers in capital cities in 2013 [1].

Country	AT	CZ	DE	DK	FI	FR	HU	NO	PL	SE	SK	UK
Number of offers (suppliers) in 2013	40 (25)	61 (32)	376 (146)	124 (23)	204 (43)	29 (11)	4 (4)	100 (35)	77 (21)	368 (91)	19 (19)	59 (22)

Table 1 Number of products offered by the suppliers in 2013 in capital cities of the countries [1]

The complete list of suppliers, their products and prices is changing every year. The selection of the suppliers in the Czech Republic in case of households is encouraged by various web calculators and by the possibility to take part in the electricity auctions. The final decision depends on the contract conditions but mainly on the prices. With respect to all these it is a hard task for consumer to find the best supplier. The models and methods how to cope with this problem were described in [7], [8] or [9] and it can be also find in [11].

This paper continues in the analyses of the suppliers' selection for the households in the Czech electricity market with respect to the level of consumption and annual cost influenced by the suppliers' prices. In the year 2016 some changes happened as in previous years but the main event was the declaration of the Energy Regulatory Office (ERU) to completely change the price policy since 2017. ERU has announced that the tariff structure nowadays used increases the cost of the system as it is necessary to reserve double energy input in high voltage and in very high voltage than the customers really take from the network. Therefore in this paper the optimization model [7] that looks for the consumption where the annual cost would be lower in the new conditions (when D25d tariff rate is used) is applied. Other model that looks for the height of the annual consumption of household to have the same cost as in the new conditions case has been created.

[1] University of Economics, Prague, Department of Econometrics, W.Churchill Sq. 4, 13067, Prague 3, e-mail: martina.kuncova@vse.cz .

2 Situation in the Czech Electricity Market

The electricity market has been one of the last remains of former product markets limited by various regulations, restrictions or barriers to entry [2]. The liberalization process in EU has started in 1999, the first country that changed the conditions was Great Britain. The Czech Republic joined in 2002 when the companies could enter the market but since 2006 also the households can choose the electricity supplier on the retail market. As it is still necessary to regulate this market, especially the prices of the transfer and distribution of the electricity, the Energy Regulatory Office (ERU) is taking this role [4] together with three distributors (PRE, E.ON, CEZ) that operate their networks and regions. Except of ERU there is other special company - the Operator of the market (OTE) that predicates the whole market consumption and analyses the differences. Each household has its tariff rate given according to the distributor's and the supplier's conditions. The high number of suppliers and their products on the retail market embarrasses the position of the households. According to this situation it is hard to follow the rules and the price changes on the market and so it is hard to choose the best (cheapest) product.

The final price for the electricity consumption of the household is given by more factors such as consumption, fixed fees or taxes. Generally the price can be divided into two components. The first one is the controlled charge for services related to electricity transport from the generator to the final customer. This charge is annually given by ERU [5]. It covers:

- monthly lease for the circuit breaker,
- price per megawatt hour (MWh) in high tariff (HT),
- price per megawatt hour in low tariff (LT),
- price per system services,
- price for the support of the renewable energy purchase,
- charges for the electricity market operator,
- electricity ecological tax (28,30 CZK per 1 MWh).

The second part of the total price is given by the electricity supplier. It covers:

- fixed monthly fee for the selected product,
- price per megawatt hour (MWh) in high tariff (HT),
- price per megawatt hour in low tariff (LT).

The final price is increased by VAT that is 21% from 2013.

ERU as the market regulator described the problem in the necessity of having very high installed capacity to satisfy potential demand of the customers. As the consumption has been decreasing since 2010 (Figure 1) and the production is much higher (part of it is exported but also imported), ERU decided to change the rules on the market from 2017 for the households (the households consumption makes about 25% of the domestic consumption). The households should pay more for the circuit breaker almost regardless of the electricity consumption. The idea was that the households should changed the circuit breaker into a lower amperage one not to pay so much and on this basis the installed capacity for the whole network could be lower. But there are two main problems: first – the type of the circuit breaker is usually given by the distributor according to the equipment of the household and it is not possible to lower it. Second – when the final costs are not so sensitive on the amount of the consumption, households could waste the electricity. The first problem started to be more important as in the Czech Republic a lot of consumers own a cottage where the electricity consumption is not so high but it is necessary to have higher amperage circuit breaker.

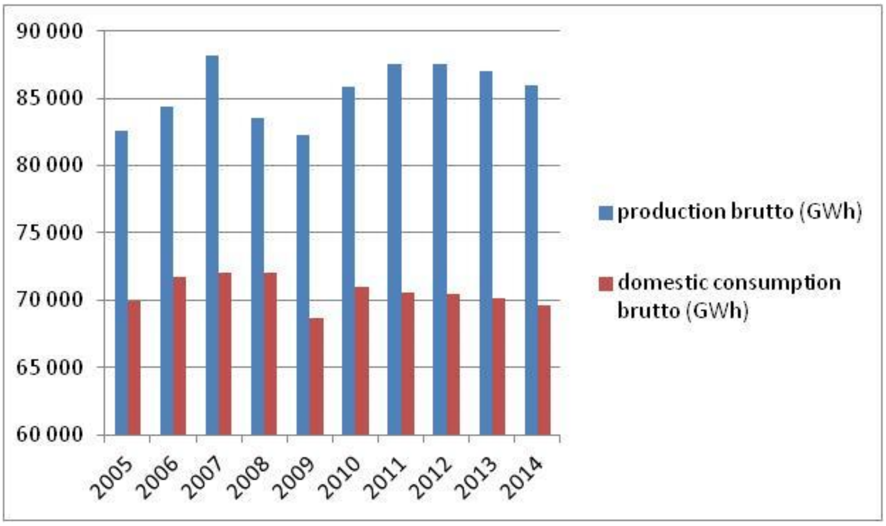


Figure 1 Electricity production and domestic consumption in the Czech Republic [4]

3 Data and Methods

The formula for the annual cost calculation for each supplier's product till the year 2015 was following:

$$COST_{ij} = (1+VAT)\begin{bmatrix} 12\left(mf_{ij} + mf_j\right) + p_{HT}c\left(ph_{ij} + ph_j\right) + \\ + p_{LT}c\left(pl_{ij} + pl_j\right) + c(os + t) \end{bmatrix}, \tag{1}$$

where

i … product, i =1, …, m,
j … distributor, j = 1, …, 3,
VAT … value added tax (VAT = 0.21),
mf … fix monthly fee,
c … annual consumption in MWh,
p_h … price in high tariff per 1 MWh,
p_l … price in low tariff per 1 MWh,
p_{HT} … percentage of the consumption in high tariff
p_{LT} … percentage of the consumption in low tariff
o_s … price for other services per 1 MWh,
t … electricity tax per 1 MWh (t = 28.3 CZK).

For the year 2016 there is a small chase in the formula (1) when the price for other services is not paid per 1 MWh but part of it is paid monthly. In our previous analysis in [8] and [7] the optimization model to compare products for the tariff rate D25d with the electricity consumption about 10 MWh annually was used. The 45% of the consumption was used in high tariff and 55% in low tariff, the circuit breaker type was from 3x20A to 3x25A. This tariff rate is given to household when the electricity is used also for the accumulative heating and hot water heating for lower and middle yearly offtake with operative management of the validity period of the low tariff for 8 hours. It is so-called dual tariff rate as it has 2 periods (high tariff, low tariff) during the day. For the analysis the same type of household is used to be able to compare the results. Since 2013 the annual cost paid by this type of households are falling down [9] but the differences between the highest and lowest costs are higher than in 2010. During the whole period the CEZ distribution area was the most expensive one, the distributor PRE is not the cheapest as before as in the last years E.ON distributor has the lowest costs. The number of products offered by suppliers circulates around 60 (in 2015 60 products for D25d, in 2016 there are 30 companies offering 57 products).

The suggested ERU model for the new conditions can be describes as follows (according the results obtain at [6], some of the values are not known separately but only in the sum):

$$COST_{ij} = (1+VAT)\begin{bmatrix} 12\cdot\left(mf_{ij} + mf_j + cp + ss + ote\right) + p_{HT}\cdot c\cdot\left(ph_{ij} + ph_j\right) + \\ + p_{LT}\cdot c\cdot\left(pl_{ij} + pl_j\right) + c\cdot(t + sr) \end{bmatrix}, \tag{2}$$

where

cp … monthly rate for the consumption place in CZK,
ss … monthly rate for the system services in CZK,
ote … monthly rate for the OTE services in CZK,
sr … price for the supported electricity resources per 1 MWh.

The rest of the values is similar to the model (1) where o_s consisted of ss, ote and sr. ERU changed the prices and rates so as the difference between the old (in 2016) and new (in 2017) annual costs would be the same for all products offered.

In the first step the value of the old and new conditions was calculated for each product i = 1, …, 57 and each distributor j according to the models. As we know that the difference in the annual costs is the same for all products in given area [6], only three optimization models were necessary to find the level of the consumption from what the new conditions give lower annual cost than the old ones. The optimization model using the previous formulas is following:

$$\max Z = c$$
$$\text{subject to } COST_{ij}\,(2016) \le COST_{ij}\,(2017)$$

Afterwards the differences in 3 consumption levels were compared – firstly with 0 consumption, secondly with the consumption lower than the founded limit and thirdly with the consumption higher than the limit.

As the results show that the new costs are decreasing with the consumption increase but the starting cost for 0 consumption are about 6 times higher than in old condition, the second optimization model is aimed at the searching for the maximal consumption c for each product where the annual costs in 2016 are equal to the annual costs in 2017 with 0 consumption. In this case 3x57 linear optimization models were solved.

4 Results

The model was created according to the information in ERU calculators [5] and [6] but the calculator for the new conditions [6] when comparing old and new prices gives a little bit different results for the old prices (the difference is about 5-10 Czech koruna - CZK) than the old calculator [5]. As the old one is valid for several years and the results of the model (1) were always equal to the results of this calculator it is probable that there might be some small mistake in the new calculator. That is why the model (1) with small change in o_s and model (2) were used.

The results of the first optimizations are different only in the distribution areas. For all of them the cheapest supplier for D25d tariff rate is ST Energy company with the product Standard AKU 8 and the limit for the annual consumption when the costs in 2016 are lower than in 2017 is about 6 or 7 MWh depending on the distributor (Table 2). For the higher consumption the new conditions are better.

Distributors	**E.ON**	**PRE**	**CEZ**
Max. consumption limit with lower prices in 2016 in MWh	6.3	7.03	6.21

Table 2 Consumption limit in MWh per year where the annual costs in 2016 are lower than in 2017

The next part was aimed at the comparison of the costs for different annual consumption. According to the results presented in Table 2, three consumption levels were selected: 0 consumption, consumption of 5 MWh per year (lower than the limit) and 10 MWh per year (higher than the limit, comparable with the previous analysis in [9]). Zero consumption is a specific case and probably unreal but on this example we see the main and big differences that influenced the values founded as the limits in the analysis above and describe the main changes in fixed fees per circuit breaker and the place of living. The difference between the situation in 2016 and the plan for 2017 is very high (Table 3) and the annual cost in 2017 would be around 6 times higher.

Consumption 0 MWh Distributor (supplier)	**Annual costs 2016 CZK**	**Annual costs 2017 CZK**	**Difference CZK**	**Increase %**
E.ON (FOSFA, Fonergy)	1605.62	11504.27	9898.65	616.50%
PRE (FOSFA, Fonergy)	1736.30	12722.34	10986.04	632.73%
ČEZ (FOSFA, Fonergy)	1852.46	12145.35	10292.89	555.63%

Table 3 Comparison of annual costs for zero consumption and the cheapest products

In the next case the annual consumption level was set to 5 MWh (to be lower than all the limits showed in Table 2 and to be equal to half of the consumption tested in our previous analyses [7] and [8]). As it was written above the cheapest region is E.ON, the most expensive is CEZ but in the new model the most expensive starts to be Prague region (PRE) where the raise of the costs is highest and it is more than 20 % higher compared with the 2016 data (Table 4).

Consumption 5 MWh Distributor (supplier)	**Annual costs 2016 CZK**	**annual costs 2017 CZK**	**Difference CZK**	**Increase %**
E.ON (ST Energy)	15031.66	17406.65	2374.99	15.80%
PRE (ST Energy)	15451.86	18624.72	3172.86	20.53%
ČEZ (ST Energy)	16041.11	18047.73	2006.62	12.51%

Table 4 Comparison of annual costs for 5 MWh consumption and the cheapest products

Consumption 10 MWh Distributor (supplier)	**Annual costs 2016 CZK**	**Annual costs 2017 CZK**	**Difference CZK**	**Increase %**
E.ON (ST Energy)	28798.25	22989.59	-5808.66	-20.17%
PRE (ST Energy)	28847.99	24207.66	-4640.33	-16.09%
ČEZ (ST Energy)	29910.32	23630.67	-6279.65	-20.99%

Table 5 Comparison of annual costs for 10 MWh consumption and the cheapest products

The last tested case was the same as in previous studies [7], [8] and [9] – that is the annual consumption 10 MWh. The consumption was taken according to the real data for one family house with a small shop. This consumption is higher than the limits and we see (Table 5) that the cheapest region would be again E.ON with about 20 % lower annual costs than in actual conditions.

All the results described above support the apprehension that the new conditions are not good for the households with small consumption and the costs are much more higher than those in 2016. As the consumption of 10 MWh yearly was taken from the real case of the family house with a shop it is possible to assume that the consumption of the usual house or flat would be lower than this (including the cottages) and so for most of the consumers the new conditions would signify an increase in final costs.

According to the results showed in Table 3 with the highest differences between old and new rates the optimization models for each supplier's product (57 cases) and each of the three region were applied to find the level of the annual consumption for the year 2016 with the same costs equal to the new case zero consumption. The results are different not only among the regions but also among the products inside the region (Table 6). The best (cheapest) supplier is ST Energy. With this supplier and its one product the annual consumption for this year could be more than 3.6 MWh in any distribution area to have the annual cost equal to the situation with zero consumption in 2017 (under new conditions). The "worst" product is E.ON Energie Elektrina AKU (here E.ON is in the role of supplier) but also this product guarantee the equal costs to zero consumption in 2017 for the consumption higher than 3.15 MWh annually. The average, median and mode of all 57 optimization models results (consumptions in MWh) with corresponding costs in each region shows that the most expensive region PRE offers the possibility of high consumption for this year that has the same costs as in the new model with zero consumption. This consumption is not so small to be able to say that it does not influence the possible increase of costs in case new conditions would be valid.

	E.ON	**PRE**	**CEZ**
min MWh	3.15	3.5	3.19
Product	E.ON Energie Elektrina AKU	E.ON Energie Elektrina AKU	E.ON Energie Elektrina AKU
annual costs	12346.43	13564.,5	12987.51
max MWh	3.68	4.1	3.71
Product	ST Energy Standard AKU	ST Energy Standard AKU	ST Energy Standard AKU
annual costs	11823.71	13041.78	12464.79
average MWh	3.4554	3.8568	3.4977
average costs	12101.53	13357.78	12757
median MWh	3.48	3.88	3.51
median costs	12201.53	13419.3	12842.31
mode MWh	3.58	3.98	3.61
mode costs	12201.53	13419.3	13016.55

Table 6 Results of the optimization models: MWh and costs for 2016 from 57 products with the costs equal to zero consumption in 2017

5 Conclusion

The differences among the regions and among the suppliers according to the annual electricity costs of households (for the tariff rate D25d) exist not only in the conditions of up to date electricity prices but also in the conditions of the new suggested ERU plan. According to the results of all optimization models it is possible to say that the annual electricity consumption of the household must be higher than 6.21 MWh (CEZ, E.ON regions) or 7 MWh (PRE) in 2017 to have the annual costs lower or equal to the ones in 2016. The biggest differences between the old and new situation can be seen on the hypothetical example of zero consumption where the annual costs in 2017 would be 6 times higher than in 2016. The consumption this year could rise to the level around 3-4 MWh with nearly the same costs that are equal to the zero consumption costs in 2017. These results show that the new model is not good (it is more expensive) for the households with lower consumption. During the time of the preparation of this article ERU concluded (according to the call-up of a lot of dissatisfied consumers and also of the government) not to use the new model for 2017 and nowadays they prepare another one that would be much better in sense of not discriminating the low level consumers.

Acknowledgements

This work was supported by grant No. F4/54/2015 of the Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] ACER/CEER: Annual Report on the Results of Monitoring the Internal Electricity and Natural Gas Markets in 2013, [online], available at: http://www.europarl.europa.eu/meetdocs/2014_2019/documents/itre/dv/acer_market_monitoring_report_2014_/acer_market_monitoring_report_2014_en.pdf [cit. 2016-04-20].

[2] Boltz, W. The Challenges of Electricity Market Regulation in the European Union. *Evolution of Global Electricity Markets*. Elsevier Inc. 2013, 199-224.

[3] Energostat – *Elektroenergetika v ČR-Elektřina*, [online], available at: http://energostat.cz/elektrina.html [cit. 2016-02-15].

[4] ERU: *Energy Regulatory Office – Annual report of Operation ES*, ERU 2015[online], available at: https://www.eru.cz/documents/10540/462820/Rocni_zprava_provoz_ES_2014.pdf/933fc41a-ad79-4282-8d0f-01eb25a63812 [cit. 2016-04-20].

[5] ERU: *Energy Regulatory Office - Price calculator*, ERU 2015, [online], available: http://kalkulator.eru.cz/VstupniUdaje.aspx [cit. 2016-02-28].

[6] ERU: *Energy Regulatory Office – New calculator*, ERU 2016, [online], available: http://www.novatarifnistruktura.cz/nn.php?typ=doma , [cit. 2016-02-28].

[7] Kuncová, M.: Methods for the Electricity Supplier Selection – Case Study of the Czech Republic. *International Journal of Mathematical Models and Methods in Applied Sciences* **9** (2015), 714-720.

[8] Kuncová, M., and Sekničková, J.: Analysis of the efficiency of the electricity supplier selection depending upon the price changes. *Proceedings of the 34th Mathematical Methods in Economics Conference*. Palacký University in Olomouc, 2014, 542–547.

[9] Kuncová, M., and Sekničková, J.: Optimization Models Used for the Electricity Supplier Selection, *Proceedings of the 33rd Mathematical Methods in Economics 2015 Conference.* West-Bohemia University, Cheb, 2015, 443–448.

[10] Newbery, D.: Evolution of the British Electricity Market and the Role of Policy for the Low-Carbon Future. Evolution of Global Electricity Markets. Elsevier Inc. 2013, 3-30.

[11] Ventosa, M., Baíllo, Á., Ramos, A. and Rivier, M.: Electricity Markets modeling trends, *Energy Policy* **33** (2005), 897–913.

[12] Wilson, Ch., M., Waddams Price C.: Do consumers switch to the best supplier? *Oxford Economic Papers* **62/4** (2010), 647-668.

Simulation framework for testing Piketty's predictions

Michal Kvasnička[1]

Abstract. It is well known that Piketty's claim that inequality rises whenever interest rates are higher than economic growth does not hold true in general. However, two of his conjectures are worth of testing: 1) that inequality can grow by a simple capital accumulation, i.e. the rich can get even richer both in absolute and in relative terms just by saving a part of their interest incomes, and 2) that slow economic growth can raise inequality. This paper develops a general simulation framework for testing which assumptions are necessary to get these outcomes. It also tests the conjectures in a simulated heterogeneous Solow model. The simulation shows that inequality does not grow by capital accumulation when there is no economic growth in the model.

Keywords: inequality, economic growth, capital accumulation, simulation.

JEL classification: E21, E27, D63, O11

1 Introduction

In his seminal book [6], Piketty claims that inequality rises whenever interest rates r are higher than economic growth g. However, there are many problems with this theory. First, it ignores that inequality may be caused by other reasons than capital accumulation through savings. For instance, inequality may be raised or lowered by changes of technology, outsourcing abroad, successful entrepreneurs' innovations, etc. The theory also ignores that there are forces that alleviate inequality caused by capital accumulation, e.g. social mobility. Second, the theory completely ignores the role of institutions, see [1]. Third, the theory in its formal setting ignores the fact that both interest rates and the rate of economic grows are endogenous variables (i.e. it is not the ultimate explanation) and that the inequality depends on many exogenous forces as rate of saving etc., see on-line appendix to [1]. For technical criticism of Piketty's claims, see papers quoted in [1]; for a more philosophical criticism see e.g. [5].

However, even though Piketty's theory is incomplete and its major claim that the inequality rises whenever $r > g$ does not hold true in general, it has two important predictions that may still hold true: 1) Inequality can grow by a simple capital accumulation, i.e. the rich can get even richer both in absolute and in relative terms just by saving a part of their interest incomes and the differences in capital holding can explode. 2) Slow economic growth can raise inequality. Even though these two predictions may not hold true under all possible conditions, there may still exist some conditions under which they hold true. Piketty's own model, however, is insufficient to provide these conditions and to test their consequences. The goal of this paper is to provide a general simulation framework that allows testing which model specifications produce results compatible with Piketty's (generalized) predictions, and which do not. In the next section I describe the general framework. In the third section I describe, simulate, and analyze one possible model specification (heterogeneous Solow model) that shows how the framework can be used for testing Piketty's predictions given a particular model assumptions. The simulation results and suggestions for future research are summarized in the last section.

2 General simulation framework for testing Piketty's predictions

In this section, I propose a general framework for testing Piketty's predictions. The framework is based on methodology of agent-based computational economics (also called multi-agent models), see [9]. In

[1]Masaryk University, School of Economics, Department of Economics, Lipová 41a, Brno, 602 00, the Czech Republic, michal.kvasnicka@econ.muni.cz

these models, behavior of individual agents is specified and their interactions are simulated. This way, one can try to "grow the expected phenomena from the bottom up", i.e. test whether a particular set of assumptions leads to the expected outcomes. There are several reasons why multi-agent simulation is a more convenient approach to testing Piketty's theory than standard macroeconomic models. First, it allows populating the model with completely heterogeneous agents, i.e. there is no need to limit the agents to few classes of representative or homogeneous agents (e.g. worker, capitalists, and the like). Every agent can be completely different from others in these models. Second, since the models are simulated, no restrictions on functional form of the models have to be imposed to ensure their solvability. Moreover, the models are easily scalable and extensible. One can start with an extremely simple model (as is the model proposed in the next section) and later extend it easily. Third, it is very easy to implement bounded rationality, learning, and local interactions among agents if necessary. Disadvantage of this approach is that the models are not analytically tractable. This, however, is the case with almost all modern macroeconomic models. Moreover, properties of multi-agent models can be easily analyzed by standard statistical procedures—both on macro-scale (where the behavior of the model as a whole is analyzed) and on micro-scale (where the behavior of individual agents is analyzed). There are already some agent-based macro-models that inquire impacts of various policies on inequality, see e.g. [2, 4, 10]. However, as far as I know, there is no general framework or a particular model developed to test Piketty's theory. Closest comes [3] but even this paper is not specifically targeted to test Piketty's conjectures.

The proposed framework consists 1) of the description of the minimal set of heterogeneous agents and their properties and 2) of general recipe for simulating the model. The framework includes three types of agents: 1) households that provide factors of production, consume, save, and possibly learn, reproduce, and die, 2) firms that acquire factors of productions and produce and distribute goods, and 3) government that redistributes the households' incomes. There is no need for money in the model—everything can be calculated in terms of a selected numéraire. The model is first initialized and then simulated in discrete steps until a stop condition is met.

In the model initialization stage,

- initial households are created and endowed with initial amount of physical and human capital, initial labor supply functions and initial consumption / saving functions and
- initial firms are created and endowed with the initial production functions.

In each step of simulation,

1. households supply labor and capital to firms,
2. factors of production (labor, physical and human capital) are allocated among firms,
3. firms produce consumption goods and distribute it among households,
4. the government redistributes households' incomes,
5. capital depreciates,
6. households consume part of their net incomes and previous capital and save the rest, and this way they update their stocks of capital, and
7. firms update their production functions, may go bankrupt or be created; households update their labor supply functions and consumption / saving functions, may be born or die.

Various characteristics of the model's state (such as inequality, economic growth, interest rates, etc.) can be measured either in every simulation step (if one cares how the inequality changes over time) or in the steady state of the model (if one cares about equilibrium values and if the model converges to a steady state).

To implement the model, several things have to be specified:

- initial distribution of capital among households,
- each household's labor supply function,

- each household's consumption / saving function,
- each household's human capital adaptation over time,
- capital depreciation function,
- each firm's production function and its changes over time,
- process of allocation of factor of production among firms,
- distribution of firms' product,
- process of closing firms in loss and creating new firms,
- process of creating new households, eliminating the old ones, and distributing the inheritance,
- redistribution, and
- stop condition of the model.

3 Example: heterogeneous Solow model

In this section I describe one particular example of the model specification. It has been chosen such that its assumptions are favorable to Piketty's theory, it is simple, and yet is still compatible with standard macroeconomic assumptions. In fact, the model presented here is a simple extension of Solow model [8]. The only difference between the standard Solow model and this model is that many heterogeneous agents are explicitly assumed: each household has a different saving rate and different initial stock of capital.

3.1 Model setup

The model specification is following: I assume there are N heterogeneous households. Each household lives forever, i.e. the wealth accumulation is not diluted by inheritance. It supplies $l_{it} = 1$ unit of labor at time t, i.e. "capitalists" do not work less than "workers" and do not earn lower wages than "workers", and hence the income inequality is not ameliorated this way. There is no human capital, i.e. all inequality is due to accumulated physical capital. Household i's initial physical capital k_{i0} is non-negative. The total amount of the initial capital is the same in all simulations, $\sum k_{i,0} = N$. Household i's income is $y_{it} = w_t l_{it} + r_t k_{it}$ where w_t is wage rate and r_t is interest rate at time t. Household i's saving at time t is $s_{it} = \sigma_i y_{it}$ where σ_i is household i's marginal propensity to save (constant over time). Each household draws its σ_i independently from uniform distribution $U(\underline{\sigma}, \bar{\sigma})$ where $\underline{\sigma}$ is the lower range and $\bar{\sigma}$ is the upper range of the support. Household i's consumption at time t is then $c_{it} = y_{it} - s_{it}$. Thus at time $t+1$, household i will hold the amount of capital $k_{i,t+1} = (1-\delta)k_{it} + s_{it}$ where δ is depreciation rate (the same for all households and all times).

There is only one competitive firm representing the whole economy that produces a homogeneous product that is then distributed among the households in form of wages and interest payments. The firm hires all labor and capital supplied and produces the aggregate product Y_t with Cobb-Douglas production function $Y_t = A_t L_t^{\alpha} K_t^{1-\alpha}$ where $L_t = \sum l_{it}$, $K_t = \sum k_{it}$, $\alpha \in (0,1)$ is the labor share of the product, and $A_t > 0$ is the state of technology at time t. The technology improves over time such that $A_{t+1} = (1+\gamma)A_t$ where $\gamma \geq 0$. Since the firm behaves as a competitive firm, the wage rate $w_t = \alpha A_t (K_t/L_t)^{(1-\alpha)}$ and interest rate $r_t = (1-\alpha)A_t(L_t/K_t)^{\alpha}$ are at time t equal to marginal products of labor and capital, respectively, and the firm has no profit.

There is no government, and hence no redistribution that could have mitigated the inequality. The inequality is measured by Gini coefficient of income, capital (i.e. wealth), and consumption. The simulation stops when the absolute change of every measured Gini coefficients between the two following steps is lower than 10^{-7}.

For the simulation in this paper I use the following parameters. The number of households is $N =$ 1 000. The initial level of technology is $A_0 = 1$. The labor share is $\alpha = 0.7$. There is no economic growth, i.e. $\gamma = 0$. Households' marginal propensities to save are drawn independently from uniform distribution with support $[0, 0.4]$ The depreciation rate $\delta = 0.04$. There are four treatments that govern

treatment	equal	uniform	power law	one rich	all
initial capital		0.0003	0.002	0.00002	0.00003
		(0.003)	(0.002)	(0.0001)	(0.0001)
saving rate	74.168***	74.168***	74.168***	74.168***	74.168***
	(0.019)	(0.019)	(0.019)	(0.019)	(0.009)
intercept	−3.031***	−3.031***	−3.032***	−3.031***	−3.031***
	(0.004)	(0.005)	(0.004)	(0.004)	(0.002)
observations	1,000,000	1,000,000	1,000,000	1,000,000	4,000,000
R^2	0.962	0.962	0.962	0.962	0.962
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 1 Regression of the stock of capital held by an individual household in steady-state. Robust standard errors are reported in parentheses.

the initial distribution of physical capital k_{i0}: it can be distributed among households 1) equally, i.e. each household gets 1 unit of capital, 2) uniformly, i.e. each household draws its initial capital randomly from uniform distribution $U(0, 2)$, and then the total amount of capital is re-normalized to N, 3) by power law, i.e. each household draws its initial capital randomly from distribution created by preferential attachment[1], or 4) extremely unequally where one randomly chosen household gets all N units of capital and other households get no capital at all. The model has been simulated 1 000 times for each treatment, i.e. the total number of simulations is 4 000. Each four simulations, one for every treatment, start with the same random seed which secures that each one of the four treatments starts with the same state of the model, i.e. with the same values of individual households' marginal rate of saving.[2] The model was implemented and simulated in NetLogo 5.3.1 [11] and analyzed in R 3.2.5 [7].

3.2 Results

The simulation shows that Piketty's original conjecture that inequality rises whenever $r > g$ is incorrect within the present model specification. All three inequality measures converge to steady-state values even though in the model's steady state there is no economic growth and the interest rate is strictly positive (it lies in range from 5.15% to 5.62% with mean value of about 5.3%), i.e. $r > g$ indeed.

There is also no tendency for the initially rich households to get richer in relative terms over time. On the contrary, there is even no persistence in capital holding, and the households' steady-state stocks of capital are independent from stocks of capital they held initially. This can be shown in three ways. First, the steady-state distributions of all measured Gini coefficients from the simulations are independent of the way the initial capital was distributed among the households. Kruskal-Wallis rank sum tests' χ^2 statistics are 0.0004, 0.0004, and 0.0001 for Gini coefficients for capital holding, income, and consumption respectively, all with p-values indistinguishable from 1. Second, the capital held in the steady state by an individual household is independent from the stock of capital the household held initially. Table 1 shows that the steady state individual stock of capital can be almost fully explained in terms of the individual marginal rate of saving while the parameter for the initial stock of capital is negligible and statistically insignificant. Third, given the initial values of individual households' marginal rates of saving, the resulting Gini coefficients converge to the same level despite the differences in the initial distribution of capital among the households; for example see Figure 1.

[1] The algorithm for preferential attachment used here works in the following way. Initially, each household is given 0.01 units of capital. Then one household is selected randomly and is given another 0.01 units of capital—each household is selected with probability proportional to the amount of capital the household already has. This step is repeated until all households have together N units of capital.

[2] This is possible because the only stochastic part of the model is its initialization; the simulation steps are fully deterministic. In the initialization, all households are first assigned their marginal rates of saving σ_i, and only then are given their initial stocks of capital. Thus the same random seed secures the same values of all households' marginal rate of saving. Households' initial stocks of capital are different in each treatment because different algorithms are used to assign them.

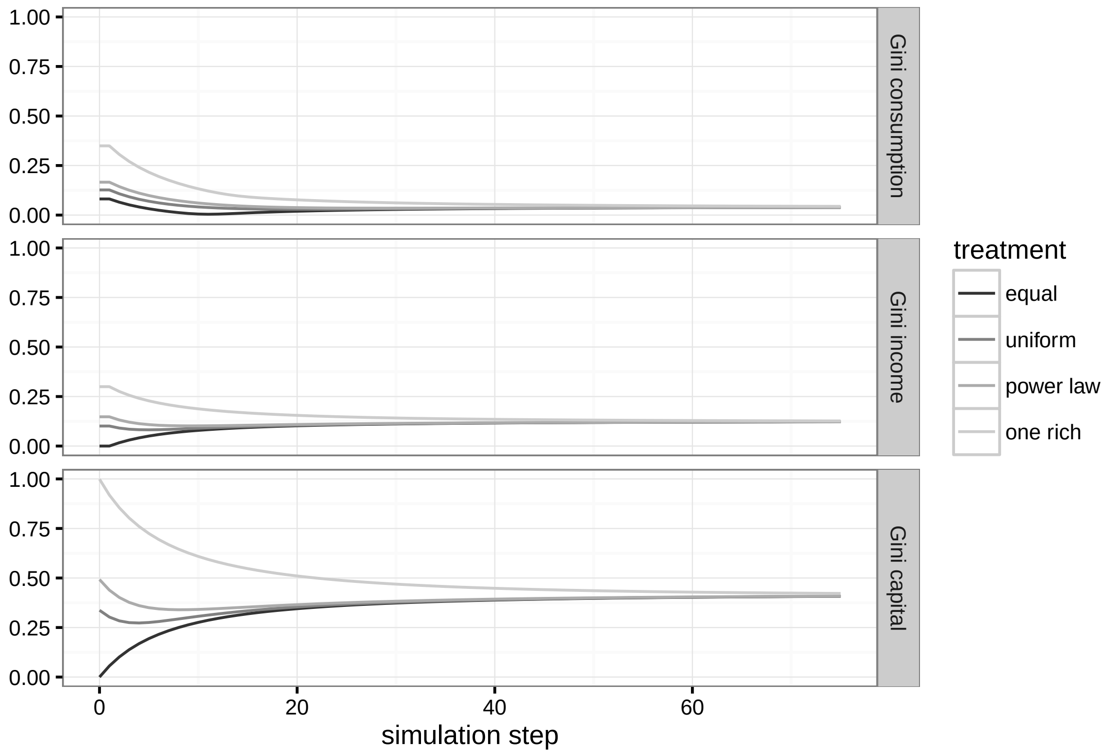


Figure 1 Example of convergence of Gini coefficients over time for one particular set of individual households' marginal rates of saving. Only first seventy steps are shown; the lines are indistinguishable later.

The simulation allows us to study causality in the model too. It is often assumed that richer households save a higher part of their incomes. It can be seen within the model too: Table 1 shows a positive correlation between the households' marginal rates of saving and their steady state stocks of capital. However, the causation is reverse here: the households that save more are richer in the model, and not vice versa—the households' marginal rate of saving are exogenous in the model.

The simulation also shows that Piketty's preoccupation with the inequality in capital holding (which is the model proxy for wealth) may be misleading. Figure 2 shows the distribution of steady-state Gini coefficients for capital holding, incomes, and consumption. The wealth inequality is much higher than income inequality which is higher than consumption inequality. If one cares about equality of households' standards of living, the inequality of capital holding is a fairly poor measure of it. Moreover, if one cares about wage-earners' incomes, he should encourage savings rather than lower it with taxation. The reason is straightforward: equilibrium wages $w_t = \alpha A_t(K_t/L_t)^{(1-\alpha)}$ clearly rise with the stock of available capital K_t in the model.

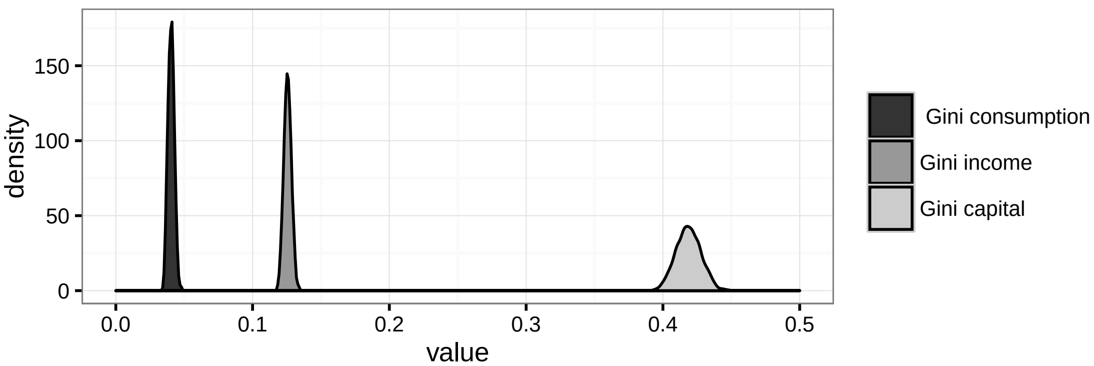


Figure 2 Densities of steady-state Gini coefficients of consumption, income, and capital stock from all treatments. All treatments are pooled because their distributions are indistinguishable from each other.

4 Conclusions

The simulation shows that interest rates higher than economic growth is not a sufficient condition for growing inequality. Moreover, the initial distribution of capital among the households has no impact on the steady-state distribution of capital, i.e. there is no persistence in capital holding within the simple heterogeneous Solow model presented above. (Testing impact of economic growth and changes in the support of the distribution of the marginal rate of saving is beyond the scope of this paper.) However, there still may be some reasonable assumptions that would yield outcomes compatible with Piketty's (generalized) predictions. The supporters of the theory should provide the list of these conditions. The general simulation framework proposed above can be then used as an easy way to verify true results of these assumptions.

Acknowledgements

This paper has been created as a part of grant project of Technology agency of the Czech Republic TAČR Omega TD03000121.

Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the program "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005), is greatly appreciated.

References

[1] Acemoglu, D. and Robinson, J. A.: The rise and decline of general laws of capitalism. National Bureau of Economic Research, No. w20766, 2014.

[2] Berman, Y., Shapira, Y., and Ben-Jacob, E.: Middling the origin and possible controls of the wealth inequality surge. *P*LoS ONE **10** (2015).

[3] Biondi, Y. and Righi, S.: Inequality and the financial accumulation process: A computational economic analysis of income and wealth dynamics. Working paper, 2015. Available at SSRN 2557883 (2015).

[4] Dosi, G., Fagiolo, G., Napoletano, M., and Roventini, A.: Income distribution, credit and fiscal policies in an agent-based Keynesian model. *Journal of Economic Dynamics & Control*, **37** (2013), 1598–1625.

[5] McCloskey, Deirdre N.: Measured, unmeasured, mismeasured, and unjustified pessimism: a review essay of Thomas Pikettys Capital in the twenty-first century. *Erasmus Journal for Philosophy and Economics* **7** (2014), 73–115.

[6] Piketty, T.: *Capital in the Twenty-First Century.* Harvard University Press, 2014.

[7] R Core Team: *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2014. http://www.R-project.org/.

[8] Solow, R. M.: A contribution to the theory of economic growth. *The quarterly journal of economics* (1956), 65–94.

[9] Tesfatsion, L., 2006. Agent-based computational economics: a constructive approach to economic theory. In: *Handbook of Computational Economics* (Tesfatsion, J., and Judd, K. L., eds.), vol. 2, 2006, Elsevier, 831–880.

[10] Varga, G. and Vincze, J.: Ants and crickets: arbitrary saving rates in an agent-based model with infinitely lived-agents. No. 1504. Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences, 2015.

[11] Wilensky, U.: NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, 1999. http://ccl.northwestern.edu/netlogo/.

Lexicographic Fair Design of Robust Emergency Service System

Marek Kvet [1], Jaroslav Janáček[2]

Abstract. Emergency system design locates limited number of service centers at a given set of possible locations to satisfy system users' demands for service. Users of such system claim equal or fair access to the provided service. The fairness emerges whenever limited resources are to be fairly distributed among participants. Plethora of fairness schemes were studied, but the strongest one is so called lexicographic min-max criterion. We study here the case of the lexicographic optimal design of the emergency system. A robust service system design is usually performed so that the design complies with specified scenarios by minimizing the maximal objective function of the individual instances corresponding with particular scenarios. The min-max link-up constraints represent an undesirable burden in any integer programming problem due to bad convergence of the branch-and-bound method. Within this paper, we try to overcome the drawback following from the link-up constraints by usage of the homogenous radial formulation of the problem, which can considerably accelerate the associated solving process especially when combined with bisection process.

Keywords: emergency service system design, radial approach, robust lexicographic fair design.

JEL Classification: C61
AMS Classification: 90C06, 90C10, 90C27

1 Introduction

As most of emergency service systems design work in real road networks, which are exposed to various random events following weather or traffic, the system designers must comply with system resistance to such critical events [3], [11]. Most of the approaches to making a system more resistant are based on making its design resistant to possible failure scenarios, which can appear in the road network as a consequence of random failures due to congestion, disruptions or blockages. An individual scenario is characterized by particular time distances between users' locations and possible service center locations. A robust service system design has to comply with all the specified scenarios. The usual way of taking into account all scenarios is based on minimizing the maximal objective function of the individual instances corresponding with the particular scenarios.

In contrast to the private systems, users of an emergency public service system claim fair access to the provided service. Generally, the fairness emerges whenever limited resources are to be fairly distributed among participants [2], [8], [9]. Many fairness schemes were studied, but the strongest one applicable in the public service system design is so called lexicographic min-max criterion [10]. If the lexicographical min-max fair problem is solved, then the disutility perceived by the worst situated users from the nearest located center is minimized first, and then the disutility of the second worst situated users is minimized unless the minimal reached disutility perceived by the previously processed users gets worsened. This process is repeated until no users' disutility following from the nearest located center can be reduced.

Within this paper, we focus on the fair emergency service system design, which is robust considering given finite set of scenarios. The basic design problem is formulated as the lexicographic min-max emergency service system.

[1] Ing. Marek Kvet, PhD., University of Žilina, University Science Park, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, marek.kvet@uvp.uniza.sk
[2] Prof. RNDr. Jaroslav Janáček, CSc., University of Žilina, Faculty of Management Science and Informatics, Univerzitná 8215/1, 010 26 Žilina, Slovak Republic, e-mail: jaroslav.janacek@fri.uniza.sk

Complexity of the underlying location problems led to searching for a suitable algorithm. It was found that the radial formulation of the problem can considerably accelerate the associated solving process [1], [4], [5]. As the approximate approach used for the system optimal public service system design proved to be a suitable tool for this purpose, we decided to apply the radial formulation with homogenous system of radii [6], [7] also on the robust emergency system design.

The remainder of the paper is organized as follows: Section 2 is devoted to explanation of the Ogryczak´s approach to lexicographic min-max optimization including the radial formulation of the service system design problem. The reformulation of suggested method to easily solvable form and its enhancement to multi scenario case is described in Section 3 and the associated numerical experiments are performed in Section 4. The results and findings are summarized in Section 5.

2 Lexicographic min-max optimization and radial formulation

The lexicographic min-max emergency service system design problem is defined on a transportation network, where two subsets of nodes are distinguished. There are specified set I of possible service center locations and set J of users' locations, where b_j denotes the number of users sharing the location j. The disutility perceived by a system user located at j is proportional to the distance of the user location from the nearest located service center. The disutility for a user located at j is denoted by d_{ij}, when the service providing center is located at i. The values of d_{ij} may be proportional to the network distances between the users' location j and the center location i. Especially, the disutility perceived by a user of emergency system may correspond with estimated travelling time in minutes. The objective of the emergency service system design problem is to place p service centers at the possible locations from the set I so that the disutility perceived by the worst situated users is minimized first, and then the disutility perceived by the second worst situated users is minimized unless the minimal reached disutility of the previously processed users gets worsened. This process is repeated until no users' disutility can be reduced.

As disutility perceived by an individual user corresponds to the distance from the user location to the nearest located service center, thus the disutility value must equal to some value from the upper subscripted sequence $d^0<d^1<\ldots<d^{w+1}$ of all possible w+2 disutility values, which occur in the matrix $\{d_{ij}\}$, $i\in I$, $j\in J$. Let d^w denote the highest but one member of the sequence. Then, the range of all disutility values can be represented by a finite set of ordered values $G_0=d^{w+1}$, $G_1=d^w$ … $G_w=d^1$ and $G_{w+1}=d^0$. To model the decisions on individual center locations, we introduce a zero-one variable $y_i\in\{0,1\}$ for each possible center location from I. Then the vector $\boldsymbol{y}$ of y_i with m components where m corresponds to the cardinality of I, corresponds to a feasible design if at most p components equal to one. Quality of the solution $\boldsymbol{y}$ can be characterized by distribution vector $[B_0(\boldsymbol{y}), B_1(\boldsymbol{y}) \ldots B_w(\boldsymbol{y})]$, where the t-th component of the vector is defined as the number of users, whose disutility belongs to the right-closed interval $(G_{t+1}, G_t]$. According to [10], the lexicographic min-max problem consists in lexicographic minimizing of the vector $[B_0(\boldsymbol{y}), B_1(\boldsymbol{y}) \ldots B_w(\boldsymbol{y})]$ subject to $\boldsymbol{y}\in\{0, 1\}^m$ and the condition that the vector $\boldsymbol{y}$ contains at most p ones.

The following lexicographical minimization of the distribution vector processes the components $B_0(\boldsymbol{y})$, $B_1(\boldsymbol{y})$ … $B_w(\boldsymbol{y})$ step-by-step. In accordance with [10], we suggested an iterative process of lexicographic minimization based on solving the problem (1) – (7) for the components $B_t(\boldsymbol{y})$, where t=0 … w. Additionally to the zero-one location variables $y_i\in\{0, 1\}$ for $i\in I$, we introduce the variables x_{js}, to indicate, whether the disutility of the users located at $j\in J$ following from the nearest located center is greater than d^s. In such case, the variable takes the value of 1, and it takes the value of 0 otherwise. To complete the model, we introduce a zero-one constant a_{ij}^s for each triple $[i, j, s]$, where $i\in I$, $j\in J$, $s\in[0..w]$. The constant a_{ij}^s is equal to 1, if the disutility d_{ij} between the users´ location j and the possible center location i is less than or equal to d^s, otherwise a_{ij}^s is equal to 0. In the associated model, the symbol e_s denotes the difference $d^{s+1} - d^s$ and the value $\underline{B}^*_k$ corresponds to the objective function value of (1) obtained in the steps preceding the step t. Then the radial-type min-max public service system design problem for the t-th step can be formulated as follows:

$$Minimize \quad \sum_{j\in J} b_j \sum_{s=w-t}^{w} e_s x_{js} \tag{1}$$

$$Subject\ to: \quad x_{js} + \sum_{i\in I} a_{ij}^s y_i \ge 1 \quad for\ j\in J,\ \ s=0,1,\ldots,w \tag{2}$$

$$\sum_{i\in I} y_i \le p \tag{3}$$

$$\sum_{j\in J} b_j \sum_{s=w-k}^{w} e_s x_{js} \le \underline{B}_k^* \quad for\ k=0,\ldots t-1 \tag{4}$$

$$\sum_{i\in I} a_{ij}^{w+1} y_i \geq 1 \quad for\ j \in J \tag{5}$$

$$x_{js} \geq 0 \quad for\ j \in J,\ \ s = 0, 1, \ldots, w \tag{6}$$

$$y_i \in \{0,1\} \quad for\, i \in I \tag{7}$$

In this model, the constraints (2) ensure that the variables x_{js} are allowed to take the value 0, if there is at least one service center located in the disutility d^s radius from the users' location j. The constraint (3) puts the limit p on the number of located service centers. The constraints (4) and (5) prevent the components $B_0(\boldsymbol{y})$ ... $B_{t-1}(\boldsymbol{y})$ from worsening. As concerns the obligatory constraints (6), only values zero and one are expected in a feasible solution, but it can be seen that the model has integrality property regarding the variables x_{js}. It can be noticed that in the optimization process all relevant values of x_{js} are "pushed down" and the constraints (2) and (6) bound the variable x_{js} from below by value of one or zero. It follows that the relevant values of x_{js} stay at one of these values.

3 Robust design of the lexicographic min-max system

A robust service system design has to comply with all given scenarios. The usual way of taking into account all scenarios is based on minimizing the maximal objective function of the individual instances corresponding with the particular scenarios. Let symbol U denote the set of possible failure scenarios. Disutility following from the distance between locations i and j under a specific scenario $u\in U$ is denoted as d_{iju}. Similarly as above, variable $y_i\in\{0,1\}$ models the decision on service center location at the place $i\in I$.

To formulate the distribution vector $[B_0(\boldsymbol{y}),\ B_1(\boldsymbol{y})\ \ldots\ B_{\underline{w}}(\boldsymbol{y})]$ mentioned above, we consider the sequence $d^0<d^1<\ldots<d^{\underline{w}+1}$ of all possible $\underline{w}$+2 disutility values, which occur in the matrix $\{d_{iju}\}$, $i\in I$, $j\in J$ $u\in U$. Let $d^{\underline{w}}$ denote the highest but one member of the sequence. Then, the range of all disutility values can be represented by a finite set of ordered values $G_0=d^{\underline{w}+1}$, $G_1=d^{\underline{w}}$ … $G_{\underline{w}}=d^1$ and $G_{\underline{w}+1}=d^0$.

In contrast to the previously presented model describing the standard lexicographic min-max emergency system design, we introduce the variables x_{jsu} for the robust design, to indicate, whether the distance from the users' location $j\in J$ to the nearest located center under the scenario u is greater than d^s. In this case, the variable takes the value of 1, and it takes the value of 0 otherwise. To complete the enhanced model, which ensures the described property of variables x_{jsu}, we introduce a zero-one constant $a_{iju}{}^s$ under the scenario $u\in U$ for each triple $[i, j, s]$, where $i\in I$, $j\in J$, $s \in [0..\underline{w}]$. The constant $a_{iju}{}^s$ is equal to 1, if the disutility d_{iju} of the users located at j following from the possible center location i is less than or equal to d^s, otherwise $a_{iju}{}^s$ is equal to 0. In the associated model, the symbol e_s denotes the difference $d^{s+1} - d^s$ as above and the value $\underline{B}^*{}_k$ corresponds to the objective function value of (8) obtained in the steps preceding the step t. The radial-type lexicographical min-max robust emergency service system design problem for the t-th step can be formulated as follows:

$$Minimize\quad \sum_{u\in U}\ \sum_{j\in J} b_j \sum_{s=\underline{w}-t}^{w} e_s x_{jsu} \tag{8}$$

$$Subject\ to:\quad x_{jsu} + \sum_{i\in I} a_{iju}^{s} y_i \geq 1 \quad for\ j \in J,\ \ s = 0, 1, \ldots, \underline{w},\ u \in U \tag{9}$$

$$\sum_{i\in I} y_i \leq p \tag{10}$$

$$\sum_{u\in U}\ \sum_{j\in J} b_j \sum_{s=\underline{w}-k}^{w} e_s x_{jsu} \leq \underline{B}_k^* \quad for\ k = 0, \ldots t-1 \tag{11}$$

$$\sum_{i\in I} a_{iju}^{\underline{w}+1} y_i \geq 1 \quad for\ j \in J,\ u \in U \tag{12}$$

$$x_{jsu} \geq 0 \quad for\ j \in J,\ \ s = 0, 1, \ldots, \underline{w},\ u \in U \tag{13}$$

$$y_i \in \{0,1\} \quad for\, i \in I \tag{14}$$

In this model, the constraints (9) ensure that the variables x_{jsu} are allowed to take the value 0, if there is at least one service center located in the disutility d^s radius from the users' location j. The constraint (10) puts the limit p on the number of located service centers. The constraints (11) and (12) prevent the components $B_0(\boldsymbol{y})$, $B_1(\boldsymbol{y})$ ... $B_{t-1}(\boldsymbol{y})$ from worsening.

4 Numerical experiments

Within this section, we present the results of numerical experiments, which are aimed at ascertainment of the characteristics of the suggested robust lexicographic approach from the viewpoint of computational time and the resulting system design.

To compare studied lexicographic approach based on the radial formulation and its usage for basic and robust design of the emergency service system, we performed the series of numerical experiments. To solve the problems described in the previous sections, the optimization software FICO Xpress 7.9 (64-bit, release 2015) was used and the experiments were run on a PC equipped with the Intel® Core™ i7 5500U processor with the parameters: 2.4 GHz and 16 GB RAM.

The used benchmarks were derived from the real emergency health care system, which was originally implemented in eight regions of Slovak Republic. For each self-governing region, i.e. Bratislava (BA), Banská Bystrica (BB), Košice (KE), Nitra (NR), Prešov (PO), Trenčín (TN), Trnava (TT) and Žilina (ZA), all cities and villages with corresponding number of inhabitants b_j were taken. The coefficients b_j were rounded to hundreds. These sub-systems cover demands of all communities - towns and villages spread over the particular regions by given number of ambulance vehicles. In the benchmarks, the set of communities represents both the set J of users' locations and also the set I of possible center locations. The cardinalities of these sets vary from 87 to 664 according to the considered region. The number p of located centers was derived from the original design and it varies from 9 to 67. The network distance from a user to the nearest located center was taken as user´s disutility.

An individual experiment was organized so that the model (1) - (7) was used to obtain the basic design. This model was solved for the basic scenario, which corresponds to the usual situation on the transportation network. After that, the robust design using the model (8) - (14) was computed concerning all studied scenarios. Due to the lack of common benchmarks for study of robustness, the scenarios used in our computational study were created in the following way. We chose 25 percent of matrix rows so that these rows correspond to the biggest cities concerning the number of system users. Then we chose randomly from 5 to 15 rows and the associated disutility values in the individual rows were multiplied by the randomly chosen constant from the range 2, 3 and 4. This way, 20 different scenarios were generated for each self-governing region.

The first set of numerical experiments was performed to compare both basic and robust designs of the emergency service system applied on all studied scenarios. It must be noted, that the scenario 0 corresponds to the usual situation on the transportation network. For each scenario and each resulting vector $\boldsymbol{y}$ (basic and robust) several characteristics were computed. The obtained results for the self-governing region of Žilina are summarized in Table 1. The disutility obtained by the worst situated users is denoted by *minMax*. The columns *minSum* are used for the total disutility perceived by all system users. Particular value can be computed according to (15).

$$minSum = \sum_{j \in J} b_j \min\left\{d_{ij} : i \in I, y_i = 1\right\} \tag{15}$$

Finally, some comparison of the distribution vectors must be performed. The lexicographic ordering of the distribution vectors enables to decide on which of two different deployments is better from the point of fairness, but it does not enable to quantify the difference between them. That is why we introduce the following gauge of the min-max lexicographic fairness. First, we extend the distribution vector by the component $w+1$, which gives the number of users, whose distance from the nearest service center equals to $G_{w+1}=d^0$. After these preliminaries, the sum of the distribution vector components is equal to the number B of all users for any solution $\boldsymbol{y}$. The suggested gauge $E(\boldsymbol{B}(\boldsymbol{y}))$ of the extended vector $\boldsymbol{B}(\boldsymbol{y}) = [B_0(\boldsymbol{y}), B_1(\boldsymbol{y}), \ldots, B_w(\boldsymbol{y}), B_{w+1}(\boldsymbol{y})]$ is defined by (16).

$$E(\mathbf{B}(\mathbf{y})) = \log_B\left(\left(\sum_{t=0}^{w+1} B_t(\mathbf{y}) * (B)^{w+1-t}\right)^{\frac{1}{w+2}}\right) \tag{16}$$

Scenario	BASIC DESIGN			ROBUST DESIGN		
	minMax	*minSum*	*E(**B**(**y**))*	*minMax*	*minSum*	*E(**B**(**y**))*
0	14	29576	0.3668	15	39209	0.3839
1	14	29576	0.3668	15	39209	0.3839
2	14	29576	0.3668	15	39209	0.3839
3	21	30592	0.5768	15	39209	0.3839
4	33	35885	0.9260	15	39209	0.3839
5	14	29576	0.3668	15	39209	0.3839
6	22	31556	0.5916	15	39209	0.3839
7	16	31012	0.4404	15	39209	0.3839
8	14	29576	0.3668	15	39209	0.3839

9	26	33486	0.7098	15	39209	0.3839
10	16	31084	0.4405	15	39209	0.3839
11	15	30159	0.4068	15	39209	0.3839
12	22	31556	0.5916	15	39209	0.3839
13	14	30304	0.3843	15	39209	0.3839
14	30	31116	0.8435	15	39209	0.3839
15	26	31605	0.7098	15	39209	0.3839
16	22	31608	0.6028	15	39209	0.3839
17	28	33863	0.7738	15	39209	0.3839
18	14	29576	0.3668	15	39209	0.3839
19	26	31934	0.7098	15	39209	0.3839
20	14	29576	0.3668	15	39209	0.3839

Table 1 Comparison of the basic and robust lexicographic emergency system design applied on individual scenarios for the self-governing region of Žilina (315 possible service center locations and 32 centers to be located)

It can be observed that the robust design (vector of location variables y_i) applied on individual scenarios produces the same solution. It can be explained by the fact that the robust design takes into account all scenarios whereas the basic design concerns only the basic scenario, which corresponds to the standard situation.

The studied approaches to emergency service system design were tested also on other benchmarks corresponding with the self-governing regions of Slovakia. The obtained results are reported in Table 2 and Table 3. Here, both designs were applied on the basic scenario 0. The computational time in seconds is denoted by CT. It must be realized, that the robust approach performs much longer due to the set of scenarios, which makes the size of the solved problem significantly bigger. The basic and robust designs can be compared by the Hamming distance HD, which is defined as follows. Let $\boldsymbol{y^b}$ denote the vector of location variables for the basic design and let $\boldsymbol{y^r}$ denote the vector for the robust one. Then the Hamming distance HD takes the form of (17).

$$HD = \sum_{i \in \hat{I}} \left(y_i^r - y_i^b \right)^2 \tag{17}$$

	$\lvert I \rvert$	p	BASIC DESIGN				ROBUST DESIGN					
			CT [s]	*minMax*	*minSum*	*E(**B**(**y**))*	CT [s]	*minMax*	*minSum*	*E(**B**(**y**))*	**PoR**	**HD**
BA	87	9	0.4	14	35189	0.38	40.5	15	56090	0.43	13.05	8
BB	515	52	5.8	13	31209	0.39	7433.6	14	39058	0.41	5.07	62
KE	460	46	11.5	12	35421	0.39	3784.1	12	41942	0.41	6.21	58
NR	350	35	58.7	13	37370	0.53	2377.5	14	42404	0.56	4.30	54
PO	664	67	24.3	12	41155	0.46	1278.9	13	43594	0.49	6.78	92
TN	276	28	13.4	12	32120	0.45	1822.7	14	38341	0.51	12.11	28
TT	249	25	7.9	13	27469	0.50	2643.8	13	35879	0.53	5.85	36
ZA	315	32	3.3	14	29576	0.37	405.0	15	39209	0.38	4.67	34

Table 2 Results of numerical experiments for the self-governing regions of Slovakia

$\lvert I \rvert$	p	BASIC DESIGN				ROBUST DESIGN					
		CT [s]	*minMax*	*minSum*	*E(**B**(**y**))*	CT [s]	*minMax*	*minSum*	*E(**B**(**y**))*	**PoR**	**HD**
315	158	0.0	4	6115	0.33	1.3	4	8240	0.35	4.40	38
315	105	0.3	6	19339	0.36	27.2	6	20061	0.38	3.89	56
315	79	0.9	7	19979	0.31	569.6	8	25694	0.35	11.82	70
315	63	1.2	8	24083	0.33	1336.5	10	23738	0.36	11.55	52
315	32	3.3	14	29576	0.37	397.7	15	41301	0.39	6.90	46
315	21	4.6	16	41488	0.42	806.2	18	51920	0.46	10.28	30
315	16	10.7	20	55001	0.52	1432.5	21	63134	0.54	2.34	14

Table 3 Results of numerical experiments for the self-governing region of Žilina.

Hamming distance evaluates the structural difference between two designs in the sense that it informs of number of locations, in which the designs differ, but it does not refer to the quality of the designs. Therefore, we have compared the basic and robust design also from the viewpoint of other characteristics. Similarly to the price of fairness introduced and studied in [2], [7] to evaluate the loss of min-sum objective function value caused by application of measures for fairness improvement, we introduce here so-called *price of robustness PoR*. It expresses the difference between the evaluations of the distribution vectors for the basic and robust design. Let $E_{basic}(\boldsymbol{B}(\boldsymbol{y}))$ denote the value of (16) for the basic design and let $E_{robust}(\boldsymbol{B}(\boldsymbol{y}))$ be used for the robust design. The value of *PoR* is given in percentage and it can be computed according to (18).

$$PoR = 100 * \frac{E_{robust}\left(\mathbf{B}(y)\right) - E_{basic}\left(\mathbf{B}(y)\right)}{E_{basic}\left(\mathbf{B}(y)\right)} \tag{18}$$

5 Conclusions

This paper is focused on mastering large instances of the lexicographic emergency service system design problem using commercial IP-solver. The original Ogryczak´s approach has been adjusted here to comply with the set of studied scenarios, which are used to make the system resistant to various catastrophic events. The basic design problem is formulated as lexicographic min-max problem, which means that the accented objective is to minimize the disutility perceived by the worst situated users first and then iteratively minimize the disutility of other users under the condition that the previously obtained users´ disutility cannot get worsened. The basic design is computed for the original scenario, which corresponds to the usual situation in the network, i.e. situation, which does not correspond to any of the above-mentioned possible failure scenarios. To find the value paid for making the system resistant to catastrophic events, a new conception called the price of robustness is introduced. The robustness measure has similar meaning as the price of fairness commonly used to evaluate the loss of min-sum objective function value caused by application of measures for fairness improvement. It has been found that the price of robustness in real instances does not exceed 12 percent on the average. Usage of the radial formulation proved its usefulness especially in large instances, where the lexicographic approach takes into account too many structural constraints. Thus we can conclude that we have presented a useful tool for the robust lexicographic emergency service system design problem, which can be easily implemented using common commercial optimization software. The future research in this field may be aimed at finding relevant scenarios, which can significantly impact the performance of emergency service system. Since the robust design can significantly differ from the basic one, it would be useful to find such method, which allows to change only limited number of service center locations in comparison with the standard solution.

Acknowledgement

This work was supported by the research grants VEGA 1/0518/15 "Resilient rescue systems with uncertain accessibility of service", VEGA 1/0463/16 "Economically efficient charging infrastructure deployment for electric vehicles in smart cities and communities", APVV-15-0179 "Reliability of emergency systems on infrastructure with uncertain functionality of critical elements" and by the project University Science Park of the University of Žilina (ITMS: 26220220184) supported by the Research & Development Operational Program funded by the European Regional Development Fund.

References

[1] Avella, P., Sassano, A., Vasil'ev, I..: *Computational study of large scale p-median problems*. In: Mathematical Programming 109, pp. 89-114, 2007.

[2] Bertsimas, D., Farias, V. F., Trichakis, N.: *The Price of Fairness*. Oper. Res., 59, pp. 17-31, 2011.

[3] Correia, I., Saldanha da Gama, F.: Facility locations under uncertainty. *Location Science*, eds. Laporte, Nikel, Saldanha da Gama, pp. 177-203, 2015.

[4] García, S., Labbé, M., Marín, A.: *Solving large p-median problems with a radius formulation*. In: INFORMS Journal on Computing 23 (4), pp. 546-556, 2011.

[5] Janáček, J.: *Approximate Covering Models of Location Problems*. In: Lecture Notes in Management Science: Proceedings of the 1st International Conference ICAOR ´08, Vol. 1, Sept. 2008, Yerevan, Armenia, pp. 53-61, ISSN 2008-0050, 2008.

[6] Janáček, J., Kvet, M.: Emergency system design with temporarily failing centers. *SOR 15: Proceedings of the 13th International Symposium on Operational Research*, Ljubljana: Slovenian Society Informatika, Section for Operational Research, ISBN 978-961-6165-45-7, pp. 490-495, 2015.

[7] Kvet, M., Janáček, J.: *Price of fairness in public service system design*. Mathematical Methods in Economics 2014, Olomouc, Czech Republic, September 10-12, ISBN 978-80-244-4209-9, pp. 554-559, 2014.

[8] Marsh, M., Schilling, D.: Equity Measurement in Facility Location Analysis. *European Journal of Operational Research*, 74, pp. 1-17, 1994.

[9] Nash, J.: The Bargaining Problem. *Econometrica*, vol. 18, No. 2, pp. 155-162, 1950.

[10] Ogryczak, W., Sliwinski, T.: *On Direct Methods for Lexicographic Min-Max Optimization*. In: Gavrilova M. et al. (Eds.): ICCSA 2006, LNCS 3982, (pp. 802-811). Berlin: Heidelberg: Springer, 2006.

[11] Pan, Y., Du, Y., Wei, Z.: Reliable Facility System Design Subject to Edge Failures. *American Journal of Operations Research*, 4, pp. 164-172, 2014.

Discussion on implicit econometric models

Petr Lachout [1]

Abstract. Random processes with convenient properties are employed to model observed data, particularly, coming from economy and finance. We will focus our interest in random processes given implicitly as a solution of a functional equation. For example, random processes MA, AR, ARMA, ARCH, GARCH are belonging in this wide class. Their common feature can be expressed by requirement that stated random process together with incoming innovations must fulfill a functional equation.

Functional dependence is linear for MA, AR, ARMA. We consider a general functional dependence, but, existence of a forward and a backward equivalent rewritings of the given functional equation is required. We present a concept of solution construction giving uniqueness of assigned solution. The procedure works as a symbolic solution.

I present in my contribution the class of implicit models where forward and backward rewritings are possible. Three illustrative examples are included.

Keywords: Econometric models. ARMA process, implicit definition.

JEL classification: C22
AMS classification: 62M10; 91B70

1 Introduction

Linear models are based on properties of the Hilbert space L_2 and the calculus of polynomials in the backward shift operator, see e.g. [1], [2], [3]. We will consider more complex models written as a functional equation; see e.g. [4], [6], [7], [8], [9], [10]. This theoretical description and consequent properties possess relevant impact to practice. Particularly, modern analysis of financial and economic data is based on this setup.

Our realization is focused to models given as an implicit solution of a functional equation. We start treating of this setting in [5]. In the present paper, we construct a solution in pointwise sense. The construction is actually a numerical algorithm based on recurrent plugging in forward or backward formula. Three examples are added.

2 Description of the model

We consider a random real vector time-series of dimension $J \in \mathbb{N}$ explained by random real vector innovations of dimension $K \in \mathbb{N}$. Spaces of their values will be denoted $\mathcal{Y} = \mathbb{R}^J$, $\mathcal{Z} = \mathbb{R}^K$. More precisely, we consider a random real vector time-series $\mathsf{Y} = (\mathsf{Y}(t), t \in \mathbb{Z})$, where for each time $t \in \mathbb{Z}$ we have $\mathsf{Y}(t) = (\mathsf{Y}_1(t), \mathsf{Y}_2(t), \dots, \mathsf{Y}_J(t))^\top$; i.e. $\mathsf{Y}(t) \in \mathcal{Y}$. We suppose Y is tightly related to a series of innovations $\mathsf{Z} = (\mathsf{Z}(t), t \in \mathbb{Z})$, where for each time $t \in \mathbb{Z}$ we have $\mathsf{Z}(t) = (\mathsf{Z}_1(t), \mathsf{Z}_2(t), \dots, \mathsf{Z}_K(t))^\top$; i.e. $\mathsf{Z}(t) \in \mathcal{Z}$. The relation is implicit described by a functional equation

$$\forall\, t \in \mathbb{Z} \qquad \mathsf{f}(\mathsf{Y}(t), \mathsf{Y}(t-1), \dots, \mathsf{Y}(t-p); \mathsf{Z}(t), \mathsf{Z}(t-1), \dots, \mathsf{Z}(t-q)) = \mathbf{0} \in \mathbb{R}^r, \tag{1}$$

where $p \in \mathbb{N}_0$, $q \in \mathbb{N}_0$, $r \in \mathbb{N}$, $\mathsf{f} : \mathcal{Y}^p \times \mathcal{Z}^{q+1} \to \mathbb{R}^r$ is a measurable function.

[1] Dept. Probability Theory and Mathematical Statistics, Charles University in Praha, Faculty of Mathematics and Physics, Sokolovská 83, 186 75 Praha 8, The Czech Republic, Petr.Lachout@mff.cuni.cz

To abbreviate forthcoming formulas, we introduce a notation

$$\begin{aligned} \mathsf{Y}(t \circledcirc k) &= \mathsf{Y}(t), \mathsf{Y}(t-1), \ldots, \mathsf{Y}(t-k), \\ \mathsf{Z}(t \circledcirc k) &= \mathsf{Z}(t), \mathsf{Z}(t-1), \ldots, \mathsf{Z}(t-k). \end{aligned} \tag{2}$$

Hence, the relation (1) converts to an abbreviate form

$$\forall\, t \in \mathbb{Z} \qquad \mathsf{f}(\mathsf{Y}(t \circledcirc p); \mathsf{Z}(t \circledcirc q)) = \mathbf{0} \in \mathbb{R}^r. \tag{3}$$

3 Main results

3.1 Deterministic solution

At first, we discuss existence of a deterministic solution of

$$\forall\, t \in \mathbb{Z} \qquad \mathsf{f}(\mathsf{y}(t \circledcirc p); \mathsf{z}(t \circledcirc q)) = \mathbf{0} \in \mathbb{R}^r. \tag{4}$$

For construction of a solution, we require equivalent rewriting of (4). Forward solution needs (5) and backward solution is based on (8).

At first we specify notion of equivalency between two lists of statements.

Definition 1. Let two lists of statements $R = (R(t), t \in \mathbb{Z})$ and $\tilde{R} = (\tilde{R}(t), t \in \mathbb{Z})$ be given. We say R is equivalent to $\tilde{R}$ if and only if for all $t \in \mathbb{Z}$ we have $R(t)$ is true $\iff$ $\tilde{R}(t)$ is true.

Forward solution

To be able to construct a forward solution of (4), we need an equivalent rewriting of (4) into a form

$$\forall\, t \in \mathbb{Z} \qquad \mathsf{y}(t) = \mathsf{F}(\mathsf{y}((t-1) \circledcirc (p-1)); \mathsf{z}(t \circledcirc q)). \tag{5}$$

Let us call (5) a forward form of (4). Now, we will construct a forward solution of (4).

Fix $\mathsf{z} \in \mathcal{Z}^{\mathbb{Z}}$, $\tau \in \mathbb{Z}$ and $\xi \in \mathcal{Y}^p$. Determine initial values of a solution

$$\mathsf{y}(\tau) = \xi_1,\ \mathsf{y}(\tau+1) = \xi_2,\ \ldots,\ \mathsf{y}(\tau+p-1) = \xi_p. \tag{6}$$

Using recurrent plugging in (5), we prolong the sequence $\mathsf{y}(t)$ for all $t \in \mathbb{Z}$, $t \geq \tau + p$ by

$$\mathsf{y}(t) = \mathsf{F}(\mathsf{y}((t-1) \circledcirc (p-1)); \mathsf{z}(t \circledcirc q)). \tag{7}$$

Hence for given $\mathsf{z} \in \mathcal{Z}^{\mathbb{Z}}$, $\tau \in \mathbb{Z}$ and $\xi \in \mathcal{Y}^p$, a sequence $\mathsf{y}(t)$ is uniquely determined for all $t \in \mathbb{Z}$, $t \geq \tau$ such that (5) and, also, (4) are fulfilled for all $t \in \mathbb{Z}$, $t \geq \tau + p$.

Backward solution

To be able to construct a backward solution of (4), we need an equivalent rewriting of (4) into a form

$$\forall\, t \in \mathbb{Z} \qquad \mathsf{y}(t-p) = \mathsf{G}(\mathsf{y}(t \circledcirc (p-1)); \mathsf{z}(t \circledcirc q)). \tag{8}$$

Let us call (8) a backward form of (4). Now, we will construct a backward solution of (4).

Fix $\mathsf{z} \in \mathcal{Z}^{\mathbb{Z}}$, $\tau \in \mathbb{Z}$ and $\xi \in \mathcal{Y}^p$. Determine initial values of a solution

$$\mathsf{y}(\tau) = \xi_1,\ \mathsf{y}(\tau+1) = \xi_2,\ \ldots,\ \mathsf{y}(\tau+p-1) = \xi_p. \tag{9}$$

Using recurrent plugging in (8), we prolong the sequence $\mathsf{y}(t)$ for all $t \in \mathbb{Z}$, $t < \tau$ by

$$\mathsf{y}(t) = \mathsf{G}(\mathsf{y}((t+p) \circledcirc (p-1)); \mathsf{z}((t+p) \circledcirc q)). \tag{10}$$

Hence for given $\mathsf{z} \in \mathcal{Z}^{\mathbb{Z}}$, $\tau \in \mathbb{Z}$ and $\xi \in \mathcal{Y}^p$, a sequence $\mathsf{y}(t)$ is uniquely determined for all $t \in \mathbb{Z}$, $t \leq \tau + p - 1$ such that (8) and, also, (4) are fulfilled for all $t \in \mathbb{Z}$, $t < \tau + p$.

Both-side solution

To be able construct a solution of (4), we require existence of both equivalent rewritings (5), (8). Then, a construction goes like that. Using constructions of forward solution (7) and backward solution (10), we are receiving a sequence $\mathsf{y}(t)$ correctly determined for all $t \in \mathbb{Z}$.

Theorem 1. *Let (4) possess equivalent rewritings (5) and (8). For given $\mathsf{z} \in \mathcal{Z}^{\mathbb{Z}}$, $\tau \in \mathbb{Z}$ and $\xi \in \mathcal{Y}^p$, our construction is giving a sequence $\mathsf{y}(t)$ uniquely determined for all $t \in \mathbb{Z}$ and fulfilling (4).*

Proof. Evidently, constructed sequence is uniquely defined. It remains to check (4). Take $t \in \mathbb{Z}$ for that.

1. The case $t \geq \tau + p$.
   According to forward construction, we have
   $$\mathsf{y}(t) = \mathsf{F}(\mathsf{y}((t-1) \circledcirc (p-1)); \mathsf{z}(t \circledcirc q)).$$
   Since (5) is equivalent with (4), we have
   $$\mathsf{f}(\mathsf{y}(t \circledcirc p); \mathsf{z}(t \circledcirc q)) = \mathbf{0} \in \mathbb{R}^r.$$
2. The case $t < \tau + p$.
   According to backward construction, we have
   $$\mathsf{y}(t-p) = \mathsf{G}(\mathsf{y}(t \circledcirc (p-1)); \mathsf{z}(t \circledcirc q)).$$
   Since (8) is equivalent with (4), we have
   $$\mathsf{f}(\mathsf{y}(t \circledcirc p); \mathsf{z}(t \circledcirc q)) = \mathbf{0} \in \mathbb{R}^r.$$

We have verified constructed sequence is a solution of (4). □

Construction is describing all solutions.

Theorem 2. *Let (4) possess equivalent rewritings (5) and (8). Let $\mathsf{z} \in \mathcal{Z}^{\mathbb{Z}}$ be given. Then the space of all solutions of (4) is isomorphic with $\mathbb{R}^{pJ}$.*

Proof. The construction is saying a solution of (4) is uniquely determined by its values at $t = 0, 1, 2, \ldots, p-1$. Since initial values can be chosen arbitrarily, the space of all solutions of (4) is isomorphic with $\mathbb{R}^{pJ}$. □

3.2 Random solution

Starting with random initials, our construction is giving all solutions of (1).

Theorem 3. *Let (4) possess equivalent rewritings (5) and (8). Let a random sequence $\mathsf{Z} \in \mathcal{Z}^{\mathbb{Z}}$ be given. Then the space of all solutions of (1) is isomorphic with $\mathcal{L}(\mathbb{R}^{pJ})$, where $\mathcal{L}(\mathbb{R}^{pJ})$ denotes the set of all real random vectors of dimension pJ.*

Proof. Statement follows immediately Theorem 2 applied to random initials. □

Examples

Here we present some examples.

Example 1. Consider an example with $J = K = r = 1$, $p = 2$, $q = 1$ given by a linear equation connecting to an ARMA process:

$$\mathsf{y}(t) + \frac{2}{3}\mathsf{y}(t-1) - \frac{1}{8}\mathsf{y}(t-2) - \mathsf{z}(t) + \frac{1}{3}\mathsf{z}(t-1) \quad = \quad 0.$$

Forms of (5), (8) are straightforward:

$$\begin{aligned}\mathsf{y}(t) &= -\frac{2}{3}\mathsf{y}(t-1)+\frac{1}{8}\mathsf{y}(t-2)+\mathsf{z}(t)-\frac{1}{3}\mathsf{z}(t-1),\\ \mathsf{y}(t-2) &= 8\mathsf{y}(t)+\frac{16}{3}\mathsf{y}(t-1)-8\mathsf{z}(t)+\frac{8}{3}\mathsf{z}(t-1).\end{aligned}$$

Example 2. Consider an example with $J=K=r=2$, $p=2$, $q=0$ given by two equations related to an ARCH process:

$$\begin{aligned}\mathsf{y}_1(t)-\mathsf{y}_2(t)\,\mathsf{z}_1(t) &= 0,\\ \mathsf{y}_2(t)-\Gamma\left(\frac{2}{3}\mathsf{y}_2(t-1)^2+\frac{1}{2}\mathsf{y}_2(t-2)^2+\mathsf{z}_2(t)\right) &= 0,\end{aligned}$$

where $\Gamma(x)=\sqrt{x}$ if $x\geq 0$, $\Gamma(x)=-\sqrt{-x}$ if $x<0$.

Form of (5) is:

$$\begin{aligned}\mathsf{y}_1(t) &= \Gamma\left(\frac{2}{3}\mathsf{y}_2(t-1)^2+\frac{1}{2}\mathsf{y}_2(t-2)^2+\mathsf{z}_2(t)\right)\mathsf{z}_1(t),\\ \mathsf{y}_2(t) &= \Gamma\left(\frac{2}{3}\mathsf{y}_2(t-1)^2+\frac{1}{2}\mathsf{y}_2(t-2)^2+\mathsf{z}_2(t)\right).\end{aligned}$$

Form of (8) is:

$$\begin{aligned}\mathsf{y}_1(t-2) &= \left(2\mathsf{y}_2(t)^2-\mathsf{sign}(\mathsf{y}_2(t))\left(\frac{4}{3}\mathsf{y}_2(t-1)^2+2\mathsf{z}_2(t)\right)\right)^{\frac{1}{2}}\mathsf{z}(t-2),\\ \mathsf{y}_2(t-2) &= \left(2\mathsf{y}_2(t)^2-\mathsf{sign}(\mathsf{y}_2(t))\left(\frac{4}{3}\mathsf{y}_2(t-1)^2+2\mathsf{z}_2(t)\right)\right)^{\frac{1}{2}}.\end{aligned}$$

Example 3. Consider an example with $J=3$, $K=2$, $r=1$, $p=3$, $q=2$ given by

$$\left\|\begin{pmatrix}\mathsf{y}_1(t)^3\\ \mathsf{y}_2(t)^3\\ \mathsf{y}_3(t)^5\end{pmatrix}-\frac{2}{5}\begin{pmatrix}\mathsf{y}_1(t-1)^2\,\mathsf{z}_1(t)\\ \mathsf{y}_2(t-1)^{-3}\,\mathsf{y}_2(t-2)^2\,\mathsf{z}_2(t)\\ \mathsf{y}_3(t-1)^4\,\mathsf{z}_1(t)^2\,\mathsf{z}_2(t-1)^{-2}\end{pmatrix}-\frac{1}{6}\begin{pmatrix}\mathsf{y}_3(t-3)^3\\ \mathsf{y}_1(t-3)^{-3}\\ \mathsf{y}_2(t-3)^{-1}\end{pmatrix}\right\|=0,$$

where $\|.\|$ denotes Euclidean norm. Here a form of (5) is:

$$\begin{aligned}\mathsf{y}_1(t) &= \left(\frac{2}{5}\mathsf{y}_1(t-1)^2\,\mathsf{z}_1(t)+\frac{1}{6}\mathsf{y}_3(t-3)^3\right)^{\frac{1}{3}},\\ \mathsf{y}_2(t) &= \left(\frac{2}{5}\mathsf{y}_2(t-1)^{-3}\,\mathsf{y}_2(t-2)^2\,\mathsf{z}_2(t)+\frac{1}{6}\mathsf{y}_1(t-3)^{-3}\right)^{\frac{1}{3}},\\ \mathsf{y}_3(t) &= \left(\frac{2}{5}\mathsf{y}_3(t-1)^4\,\mathsf{z}_1(t)^2\,\mathsf{z}_2(t-1)^{-2}+\frac{1}{6}\mathsf{y}_2(t-3)^{-1}\right)^{\frac{1}{5}}.\end{aligned}$$

Here a form of (8) is:

$$\begin{aligned}\mathsf{y}_1(t-3) &= \left(6\mathsf{y}_2(t)^3-\frac{12}{5}\mathsf{y}_2(t-1)^{-3}\,\mathsf{y}_2(t-2)^2\,\mathsf{z}_2(t)\right)^{-\frac{1}{3}},\\ \mathsf{y}_2(t-3) &= \left(6\mathsf{y}_3(t)^5-\frac{12}{5}\mathsf{y}_3(t-1)^4\,\mathsf{z}_1(t)^2\,\mathsf{z}_2(t-1)^{-2}\right)^{-1},\\ \mathsf{y}_3(t-3) &= \left(6\mathsf{y}_1(t)^3-\frac{12}{5}\mathsf{y}_1(t-1)^2\,\mathsf{z}_1(t)\right)^{\frac{1}{3}}.\end{aligned}$$

4 Conclusion

Implicit models with forward and backward equivalent rewritings are convenient for modeling. Random processes following such models always exist and are uniquely determined by initial values. Moreover, we have found a full description of solutions of these implicit formulas. Our investigation is based on a construction provided a solution given a starting time τ, initial values and innovations.

Acknowledgements

The research was supported by the grant No. P402/12/G097 of the Grant Agency of the Czech Republic.

References

[1] Box, G.E.P.; Jenkins, G.M.: *Time-series Analysis Forecasting and Control.* Holden-Day, San Francisco, 1976.

[2] Brockwell, P.J.; Davis, R.A.: *Time-series: Theory and Methods* Springer-Verlag, New York, 1987.

[3] Durbin, J.; Koopman, S.J.: *Time Series Analysis by State Space Models.* Oxford University Press, Oxford, 2001.

[4] Hommes, C.: Financial markets as nonlinear adaptive evolutionary systems. Quantitative Finance **1,1 (2001)**, 149-167.

[5] Lachout, P.: On functional definition of time-series models. In: Proceedings of the 32th International Conference on Mathematical Methods in Economics, Olomouc, (Ed.: Jana Talaov, Jan Stoklasa, Tom Talek), Palack University, Olomouc, 560-565, 2014.

[6] Liao, H-E; Sethares, W.A.: Suboptimal identification of nonlinear ARMA models using an orthogonality approach. IEEE Transactions on Circuits and Systems'**42,1** (1995), 14-22.

[7] Liu, J.; Susko, E.: On Strict Stationarity and Ergodicity of a Non-Linear ARMA Model. Journal of Applied Probability **29,2** (1992), 363-373.

[8] Masry, Elias; Tjøstheim, Dag: Nonparametric estimation and identification of nonlinear ARCH time-series. Econometric Theory **11,1 (1995)**, 258-289.

[9] Priestley, M.B.: *Non-Linear and Non-Stationary time-series.* Academic Press, London, 1988.

[10] Shumway, R.H.; Stoffer, D.S.: Time Series Analysis and Its Applications - With R Examples. EZ Green Edition, 2015.

On modelling of macroeconomic context of current economic development

Eva Litavcová[1], Sylvia Jenčová,[2] Petra Vašaničová[3]

Abstract. Unemployment and inflation are regarded as the two biggest evils facing any economy today. They have distinct negative impacts on economic growth. These macroeconomic problems present issues dealt with not only by economists, politicians and governments, but also by the residents of countries whose standard of living is negatively impacted by them. The aim of this paper is to determine the influence of inflation and unemployment on economic growth, highlighting the commonalities between them in the context of economic development from the point of view robust quantile regression technique.

We used robust quantile regression in traditional framework of Okun's law. Our study focuses on Slovakia's macroeconomic data. The same technique was used on relationship of another two macroeconomic indicators, unemployment rate and CPI. We have found evidence about relevance of classical Okun's law on Slovakia's data. On the contrary, to the relationship of unemployment rate and CPI quantile regression is more appropriate.

Keywords: economic growth, unemployment, inflation, quantile regression.

JEL Classification: C20
AMS Classification: 62J05

1 Introduction

Unemployment, price stability, balance of payments and economic growth are essential corners of a magic quadrangle. One of the objectives may be better achieved only at the expense of another. Inflation is one of the major problems that complicate the performance of the economy at the macro level. Inflation and unemployment are inversely proportional, i.e. in order to keep inflation low requires higher unemployment. The hypothesis on the relationship between unemployment and inflation, which claims that prices and wages tend to rapidly increase when unemployment is below its natural rate, is known as the natural rate hypothesis, as stated Mankiw [13].

Significant labor market indicators include: the number of people of working age; percentage of the elderly; the number of employed; the number of unemployed; the economically active portion of a population; the level of economic activity as measured by the share of the economically active portion of a population attributable to the total population of working and the elderly; the economically inactive population; employment rate; and unemployment rate. The unemployment rate is quantified using data from the Labour Force Survey by the percentage of unemployed in the total number of people willing to work, i.e. the total number of unemployed persons divided by the total number of the economically active population. If a certain percentage of people over a certain period of time leave work, then we must monitor the rate of lost work, while watching what percentage of the unemployed who, for a given time period, find work. According to Kucharčíková et al. [12], macroeconomic observations indicate that the majority of employment trends show a similar trend as the output of the gross domestic product. Consequently, the unemployment rate is usually highest during the economic crisis. Full utilization of the labor force in accordance with its qualitative and skills levels indicates the equilibrium level of employment. Full employment, as a fundamental objective of the Lisbon Treaty, is a condition that allows for some level of natural unemployment rate as the economy always contains some individuals who, for whatever reason, are not working.

Holman [7] describes inflation as "price increases or reducing the purchasing power of money, but not diminishing the purchasing power of people. Inflation reduces the amount of goods and services that you can buy for a monetary unit, but does not reduce the quantity of goods and services that we can buy for our retirement." We see it as a dynamic process, the spiral of price increases, wages and other cost factors. Beňová [3] sees the inflation rate exceeding the optimal amount of money in circulation, manifested as a disturbance of the balance

[1] University of Presov, Faculty of Management, Konstantinova 16, Presov, Slovakia, eva.litavcova@unipo.sk.
[2] University of Presov, Faculty of Management, Konstantinova 16, Presov, Slovakia, sylvia.jencova@unipo.sk.
[3] University of Presov, Faculty of Management, Konstantinova 16, Presov, Slovakia, petra.vasanicova@smail.unipo.sk.

of cash and producing an upward pressure on prices. Inflation is investigated by economists from different perspectives, through the prism of the classical quantity theory of money, monetarist and the Keynesian view.

From a theoretical point of view, inflation is classified by cause (demand-pull inflation, cost-push inflation); anticipated and nonanticipated by prediction of economist; overt, suppressed and hidden by its expression; walking, galloping and hyperinflation in quantitative terms. According to Samuelson [16], inflation is largely internal, tending to remain at the level reached before the shock and does not suffer from supply or demand. According to Holman [7] demand inflation is caused by an increase in aggregate demand, which will reduce unemployment below the natural rate causing nominal wage growth followed the rise in prices. Felderer and Homberg [4] in their work point out that when looking at crowding-out and inflation, always it comes to securing employment and production by an active fiscal and monetary policy. While classical and neoclassical authors were in favor of a policy of laissez faire, Keynesians insisted on an active, discretionary stabilization policy.

Each type of inflation is quantified by indices. The average annual inflation rate, measured by the HCPI is one of the basic economic metrics considered in the context of convergence Maastricht criteria for achieving the third stage of Economic and Monetary Union (before joining the euro area) [1]. The Laspeyres price index compares the change in the cost of purchasing consumer good last year to the prices of the current year. This index reflects the reality of having to pay for basic consumer goods this year to the prices of the reference year. This requires measurement of the increase in expenses necessary to achieve a higher degree of satisfaction than in the base year. This index is imperfect, does not account for the substitution of goods. The Paasche price index compares the cost of buying, given a purchasing plan, of the current year at various prices as scales. This underestimates the whole situation, giving lower results, i.e. the consumer is expected to buy the same quantity of goods before and after the increase in prices.

Since 2005, the Statistical Office started to calculate consumer price indices (CPI and HICP) as a chain Laspeyres-type index with an annually updated scheme. Harmonized Indices of Consumer Prices (HICP) began to be assembled in order to ensure comparability of the consumer price indices of individual countries within the European Union. Basic regulation for the introduction of calculating the harmonized index of consumer prices in the Member States of the European Union started in 1997, with Council Regulation (EC) No. 2494/1995 (Council Regulation (EC) No 2494/95) concerning harmonized indices of consumer prices [1].

The aim of the following part of this article is quantification of the relationship between unemployment and economic growth, as well as inflation and unemployment in Slovakia. The link between changes in GDP and changes in the unemployment was studied by American economist Arthur Okun [14]. Fuhrmann [6] in his study describes a version of the Okun’s relationship that is based on a regression of changes in the unemployment rate and economic growth rate. The assertion of Okun’s law to the terms and conditions of the European Union is dedicated to his scientific work as determined by Košta [11], Slušná [17] and others. Okun’s law show how important implications for macroeconomic policy relate to the rate of change in domestic product and unemployment. The law is crucial since it involves two economic variables. By Knotek [10] Okun 's Law equation is quantified by simple ordinary least squares equations

(mean of change in the unemployment rate) = 0.30 – 0.07 (growth of real output),*

(mean of change in the unemployment rate) = 0.23 – 0.07 (real output growth),*

second is the gap between potential product and real product, on the US data. The unemployment rate is a decreasing function of the rate of growth performance of the economy as measured by GDP. According to him, if the real domestic product falls below the potential output by 1%, this will lead to an increase in unemployment above the natural rate of unemployment by about 0.33%. It is clear that this relationship will change over time.

According to Táncošová [18] GDP must maintain a certain growth rate so that unemployment does not grow, and if we have to reduce unemployment, then real GDP must grow faster than potential GDP. Okun’s law is an important link between aggregate market and developments in output and the labor market and the development of unemployment. The social effects are psychological suffering, growth disorders.

Košta [11] to evaluate conditions in Slovakia applied Okun’s law, taking into account the period 1995 - 2008. The regression equation used to evaluate the change in the unemployment rate according to changes in GDP growth rate, using the values for the period 1995-2008, provided: (*Δ unemployment rate*) = 2.9976 - 0.613*(*changes in GDP growth*). Coefficient of determination is 82%, which means that the model explains 82% of the total variability of the changes in unemployment. This author and his team concluded that economic growth has been accompanied by an increase in labor productivity.

We contribute to the existing literature by using robust quantile regression in traditional framework of Okun’s law. Our study focuses on Slovakia’s macroeconomic data. The same technique was used on relationship of another two macroeconomic indicators, unemployment rate and CPI. We have found evidence about relevance of classical Okun's law on Slovakia's data. On the contrary, to the relationship of unemployment rate and CPI quantile regression is more appropriate.

2 Data and methodology

Based on the quarterly data from the first quarter 1998 to the fourth quarter 2015 from the database of the Statistical Office of the Slovak Republic on GDP, the GDP growth rate determined on the basis of GDP at current prices and constant prices (2010 prices), unemployment rate (MNZ), inflation rate as measured by CPI and HICP are set out below converted to quantify the impact of macroeconomic indicators.

The effect of change in the unemployment rate of GDP growth and the effect of CPI of unemployment rate are tested by using simple ordinary least squares method, which describes conditional mean of response variable as a linear function of explanatory variable. Additionally, we focus on these relationships of investigated variables by using quantile regression. According to Agresti [2], quantile regression models quantiles of a response variable as a function of explanatory variables. This method can be less severely affected by outliers than is ordinary least squares. When the response conditional distributions are highly skewed with possibly highly non-constant variance, the method can describe the relationship better than a simple normal model with constant variance. Quantile regression model fitting minimizes of a weighted sum of absolute residuals, formulated as a linear programming problem. However, when the normal linear model truly holds, the least squares estimators are much more efficient.

Roger Koenker primary in his work with Bassett [8] specify the τ-th regression quantile function as $Q_Y(\tau|x) = x^T\beta(\tau)$, and consider of $\hat{\beta}(\tau)$ solving

$$\min_{\beta\in\mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - x_i^T\beta) \tag{1}$$

with check function $\rho_\tau(u) = u \cdot \tau$ for $u \geq 0$ and $\rho_\tau(u) = u \cdot (\tau - 1)$ for $u < 0$. Quantile regression problem may be reformulated as a linear program.

As reported by Fenske [5], quantiles are defined based on the cumulative distribution function (*cdf*) F_Y of a continuous random variable *Y*. The $\tau\,.\,100\%$ quantile of *Y* can be written as a value y_τ, where

$$F_Y(y_\tau) = P(Y \leq y_\tau) = \int_{-\infty}^{y_\tau} f(u)du = \tau \tag{2}$$

for $\tau \in (0,1)$. It is only unique if F_Y is strictly monotonic increasing. In case that information on an additional random variable *X* is given, the quantile can similarly be expressed conditional on *X* is equal *x*:

$$F_Y(y_\tau(x)|\, X = x) = P(Y \leq y_\tau(x)|\, X = x) = \tau. \tag{3}$$

The quantile function $Q_Y(\tau|X = x)$ is defined as the smallest *y* where the quantile property is fulfilled. If F_Y is strictly increasing, the quantile function is set to the inverse of the *cdf* of *Y*. The relationship between quantile function and *cdf* can be expressed as

$$F_Y(y_\tau(x)|\, X = x) = \tau \;\Leftrightarrow\; Q_Y(\tau|\, X = x) = y_\tau(x) \tag{4}$$

for strictly increasing F_Y, which emphasizes that the quantile function describes $\tau\,.\,100\%$ quantiles of *Y* depending on covariates *x* and a quantile parameter $\tau \in (0,1)$.

Quantile regression is an approach to model the conditional quantile function of a continuous variable of interest *Y*, e.g. response variable, depending on further variables or covariates *X*. In the linear model it can be expressed as $y_i = x_i^T\beta_\tau + \varepsilon_{\tau i}$, $i = 1, \dots, n$. The index $i = 1, \dots, n$, denotes the observation, y_i is the response value and $x_i = (1, x_{i1}, \dots, x_{ip})^T$ the given covariate vector for observation i. The quantile-specific linear effects are denoted by $\beta_\tau = (\beta_{\tau 0}, \beta_{\tau 1}, \dots, \beta_{\tau p})^T$, and $\tau \in (0,1)$ indicates a quantile parameter which has to be fixed in advance. The random variable $\varepsilon_{\tau i}$ is assumed to be an unknown error term with *cdf* $F_{\varepsilon_{\tau i}}$ and density $f_{\varepsilon_{\tau i}}$ depending on quantile parameter τ and observation i. For quantile regression, no specific assumptions are made apart from $\varepsilon_{\tau i}$ and $\varepsilon_{\tau j}$ being independent for $i \neq j$, and $\int_{-\infty}^{0} f_{\varepsilon_{\tau i}}(\varepsilon_{\tau i})\, d\varepsilon_{\tau i} = F_{\varepsilon_{\tau i}}(0) = \tau$. Due to this assumption, the quantile function $Q_{Y_i}(\tau|x_i)$ of the response variable Y_i conditional on covariate vector x_i at a given quantile parameter τ is equal to $x_i^T\beta_\tau$.Thus, the parameter $\beta_{\tau 1}$, for example, can be interpreted as the change of the conditional quantile function when x_{i1} changes to $x_{i1} + 1$, given all other covariates remain constant [5].

3 Results and Conclusion

Table 1 shows coefficients of GDP growth rate, which were estimated at the 10th, 25th, 50th, 75th and 90th quantiles. The tests of significance were based on robust bootstrap estimations of standard errors [9]. We also presented OLS result as reference, the same as was reported in the classical framework of Okun's law. According to expectation due Okun's law, variable GDP growth rate significantly influence the change in the unemployment rate in OLS regression. The results of quantile estimates were allowed for richer interpretation of examined well-known relationship between GDP growth rate and change in the unemployment rate. The estimates of the effect of explanatory variable for each chosen quantile allowing us to detect different impacts of GDP growth rate depending on the level of the change in the unemployment rate. The GDP growth rate did not present significantly different effect over the conditional distribution of the change in the unemployment rate (probability of joint test of equality of slopes is 0.2387). However, the result proved that the constant effect estimated through OLS was not actually constant across the quantiles. Figure 1 shows some selected quantile regression lines and Figure 2 shows different GDP growth rate slope coefficients in more gentle division of levels of the change in the unemployment rate (without tails of quantiles due to the small number of observations, quarters was 72 only) as was reported in Table 1 and Figure 1. Every slope coefficient (except tails) is significantly different from zero, but each of them lies within confidence boundaries of OLS regression.

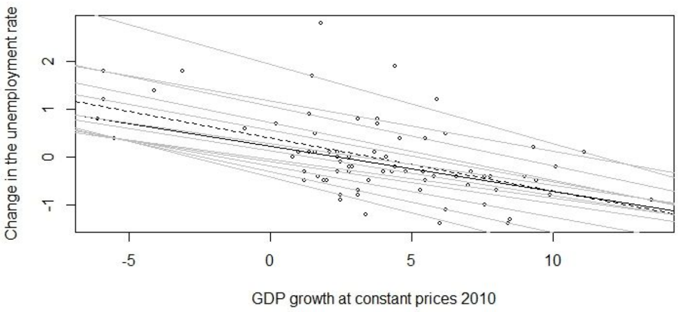


Figure 1 Quantile regression *Change in the unemployment rate ~ GDP growth rate*

Note: OLS is dashed line, median line is black and gray lines are for taus 0.1, 0.2, 0.25, 0.3, 0.4, 0.6, 0.7, 0.75, 0.8 and 0.9.

	OLS	Quantiles				
		0.1	0.25	0.5	0.75	0.9
Constant	0.4024[d]	–0.3051	–0.0873	0.2236[a]	0.7241[c]	1.1790[c]
	(0.1152)	(0.2669)	(0.1629)	(0.1168)	(0.2146)	(0.4178)
GDP growth rate	–0.1116[d]	–0.1282[b]	–0.0886[b]	–0.0945[d]	–0.1203[d]	–0.1053[a]
	(0.0214)	(0.0488)	(0.0393)	(0.0243)	(0.0270)	(0.0551)

Table 1 OLS and quantile regression estimated coefficients for *y = Change in the unemployment rate*

Note: OLS and quantile regression; for OLS standard errors, for quantile regression bootstrap standard errors in parenthesis; a, b, c, d denotes 10%, 5%, 1%, and 0.1% significance level; Joint test of equality of slopes: F value = 1.3846 and Pr(>F) = 0.2387.

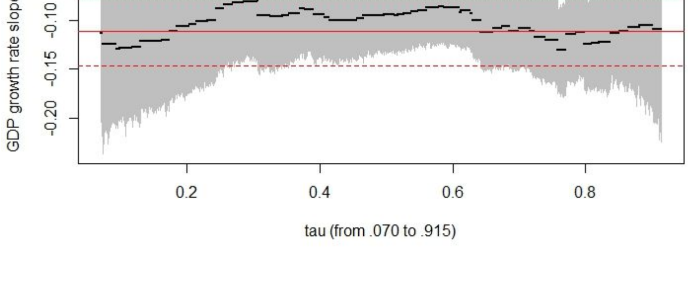


Figure 2 Quantile regression´s slope coefficient estimates at different quantiles of *Change in the unempl. rate*

Note: There are 846 estimates of slope coefficients with confidence intervals based on bootstrap SE

Slopes of the quantile regression lines are less pronounced in the middle quantile levels of the explaining variable than reports OLS regression. Appearance of the linear relationship between investigated variables is stable (in the range from the quantile 0.07 to the quantile 0.91) from the point of view quantile regression, too. The normal linear model there truly holds, with respect on this result.

Another relationship that we have dealt with in the same way as above is a linear relationship between the unemployment rate and CPI. Classical ordinary least squares method indicates positive linear relationship between these variables (Table 2). But, the next listed pattern of significance in the same table signalizes different impacts of unemployment rate depending on the level of CPI. As is shown in Figure 4, at low quantiles up to 0.38 approximately, there is no significance. Zero falls into this confidence intervals. Point estimates of slope coefficients of unemployment rate for tau values around the lower quartile falls under confidence band of OLS, however, confidence intervals are there overlapped. And vice versa, estimates of slope coefficients above quantile 0.4 are very similar to OLS estimate. But, slopes of variable unemployment rate over approximately 0.85^{th} quantile level of CPI are not significant. In summary, significance exists only in interval of quantiles which covers only around 45% of whole quantile range. And, from Figure 3 is clear no significant difference between slope coefficients on various quantile levels of CPI. Test was made for levels reported in Table 2 and is not significant. Computed probability of joint test of equality of slopes is 0.2104. The normal linear model holds there not very clear, with respect on this result.

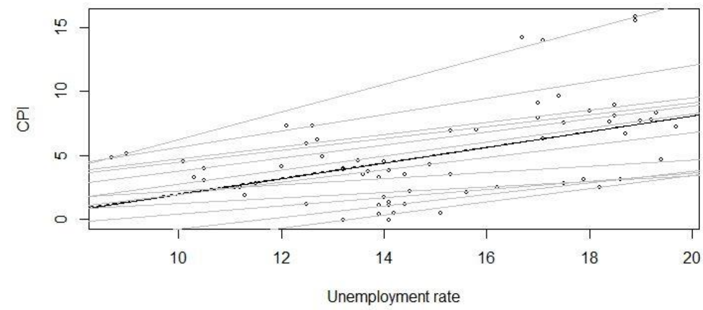


Figure 3 Quantile regression *CPI ~ Unemployment rate*

Note: OLS is dashed line, median line is black and gray lines are for taus 0.1, 0.2, 0.25, 0.3, 0.4, 0.6, 0.7, 0.75, 0.8 and 0.9.

	OLS	Quantiles				
		0.1	0.25	0.5	0.75	0.9
Constant	-4.0291^{a}	–5.3143	–0.9829	-4.2231^{a}	–0.1615	–0.7769
	(2.0479)	(6.3509)	(4.9018)	(2.2252)	(2.2908)	(3.5154)
Unemployment rate	0.6059^{d}	0.4524	0.2195	0.6154^{d}	0.4615^{c}	0.6410^{a}
	(0.1351)	(0.4502)	(0.3257)	(0.1577)	(0.1478)	(0.3236)

Table 2 OLS and quantile regression estimated coefficients for *y = CPI*

Note: OLS and quantile regression; for OLS standard errors, for quantile regression bootstrap standard errors in parenthesis; a, b, c, d denotes 10%, 5%, 1%, and 0.1% significance level; Joint test of equality of slopes: F value = 1.472 and Pr(>F) = 0.2104.

GDP growth or reducing the rate of unemployment and price stability are the basic goals of any country. Slušná [17] describes the factors determining unemployment. In addition to GDP growth, wage levels, the structure of the economy, industry structure, demography, social system, general level of education and technological changes all affect unemployment.

This contribution shows the importance and significance of macroeconomic indicators on the unemployment rate, the inflation rate and the rate of growth of gross domestic product. Okun's law is only one of the forecasting tools available for economists, politicians and the labor market in the development of gross domestic product in relation to the various factors affecting production work. In essence this is a very simplified model, but also a very effective tool in the issue of unemployment and economic growth. Nevertheless, reported results of two above stated relationships describe the situation only in part. We started with classical framework of Okun's law and we omitted that we handle with time series. One of the next possibilities is to examine bi-variate VAR model. This remains to be tested in our subsequent works.

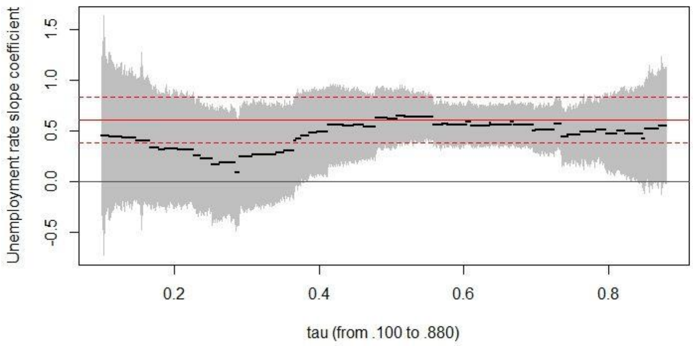


Figure 4 Quantile regression´s slope coefficient estimates at different quantiles of *CPI*

Note: There are 781 estimates of slope coefficients with confidence intervals based on bootstrap SE.

Acknowledgements

Supported by the grant No. 1/0392/15 of the Slovak Grant Agency, by the grant KEGA 037PU-4/2014 of the Cultural and Educational Grant Agency of the Slovak Republic, and by the grant VEGA 1/0412/17 of the Scientific Grant Agency of Ministry of Education SR and Slovak Academy of Sciences.

References

[1] Adamec, Š.: *Vývoj spotrebiteľských cien a inflácia*. ŠÚ SR, Bratislava, 2011.

[2] Agresti, A.: *Foundations of Linear and Generalized Linear Models*. Wiley Interscience, John Wiley & Sons, Inc., Hoboken, New Jersey, 2015.

[3] Beňová, E. a kol.: *Financie a mena*. Iura Edition, Bratislava, 2007.

[4] Felderer, B. and Homberg, S.: *Makro ekonomika a nová makro ekonomika*. Elita, Bratislava,1995.

[5] Fenske, N.: *Structured additive quantile regression with applications to modelling undernutrition and obesity of children*. Dissertation thesis. Munchen, 2012.

[6] Fuhrmann R. C.: *Okun's Law: Economic Growth And Unemployment*. (2012). http://www.investopedia.com/articles/economics/12/okuns-law.asp.

[7] Holman, R.: *Makroekonómie*. C.H.Beck, Praha, 2004.

[8] Koenker, R. and Bassett, G.: Regression Quantiles. *Econometrica* **46**.1 (1978), 33–50.

[9] Koenker, R., and Hallock, K. F:. Quantile regression. *The Journal of EconomicPerspectives* **15** (2001), 43-56.

[10] Knotek, II, E.S.: How Useful is Okun's Law? *Economic Review* **4** (2007), 73-103.

[11] Košta, J.: Aktuálne problémy trhu práce v Slovenskej republike po vstupe do Európskej menovej únie. In *Zborník EÚ SAV*. Ekonomický ústav SAV, Bratislava, 2011.

[12] Kucharčíková, A., Tokarčíková, E., Ďurišová, M., Jacková, A., Kozubíková, Z. a Vodák, J.: *Efektívní výroba - využívejte výrobní faktory a připravte se na změny na trzích*. Computer Press, Brno, 2011.

[13] Mankiw, N. G.: *Zásady ekonomie*. Grada, Praha, 2009.

[14] Okun, A.: Potential GNP: Its Measurement and Significance. In: *Proceedings of the Business and Economic Statistics Section* **7** (1962), 89-104.

[15] Parkin, M.: *Macroeconomics*. Addison-Wesley Publishing Company, USA, 1990.

[16] Samuelson, P. A a Nordhaus, W. D.: *Ekonómia 2*. Bradlo, Bratislava, 1992.

[17] Slušná, Ľ.: Vzťah ekonomického rastu a nezamestnanosti v krajinách Európskej únie. In: *Zborník príspevkov z vedeckej konferencie*. Ekonomický ústav SAV, Bratislava, 2011.

[18] Táncošová, J.: *Ekonómia*. Iura Edition, Bratislava, 2013.

American Option Pricing Problem Formulated as Variational Inequality Problem

Ladislav Lukáš[1]

Abstract. The paper deals with variational formulation of option pricing problems. We start with basic platform of the Black-Scholes model for a put option with strike price and maturity given, which assumes the underlying asset to follow a geometric Brownian motion. Pricing American options requires, due to the early exercise feature of such derivative contracts, the solution of optimal stopping problems for the price process. Unlike in the European case, the pricing function of an American option does not satisfy a partial differential equation, but a partial differential inequality, or a system of inequalities. Recasting such problem into a variational inequality problem is the next step, which is given in detail. We mention briefly the functional space which provides natural framework for weak formulation of American put option pricing problem. Both optimal exercise boundary and additive decomposition of American put option are discussed, as well.

Keywords: American option pricing, additive option price decomposition, variational formulation, variational inequality.

JEL Classification: C63, G13
AMS Classification: 91G80

1 Introduction

There is well-known, options are derivative contracts between a buyer, or holder, and a seller, writer or issuer, of the contract where future payoffs to the buyer are determined by the price of another security, called also an underlying asset, such as a common stock. An option is a contract that gives the holder the right but not the obligation to buy or sell a certain financial asset by a certain date T (expiration date, or maturity date), for a given price K (exercise price, or strike price). There are two main types of options

- call option, which is a option to buy security,
- put option, which gives the right to sell it.

The option pricing theory is available in many famous textbook, e.g. [4],[6],[7], and [11]. Since the value of the option depends on the security value at the exercise day, which is a priori unknown, the underlying asset price is to be modeled by a stochastic process. If the underlying asset price follows a geometric Brownian motion, then the value of the option is described by a deterministic backward partial differential equation (PDE) of the second order and of parabolic type, in particular.

In the classical Black-Scholes (B-S) model, we assume that the price $(S_t)_{t\geq 0}$ of the underlying risky asset, e.g. a stock, is described by geometric Brownian process, which is given by stochastic differential equation (SDE)

$$\mathrm{d}S_t = S_t(\mu \mathrm{d}t + \sigma \mathrm{d}B_t), \quad t \geq 0, \tag{1}$$

2 American option pricing problem – classical formulation

There is well-known that American options are more flexible than European options regarding the exercise time. American option can be exercised at any time between the writing and expiration of the contract, i.e. $0 < t \leq T$. As usual, one assumes also that there no commissions and/or fees within the contract, and the bid-ask spreads of asset priced S_t and the corresponding option are zero both. Sure, the exercising flexibility gives American option more profiteering opportunity than European option offers.

The exercise freedom of American options means that the pay-off functions are

$$\begin{aligned} V_C(S_t,t) &= \max(S_t - K, 0), \quad t \in \,]0,T], \quad \text{for call option,} \\ V_P(S_t,t) &= \max(K - S_t, 0), \quad t \in \,]0,T], \quad \text{for put option,} \end{aligned} \tag{2a}$$

[1] University of West Bohemia/Faculty of Economics, Dept. of Economics and Quantitative Methods, Husova 11, 306 14 Pilsen, lukasl@kem.zcu.cz.

thus yielding at $t = T$ exactly the same pay-off functions as European options, i.e.

$$V_C(S_T, T) = \max(S_T - K, 0), \quad \text{for call option,}$$
$$V_P(S_T, T) = \max(K - S_T, 0), \quad \text{for put option,} \tag{2b}$$

and giving also a-priori bounds under no-arbitrage assumptions, using denotation $\mathbb{R}_+ = [0, +\infty)$,

$$V_C(S, t) \geq \max(S - K, 0), \quad \forall\, S, t, \quad S \in \mathbb{R}_+, \ t \in\,]0, T], \quad \text{for call option,}$$
$$V_P(S, t) \geq \max(K - S, 0), \quad \forall\, S, t, \quad S \in \mathbb{R}_+, \ t \in\,]0, T], \quad \text{for put option.} \tag{2c}$$

The well-known crucial problem in option theory is following one: If at time t we have an asset priced at S_t

- What is the fair price $V(S_t,t)$ of the option.
- When is the optimal time to exercise the option.

Now, we will follow [2], [3], [5], and [6], mainly. In fact, for American put option, when underlying asset price S_t falls below a certain point, one should exercise the option immediately.

For example, if at time t, $S_t < K(1 - \exp(-r(T-t))$, and keeping in mind that the pay-off at the option expiration date will never exceed K in any case due to (2b), then the option holder can get the immediate gain at time t in amount

$$K - S_t > K - K(1 - \exp(-r(T-t)) = K \exp\big(-r(T-t)\big), \tag{3a}$$

and by depositing the gain in a saving account, the total pay-off will exceed K at $t = T$, evidently. Therefore one can conclude that there exists a point at t which is known as an optimal exercise point. Moreover, we know that the price of American put option is never less than the pay-off function $\max(K-S_t,0)$ because of non-arbitrage assumption.

In case of American call option the situation with early exercise is quite different. In general, for American call option on contrary to American put option, we need to include also dividend yields δ in order to allow rationality of an early exercise, otherwise it coincides with a European call option.

In fact, one can show using (2c) that if $\delta = 0$ then an early exercise does not pay off for the call option as follows

$$V_C(S, t) \geq S - Ke^{-r(T-t)} \quad \Rightarrow \quad V_C(S, t) > S - K, \quad S \in \mathbb{R}_+, \quad t < T, \quad \text{if } r > 0, K > 0. \tag{3b}$$

Discussion of (3a) and (3b) leads to analysis of American put option first, as usual. Following [6], we present a decomposition formula for American put option $V_P(S_t,t)$ which gives a link to European put option

$$V_P(S_t, t) = V_P^E(S_t, t) + e(S_t, t), \tag{4}$$

where $V_P^E(S_t, t)$ is the European put option price on the same asset, and $e(S_t, t)$ is the early exercise premium.

Formula for $V_P^E(S_t, t)$ is rather standard and well-known, while formula for $e(S_t, t)$ is based upon so called fundamental solution of Black-Scholes (B-S) partial differential equation (PDE). These quantities are defined by (5a) and (5b), respectively, as follows

$$V_P^E(S_t, t) = K\, e^{-r(T-t)} N(-d_2) - S_t N(-d_1),$$
$$d_1 = [\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T - t)]/(\sigma\sqrt{(T-t)}), \quad d_2 = d_1 \sigma \sqrt{(T-t)}, \tag{5a}$$

where $N(x)$ is the cumulative probability distribution function of a standard normal distribution $N(0,1)$,

$$e(S_t, t) = Kr \int_t^T d\eta \int_0^{S_\eta} G(S_t, t; \xi, \eta) d\xi, \tag{5b}$$

where $G(S_t, t; \xi, \eta)$ is the fundamental solution, sometimes called Green function, too, of the B-S PDE.

The fundamental solution of B-S PDE is defined as a solution of terminal value problem for classical Black-Scholes PDE with a very special terminal condition using the Dirac delta function

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0, \quad \text{for } (S, t) \in (0, +\infty) \times [0, T), \tag{6a}$$

$$V(S,T) = \delta(S-\xi), \quad \xi \in (0,+\infty)\ . \tag{6b}$$

In (6a) and (6b), we use simplified denotation of asset price S instead of S_t, and S_T, respectively, in order to elevate readability of the expressions.

The Dirac delta function $\delta(x)$ can be simply characterized by properties

$$\delta(x) = +\infty, \text{ for } x = 0, \quad \delta(x) = 0, \text{ for } x \neq 0, \quad \int_{-\infty}^{+\infty} \delta(x)dx = 1\ . \tag{7}$$

Finally, following [6], we get the fundamental solution of problem (6a) and (6b) in explicit form, as follows

$$G(S,t;\xi,T) = \frac{e^{-r(T-t)}}{\xi\sigma\sqrt{(2\pi(T-t))}} \exp\{-\frac{1}{2\sigma^2(T-t)}\left[\ln\frac{S}{\xi} + \left(r-\frac{\sigma^2}{2}\right)(T-t)\right]^2\}. \tag{8}$$

Now, let denote the price of non-dividend-paying American put option $V(S_t,t)$ instead of $V_P(S_t,t)$, simply dropping the sub-index P from now on.

Thorough inspection of (2c) and (3a) gives an idea that for American put option there exists two disjunctive sub-regions, say Σ_1 and Σ_2, respectively, covering $\mathbb{R}_+ \times [0,T]$, i.e. $\Sigma_1 \cup \Sigma_2 = \mathbb{R}_+ \times [0,T]$.

Let $\partial\Sigma_1$, $\partial\Sigma_2$ denote boundaries of Σ_1 and Σ_2, respectively. As usual, a common piece of their boundaries, say $\partial\Sigma_1 \cap \partial\Sigma_2 = \Gamma$, is called optimal exercise boundary.

The optimal exercise boundary is reasonably defined by mapping Γ: $[0, T] \rightarrow \Gamma(t)$ thus specifying the asset price at Γ, denoted $S_t = \Gamma(t)$, in particular, but which is not given a priori and has to be determined together with option pricing function $V(S_t,t)$ by solving the corresponding option pricing problem.

Providing (2c) holds, these sub-regions are defined as follows

$$V(S_t,t) > \max(K-S_t,0),\ \forall\ (S_t,t) \in \Sigma_1, \quad \Sigma_1 = \{(S_t,t) | \Gamma(t) \le S_t < +\infty,\ \forall\ t \in [0,T]\}, \tag{9a}$$

$$\begin{gathered} V(S_t,t) = \max(K-S_t,0),\ \forall\ (S_t,t) \in \Sigma_2, \quad \Sigma_2 = \{(S_t,t) | 0 \le S_t \le \Gamma(t),\ \forall\ t \in [0,T]\}, \\ \Gamma(t) < K,\ \forall\ t \in [0,T), \end{gathered} \tag{9b}$$

where Σ_1 is called the continuation sub-region, since the pay-off is zero when $S_t \geq K$, and the holder should continue to keep the option, therefore Σ_2 is called stopping sub-region. The asset prices S_t on the exercise boundary are denoted $S_t = \Gamma(t)$ precisely, as we have already stated above.

Noting, the continuation sub-region Σ_1 is sometimes called retained sub-region alternatively, whereas the stopping sub-region Σ_2 is called selling sub-region, or exercise sub-region, as well.

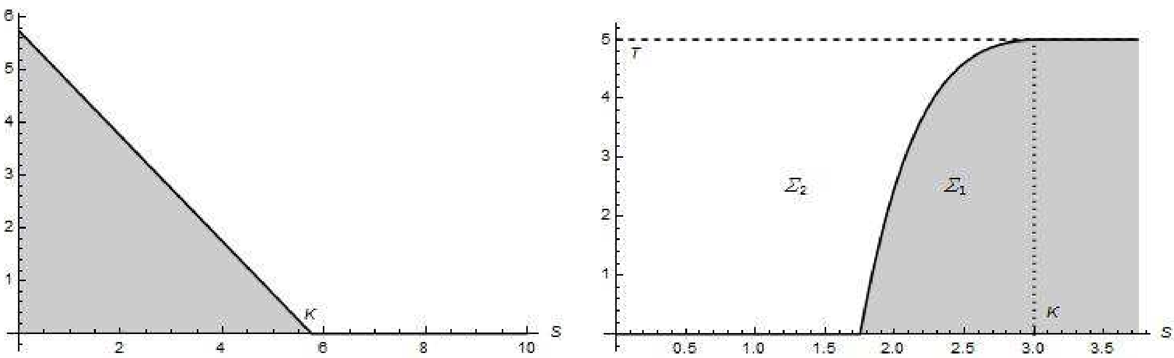


Figure 1 (Left): Pay-off function of put option: max(K – S, 0); (Right): Sub-regions $\Sigma_1 \cup \Sigma_2 = \mathbb{R}_+ \times [0,T]$

In **Fig. 1** on the left panel, we depict the pay-off function of American put option $V(S_t,t) = \max(K - S_t,0)$, while on the right panel, there are plotted both sub-regions Σ_1 and Σ_2, schematically, with horizontal axis for asset price S, and vertical axis for option contract time t, showing the expiration date T, too.

Now, following the reasoning procedure based on construction of self-financing portfolio, Δ-hedging principle, and Ito formula, which is common and typical for derivation of almost all models within the field of financial option pricing problems, referring to [2], [4], and [7], mainly, one can infer that $V(S_t, t)$ satisfies the B-S PDE in the sub-region Σ_1

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0, \quad \text{for } (S,t) \in \Sigma_1, \quad \text{assuming } V(S_t,t) \in C^{2,1}(\Sigma_1). \tag{10}$$

The boundary conditions on $\partial\Sigma_1$ are specified in following way. First, on the optimal exercise boundary Γ, it holds

$$V(S_t,t) = \max(K - S_t, 0) \quad \Rightarrow \quad \frac{\partial V(S,t)}{\partial S} = -1, \quad S_t \in \Gamma(t), \quad \forall\, t \in [0,T). \tag{11a}$$

Further, the terminal condition at $t = T$, and the condition when $S_t \to +\infty$, are following

$$V(S_T,T) = \max(K - S_T, 0), \qquad \lim V(S_t,t) = 0, \;\; S_t \to +\infty, \quad \forall\, t \in [0,T). \tag{11b}$$

Concluding, we see that American put pricing problem leads to solving a function pair $\{V(S_t,t), \Gamma(t)\}$ in sub-region Σ_1 satisfying PDE (10) and boundary-terminal conditions (11a) and (11b).

As the optimal exercise boundary Γ is not known a-priori, but has to be determined together with valuation function $V(S_t,t)$, therefore the problem is called free-boundary problem for parabolic PDE. It is already a problem formulated in mathematically elegant way. However, we have to point out that the major difficulty under this setting is that one needs to solve for function $V(S_t,t)$ along with the unknown optimal exercise boundary Γ.

In order to get more compact form of the problem and clarified its setting, too, there is reasonable to introduce so called Black-Scholes differential operator $\mathcal{L}$ as follows

$$\mathcal{L}\, G(S,t) = \frac{\partial G}{\partial t} + rS\frac{\partial G}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 G}{\partial S^2} - rG = 0, \quad \text{for } (S,t) \in \Sigma = [0,T) \times \mathbb{R}_+, \tag{12}$$

which is defined by its application on any function $G(S,t) \in C^{2,1}(\Sigma)$.

Using (9a) and the B-S operator $\mathcal{L}$, we can simply write properties of function $V(S_t,t)$ in sub-region Σ_1 in following way

$$\mathcal{L}\, V(S_t,t) = 0, \quad V(S_t,t) > \max(K - S_t, 0)\,, \quad \forall\, (S_t,t) \in \Sigma_1. \tag{13}$$

Yet, we need to discuss the situation in the stopping sub-region Σ_2 assuming (9b) to hold. It means clearly that the function $V(S_t,t)$ is coincident with pay-off function $\max(K - S_t, 0)$ thus enabling direct computation of $\mathcal{L}\, V(S_t,t)$ in following way

$$\begin{gathered} V(S_t,t) = \max(K - S_t, 0) \;\; \Rightarrow \\ \mathcal{L}\, V(S_t,t) = \mathcal{L}(K - S_t) = -rS_t - rK + rS_t = -rK < 0, \quad \forall\, (S_t,t,) \in \Sigma_2. \end{gathered} \tag{14}$$

Hence, using (9b) and (14), we get properties of function $V(S_t,t)$ in sub-region Σ_2 as follows

$$\mathcal{L}\, V(S_t,t) = -rK < 0, \quad V(S_t,t) = \max(K - S_t, 0)\,, \quad \forall\, (S_t,t) \in \Sigma_2. \tag{15}$$

Since (15) holds on Γ, being tackled as a part of $\partial\Sigma_2$, too, we can conclude by direct computation that both the pricing function $V(S_t,t)$ and its first derivative with respect to S, i.e. $\partial V(S,t)/\partial S$, as well, are continuous on the optimal exercise boundary.

3 American option pricing problem – variational formulation

In [1] and [10], there is given a nice overview of theory and numerical solution of option pricing problems, too. Further, [2], [5], [6], and [12] bring more details as for the variational formulation of option pricing problems as well as some numerical algorithms for solving these problems. Other methods are in [9], and [3], which are focused upon penalty method and finite difference method for solving complementary problem arising in American option pricing, in particular. Our previous work on the topic is in [8].

Following [6], we will combine (13) holding in Σ_1 together with (15) holding in Σ_2, and we look for formulation of American put option problem on whole domain $\Sigma = \Sigma_1 \cup \Sigma_2 = \mathbb{R}_+ \times [0,T]$.

American put option problem formulated in form of variational inequality in strong sense, see [6], p.126, is following

$$\begin{gathered} \min\{-\mathcal{L}\, V(S_t,t),\; V(S_t,t) - \max(K - S_t, 0)\} = 0, \quad \forall\, (S_t,t) \in \Sigma\,, \\ V(S_T,T) = \max(K - S_T, 0)\,,\; S_T \in \mathbb{R}_+, \quad V(S,t) \to 0,\;\; S \to +\infty,\;\; \forall\, t \in [0,T]. \end{gathered} \tag{16}$$

Further, following [3], this formulation provides also another equivalent setting which is called linear complementary formulation, being motivated by well-known algebraic expression $a \cdot b = 0$, a, $b \geq 0 \Leftrightarrow (a = 0, b > 0) \cup (a > 0, b = 0)$. Combining (13) and (15) properly and adjoining (11b), it is given in following way

$$(-\mathcal{L}\, V(S_t,t))(V(S_t,t) - \max(K - S_t, 0)) = 0, \quad \forall\, (S_t,t) \in \Sigma\,,$$
$$-\mathcal{L}\, V(S_t,t) \geq 0, \quad V(S_t,t) - \max(K - S_t, 0) \geq 0, \quad \forall\, (S_t,t) \in \Sigma\,, \tag{17}$$
$$V(S_T,T) = \max(K - S_T, 0)\,,\ S_T \in \mathbb{R}_+, \quad V(S,t) \to 0,\ S \to +\infty,\ \forall\, t \in [0,T].$$

In order to get variational formulation of American put option problem in weak sense, we need to recast (16), or (17), respectively, into weak formulation. Following [1] and [10], we try to concentrate ourselves on the main steps in theory, not being burden with technicalities here.

The theory of variational formulations of parabolic PDEs is well known, and in [1], [5], and [10], in particular, there are given in much more details. The framework is particularly useful when classical, or strong, solution do not exist either because of some singularity in the data, or the domain boundary, or the coefficients, or nonlinearity, or both. Even when the boundary value problem for PDE has a classical solution, the variational theory is interesting for several reasons:

- it provides global estimates of solution;
- it is the most natural way to study obstacle problems, e.g. American option pricing;
- it has strong connections to numerical solution based upon finite element method, in particular.

There is well-known, variational formulations of parabolic PDE rely on suitable functional spaces, known as Sobolev spaces. Following [1], [5], and [10], we are going to mention the Sobolev space useful particularly for solving problems governed by B-S PDE (10), which is posed with single asset price variable S.

Let $L^2(\mathbb{R}_+)$ denotes the Hilbert space of square integrable functions on $\mathbb{R}_+$ endowed with the norm $\|.\|$, and the inner product (.,.) as usual

$$\| v \| = (\int_{\mathbb{R}_+} v(S)^2 \mathrm{d}S)^{\frac{1}{2}}, \quad (v,w) = \int_{\mathbb{R}_+} v(S) w(S) \mathrm{d}S\,. \tag{18}$$

Functional space which is well-suited to American put option pricing problem is

$$W = \left\{ v \in L^2(\mathbb{R}_+) \mid S \frac{\mathrm{d}v}{\mathrm{d}S} \in L^2(\mathbb{R}_+) \right\}, \tag{19}$$

where the derivative dv/dS has to be understood in the sense of distributions on $\mathbb{R}_+$. The space W being endowed with the norm $\| v \|_W = \sqrt{(v,v)_W}$ is a Hilbert space, as stated in [1], provided a natural scalar product in space W has the following form

$$(v,w)_W = (v,w) + \ (S\frac{\mathrm{d}v}{\mathrm{d}S}\,, S\frac{\mathrm{d}w}{\mathrm{d}S})\,. \tag{20}$$

Since (13) and (14) yield (17), i.e. $-\mathcal{L}\, V(S_t,t) \geq 0$, on Σ, we start with this PDE. Following [10], and making time substitution $\tau = T - t$, expressing time from current date t to expiration date T, actually, on opposite to t, expressing the time elapsed since the contract was pushed in action, we convert the terminal boundary value problem (16) into an initial boundary value problem

$$\frac{\partial u}{\partial \tau} - rS\frac{\partial u}{\partial S} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 u}{\partial S^2} + ru = 0,\ \text{for}\ (S,\tau) \in \Sigma = \mathbb{R}_+ \text{x} [0,T), \quad u(S,0) = u_0(S),\ \text{for}\ S \in \mathbb{R}_+, \tag{21}$$

where we just adopted to denote $u_0(S) = \max(K - S, 0)|_{\tau = T - T = 0}$.

In general, following steps are very usual when deriving weak formulation of problems within the framework of PDE theory. Let us simply multiply (12) by a smooth real valued function w defined on $\mathbb{R}_+$, and next, integrate the product integrand being constructed in the variable S on $\mathbb{R}_+$. Further, assuming the integration by parts is allowed, one obtains

$$\frac{\partial}{\partial \tau}(\int_{\mathbb{R}_+} u(S,\tau) w(S) \mathrm{d}S) \ + \ a_\tau(u,w) \geq 0\,, \tag{22}$$

where the bilinear form a_τ is defined by following expression

$$a_\tau(v,w) = \int_{\mathbb{R}_+} \left(\frac{1}{2}\sigma^2 S^2 \frac{\partial v}{\partial S}\frac{\partial w}{\partial S} + \ rvw\right) \mathrm{d}S \ + \ \int_{\mathbb{R}_+} \left(-r + \sigma^2 + \sigma S \frac{\partial \sigma}{\partial S}\right) S \frac{\partial v}{\partial S} w \ \mathrm{d}S\,. \tag{23}$$

Now, precise theory needs to bring some technical assumptions as prospectively allowing $\sigma(S,t)$ to be even time-dependent local volatility and not the constant only:

- function $\sigma(S,\tau)$ with its values to be squeezed within a layer defined by its bottom level and upper one given by a couple of positive constants for all $\tau \in [0, T]$, and $S \in \mathbb{R}_+$;
- all non-negative values of $|S\ \partial\sigma(S,\tau)/\partial S|$ to be bounded from up with another positive constant, in similar way as well.

Set of feasible solutions Q of American put option problem in weak sense is a closed convex set of space W assuming the boundary value function $u_0(S)$ being understood in sense of traces

$$Q = \{w \in W, w \geq u_0(S),\ \text{for}\ S \in \mathbb{R}_+\}. \tag{24}$$

Finally, all these assumptions maintain desired regularity of integrands in (23), which further lead to continuity and other desired properties of bilinear form $a_\tau(v, w)$ on W, which enable us to formulate weak form of the problem (16) precisely, by recalling Theorem 6.1 from [10], as follows

$$\begin{aligned}&\text{Find } u \in C^0([0, T]; L^2(\mathbb{R}_+)) \cap L^2(0,T; Q) \text{ with } \frac{\partial u}{\partial t} \in L^2(0,T; W'), \text{ and } u_{|t=0} = u_0 \text{ in } \mathbb{R}_+,\\ &\text{and for a.e. } t \in (0, T), \text{ holds } \left(\frac{\partial u}{\partial t}(t), w - u(t)\right) + a_t\big(u(t), w - u(t)\big) \geq 0,\ \forall w \in Q.\end{aligned} \tag{25}$$

4 Conclusions

Framework of both classical and variational formulations of American put option pricing problem was briefly discussed. Further, additive decomposition of American put option price is presented with corresponding details.

The optimal early exercise boundary must be determined together with pricing function by solving variational inequality numerically. Weak formulation of American option pricing problem leads to variational inequality.

Near future research will be focused on numerical solution of option pricing problems using finite element method with higher order polynomial shape functions and triangular elements with curved boundary for approximation of early exercise boundary.

Acknowledgements

The research project was supported by the grant no. 15-20405S of the Grant Agency, Prague, Czech Republic.

References

[1] Achdou, Y., Pironneau, O.: *Computational methods for option pricing*. Frontiers in Applied Mathematics, vol. 30, SIAM, Philadelphia, 2005.

[2] Feng, L., Kovalov, P., Linetsky, V., Marcozzi, M.: Chapter 7, Variational Methods in Derivatives Pricing. In: *Handbooks in OR & MS*. Vol. 15 (Birge, R. and Linetsky, V., eds.), Elsevier, New York, 2008.

[3] Feng, L., Linetsky, V., Morales, J.L., Nocedal, J.: On the Solution of Complementarity Problems Arising in American Options Pricing. *Optimization Methods and Software* **26 (4-5)** (2011), 813–825.

[4] Fries, Ch.: *Mathematical finance – theory, modeling, implementation*. John Wiley & Sons, Hoboken, New Jersey, 2007.

[5] Jaillet, P., Lamberton, D., Lapeyre, B.: Variational inequalities and the pricing American options. *Acta Applicandae Mathematicae* **21** (1990), 263–289.

[6] Jiang, L.: *Mathematical Modeling and Methods of Option Pricing*. World Scientific Publ., Singapore, 2005.

[7] Karatzas, I., Shreve, S.E.: *Methods of mathematical finance*. Springer-Verlag, New York, 1998.

[8] Lukáš, L.: Variational Formulation of Option Pricing Problem a Platform for Finite Element Method in Finance. In: *Conference Proceedings of 33rd Int. Conf. Mathematical Methods in Economics 2015* (Martinčík, D., Ircingová, J., and Janeček, P., eds.). University of West Bohemia, Pilsen, 2015, 467–472.

[9] Memon, S.: Finite element method for American option pricing: a penalty approach. *Int. J. of Numer. Analysis and Modeling, Series B* **3** (2012), 345–370.

[10] Pironneau, O., Achdou, Y.: Partial Differential Equations for Option Pricing. In: *Mathematical Modeling and Numerical Methods in Finance*. Special Volume (Bensoussan, A., Zhang, Q., eds.) of *Handbook of Numerical Analysis*. Vol. XV (Ciarlet, P.G., ed.), 369-495.

[11] Shiryaev, A.N.: *Essentials of Stochastic Finance – Facts, Models, Theory*. World Scientific Publ. Co., Singapore, 2008.

[12] Zhang, Ch.-S.: *Adaptive Methods For Variational Inequalities – Theory And Applications in Option Pricing*. Lambert Acad. Publ., Saarbruecken, 2010.

Rescue System Resistance to Failures in a Transport Network

Tomáš Majer[1], Stanislav Palúch[2]

Abstract. The rescue system is represented by few service centers (fire and rescue stations) and many customers (probable locations of first aid or fire), which are some nodes of a transportation network. Service centers were located so that each customer was served from his associated center with respect to the time limit. The unlikely events (excessive snow, black ice, accidents, traffic jams and so on) are increasing the traveling time between the customers and associated centers.

In this article we study structure of transport network. We want to find some critical ways (arcs in directed graph) which makes the current (specified) location of centers unusable. Several failures will be modeled by increasing the traveling time simultaneously on multiple ways. This will result not only to an increase of the traveling time between customer and his center, but also to change the route or even change the associated center.

An algorithm to find critical ways in transport network for specified service centers locations, based on Palúch's K-shortest path algorithm, will be presented.

Keywords: rescue system, reliability, location problem, shortest path.

JEL classification: R53
AMS classification: 90B06

1 Introduction

The rescue or emergency system is special case of public facility system. The task of designing the rescue system meant to place rescue stations to transport network nodes so that in case of requirements for attendance at any point, the time required to move from rescue station to the place of intervention did not exceed the legal limit. In other words, the time distance between each node of the transport network and the nearest rescue station must be less than the specified limit. In the case of rules recognized in the Slovak Republic within 15 minutes.

This task can be solved using some of the number of mathematical models designed in many scientific articles. The problem can be formulated as Set Covering Location Problem, Maximal Covering Location Problem, P-Median or other Location Problem.

We chose as the basis of our work model published by Janosikova in [1]. Other location problems are described in [2].

Let us denote the set of location candidates by symbol J. Its elements are indexed by symbol j. I states for the set of all municipalities in the region under consideration and i denotes a particular municipality. Both sets I and J correspond to nodes of a transport network.

$N_i \subset J$ is the set of potential locations which can cover municipality i. Bivalent variable $x_j \in \{0, 1\}$ models the decision if a station is located at node j ($x_j = 1$) or not ($x_j = 0$).

Further, constants A_i are the population of municipality i and constant P is the predefined number of stations to be deployed. The coverage of node i is conditional and depends on if a station is located

[1]University of Žilina, Faculty of management science and informatics, Departement of mathematical methods and operation research, tomas.majer@fri.uniza.sk.

[2]University of Žilina, Faculty of management science and informatics, Departement of mathematical methods and operation research, stanislav.paluch@fri.uniza.sk.

in its set N_i. This is modelled by decision variables $y_i \in \{0,1\}$. Variable y_i takes value 1 if at least one station is located within the time standard, otherwise $y_i = 0$.

Instead of full coverage of all demands, this model seeks the location of a fixed number of facilities in such a way that population covered by the service is maximized. The integer programming formulation of the problem is the following:

$$\text{Maximize} \qquad z = \sum_{i \in I} \mathrm{A}_i y_i \tag{1}$$

$$\text{subject to:} \qquad \sum_{j \in N_i} x_j \geq y_i \qquad \forall i \in I \tag{2}$$

$$\sum_{j \in J} x_j = P \tag{3}$$

$$x_j, y_i \in \{0,1\} \qquad \forall i \in I, j \in J \tag{4}$$

The set N_i depends on distances between nodes of transport network and can be defined as

$$N_i = \{j \mid d(i,j) \leq \mathrm{T}_{\max}\}, \tag{5}$$

where $\mathrm{T}_{\max}$ is time limit and $d(i,j) = \mathrm{D}_{ij}$ is traveling time between nodes i and j – the length of the fast i–j path, fixed in this model.

We do not want to seek location of centers in this paper, but to examine the impact of traffic restrictions on the quality of coverage. So the locations of centers are fixed in our work and we expect that $|J| = P$ and $x_j = 1$ for all $j \in J$.

Let transport network be modeled using undirected edge weighted graph $G = (V, E, c)$, where V is set of vertices or nodes of transport network, E is set of edges or streets in transport network and $c: E \to \mathbb{R}$ is cost function and denotes traveling time over every edge. Of course, $I \subset V$ and $J \subset V$. Then time distance or traveling time between vertices i and j is the length of shortest path $\mu(i,j)$ – the sum of cost of all edges on this shortest path

$$d(i,j) = \sum_{e \in \mu(i,j)} c(e) = \sum_{e \in \mu(i,j)} \mathrm{C}_e. \tag{6}$$

Due to adverse circumstances (weather, traffic jams, accidents, etc.) the time needed to move from rescue station to the place of intervention may be extended. We will model this extension separately on each edge, so the length of not necessary shortest path $\mu(i,j)$ will be

$$\overline{d}(i,j) = \sum_{e \in \mu(i,j)} \overline{c}(e) = \sum_{e \in \mu(i,j)} \mathrm{C}_e + \sum_{e \in \mu(i,j)} t_e, \tag{7}$$

where t_e is latency on edge e. These delays may cause:

- increase of the length of shortest path,
- change the shortest path (used edges),
- change the associated rescue station.

If the length of the shortest path to nearest station exceeds time limit $\mathrm{T}_{\max}$ then municipality i will be not covered and N_i will be empty set. We want to find worst cases of minimal failures in the transport network, where the number of municipalities covered by drops below a critical value.

Authors in [4] present another model, in which failure of edge makes the edge useless. So affected edges are then deleted from the graph of transportation network and shortest paths are recalculated.

2 Mathematical model

We want to find a small failure on edges with catastrophic impact on service coverage. So the objective function (8) is minimizing the sum of latencies t_e on edges of transport network.

Each municipality can be served from any rescue station when length of any path between municipality and station satisfies time criteria. Set K_{ij} represents the number of those paths between municipality i and station j. Binary variable $r_{ijk} = 1$ if kth path between i and j is within time limit and $r_{ijk} = 0$ if kth path is longer as time limit. This is ensured by conditions (9) and (10).

Conditions (11) and (12) ensures that $y_i = 1$ if and only if there is at least one path which satisfies time criteria.

Condition (13) prescribes the size of a service outage. $\mathrm{Z}_{\text{critical}}$ denotes the maximum permissible number of people who may not be served by the specified time limit.

$$\text{Minimize} \qquad t = \sum_{e \in E} t_e \tag{8}$$

$$\text{subject to:} \qquad (1+\mathrm{M})r_{ijk} + \sum_{e \in \mu_k(i,j)} (\mathrm{C}_e + t_e) \leq 1 + \mathrm{T}_{\max} + \mathrm{M} \qquad \forall i \in I, j \in J, k \in K_{ij} \tag{9}$$

$$\mathrm{M}r_{ijk} + \sum_{e \in \mu_k(i,j)} (\mathrm{C}_e + t_e) \geq 1 + \mathrm{T}_{\max} \qquad \forall i \in I, j \in J, k \in K_{ij} \tag{10}$$

$$y_i \leq \sum_{j \in J} \sum_{k \in K_{ij}} r_{ijk} \tag{11}$$

$$\mathrm{M}y_i \geq \sum_{j \in J} \sum_{k \in K_{ij}} r_{ijk} \tag{12}$$

$$\sum_{i \in I} \mathrm{A}_i y_i \leq \mathrm{Z}_{\text{critical}} \tag{13}$$

$$y_i, r_{ijk} \in \{0,1\} \qquad \forall i \in I, j \in J, k \in K_{ij} \tag{14}$$

$$t_e \in \mathbb{R}_0^+ \qquad \forall e \in E \tag{15}$$

Constant M is large positive integer used for produce binary variables. It must be greater than the longest path used.

3 Example

Let transport network be represented by graph $G = (V, E)$, where $V = \{1, 2, 3, 4, 5, 6\}$, $E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{3, 4\}, \{4, 5\}, \{4, 6\}, \{5, 6\}\}$. Diagram of network with edge length is shown in Figure 1.

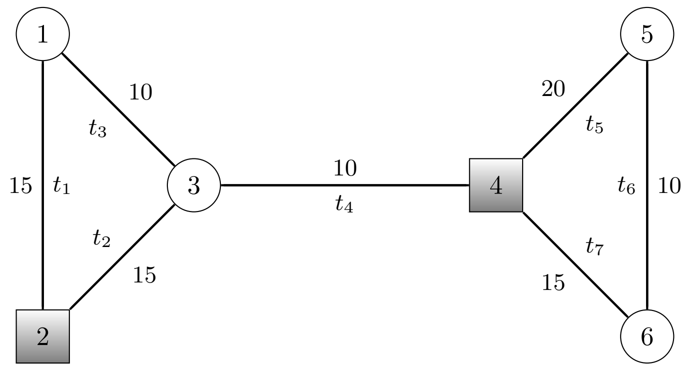


Figure 1 Diagram of transport network

Set of municipals $I = \{1, 3, 5, 6\}$ and set of stations $J = \{2, 4\}$. We choose $\mathrm{T}_{\max} = 30$, $\mathrm{Z}_{\text{critical}} = 2$ and $\mathrm{M} = 999$.

Paths between municipals and rescue stations are shown on Table 1. Some paths are not needed in the model, since they contain all the edges of other path. In example path $5-4-3-2$ is not needed because municipal 5 can be served from station 4 by path $5-4$ which is in any condition shorter than path $5-4-3-2$.

Model after substitution of input values is:

$$
\begin{array}{rl}
\text{Minimize} & t = t_1 + t_2 + t_3 + t_4 + t_5 + t_6 + t_7 \\
\text{subject to:} & 1000r_{1,2,1} + 15 + t_1 \leq 1030 \\
& 999r_{1,2,1} + 15 + t_1 \geq 31 \\
& 1000r_{1,2,2} + 25 + t_2 + t_3 \leq 1030 \\
& 999r_{1,2,2} + 25 + t_2 + t_3 \geq 31 \\
& 1000r_{1,4,1} + 20 + t_3 + t_4 \leq 1030 \\
& 999r_{1,4,1} + 20 + t_3 + t_4 \geq 31 \\
& 1000r_{3,2,1} + 15 + t_2 \leq 1030 \\
& 999r_{3,2,1} + 15 + t_2 \geq 31 \\
& 1000r_{3,2,2} + 25 + t_1 + t_3 \leq 1030 \\
& 999r_{3,2,2} + 25 + t_1 + t_3 \geq 31 \\
& 1000r_{3,4,1} + 10 + t_4 \leq 1030 \\
& 999r_{3,4,1} + 10 + t_4 \geq 31 \\
& 1000r_{5,4,1} + 20 + t_5 \leq 1030 \\
& 999r_{5,4,1} + 20 + t_5 \geq 31 \\
& 1000r_{5,4,2} + 25 + t_6 + t_7 \leq 1030 \\
& 999r_{5,4,2} + 25 + t_6 + t_7 \geq 31 \\
& 1000r_{6,4,1} + 15 + t_7 \leq 1030 \\
& 999r_{6,4,1} + 15 + t_7 \geq 31 \\
& 1000r_{6,4,2} + 30 + t_5 + t_6 \leq 1030 \\
& 999r_{6,4,2} + 30 + t_5 + t_6 \geq 31 \\
& y_1 \leq r_{1,2,1} + r_{1,2,2} + r_{1,4,1} + r_{1,4,2} \\
& 999y_1 \geq r_{1,2,1} + r_{1,2,2} + r_{1,4,1} + r_{1,4,2} \\
& y_3 \leq r_{3,2,1} + r_{3,2,2} + r_{3,4,1} \\
& 999y_3 \geq r_{3,2,1} + r_{3,2,2} + r_{3,4,1} \\
& y_5 \leq r_{5,4,1} + r_{5,4,2} \\
& 999y_5 \geq r_{5,4,1} + r_{5,4,2} \\
& y_6 \leq r_{6,4,1} + r_{6,4,2} \\
& 999y_6 \geq r_{6,4,1} + r_{6,4,2} \\
& y_1 + y_3 + y_5 + y_6 \leq 2
\end{array}
$$

We solve this model using *Gurobi MILP Solver* and we get the result that municipalities 5 and 6 are not covered ($y_5 = 0$ and $y_6 = 0$) by traffic restriction on edges $\{4,5\}$ and $\{4,6\}$ ($t_5 = 11$ and $t_7 = 16$).

If we fix in model variable t_5 to zero, we get the result that municipalities 1 and 6 are not covered ($y_1 = 0$ and $y_6 = 0$) by traffic restriction on edges $\{1,2\}$, $\{1,3\}$, $\{3,4\}$, $\{5,6\}$ and $\{4,6\}$ ($t_1 = 16$, $t_3 = 6$. $t_4 = 5$, $t_6 = 1$ and $t_7 = 16$).

4 Conclusion

We show in this paper, how to identify bottlenecks in transport infrastructure by given locations of emergency services. It is important to proove this model and his effectivity on larger transport networks. We want to use modified Palúch's multilabel K–shortest path algorithm [3] to enumerate all paths which satisfy time criteria and automatically generate conditions (9) and (10).

Path	Length
1–2	$15 + t_1$
1–3–2	$25 + t_2 + t_3$
1–3–4	$20 + t_2 + t_4$
3–2	$15 + t_2$
3–1–2	$25 + t_1 + t_3$
3–4	$10 + t_4$
5–4	$20 + t_5$
5–6–4	$25 + t_6 + t_7$
6–4	$15 + t_7$
6–5–4	$30 + t_5 + t_6$

Table 1 Paths between municipals and stations

Acknowledgements

The authors are pleased to acknowledge the financial support of the Scientific Grant Agency of the Slovak Republic VEGA under the grant No. 1/0518/15.

References

[1] Jánošíková Ľ.: Emergency Medical Service Planning. *Communications* **2** (2007), 62–66.

[2] Laporte G., Nickel S., Gama F.S.: *Location Science*. Springer, Switzerland, 2015.

[3] Palúch S.: A Multi Label Algorithm for k Shortest Paths Problem. *Communications* **3** (2009), 11–14.

[4] Yuangang P., Yali D., Zongtian D.: Reliable Facility System Design Subject to Edge Failures. *American Journal of Operations Research* **4** (2014), 164–172.

Evaluation of Parametric ES Tests

Marta Małecka[1]

Abstract. Since the beginning of the 21st century the theory of statistics offers a consistent, coherent methodology of risk measurement, based on the axiomatic risk definition. The central point of the axiomatic risk theory is the idea of the Expected Shortfall (ES), which is a risk measure that fulfils the set of coherency axioms. Since the inception of ES on the turn of the 21st century, numerous estimation methods has been developed and discussed, while testing procedures are still relatively few. As sample ES measure is calculated as the average value over a given threshold, the relevant distribution is unknown. Thus statistical evaluation of the ES model cannot use the natural measure of discrepancy between estimated and empirical ES. Hypothesis testing uses other means like regression approach, saddlepoint technique for approximate p-values or the goodness of fit of the truncated return density. The study presents parametric methods of statistical inference connected with ES measure and, through the simulation study, gives the comparison of the size and power of the considered tests. In order to reflect the stylized facts about real financial processes, simulation experiments are based on the GARCH processes.

Keywords: expected shortfall, ES test, test size, test power.

JEL Classification: C22, C52, D53
AMS Classification: 62M10, 91B84, 62P05

1 Introduction

Inception of the axiomatic risk definition on the turn of the 21st century gave grounds for the development of the consistent theory of risk measurement [2]. The new risk concept and accompanying coherent risk measures, especially expected shortfall (ES), being the central point in the theory, stimulated formulation of new models. Particularly broad discussion is connected with estimation of ES, whose idea is to inform about average loss in case of extreme events, defined as values reaching over a given threshold. Dynamic development and wide variety of ES-based risk models created the need for relevant testing procedures. However in the general case, the distribution of a sample average of extreme observations is unknown, thus classic statistical methods are unfeasible for ES value testing. Since scarcity of observations is inherent to extreme events, the statistical inference cannot be based on the central limit theorem either.

Since the beginning of the 21^{st} century several approaches have been proposed for ES model backtesting. One idea was to compare empirical and estimated return distribution tail in the likelihood ratio test, based on the censored normal likelihood [3]. Another ES-testing procedure circumvented the problem of the unknown distribution through the saddlepoint technique, which gives approximate p-values based on the Taylor expansion of the moment generating function [10]. Finally the regression-based approach, using the standard Fischer statistic, was proposed [5].

The aim of the paper was to provide a comprehensive comparative analysis of parametric ES backtesting procedures. The paper presents evaluation of statistical properties of the considered tests as well as discusses practical aspects of their application in real-life situations. Formal test assessment included their size and power and was conducted through the Monte Carlo method. The analysis of the test properties and their practical utility was preceded by the overview of statistical inference methods proposed in the literature for ES models. The study of the regression-based approach was complemented by the analysis of the modification of Christoffersen's statistic, aimed at achieving approximate stationarity of the error term in the linear regression.

The paper is comprised of four sections. In the second section we provided the definition of the expected shortfall and presented available testing procedures dedicated to ES model backtesting. We also introduced a modification of the regression-based ES test. This was followed by the exploration of the size and power properties through the Monte Carlo study in the third section. The final section summarizes and concludes the paper.

[1] University of Łódź, Department of Statistical Methods, 41/43 Rewolucji 1905 r., Łódź 90-214, marta.malecka@uni.lodz.pl.

2 ES backtesting methods

Let us consider the random variable X defined on $(\Omega, \mathcal{F}, P)$, such that $E(\max\{0,-X\}) < \infty$. Let $p \in (0,1)$ be a fixed real number. The definition of the expected shortfall reflects the idea of representing the average loss in case of extreme events and is given by

$$ES_p(\mathrm{X}) = -\frac{1}{p}\Big(E(X\mathbf{1}_{\{X \le q^p(X)\}}) - q^p(X)\big(P(X \le q^p(X)) - p\big)\Big), \tag{1}$$

where p denotes the level of tolerance and $q^p(\mathrm{X})$ is the upper p-quantile of X, given by $q^p(X) = \inf\{x \in \mathbf{R} : P(X \le x) > p\}$ [1]. In the continuous case the above definition reduces to the conditional expectation $ES_p(X) = E(-X \mid X \le F^{-1}(\mathrm{X}))$, where F denotes the distribution function of X.

ES tests proposed in recent literature use various ways to circumvent the problem of unknown distribution of the sample ES and verify risk models through different aspects (another way to manage the problem of unknown distribution is to use bootstrap methods described eg. in [11]). The most restrictive procedure is the exception magnitude test, which verifies the goodness-of-fit of the return distribution tail [3]. Firstly it uses the Rosenblatt transformation to convert return data into normal random variables (various possibilities to construct the rate of return and average return in both discrete and continuous time are discussed in [4]). Then it employs the censored normal likelihood to test for the discrepancies in the distribution tails through the likelihood ratio statistic. The goodness of fit in the likelihood ratio framework is checked only through the chosen distribution moments. In particular if we consider the null $H_0 : \mu = 0,\ \sigma = 1$ tested against the alternative $H_0 : \mu \ne 0 \ \vee\ \sigma \ne 1$ the likelihood ratio statistic takes the form

$$LR_B^{ES} = -2(\log L(0,1,X_1^*,X_2^*,...,X_{T*}^*) - \log L(\hat{\mu},\hat{\sigma},X_1^*,X_2^*,...,X_{T*}^*) \tag{2}$$

where μ and σ are the expectation and the standard deviation of the return variable X_t and $X_1^*,X_2^*,...,X_{T*}^*$ denote tail observations exceeding a given threshold [3].

Another parametric ES testing procedure is based on the saddlepoint technique, which allows for calculating approximate p-values for the sum of random variables [10]. It uses Taylor expansion of moment and cumulant generating functions. Originally this method was proposed for the iid normal series, however, through the Rosenblatt transformation, it is possible to use it for a general class of stochastic processes.

If $\bar{X}$ is the sample average of the tail observations $X_1^*,X_2^*,...,X_{T*}^*$, then through the Taylor expansion of the moment generating function it can be shown that

$$P(\bar{X} \le \bar{x}) = \begin{cases} \Phi(\xi) - \phi(\xi)\big(\frac{1}{\eta} - \frac{1}{\xi} + \mathbf{O}(T^{*-\frac{3}{2}})\big) & \text{for } \bar{x} < q_p, \\ 1 & \text{for } \bar{x} \ge q_p, \end{cases} \tag{3}$$

where $\eta = s\sqrt{T^* K_{X'}(s)},\ \xi = sgn(s)\sqrt{2T^*(s\bar{x} - K_X(s))},\ \bar{x} \ne \mu_x$ and s is the saddlepoint satisfying $K_{X'}(s) = \bar{x}$ [6]. In case of normally distributed variable the saddlepoint s can be obtained as a solution to the equation

$$K_{X'}(t) = \frac{M_{X'}(t)}{M_X(t)} = t - e^{\frac{t^2}{2}}\frac{\phi(q_p) - t}{\Phi(q_p) - t} = \bar{x}. \tag{4}$$

The third ES test proposed in the recent literature uses linear regression to verify the potential of additional available random variables to explain ES value. Let us consider the regression

$$-X_{t+1}^* - ES_p(X_{t+1}) = a + b\mathbf{X}_t + \epsilon_{t+1}, \tag{5}$$

where ε_t are iid for $t = 1,...,T$ and $\mathbf{X}_t$ denotes the set of explanatory random variables available at time t. The test hypothesis is formulated as $H_0 : a = 0,\ H_0 : b = 0$ or jointly as $H_0 : a = b = 0$ and can be verified by the standard Fischer statistic, denoted here as F_{Ch} [5].

In the general case, the X_t and $ES_p(X_t)$ variables, are stochastic processes where the variable distributions change over time. In particular, when X_t represents a rate of return from a financial variable, due to volatility clustering it is characterized by a time-varying variance. In the case of the non-constant variance, the stationarity assumption about the error term ε_t is not satisfied, therefore statistical inference based on the regression (5) may involve serious type-one error. To reduce the error we considered the standardization of the dependent variables in the regression (5) with respect to the standard deviation, using its estimate $\hat{\sigma}_t$. After the standardization, the underlying regression takes the form

$$\frac{-X_{t+1}^{*} - ES_p(X_{t+1})}{\hat{\sigma}_t} = a^{*} + b^{*}\mathbf{X}_t + \epsilon_{t+1}^{*}, \tag{6}$$

and ϵ_t^{*} is the approximately stationary variable. As in case of F_{Ch}, the modified test F_{Ch}^{*}, based on the regression (6), uses the Fischer statistic.

3 Monte Carlo test evaluation

The size and power evaluation experiments were designed in a way that they reflected volatility clustering phenomenon, which hinders volatility prediction and is commonly regarded as a key issue in risk control. Volatility clustering was represented through inclusion of a GARCH process in the data generating algorithm. For the ES test size assessment the X_t values were generated from the GARCH(1,1) process:

$$\begin{aligned} X_t &= \sigma_t Z_t, \quad Z_t \sim N(0,1), \\ \sigma_t^2 &= \omega_1 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \end{aligned} \tag{7}$$

with parameters $\omega_1 = 0,05,\ \alpha = 0,14,\ \beta_1 = 0,85$ (the parameter values were fixed on the basis of the initial study for six stock market indices [7]). ES values were calculated as sample averages of the $p-$tail values of X_t, $t = 1,2,...,T$ (more about the estimation of ES value can be found in recent literature [8, 9]).

For the power comparison, we assumed the data generating process in the form of the GARCH(1,1) model given by (7), while ES estimates were obtained from processes with incorrect parameters. The ES values were generated from the model with systematically underestimated volatility. We used GARCH(1,1) with parameters chosen in such a way that we obtained the standard deviation on levels $0.9\sigma_t$, $0.7\sigma_t$ and $0.5\sigma_t$, where σ_t denotes the true parameter value.

The explanatory variables in the F_{Ch} and F_{Ch}^{*} tests included five lags of X_t and past five ES predictions. Since the testing procedures LR_B^{ES} and F_{Ch}^{*} presented in the paper are based on asymptotic distributions, the Monte Carlo test technique was employed for the power comparison. Based on simulated distributions, it provided the empirical quantiles for a given finite sample size and guaranteed the exact test sizes. Hence the power estimates were comparable among different tests.

The tests were conducted for the 5% significance level. The rejection frequencies were computed for sample sizes $T = 250, 500, 750, 1000$ over 10000 replications.

The results of the Monte Carlo study showed that the tests LR_B^{ES} and S were comparable in terms of the test size. The rejection frequencies obtained under the null for both procedures were similar and close to the desired nominal level of 5% [Tab. 1].

The power study, based on the rejection rates under the alternative, showed that the LR_B^{ES} test outperformed the S procedure in the ability to detect the incorrect ES models [Tab. 2]. The LR_B^{ES} statistic was characterized by the rejection frequencies of over 60% in all variants of Monte Carlo experiments, which showed its excellent ability to detect even minor departures from the null. Thus the Berkowitz test LR_B^{ES} emerged as a very restrictive procedure, which verifies goodness of fit of the distribution tails through chosen moments of the return distribution. However its construction requires the estimate of the distribution function, which largely reduces its practical utility. It is moreover restricted to normally distributed variables.

Test	Series length			
	250	500	750	1000
LR_B^{ES}	0.041	0.046	0.052	0.051
S	0.053	0.050	0.057	0.048
F_{Ch}	0.121	0.203	0.265	0.309
F_{Ch}^*	0.050	0.050	0.049	0.049

Table 1 Size estimates of ES tests

The S test power performance was worse than LR_B^{ES} for small departures from the null, with the rejection frequencies only slightly exceeding 30%. However the power estimates showed dynamic growth with lengthening the series, reaching over 50% for 500 observations and 70% for 1000 observations. The test also turned out powerful for larger departures from the null. For undersized volatility of 70% or 50% of its original value, power estimates were over 50% and grew fast with increasing the number of observations.

Test	σ_t^*	Series length			
		250	500	750	1000
LR_B^{ES}	$0,9\sigma_t$	0.64	0.80	0.88	0.92
	$0,7\sigma_t$	0.67	0.83	0.91	0.94
	$0,5\sigma_t$	0.93	0.99	0.99	1.00
S	$0,9\sigma_t$	0.35	0.53	0.63	0.70
	$0,7\sigma_t$	0.54	0.74	0.85	0.92
	$0,5\sigma_t$	0.88	0.97	0.99	1.00
F_{Ch}^*	$0,9\sigma_t$	0.05	0.05	0.05	0.05
	$0,7\sigma_t$	0.06	0.04	0.05	0.05
	$0,5\sigma_t$	0.06	0.04	0.05	0.05

Table 2 Power estimates of ES tests

The size estimates for the regression-based F_{Ch} test showed that the procedure is burdened with a large type-one error, which may be the result of non-stationarity of the error term in the linear regression [Tab. 1]. The rejection frequencies for the F_{Ch} test more than doubled the nominal level of 5%. Moreover the results gave no evidence of convergence to the theoretical Fischer distribution. The corrected version of the Christoffersen's test statistic F_{Ch}^*, which ensured the approximate stationarity of the error term, gave the desired rejection rates of around 5% in the size exercise. However, according to the power results, the modified regression was not capable of explaining the differences between the true and the incorrect ES value [Tab. 2]. The standardization procedure, conducted with respect to the standard deviation, guaranteed the stationarity of the dependent variables and error term, at the same time reducing the explanatory value of the regression. The obtained rejection frequencies of below 10% showed that the test had no power of detecting incorrect ES estimates.

4 Conclusion

The study presented in the paper was dedicated to evaluation of statistical properties of the parametric ES tests. The comparative analysis of three considered procedures indicated satisfactory statistical properties of the Berkowitz LR_B^{ES} and saddlepoint S test statistics. While LR_B^{ES} procedure turned out superior in terms of the power against small departures from the null, the S technique offers advantages connected with practical applicability of the test.

Acknowledgements

The research was supported by the Polish National Science Centre grant DEC-2013/11/N/HS4/03354.

References

[1] Acerbi C., Tasche D.: On the coherence of Expected Shortfall. *Journal of Banking and Finance* **26**, 7 (2002), 1487–1503.
[2] Artzner, P., Delbaen, F., Eber, J. M., and Heath, D.: Coherent measures of risk. *Mathematical Finance* **9**, 3 (1999), 203–228.
[3] Berkowitz, J.: Testing density forecasts with applications to risk management. *Journal of Business and Economic Statistics* **19** (2001), 465–474.
[4] Białek, J.: Average rate of return of pension or investment funds based on original, stochastic and continuous price index. In: *Proceedings of the 31st International Conference Mathematical Methods in Economics 2013* (Vojáčková, H., eds.). College of Polytechnics Jihlava, Jihlava, 2013, 37–42.
[5] Christoffersen, P. F.: *Elements of Financial Risk Management*. (2nd ed.). Elsevier Inc., Oxford, 2012.
[6] Lugannani, R., and Rice, S. O.: Saddlepoint approximation for the distribution of the sum of independent random variables. *Advanced Applied Probability* **12** (1980), 475–490.
[7] Małecka, M.: Prognozowanie zmienności indeksów giełdowych przy wykorzystaniu modelu klasy GARCH. *Ekonomista* **6** (2011), 843–860.
[8] Pietrzyk, R.: Szacowanie miary zagrożenia expected shortfall dla wybranych instrumentów polskiego rynku kapitałowego. *Inwestycje finansowe i ubezpieczenia – tendencje światowe a polski rynek* 1037 (2004), 118–127.
[9] Trzpiot, G.: *Wielowymiarowe metody statystyczne w analizie ryzyka inwestycyjnego*. Polskie Wydawnictwo Ekonomiczne, Warszawa, 2010.
[10] Wong, W.: Backtesting trading risk of commercial banks using expected shortfall. *Journal of Banking and Finance* **32**, 7 (2008), 1404–1415.
[11] Żądło, T.: On parametric bootstrap and alternatives of MSE. In: *Proceedings of the 31st International Conference Mathematical Methods in Economics 2013* (Vojáčková, H., eds.). College of Polytechnics Jihlava, Jihlava, 2013, 1081–1086.

Evaluating the ability to model the heavy tails of asset returns of two alternative distributions

Jiří Málek[1] , Quang Van Tran [2]

Abstract. Financial asset returns tend to be distributed with heavier tails than the normal one does which makes the normal less suitable for characterizing returns of financial instruments. Alpha stable distribution and generalized hyperbolic distribution are suggested to be a good replacement for the normal distribution. In this paper we examine and compare the ability of these two distributions and the special cases of the generalized hyperbolic distribution (hyperbolic distribution and normal inverse Gaussian distribution) to model the distinctive feature of asset returns. For this purpose, first we use maximum likelihood estimation method to estimate parameters of these distributions for daily returns of currency pairs CZK/EUR and CZK/USD. After that a global comparison of whole distributions as well as their selected quantiles are performed to assess how accurately they can approximate returns distribution as the whole and in segments.

Keywords: heavy tails, alpha stable distribution, generalized hyperbolic distribution, Czech financial assets

JEL classification: 91B82, 91B28
AMS classification: G10, G120

1 Introduction

There are two types of distributions, which are currently used for modeling of financial returns. They are stable distribution and generalized hyperbolic distribution. Both types have their advantages and disadvantages. The stable distribution has a very interesting property that the sum of several random variables from the same distribution is a stable distribution again. In addition, according to the Generalized Central Limit Theorem these distributions may be the standardized limit of the sum of independent identical distributed random variables (with the same tail index). On the other hand, modeling tail behavior is often a difficult task as the distribution does not have the variance and some even the mean value. Further, there is no explicit formula for the probability density function and the only known form is its characteristic function, which makes parameter estimation even more complicated. The alternative to stable distribution is generalized hyperbolic distribution and its variants. They are not generally closed under convolution (the sum of the two GHD may not be the GHD). But their advantage is that their tails are "semi-heavy" and thus it seems to be more appropriate for modeling financial assets returns. Moreover, their probability density function exists with finite mean and variance. In this paper, we will examine these properties on returns of two exchange rate of daily series USD/CZK and EUR/CZK from 2003-10 to 2016-4.

2 Alpha stable distribution

The standard alpha stable distribution is characterized by its characteristic function $\phi_0(t)$

$$\log \phi_0(t) = \begin{cases} -\sigma^\alpha \mid t \mid^\alpha \{1 + i\beta \text{sign}(t) \tan \frac{\alpha\pi}{2} [(\sigma \mid t \mid)^{(1-\alpha)} - 1]\} + i\mu_0 t, \ \alpha \neq 1 \\ -\sigma \mid t \mid \{1 + i\beta \text{sign}(t) \frac{2}{\pi} \log(\sigma \mid t \mid)\} + i\mu_0 t, \ \alpha = 1 \end{cases} \tag{1}$$

[1]University of Economics in Prague, Department of Banking and Insurance, nám. W. Churchilla 4, Praha 3, Czech Republic, GMalek@post.cz
[2]University of Economics in Prague, Department of Statistics and Probability, nám. W. Churchilla 4, Praha 3, Czech Republic, tran@vse.cz

There are four parameters $\alpha, \beta, \mu, \sigma$ included in this characteristic function. α, β are shape parameters, where α is the tail power as α decreases tail thickness increases and β is skewness parameter. The meaning of two remaining parameters is: μ is location parameter and σ is scale parameter. As the density function of the stable distribution is not generally known in the explicit form, the integral expression [5], [3] is often used as an alternative. After substitution $\zeta = -\beta \tan \frac{\pi\alpha}{2}$ the density of standard α - stable random variable ($\mu = 0, \sigma = 1$), can be expressed as:

- when $\alpha \neq 1$ and $x \neq \zeta$

$$f(x;\alpha,\beta) = \frac{\alpha(x-\zeta)^{\frac{1}{1-\alpha}}}{\pi \mid \alpha - 1 \mid} \int_{-\theta_0}^{\frac{\pi}{2}} V(\theta;\alpha,\beta) \exp\{-(x-\zeta)^{\frac{\alpha}{\alpha-1}} V(\theta;\alpha,\beta)\} \mathrm{d}\theta, \tag{2}$$

for $x > \zeta$ and $f(x;\alpha,\beta) = f(-x;\alpha,-\beta)$ for $x < \zeta$.

- when $\alpha \neq 1$ and $x = \zeta$

$$f(x;\alpha,\beta) = \frac{\Gamma\left(1+\frac{1}{\alpha}\right)\cos(\xi)}{\pi(1+\zeta^2)^{\frac{1}{2\alpha}}}, \tag{3}$$

- when $\alpha = 1$

$$f(x;1,\beta) = \begin{cases} \frac{1}{2|\beta|} \exp(\frac{x\pi}{2\beta}) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V(\theta;1,\beta) \exp\{-\exp(\frac{x\pi}{2\beta}) V(\theta;1,\beta)\} \mathrm{d}\theta, & \beta \neq 0 \\ \frac{1}{\pi(1+x^2)}, \ \beta = 0 & \text{where} \end{cases} \tag{4}$$

$$\xi = \begin{cases} \frac{1}{\alpha}\arctan(-\zeta), & \alpha \neq 1 \\ \frac{\pi}{2}, & \alpha = 1 \end{cases} \quad \text{and}$$

$$V(\theta;\alpha,\beta) = \begin{cases} (\cos(\alpha\xi))^{\frac{1}{\alpha-1}} \left(\frac{\cos(\theta)}{\sin(\alpha(\xi+\theta))}\right)^{\frac{\alpha}{\alpha-1}} \frac{\cos[\alpha\xi+(\alpha-1)\theta]}{\cos(\theta)} \\ \frac{2}{\pi}\left(\frac{0.5\pi+\beta\theta}{\cos(\theta)}\right) \exp\{\frac{1}{\beta}(0.5\pi+\beta\theta)\tan(\theta)\} \end{cases}$$

The cumulative distribution function $F(x;\alpha,\beta)$ of a standard stable random variable (in representation S^0) is

- when $\alpha \neq 1$ and $x \neq \zeta$

$$F(x;\alpha,\beta) = c_1(\alpha,\beta) + \frac{\mathrm{sign}(1-\alpha)}{\pi} \int_{-\zeta}^{\frac{\pi}{2}} \exp\{-(x-\zeta)^{\frac{\alpha}{\alpha-1}} V(\theta;\alpha,\beta)\} \mathrm{d}\theta \tag{5}$$

if $x > \zeta$ and $F(x;\alpha,\beta) = 1 - F(-x;\alpha,-\beta)$ if $x < \zeta$, where

$$c_1(\alpha,\beta) = \begin{cases} \frac{1}{\pi}\left(\frac{\pi}{2}-\xi\right) & \text{if } \alpha < 1 \\ 1 & \text{otherwise} \end{cases}$$

- when $\alpha \neq 1$ and $x = \zeta$

$$F(x;\alpha,\beta) = \frac{1}{\pi}\left(\frac{\pi}{2}-\xi\right), \tag{6}$$

- when $\alpha = 1$

$$F(x;1,\beta) = \begin{cases} \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp\{-\exp(-\frac{x\pi}{2\beta})V(\theta;1,\beta)\}\mathrm{d}\theta & \text{if } \beta > 0 \\ \frac{1}{2}+\frac{1}{\pi}\arctan(x) & \text{if } \beta = 0 \\ 1 - F(x;1,-\beta) & \text{if } \beta < 0 \end{cases} \tag{7}$$

These pdf and cdf functions are used for the estimation of parameters of this distribution.

3 Generalized hyperbolic distribution

The generalized hyperbolic distribution (GHD) is characterized by five parameters $\theta = (\lambda, \alpha, \beta, \delta, \mu)$ and its probability density function is

$$f_{\mathrm{GH}}(x;\theta) = \kappa[\delta^2 + (x-\mu)^2]^{\frac{1}{2}(\lambda - \frac{1}{2})} K_{\lambda-\frac{1}{2}} \left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \exp(\beta(x-\mu)), \tag{8}$$

where $\kappa = \frac{(\alpha^2-\beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^\lambda K_\lambda\left(\delta\sqrt{\alpha^2-\beta^2}\right)}$, and K_λ is the modified Bessel function of the third kind with index λ. All moments of a random variable from a generalized hyperbolic distribution exist and the first two of them are

$$\mathbb{E}(X) = \mu + \frac{\beta\delta^2}{\zeta}\frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} \tag{9}$$

$$\mathrm{Var}(X) = \delta^2 \left\{ \frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)} + \left(\frac{\beta\delta}{\zeta}\right)^2 \left[\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \left(\frac{K_{\lambda+1}(\zeta)}{\zeta K_\lambda(\zeta)}\right)^2\right]\right\} \tag{10}$$

where $\zeta = \delta\sqrt{\alpha^2 - \beta^2}$.

When $\lambda = -\frac{1}{2}$, we get the first special case called the normal inverse gaussian distribution (NIG) and its density becomes

$$f_{\mathrm{NIG}}(x;\theta) = \frac{\alpha\delta}{\pi}\exp[\delta\sqrt{\alpha^2-\beta^2} + \beta(x-\mu)]\frac{K_1(\alpha\sqrt{\delta^2 + (x-\mu)^2})}{\sqrt{\delta^2+(x-\mu)^2}}. \tag{11}$$

If $\lambda = 1$, it becomes the second special case called the hyperbolic distribution (HD) and its density is

$$f_{\mathrm{H}}(x;\alpha,\beta,\delta,\mu) = \frac{\sqrt{\alpha^2-\beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2-\beta^2})}\exp(-\alpha\sqrt{\delta^2+(x-\mu)^2} + \beta(x-\mu)) \tag{12}$$

The fat tail property of the GHD and the stable distribution comes from the fact that for the stable distribution: $P(X < x) \approx c_\alpha|x|^{-\alpha}$ as $x \to \pm\infty$[3] and for the GHD: $P(X < x) \approx |x|^{\lambda-1}\exp[(\alpha+\beta)x]$ as $x \to \pm\infty$. Regarding the role of each parameters in the GDH, it is not so straightforward compared to those of α-stable distribution. The mean and variance of the GDH are expressed by (9) and (10), but the expressions of skewness and kurtosis are far more complicated. So we use the following transformation $\tau = \frac{\beta}{\gamma},\ \zeta = \delta\gamma,\ \chi = \frac{\frac{\beta}{\alpha}}{\sqrt{1+\zeta}},\ \xi = \frac{1}{\sqrt{1+\zeta}}$, where $\gamma = \sqrt{\alpha^2+\beta^2}$, then according to [2] the parameters χ and ξ are natural measures of the skewness and kurtosis.

4 Empirical analysis and results

We verify the ability of these two types of distributions described in the previous part to model the behavior of financial asset returns on exchange rate of Czech crown to US Dollar and Euro. This is daily exchange rates EUR/CZK and CZK/USD from 2013-10 to 2016-04. The original data series are converted into the so called logarithmic returns series. The descriptive statistics of both original series as well as their corresponding returns are shown in Table 1.

First, we use the returns series to estimate the four parameters of α-stable distribution using maximum likelihood estimation method [1]. All estimation procedures are performed in Matlab. The results are reported in Table 2. In both cases, the value of parameter α is less than 2 and δ is less than the value of the sample standard deviation which means that α stable distribution captures leptokurtic behavior of returns. On the hand, the value of α is higher than those values obtained by other methods proposed to be used to estimate this parameter.

The same estimation technique is employed to estimate the values of parameters of generalized hyperbolic distribution and its two special cases. The results are reported in Table 3. Regarding the estimation process, the convergence of the estimation procedure of hyperbolic and normal inverse gaussian distributions are quite straightforward. The estimation of parameters of the general case is far more complicated.

[3]The exact value of c_α may be found in [7].

Characteristic	EURO/CZK rate	USD/CZK rate	USD/CZK returns	EURO/CZK returns
Mean	27.18219	21.03019	-4.36e-05	-5.31e-05
Median	27.04400	20.30820	-0.000192	-7.28e-05
Maximum	33.32600	27.97430	0.055401	0.044700
Minimum	22.99900	14.49940	-0.052245	-0.028391
Std. Dev.	2.229902	2.963258	0.008083	0.003879
Skewness	0.779875	0.288973	-0.002037	0.386582
Kurtosis	2.969720	2.145652	6.812545	13.64204

Table 1 Descriptive statistics of analyzed data

Series	α	β	δ	γ
EUR/CZK	1.5552	1.737e-5	1.932e-3	5.51e-5
USD/CZK	1.7366	-0.0284	4.744e-3	1.031e-4

Table 2 Estimation results of α-stable distribution

We have encountered the similar issue as mentioned in [8]. To mitigate the problem of flat and non-smooth objective function, first we localize several admissible centers by using direct search methods. Then we look for the optimal solution by seeking for local optima around these centers and the one with the highest value function is the global optimum.

Series	Distribution	γ	α	β	δ	μ
EUR/CZK	Hyperbolic	1	190.488	-5.1594	0.00285	-3.726e-4
	Norm. Inv. Gaussian	-0.5	119.108	-4.6901	0.00772	3.480e-4
	Gen. Hyperbolic	0.5269	163.2829	-3.3686	0.0044	2.087e-4
USD/CZK	Hyperbolic	1	386.020	-1.5563	2.683e-7	7.406e-5
	Norm. Inv. Gaussian	-0.5	168.062	-0.4960	0.0025	6.044e-5
	Gen. Hyperbolic	0.4915	291.120	-1.3795	8.222e-4	7.414e-5

Table 3 Estimation results for generalized hyperbolic distributions

Next, we use our estimation results to calculate the values of pdf function for the two series. The results are shown in Figures 1 and 2, where the left tails of these distributions are zoomed in the right-hand side panels. The results show that the examined distributions are better alternatives than the normal distribution. We also see that α distribution has heavier tails than those from the GHD family.

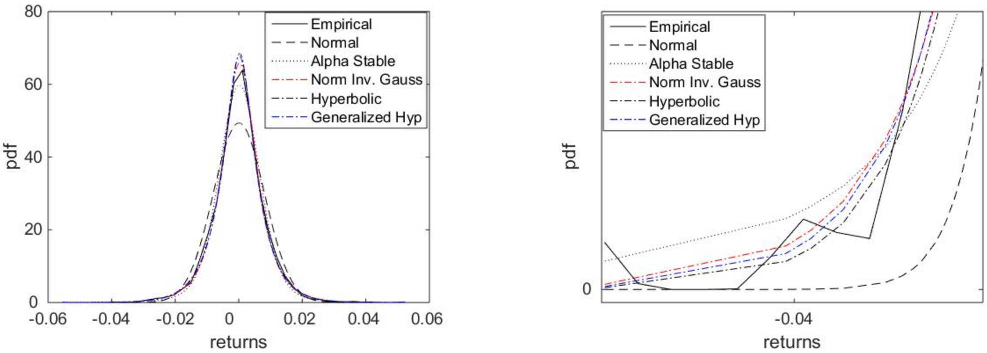


Figure 1 The estimated PDFs of analyzed distributions of USD/CZK returns

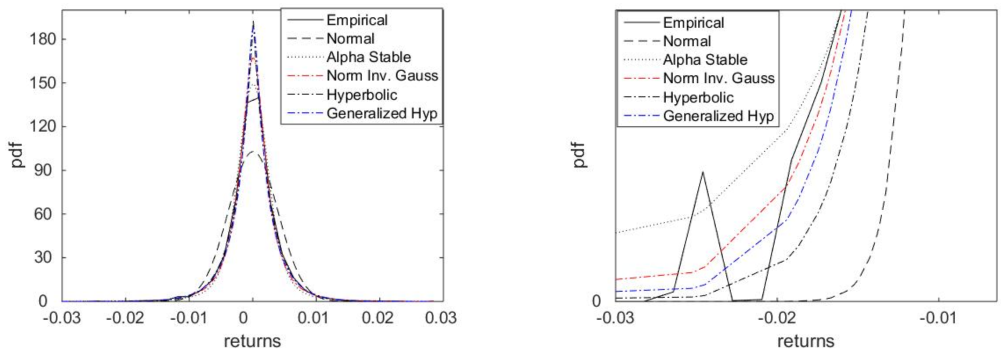


Figure 2 The estimated PDFs of analyzed distributions of EUR/CZK returns

Series	Quantile	Normal	α-stable	Hyperbolic	NIG	Gen. Hyp.
USD/CZK	0 - 10	0,0848	**0,0619**	0,0758	0,0698	0,0737
	10 - 20	0,2114	**0,0568**	0,0990	0,0800	0,0924
	20 - 30	0,2661	**0,1946**	0,3600	0,2901	0,3135
	30 - 40	1,5072	1,0615	0,8618	**0,6432**	0,5775
	40 - 50	5,0318	1,5860	**1,2954**	1,7940	1,9146
	50 - 60	10,1598	4,7721	**2,4196**	3,1490	2,5216
	60 - 70	4,0381	1,4434	2,3385	**1,1084**	1,1802
	70 - 80	0,5816	0,5547	0,3350	0,2072	**0,1901**
	80 - 90	0,3122	0,1363	0,1444	0,1025	**0,0929**
	90 - 100	0,0645	0,0700	0,0468	**0,0380**	0,0390
	Total	3,7209	1,6454	1,1467	1,1830	**1,0567**
EUR/CZK	0 - 10	0,1640	0,1502	0,1498	**0,1092**	0,1301
	10 - 20	0,4039	0,2238	0,2212	0,2075	**0,2028**
	20 - 30	2,7690	1,4169	0,6012	**0,5723**	0,5808
	30 - 40	30,1321	10,0006	5,4790	**3,9904**	6,4670
	40 - 50	19,0677	7,1071	**3,169**	6,2096	6,3759
	50 - 60	1,7618	1,0085	0,8575	**0,7064**	0,9213
	60 - 70	0,4213	0,2305	0,2042	0,1138	**0,0976**
	70 - 80	0,1488	0,1676	0,1415	**0,1170**	0,1301
	80 - 90	$\mathbf{1{,}9.10^{-12}}$	0,1083	0,0008	0,0144	0,0037
	90 - 100	0,0779	0,0799	0,0779	**0,0774**	0,0779
	Total	11,2632	3,8991	**2,0211**	2,3407	2,8780

Table 4 RMSE of PDFs wrt empirical PDF

Besides visual comparison, we also quantitatively evaluate the differences among them by calculating the distances between them and an empirical density estimated by the kernel smoothing method. The measure for the distance is the so call root mean squared errors (RMSE). The comparison is conducted for the whole series as well as for each 10% observations. The results are displayed in Table 4. The results strongly confirm the fact that normal distribution is the worst option among all alternatives taken into account in this research. α-stable distribution as the whole is the second best option and in the case of USD/CZK exchange rate, it models better the left tail behavior of returns. Finally, semi-heavy tail distributions as the whole come as the best choice for this purpose.

Within this group, however, the results are difficult to be unambiguously interpreted. The generalized hyperbolic distribution as the general case with all five parameters is expected to perform best. But it is true only in the case of USD/CZK exchange rate as the whole. But for the left tail of this returns series, the normal inverse gaussian, which is its special case when $\lambda = -0.5$, is a better option. For EUR/CZK exchange rate returns series, hyperbolic distribution is the better choice when the whole distributions are compared and the normal inverse distribution is better option when we want to model only the left tail of distribution of returns. For the time being, we are unable to decide whether this result comes from the difficulties connected with estimation procedure of parameters of the generalized hyperbolic distribution or it originates from irrelevance of value of λ. The answer to the question will be the subject of our future research.

5 Conclusion

In this work we try to investigate the ability of α-stable distribution and the distributions from generalized hyperbolic distribution group to capture characteristic leptokurtic property of returns of financial assets. Our results confirm the fact that normal distribution is an unsuitable for this task and MLE seems to be the most appropriate estimation method. Our results obtained for two exchange rate series indicate the α-stable distribution seems to be generally less suitable for modeling returns distribution than any distribution from generalized hyperbolic distribution group. But the stable distribution provides the most accurate approximation left quantiles for the USD/CZK, although the remaining three distributions also give very good results that are only slightly worse than those of the stable one. However, all distributions show the greatest variation in middle quantiles around the peak. For the EUR/CZK NIG distribution gives the best results in two cases in the first three left quantiles. It seems that GHD fares well, but as HD and NIG are special cases of GHD, GHD should always give better results, which generally have not been confirmed. Many authors ([4],[6],[8]) consider GHD as inappropriate one due to its overfitting tendency as well as the non-convexity of its likelihood function. Our results have confirmed these problems. However, the answer to the question, which type of distribution from this distribution family should be used to model financial asset returns, still requires further research.

Acknowledgements

The support from the Czech Science Foundation under Grant P402/12/G097 and IP 100040/1020 is gratefully acknowledged.

References

[1] Akgiray, A. and Lamoureux, C.: Estimation of stable law parameters: A comparative study. *Journal of Business and Economic Statistics* 7, 1977.

[2] Barndorff-Nielsen, O. E., Blæsild, P., Jensen, J. L., and Sørensen, M.: The fascination of sand. In:*A Celebration of Statistics*, (Atkinson, A. C., Fienberg, S. E., ed.) 1985, Springer Verlag, New York, pp. 57—87.

[3] Belov, A.: On the Computation of the Probability Density Function of Stable Distribution Chambers, *Mathematical Modelling and Analysis, Proceedings of the 10th International Conference MMA, Trakai*, 2005. 333–341.

[4] Bibby, B. M. and Sorensen, M.: Hyperbolic processes in finance. In: *Handbook of Heavy Tailed Distributions in Finance*, (S. Rachev, ed.), Elsevier Science 2003 Amsterdam, 211—248.

[5] Borak, S., Hardle, W., and Weron, R.: Stable Distributions. *SFB 649 Discussion Paper 2005-008*, Center for Applied Statistics and Economics, Humbolt-Universitat zu Berlin, 2005.

[6] Borak, S., Misiorek, A., and Weron, R.: Models for Heavy-tailed Asset Returns, *MPRA paper No. 25494*, September 2010.

[7] Nolan, J. P.: Modeling Financial Data with Stable Distribution. In: *Handbook of Heavy Tailed Distributions in Finance*, (S. Rachev, ed.), Elsevier Science Amsterdam 2003, 105–130.

[8] Prause, K.: *The generalized hyperbolic model: Estimation, financial derivatives and risk measures.* University of Freiburg, Doctoral Thesis, 1999.

Comparison of wages in the Czech Regions

Lubos Marek[1]

Abstract. The aim of paper is the comparison of wages in Czech regions. We have a time series of wages from 2000 to 2014, ie for 15 years. We work with interval distribution of wages. The data are observed in great detail. The length of interval is 500 CZK. In addition to the absolute frequencies in individual intervals, so we can calculate the relative frequency and we are able to construct the empirical distribution of wages. The data file is progressively increased over time, in 2014 contains some 2 million observations. We work with basic descriptive characteristics - average wage, variance, quantile measures (10%, 25%, 50%, 75% a 90% quantiles) and working time fund. Data are presented for the second quarter of the year, because this quarter has the most stable working time fund. Because our data are observed in time, we will also be interested in the trend of the characteristics over all 15 years. We compute the trend function and growth rate over time for each individual time series. The Gini index will be very interesting for us, too. This index allows us to compare the rate of wage inequality in individual regions.

Keywords: wage, region, average wage, quantile, Gini index.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

One of the most frequently discussed issues not only in the economy is the amount of wages. There is a general opinion that wages are not the same across the country and varies from region to region. We will try to support this feeling and compare wages in individual regions of the Czech Republic. We work with data over years 2000-2014. The data source is firm Trexima, which conducts regular surveys regarding wages - see [10]. Data are published as interval frequency table. We work with very detailed data. Our interval frequency table has the length of interval 500 CZK. The sample size is great – over two million observations. The basic statistical characteristics were computed form data – mean, standard deviation and quantile measures (10%, 25%, 50%, 75% and 90% quantile). The very important for us is median (50% quantile). Because we work with interval frequency distribution, we can construct a frequency polygon, as an empirical counterpart of the probability density function. Since the data have the character of a time series, it is possible to study behavioral trends or the dynamism of these time series. We are also interested in the Gini index, which enabled the comparison of redistribution of wages in individual regions.

2 Methodology

Whole analysis is performed in MS Excel. We used the basic statistical functions and procedures for our computing. The theory of descriptive statistics is used, specifically working with interval frequency tables – details in Čermák and Vrabec [3].

The calculation of trend function was made in *Regression* procedure. This procedure is a part of *Data Analysis* in *Add Ins* menu. The trend function were computed directly in procedure *Graphs.* We computed the trend equaiton and *R*-square in this procedure, too – details in Cipra [2]. *R*-square we use as a tool for quality model checking. When we computed Gini index, we fitted the data by 5th degree polynomial. Then was needed to compute the relevant integral (so we have calculated the area under the curve of the polynomial). Details on the methodology of calculating the Gini index can be found eg. in Gini [4] and in Marek [8].
Our data have the character of time series. The level of wages is affected by inflation. The level of inflation is not considered in our calculations. So, we work with data in current prices of the year. It is not a problem to include the inflation in the calculations – but general view of the results of the comparison, however, does not change. Empirical frequency distributions can be modeled either directly or through a mixture of probability distributions. We chose the first approach, the second is described in other articles or in papers Mala [7] or Bartosova [1].

[1] University of Economics, Prague/Department of Statistics and Probability, W. Churchill sq. 4, Prague 3, Czech Republic, marek@vse.cz.

3 Data analysis

3.1 Average wage

Let's first look, how it developed average wage in individual regions. This issue is very closely involved in the work Marek [8] or [9]. The average wages (in CZK) are included in Table 1 and Figure 1. At the firs look we can see that a special position has the region Prague. The average wage here is much higher than average wages in other regions. The second highest wage has region Středočeský. The average wages are significantly affected by the production of cars in Mladá Boleslav and in Kolín. Among other region has such differences are not so significant. The average wages are comparable, the differents are small. Long-term low wages are in region Karlovarský.

	Praha	Středočeský	Jihočeský	Plzeňský	Karlovarský	Ústecký	Liberecký	Královehradecký	Pardubický	Vysočina	Jihomoravský	Olomoucký	Zlínský	Moravskoslezský
2000	15,570	12,977	11,762	11,883	10,969	12,436	11,721	11,436	10,848	10,272	11,271	10,897	11,395	12,454
2001	17,752	14,916	13,421	13,971	12,997	14,515	13,660	13,601	12,879	12,592	12,931	12,916	13,084	14,723
2002	19,592	15,821	14,581	15,558	14,362	15,565	14,572	14,312	14,010	13,583	14,344	14,242	13,716	15,793
2003	21,144	16,718	15,175	16,523	15,013	16,599	15,716	15,419	15,001	14,136	15,659	15,346	15,121	16,870
2004	21,289	17,130	15,960	16,499	15,706	17,124	16,453	15,903	15,558	15,258	16,054	15,778	15,530	17,099
2005	23,173	18,402	16,914	17,569	16,542	17,756	17,194	16,634	16,393	16,528	17,032	16,599	16,381	17,940
2006	25,463	19,281	17,550	18,059	17,414	18,206	18,436	17,349	17,248	17,485	18,129	17,344	17,084	18,627
2007	27,726	20,662	19,003	19,781	18,678	19,681	19,521	19,024	18,478	18,986	19,704	18,792	18,452	19,925
2008	29,780	22,184	20,186	20,953	19,482	20,567	20,492	20,203	19,776	20,427	20,997	19,662	19,489	21,262
2009	31,411	23,443	20,988	21,834	20,775	21,641	21,593	21,305	20,616	20,785	22,203	20,647	20,245	21,710
2010	31,827	24,174	21,720	22,554	21,310	22,211	22,547	21,785	20,988	21,454	22,900	21,217	20,864	22,475
2011	32,257	24,639	21,966	22,697	21,146	22,521	22,963	22,160	21,244	21,955	23,689	21,633	21,256	23,063
2012	32,506	25,449	22,357	23,389	21,441	22,909	23,341	22,827	21,742	22,381	24,130	21,936	21,485	23,515
2013	33,120	25,801	22,705	23,863	21,697	23,188	23,589	23,286	21,976	22,687	24,867	22,495	21,892	23,837
2014	33,308	26,282	23,130	24,310	21,970	23,427	23,972	23,513	22,428	23,029	25,122	22,704	22,383	23,957

Table 1 Average wages in regions

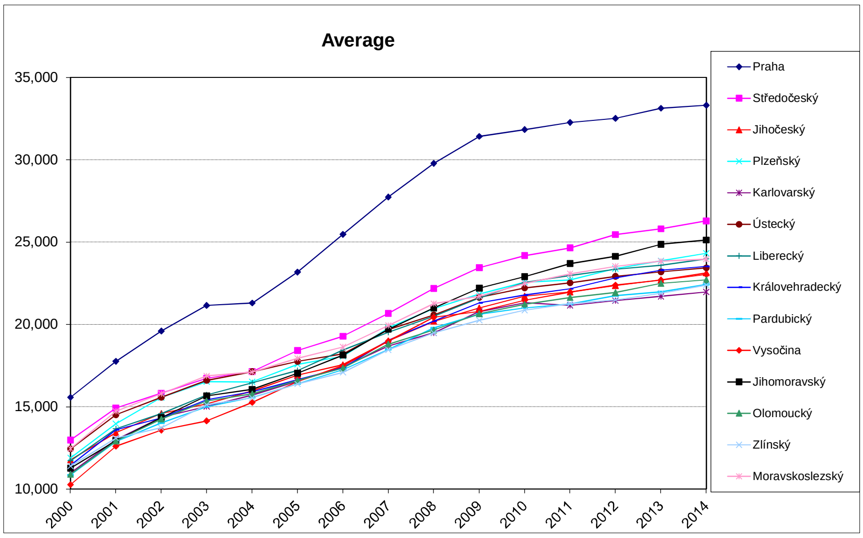


Figure 1 Average wages in regions

3.2 Quantile measures of wages

We work with 14 regions, for each region we calculated 5 quantiles for each year (2000 – 2014) – details in Malá [6] or in Hindls [5]. So, we have 14 x 5 x 15 = 1050 different quantiles. It is not possible to show all result in our paper. We show only a few selected quantiles in Figure 2. We choiced regions Praha (indication Pr in figure) and Karlovarský (indication Ka in figure) and we compare these two regions only. The symbols have the following meaning: D1 – 10% quantile, Q1 – 25% quantile, Median – 50% quantile, Q3 – 75% quantile, D9 – 90% quantile. Quantiles are sorted from top to bottom as listed in the legend. The PR D9 is 90% quantile for the Prague region etc. At first view it is obvious that the 90% quantile and the 75% quantile for Prague are more greater than the 90% quantile for the Karlovarský region. This means that 25% of the highest wages in Prague is equal or higher than the 10% of the highest wages in the Karlovarský region. The median for the Prague and 75% quantile for Karlovy Vary are almost the same. Thus, 50% of the highest wages in Prague corresponds to its level of 25% of the highest wages in Karlovarský region. Furthermore, it is almost the same 25% quantile for the Prague and the median for Karlovarský region. This means that 75% of the highest wages in Prague is equal to or higher than 50% of the highest wages in the Karlovarský region. If we compare the Prague percentiles with other regions, the differences were not already so significant, but there are still great.

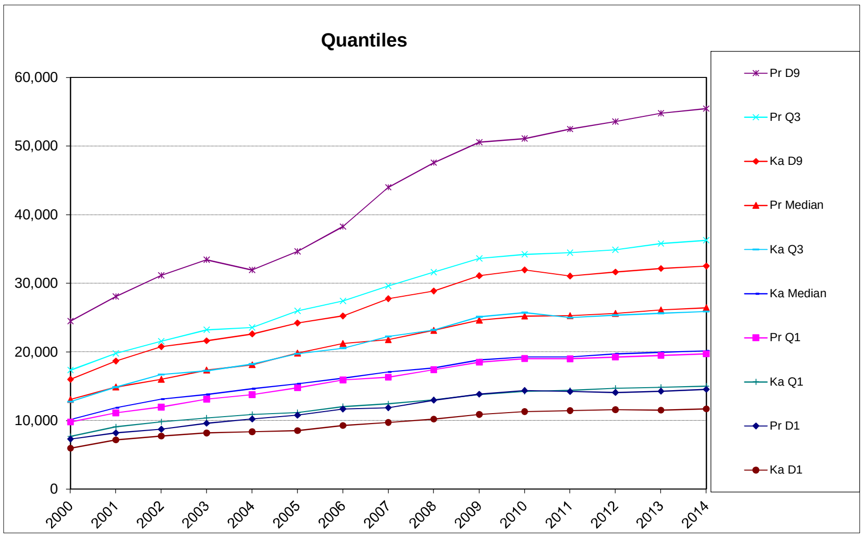


Figure 2 Quantiles – Praha and region Karlovarský

3.3 Trend and growth rate

Because we have a large number of results, it is not possible to show all trend functions for all regions. We choiced 3 regions only – Praha (the highest wages), Jihočeský region (the lowest wages) and Jihomoravský region, which represents all other regions. Previously wages grew by linear trend but in the last years the growth rate slows and the trend is rather quadratic. But we have data only up to 2014 and it is expected that due to the favorable development of the economy will grow wages much faster in the next years. The R-square is very near to 1, the quality of trend function is very good.

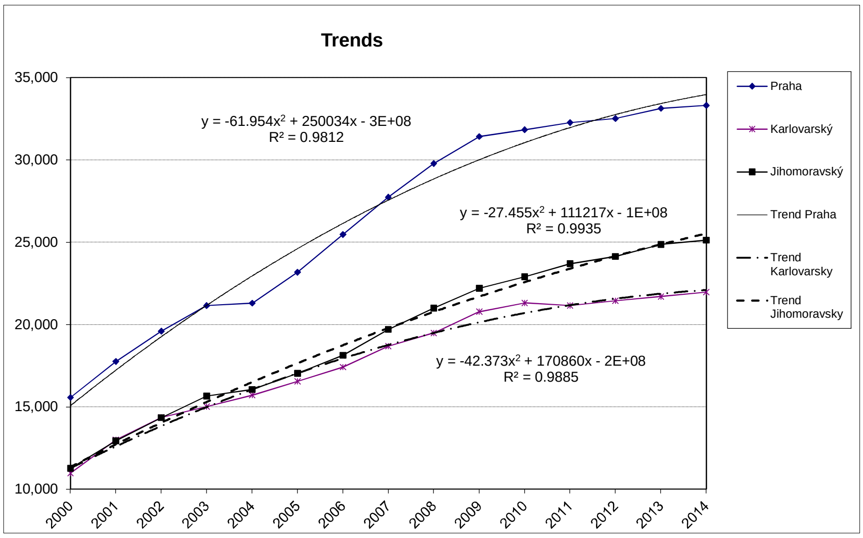


Figure 3 Trend for selected regions

	Praha	Středočeský	Jihočeský	Plzeňský	Karlovarský	Ústecký	Liberecký	Královehradecký	Pardubický	Vysočina	Jihomoravský	Olomoucký	Zlínský	Moravskoslezský
2000	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2001	1.140	1.149	1.141	1.176	1.185	1.167	1.165	1.189	1.187	1.226	1.147	1.185	1.148	1.182
2002	1.104	1.061	1.086	1.114	1.105	1.072	1.067	1.052	1.088	1.079	1.109	1.103	1.048	1.073
2003	1.079	1.057	1.041	1.062	1.045	1.066	1.078	1.077	1.071	1.041	1.092	1.078	1.102	1.068
2004	1.007	1.025	1.052	0.999	1.046	1.032	1.047	1.031	1.037	1.079	1.025	1.028	1.027	1.014
2005	1.088	1.074	1.060	1.065	1.053	1.037	1.045	1.046	1.054	1.083	1.061	1.052	1.055	1.049
2006	1.099	1.048	1.038	1.028	1.053	1.025	1.072	1.043	1.052	1.058	1.064	1.045	1.043	1.038
2007	1.089	1.072	1.083	1.095	1.073	1.081	1.059	1.097	1.071	1.086	1.087	1.084	1.080	1.070
2008	1.074	1.074	1.062	1.059	1.043	1.045	1.050	1.062	1.070	1.076	1.066	1.046	1.056	1.067
2009	1.055	1.057	1.040	1.042	1.066	1.052	1.054	1.055	1.042	1.017	1.057	1.050	1.039	1.021
2010	1.013	1.031	1.035	1.033	1.026	1.026	1.044	1.023	1.018	1.032	1.031	1.028	1.031	1.035
2011	1.014	1.019	1.011	1.006	0.992	1.014	1.018	1.017	1.012	1.023	1.034	1.020	1.019	1.026
2012	1.008	1.033	1.018	1.030	1.014	1.017	1.016	1.030	1.023	1.019	1.019	1.014	1.011	1.020
2013	1.019	1.014	1.016	1.020	1.012	1.012	1.011	1.020	1.011	1.014	1.031	1.025	1.019	1.014
2014	1.006	1.019	1.019	1.019	1.013	1.010	1.016	1.010	1.021	1.015	1.010	1.009	1.022	1.005
Average	**1.056**	**1.052**	**1.049**	**1.052**	**1.051**	**1.046**	**1.052**	**1.053**	**1.053**	**1.059**	**1.059**	**1.054**	**1.049**	**1.048**

Table 2 Growth rate

The growth rate is in Table 2. This rate is more or less similar in all regions. Wages are growing by an average of 5% annually (impact 2005-2008 period), in recent years, growth slowed. We expect a faster growth of wages in the next year, because the economy is growing rapidly.

3.4 Gini index

The Gini index is a measure of statistical dispersion intended to represent the income distribution of a nation's residents, and is the most commonly used measure of inequality. A Gini index of zero expresses perfect equality, where all wages are the same. A Gini coefficient of one (or 100%) expresses maximal inequality among wages. In our country, the rate of redistribution classically is low. In most western economies, this index is higher, but it is not valid for the Scandinavian countries.

This index is mostly show for the whole country. We computed the values of this index for each region for yeach year. The values of Gini index are in Table 3. As expected, the highest value are reached in Prague. Prague is comparable to western economies. Gini index values in other regions are significantly smaller. This means that in Prague is the greatest degree of redistribution. The Gini index values for year 2014 are in Figure 4.

	Praha	Středočeský	Jihočeský	Plzeňský	Karlovarský	Ústecký	Liberecký	Královehradecký	Pardubický	Vysočina	Jihomoravský	Olomoucký	Zlínský	Moravskoslezský
2000	0.293	0.236	0.246	0.234	0.232	0.241	0.242	0.225	0.235	0.229	0.242	0.223	0.240	0.233
2001	0.295	0.235	0.248	0.230	0.230	0.234	0.237	0.230	0.230	0.236	0.240	0.239	0.224	0.232
2002	0.232	0.303	0.245	0.228	0.228	0.238	0.241	0.232	0.229	0.235	0.235	0.228	0.228	0.225
2003	0.298	0.235	0.235	0.228	0.230	0.230	0.224	0.224	0.225	0.237	0.235	0.235	0.222	0.236
2004	0.227	0.275	0.224	0.224	0.230	0.230	0.236	0.218	0.215	0.228	0.224	0.218	0.218	0.229
2005	0.223	0.279	0.242	0.232	0.235	0.235	0.241	0.222	0.222	0.230	0.237	0.237	0.234	0.234
2006	0.232	0.285	0.243	0.236	0.229	0.229	0.245	0.226	0.234	0.234	0.247	0.247	0.224	0.237
2007	0.237	0.306	0.237	0.237	0.236	0.236	0.243	0.230	0.230	0.231	0.231	0.227	0.236	0.236
2008	0.306	0.306	0.237	0.237	0.226	0.234	0.243	0.234	0.228	0.234	0.235	0.252	0.225	0.235
2009	0.304	0.304	0.248	0.232	0.240	0.240	0.243	0.231	0.230	0.232	0.256	0.256	0.230	0.234
2010	0.234	0.299	0.241	0.241	0.231	0.237	0.232	0.232	0.228	0.233	0.254	0.254	0.229	0.235
2011	0.304	0.304	0.249	0.241	0.231	0.231	0.244	0.229	0.229	0.234	0.236	0.258	0.230	0.233
2012	0.307	0.307	0.253	0.236	0.236	0.231	0.242	0.229	0.231	0.232	0.235	0.259	0.234	0.231
2013	0.307	0.258	0.239	0.240	0.235	0.246	0.231	0.233	0.236	0.237	0.264	0.237	0.235	0.241
2014	0.305	0.261	0.238	0.239	0.235	0.247	0.234	0.235	0.234	0.237	0.265	0.238	0.237	0.239

Table 3 Gini index

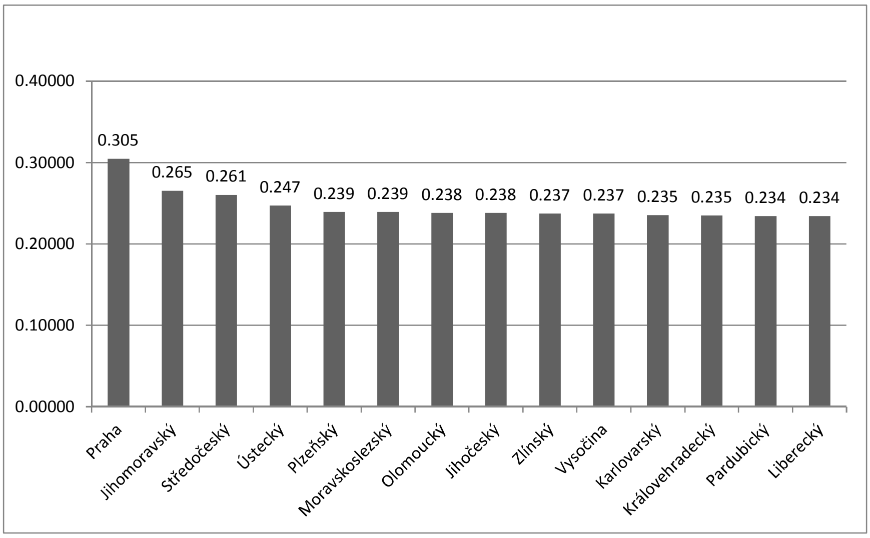


Figure 4 Gini index in regions 2014

3.5 Frequency polygon

Frequency polygon is the empirical counterpart probability density function. Wage developments in individual regions have changed over time in terms of basic measures – level, variability, skewness and kurtosis. Generally, the measure of the level increases, the variability of distribution increases over time, increases skewness and significantly decreases kurtosis. All this clearly points to the fact that the wages grow significantly over time and individual differences between wages are more and more greater. The wage development over time we can be relatively good describe by the various probabilistic models – more in articles Marek [8] or [9].

We chose to show only a few polygons. At the Figures 5 and 6 we compare the polygons for Praha and region Karlovarský over years 200-2014. We can see the basic differents between these two regions in all statistical characteristics. At Figure 7 is comparison of all regions in year 2014. The last Figure 8 is subset of

Figure 7. Three choiced regions are displayed. In terms of time, the development in all regions is very similar, but in different measures. The difference between regions are in all observed charasteristics. Of all the regions is significantly different Praha. The wages in Praha are significantly highest, there is the greatest variability and skewness, kurtosis is smallest.

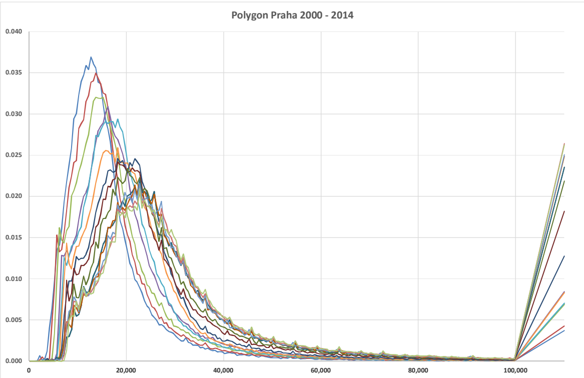


Figure 5 Polygon – Prague 2000 - 2014

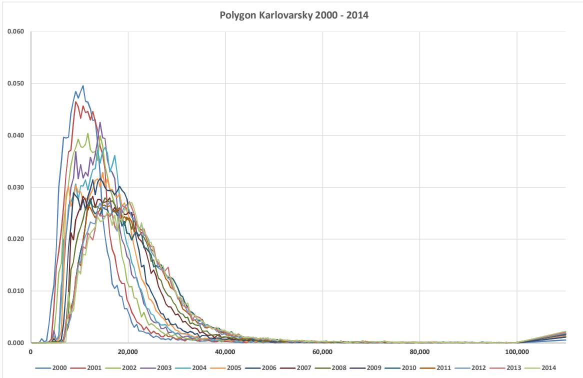


Figure 6 Polygon – region Karlovarsky 2000 - 2014

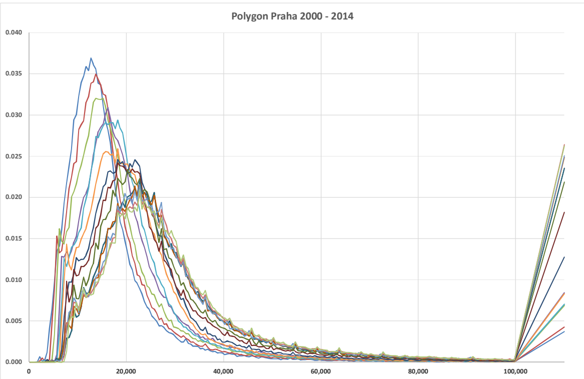


Figure 7 Polygon - all regions 2014

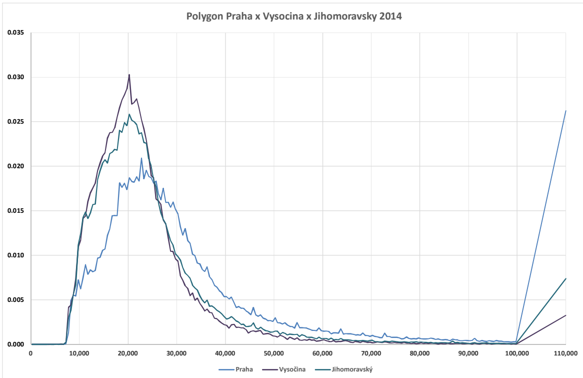


Figure 8 Polygon - 3 regions 2014

4 Conclusions

Based on the results we can make some conclusions.

- The average wage in individual region is different. Praha is quite different, wages here are completely different from other regions. The average wage in Praha is the highest and from other regions differs very significantly. High wages achieves even region Středočeský, which is due to production of cars in Mladá Boleslav and in Kolín. Among other regions, already is not a significant difference. The wages are lower and most regions are comparable. The lowest wages are in the region Karlovarský.
- To quantile measures are valid similar conclusions as to the average. Interesting is the fact that the median in Prague reaches an average for the whole country. This means that 50% of wages in Prague is larger than the national average. Normally, the average is 68% quantile of the whole country. Thus, only approximately 32% wages are greater than average.
- Development in time is described by trend function. The best trend function is 2nd degree polynomial. The R-square is very near to 1 for all our models. This means that the quality of trend function is very good. The growth rate is more or less similar in all regions. Wages are growing by an average of 5% annually, in the last years is growth smaller. We expect a faster growth of wages in the next year, because the economy is growing rapidly. The highest growth rate is in region Praha the lowest in region Ústecký.

- Gini index reached its highest level for the region Praha. From this perspective, Prague is comparable with developed economies. In other regions is markedly lower value. In long term, this index is approximately constant in all regions.
- Frequency polygon changes significantly over time in all regions. The difference is between individual regions, too. Overall, the level of statistical increases. The variability of distribution and skewness increase over time, and significantly decreases kurtosis.
- The differences among regions exist in other basic economic statistics, such are GDP per capita, unemployment rate etc. It will be very interesting to compare the regions by more statistics at one time. This analysis goes beyond the scope of this article. We will publish it in other articles.

Acknowledgements

This paper was written with the support of the Czech Science Foundation project No. P402/12/G097 „DYME – Dynamic Models in Economics“.

References

[1] Bartosová, J. and Longford, N. T.: *A study of income stability in the Czech Republic by finite mixtures*, Prague Economic Papers, vol. 23, no. 3, pp. 330–348, 2014.
[2] Cipra, T.: *Analýza časových řad s aplikacemi v ekonomii*. Praha, SNTL/Alfa., 1986.
[3] Čermák, V. and Vrabec, M.: *Teorie výběrových šetření*, 3. díl. VŠE Praha, 1999.
[4] Gini, C.: *Italian: Variabilità e mutabilità, Variability and Mutability*', C. Cuppini, Bologna, 156 pages. Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T). Rome: Libreria Eredi Virgilio Veschi (1955).
[5] Hindls, R. and Hronová, S. and Seger, J.: Statistika pro ekonomy. 5. vyd. *Professional Publishing*, Praha 2003.
[6] Malá, I.: Statistické úsudky, *Professional Publishing*, Praha 2013.
[7] Malá, I.: *Vícerozměrný pravděpodobnostní model rozdělení příjmů českých domácností*, Politická ekonomie, vol. 63., 2015.
[8] Marek, L.: *Analýza vývoje mezd v ČR v letech 1995-2008*. Politická ekonomie, vol 58, no. 2, pp. 186–206., 2010.
[9] Marek, L. *Some Aspects of Average Wage Evolution in the Czech Republic*, International Days of Statistics and Economics, Prague, Slaný: Melandrium, pp. 947–958, 2013.
[10] Trexima, (2016) http://www.trexima.cz

Efficient Distribution of Investment Capital

Patrice Marek[1], Tomáš Ťoupal[2], František Vávra[3]

Abstract. Kelly showed in his well-known paper that if a result of a bet or an investment is uncertain, it is not advisable to bet or invest the whole capital as this leads, with certainty, to bankruptcy. Instead of investing the whole capital Kelly proposed to invest a fraction of the capital. More on a proportional gambling, also known as Kelly gambling, was investigated by Cover and Thomas. Our paper uses principles of a log-optimal portfolio from both sources and an approximation of the main criteria is used instead. Doing this allows us effective statistical inference.
Usual procedure is to maximize expected value of the logarithm of the capital after an investment. The obtained solution is not comfortable for use in real-life situations; therefore, we propose alternative approach where the logarithm is substituted by the second-order Taylor polynomial. Same as in the case of a log-optimal portfolio we can obtain trivial solution, i.e. to invest all or nothing but usually a fraction of the capital will be invested. This fraction is based on simple characteristics that can be easily estimated from existing data: expected value and variance.

Keywords: Investment, optimal capital distribution, Kelly gambling, proportional gambling, log-optimal portfolio, approximation, Taylor polynomial.

JEL Classification: G110
AMS Classification: 91G10

1 Introduction

It is well known that to invest the whole capital is risky and therefore an investor wants to invest only some part of his capital. This paper answers the question how to find an optimal part of the capital that is invested. The findings obtained by Kelly [1] and Cover and Thomas [3] are used in the second section to find a log-optimal portfolio and an approximation of the main criteria is suggested to ensure that characteristics necessary to determine the optimal part of the capital can be estimated easily enough.

The theory is presented in the second section of this paper where all possible situations determining the optimal part of the invested capital are investigated. Simple rules that can be used to decide what is the optimal invested part of the capital are formulated in the third section of this paper. Two examples – test case and real-life example – are presented in the fourth section with aim to demonstrate usage of obtained theoretical results. The fifth section summarizes the paper and suggests the future work.

2 Optimizing Investment

Let us consider a recurring investment with an initial capital F_0 in an asset A. In the first round a part $s_1, s_1 \in [0,1]$, of the initial capital is invested in the asset A and the rest is left uninvested, i.e. without appreciation. $Z, Z > 0$ will denote a relative part of the original invested capital that is obtained after one round of investment ($Z > 1$ means that the initial invested capital is increased and $Z < 1$ means that the initial invested capital is decreased). Z is a random variable described by a probability density function $f_Z(x)$. Further, F_i will denote the capital after the ith round of an investment, $s_i, s_i \in [0,1]$ will denote a relative part of the capital invested in the ith round of an investment and Z_i will denote the relative part of the original invested capital that is obtained in the ith round of investment.

The capital after the first round of an investment is

$$F_1 = F_0 \cdot (1 - s_1) + F_0 \cdot s_1 Z_1 = F_0[(1 - s_1) + s_1 Z_1], \;\; 0 \le s_1 \le 1 \tag{1}$$

[1] NTIS – New Technologies for the Information Society, Faculty of Applied Sciences, University of West Bohemia, Technická 8, Plzeň, Czech Republic, e-mail: patrke@kma.zcu.cz.
[2] NTIS – New Technologies for the Information Society, Faculty of Applied Sciences, University of West Bohemia, Technická 8, Plzeň, Czech Republic, e-mail: ttoupal@kma.zcu.cz.
[3] Department of Mathematics, Faculty of Applied Sciences, University of West Bohemia,
Technická 8, Plzeň, Czech Republic, e-mail: vavra@kma.zcu.cz.

the capital after the second round of an investment is

$$F_2 = F_1 \cdot (1 - s_2) + F_1 \cdot s_2 Z_2 = F_0[(1 - s_1) + s_1 Z_1] \cdot [(1 - s_2) + s_2 Z_2],\ \ 0 \le s_i \le 1 \tag{2}$$

and the capital after the nth round of an investment is

$$F_n = F_0 \prod_{i=1}^{n} [(1 - s_i) + s_i Z_i] = F_0 \prod_{i=1}^{n} [1 + s_i(Z_i - 1)]\,,\ \ 0 \le s_i \le 1. \tag{3}$$

In the following parts the logarithm of (3) will be used. The reason for this modification is a computational complexity (it is more convenient to work with summation obtained by logarithm) and the fact that logarithm is a continuous strictly increasing function and therefore this change will have no effect on the position of the maxima. Logarithm of (3) is

$$\ln F_n = \ln F_0 + \sum_{i=1}^{n} \ln[1 + s_i(Z_i - 1)]\,,\ \ 0 \le s_i \le 1. \tag{4}$$

Both F_n and $\ln F_n$ are random variables. The expected value can be used to convert (4) into a deterministic form given by

$$E(\ln F_n) = \ln F_0 + \sum_{i=1}^{n} E(\ln[1 + s_i(Z_i - 1)])\,,\ \ 0 \le s_i \le 1. \tag{5}$$

Logarithm in (5) provides a condition that $1 + s_i(Z_i - 1) > 0$, i.e. $Z_i > 0, i = 1,2, \dots, n$ but this is the assumption made in the previous paragraphs and therefore we can consider it as fulfilled. Now, it is possible to use the classical procedure to maximize (5), i.e. using partial derivatives and investigation of the interval endpoints. The first and second derivatives of (5) are

$$\frac{\partial}{\partial s_i} E(\ln F_n) = E\left(\frac{\partial}{\partial s_i} \ln[1 + s_i(Z_i - 1)]\right) = E\left(\frac{Z_i - 1}{1 + s_i(Z_i - 1)}\right), \tag{6}$$

$$\frac{\partial^2}{\partial s_i^2} E(ln\, F_n) = -E\left(\left(\frac{Z_i - 1}{1 + s_i(Z_i - 1)}\right)^2\right) \le 0. \tag{7}$$

It is clear that the decision made in the previous rounds of an investment will have no effect on the current round of investment and the current round will have no effect on the following rounds (from the point of choosing an optimal part of a capital to invest). Next, we are looking for the optimal fraction of a capital that will be invested and an amount of the current capital is therefore insignificant for the analysis. It can be seen from (5) that to maximize the sum we have to maximize each single term of the sum, i.e. to maximize outcome in each round of investment individually. Therefore, the following analysis will be restricted only on the outcome from one round of investment.

Solving (6) for s_i can be problematic in a practical task and therefore, in some parts, an approximation of the logarithm by the second-order Taylor polynomial at $x = 1$ will be used instead, i.e.

$$\ln(x) \approx g(x) = (x - 1) - \frac{1}{2}(x - 1)^2. \tag{8}$$

We can use following information to find a point s_i, $s_i \in [0,1]$, where (5) reaches the global maximum (we recall that $Z_i > 0$):

- The function given by (5) is continuous in $s_i, s_i \in [0,1]$;
- The first derivative given by (6) exists for each $s_i \in [0,1]$;
- The second derivative given by (7) is nonpositive, i.e. the $E(\ln F_n)$ as the function of s_i is concave.

Now, the situation will be examined in the following subsections for several possible cases that arise from the previous list.

2.1 Extreme in Endpoints

If the derivative in (6) is positive for each s_i in $[0,1]$, then $E(\ln F_n)$ as the function of s_i is decreasing. The maximum is therefore achieved in the point $s_i = 0$.

If we use the information that the function is continuous, the first derivative exists for each s_i and the function of s_i is concave then we are able to make the same decision from a simple situation where only the endpoints of interval for s_i are investigated. The situation described above corresponds to the situation where (6) is nonpositive for $s_i = 0$ and negative for $s_i = 1$. This gives us the following conditions which have to be fulfilled simultaneously.

$$\frac{\partial}{\partial s_i} E(\ln F_n)\bigg|_{s_i=0} = E(Z_i - 1) \leq 0 \tag{9}$$

$$\frac{\partial}{\partial s_i} E(\ln F_n)\bigg|_{s_i=1} = E\left(1 - \frac{1}{Z_i}\right) < 0 \tag{10}$$

In other words, the conditions (9) and (10) ensure that the function with given properties is decreasing in s_i and therefore the maximum is obtained for $s_i = 0$.

Similar procedure leads to the conclusion that the maximum of (5) is obtained for $s_i = 1$ when the following conditions are fulfilled simultaneously.

$$\frac{\partial}{\partial s_i} E(\ln F_n)\bigg|_{s_i=0} = E(Z_i - 1) > 0 \tag{11}$$

$$\frac{\partial}{\partial s_i} E(\ln F_n)\bigg|_{s_i=1} = E\left(1 - \frac{1}{Z_i}\right) \geq 0 \tag{12}$$

Here, the conditions (11) and (12) ensure that the function with given properties is increasing in s_i and therefore its maximum is obtained for $s_i = 1$.

In a special case the function is constant if both conditions, (11) and (12), are equal to zero. This also means that the derivative in (6) is zero on the whole interval. Using $s_i \in [0,1]$ in (6) we obtain that the only way how to fulfill (6) is that Z_i have to be deterministic and it is equal to 1 (however, this situation is almost impossible in a real-life case). Therefore, the capital does not change after the investment and this result is independent of s_i. Thus, any value of s_i can be used, e.g. $s_i = 0$. This value is reasonable because we know for sure that the result will be exactly the same as the invested capital and therefore we choose to safe our time and not to invest.

2.2 Extreme Inside Interval

The extreme is inside the considered interval if the following conditions are fulfilled simultaneously.

$$\frac{\partial}{\partial s_i} E(\ln F_n)\bigg|_{s_i=0} = E(Z_i - 1) > 0 \tag{13}$$

$$\frac{\partial}{\partial s_i} E(\ln F_n)\bigg|_{s_i=1} = E\left(1 - \frac{1}{Z_i}\right) < 0 \tag{14}$$

Remark 1. The rest situations that were not yet examined are not achievable because the function is convex, i.e. it is not possible to obtain these combinations of conditions:

- $E(Z_i - 1) \leq 0$ and $E\left(1 - \frac{1}{Z_i}\right) > 0$;
- $E(Z_i - 1) < 0$ and $E\left(1 - \frac{1}{Z_i}\right) = 0$.

For the situation where (13) and (14) are fulfilled the approximation of (5) with (8) is used. This results in

$$E(\ln F_n) \approx \ln F_0 + \sum_{i=1}^{n} E(g[1 + s_i(Z_i - 1)]), \quad 0 \leq s_i \leq 1. \tag{15}$$

The same approximation is used for (6) which is set to zero and solved

$$E\left(\frac{\partial}{\partial s_i} g[1 + s_i(Z_i - 1)]\right) = E\left(\frac{\partial}{\partial s_i}\left(s_i(Z_i - 1) - \frac{1}{2}s_i^2(Z_i - 1)^2\right)\right) = 0. \tag{16}$$

Using some algebra, it is obtained form (16)

$$E(Z_i - 1) - s_i E((Z_i - 1)^2) = 0 \tag{17}$$

and the solution is

$$s_i^{opt} = \frac{E(Z_i - 1)}{E((Z_i - 1)^2)}. \tag{18}$$

Next, the approximation of (7) gives

$$\frac{\partial^2}{\partial s_i^2} E(g[1 + s_i(Z_i - 1)]) = -E((Z_i - 1)^2) < 0. \tag{19}$$

We do not consider (19) to be equal to 0 as this would mean that $Z_i = 1$ with probability 1, i.e. it would mean that Z_i is not a random variable and after each round of investment we would have the same amount of the capital as before the investment. This result means that (19) guaranties that s_i^{opt} in (18) is the point where the maximum is realized. Using condition $s_i \in [0,1]$ and (18) we obtain

$$0 \le \frac{E(Z_i - 1)}{E((Z_i - 1)^2)} \le 1. \tag{20}$$

In the case where $E(Z_i - 1) > E((Z_i - 1)^2)$ we set $s_i^{opt} = 1$. Obviously, if $E(Z_i - 1) < 0$ then it is not reasonable to invest and we set $s_i^{opt} = 0$.

Obtaining a point estimate of $E(Z_i - 1)$ and $E((Z_i - 1)^2)$ in (18) is a standard statistical task whereas to solve (6) can be problematic.

3 Decision Summary

Using the theory described in the section 2 it is possible to formulate simple rules for the decision making. This decision is affected by random variable Z_i and it is necessary to know (or to estimate) its following characteristics:

- $E(Z_i - 1)$ which represents expected value of a net income in one round of an investment (before taxes);
- $E((Z_i - 1)^2)$ which represents variability of a net income in one round of an investment (before taxes);
- $E\left(1 - \frac{1}{Z_i}\right) = E\left(\frac{Z_i - 1}{Z_i}\right)$ which represents the expected value of a relative net income in one round of an investment (before taxes);

Decision is made by the rules in Table 1 (all rules were derived in the previous section).

Conditions	Decision
$E(Z_i - 1) \le 0 \wedge E\left(1 - \frac{1}{Z_i}\right) < 0$	$s_i^{opt} = 0$, i.e. not to invest
$E(Z_i - 1) > 0 \wedge E\left(1 - \frac{1}{Z_i}\right) \ge 0$	$s_i^{opt} = 1$, i.e. invest the whole capital
$E(Z_i - 1) = 0 \wedge E\left(1 - \frac{1}{Z_i}\right) = 0$	$s_i^{opt} \in [0,1]$; as $Z_i = 1$ for all $s_i^{opt} \in [0,1]$; use $s_i^{opt} = 0$
$E(Z_i - 1) > 0 \wedge E\left(1 - \frac{1}{Z_i}\right) < 0$	use Table 2 for a decision, i.e. decision based on the approximation of the logarithm

Table 1 Decision making rules

Remark 2. We recall that the following situations are not achievable because the function is convex:

- $E(Z_i - 1) \le 0$ and $E\left(1 - \frac{1}{Z_i}\right) > 0$;
- $E(Z_i - 1) < 0$ and $E\left(1 - \frac{1}{Z_i}\right) = 0$.

Table 2 contains decision for the last situation given in Table 1 ($E(Z_i - 1) > 0 \wedge E\left(1 - \frac{1}{Z_i}\right) < 0$).

Conditions	Decision
$0 \le \dfrac{E(Z_i - 1)}{E((Z_i - 1)^2)} \le 1$	$s_i^{opt} = \frac{E(Z_i - 1)}{E((Z_i - 1)^2)}$, condition guaranties that $s_i^{opt} \in [0,1]$
$E(Z_i - 1) > E((Z_i - 1)^2)$	$s_i^{opt} = 1$, i.e. invest the whole capital
$E(Z_i - 1) < 0$	$s_i^{opt} = 0$, i.e. not to invest

Table 2 Decision making rules in the case of an approximation

Standard statistical procedures can be used to test the rules given in Table 1 and Table 2 when a large enough random sample is available. This is usually not a problem when we are dealing with a financial data set that tends to be large.

We recall that strategy used in a given round of an investment is independent of strategies in the previous rounds of an investment. Therefore, it is also possible to use information on changing distribution of $f_Z(x)$ (by change it is meant change in time). This means that an asset used for the investment can be completely different from an asset used in the previous rounds. The presented procedure can be easily used and generalized for an investment of a money or a situation with more assets. Using more assets will require usage of nonlinear programing.

A demonstration that the objective function given by (5) and its approximation introduced in (15) are similar and flat around the point of maximum will be shown in the following section. The fact that functions are flat around the point of maximum means that the obtained solution is robust and it can be used in a real-life investment where we have to deal with problems which are extremely hard to describe analytically, i.e. a delay between time of decision and time of investment (the price can differ from the price used in analysis); different time between listing days; a difference between function (5) and its approximation; a transaction costs and taxes.

4 Results Presentation

Two examples are shown in this section, the first one where generated values are used and the second one where real data set is used.

4.1 Test Case

We will use model situation to present course of the criterion (5) for one round of investment where the initial capital is set to one, i.e. $F_0 = 1$, this results in

$$E(\ln F_1) = E(\ln[1 + s_1(Z_1 - 1)]), \;\; 0 \le s_1 \le 1 \tag{21}$$

and the approximation of the criterion based on (8) results in

$$E(\ln F_1) \approx E\left(s_1(Z_1 - 1) - \frac{1}{2}s_1^2(Z_1 - 1)^2\right), \;\; 0 \le s_1 \le 1. \tag{22}$$

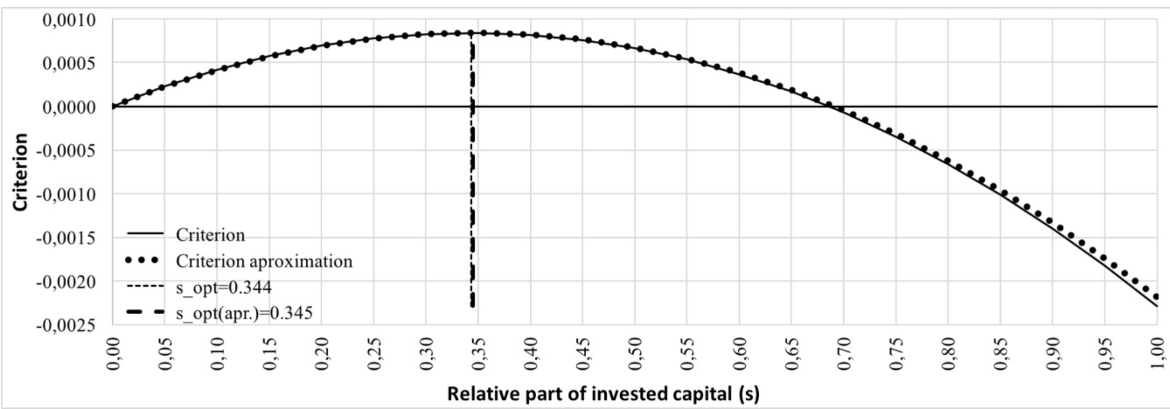


Figure 1 Demonstration of results – test case

Figure 1 shows comparison of (21) and (22). Optimal relative part of invested capital is 0.344 when the original criterion is used and 0.345 when the approximation is used (difference 0.409%). The value of the original criterion function for the optimal relative part of invested capital is $8.3957 \cdot 10^{-4}$ and the value of the approximation of the criterion function for the optimal relative part of invested capital is $8.3958 \cdot 10^{-4}$ (difference -0.002%). The results suggest that the method is robust and provides similar results even when the approximation is used.

4.2 Real Data

In this part, the application of the theory from section 3 is used for an investment to Amundi Funds Equity US Relative Value (CZK), ISIN LU0568606221. Historical prices of this fund are available at [2]. We test the time frame from 2. 1. 2014 to 19. 4. 2016. The statistical inference is made each day using daily prices in the last year (first estimation is made from historical data in 2013 and investing starts on 2nd January 2014). This produces estimation of rules used in Table 1 and Table 2. According to the obtained value of s^{opt} (relative part of the invested capital) we adjust our position in this fund each day. Comparison of this strategy with *strategy buy-and-hold*, i.e. where the whole capital is invested to the fund with no change during the whole interval, is shown in Figure 2.

Remark 3. We do not use any limitation in this example, i.e. we assume that the fund is frictionless and that there are no taxes and transaction costs. On the other hand, we use a delay between time of decision and time of investment, so the fund is not bought for the price that were used for the decision but it is bought for the price that is available at the next day. We also assume that money left aside are not subject to interest.

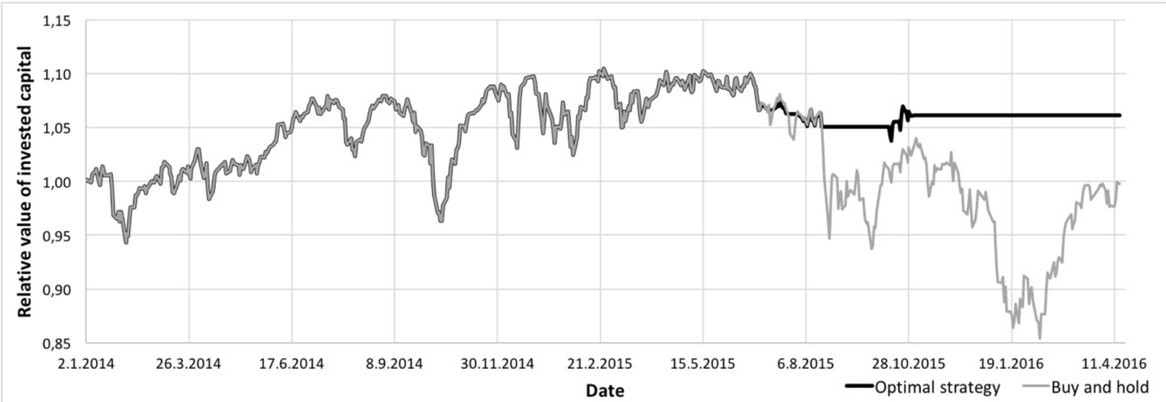


Figure 2 Demonstration of results – real example

As can be seen in Figure 2, the strategy started with $s^{opt} = 1$ and maintained this value for more than one and half year, i.e. the same results as the *strategy buy-and-hold*. Courses of both strategies went to different directions in mid-2015 and the optimal strategy produced better results.

5 Conclusion

This paper showed how to allocate investor's capital between money and a risky asset so that the log-optimal strategy is obtained. To achieve easy enough procedure an approximation of the main criteria was used. This approximation has a little impact on the solution as was presented in the fourth section of this paper and it is based on easily estimated characteristics.

In the fourth section, a test case and a real-life example were presented. In the test case, it was shown that the main criteria and its approximation has small difference. The real-life example demonstrated that the presented procedure using approximation can produce better results than *buy-and-hold strategy*.

Future works will be focused on investigation of a delay between time of a decision and time of an investment; developing simple trading strategies that will allow to avoid too frequent trading with high transaction costs. More work has to be done for a testing real-life data, i.e. more assets has to be evaluated to strongly support presented procedure.

References

[1] Cover, T., Thomas J.: Elements of Information Theory. John Wiley & Sons, Inc., New Jersey, 2006.
[2] Investiční kapitálová společnost KB a.s. [online]. Czech Republic: 2016-04-19, 2016 [cit. 2016-04-19]. Available from: http://www.iks-kb.cz/opencms/opencms/en/detail_fondu.html?Fond=Amundi_Funds_Equity_US_Relative_Value_(CZK)&FundId=10152154
[3] Kelly, J.: A New Interpretation of Information Rate. The Bell System Technical Journal **35**, 4 (1956), 917–926.

Dynamic and spatial analysis of economic and labour market characteristics in European countries

Adrianna Mastalerz-Kodzis [1], Ewa Pośpiech [2]

Abstract. The purpose of the article is to study the dynamic and spatial changes of selected variables that characterize the social, economy and labour market in EU countries. The article analyses the scale and dynamics of the changes in selected economic values with the use of taxonomic measures: the HDI indicator and the taxonomic model measure. The socio-economic development of the EU countries is inhomogeneous. Each of the countries features a certain dissimilarity, dynamic of changes. The labour market is also inhomogeneous, which is indicated by many characteristics concerning economy. The paper consists of two parts. In the first one methods used in analyses are shown, the second one presents the most important results of research. The empirical analysis was conducted on the base of selected economic indicators for the countries of the EU, a temporal-spatial analysis was conducted. The period of study was 1980 – 2015.

Keywords: dynamic analysis, multivariate comparative analysis, economic and labour market characteristics.

JEL Classification: C33, C44
AMS Classification: 91B40, 91B45

1 Introduction

Countries belonging to the European Union are, in comparison to other countries in the world, perceived as well developed socially as well as economically. This is widely confirmed by many analyses conducted by worldwide institutions as well as by independent sociologists and economists. Throughout decades Europe has gradually reached a level of development which is nowadays described as a high one. However, we can notice many differences of a socio-economic nature amongst European Union countries, which are evaluating in time.

In the 21st century most of the EU countries face a serious challenge connected with the change in the age structure of the population. The percentage of population in the past-productive age is increasing whilst the percentage of population in the pre-productive age is decreasing, which implicates many changes in the economic field, among others, it greatly influences the socio-economic indicators as well as the labour market.

The purpose of the article is to analyse the dynamics of socio-economic changes, study the selected labour market characteristics and examine the relationship between the economic development and the labour market situation in the countries belonging to the EU.

In economic literature concerning the ways of analysing panel data we can find articles of methodical nature [4, 8], as well as articles including the use of methodology in economic analyses [5, 6, 7].

2 Selected elements of research methodology

In the empirical analyses we have used selected statistic methods, especially the analysis of correlation, regression and dynamics, as well as selected methods of multivariate comparative analysis [4, 10]. The conducted analyses have been compared with the commonly used measure of socio-economic development level, which is the HDI indicator. Its way of construction and interpretation can be found in articles such as: [1, 2, 3, 9].

2.1 Multivariate comparative analysis

In the article we have used selected methods of spatio-temporal data classification. By characterising n objects (the EU countries) with the use of m variables (factors), the character of the variables has been defined (stimulants, destimulants) and their normalization has been performed [10]. In multivariate methodology of comparative analysis historical data is taken into consideration. Through this method we can compare different objects

[1] University of Economics in Katowice, Department of Mathematics, 1 Maja 50, 40-287 Katowice, Poland, adrianna.mastalerz-kodzis@ue.katowice.pl
[2] University of Economics in Katowice, Department of Mathematics, 1 Maja 50, 40-287 Katowice, Poland, ewa.pospiech@ue.katowice.pl

(e.g. the EU countries), which are described using many features, for example general economic factors or labour market factors. The methodology is based on the construction of the taxonomic model measure.

We conduct the normalization of variables . For $\bar{x}_j$ – arithmetic mean of feature j and for s_j standard deviation of feature j, standardization according to the following formula has been used: $z_{ij} = \left(x_{ij} - \bar{x}_j\right)/ s_j$.

A pattern $z_0 = \left[z_{01}, z_{02}, \ldots, z_{0m}\right]$ and anti-pattern for development $z_{-0} = \left[z_{-01}, z_{-02}, \ldots, z_{-0m}\right]$ has been defined for:

$$z_{0j} = \begin{cases} \max\limits_i z_{ij}, & Z_{ij} \quad - \quad stimulant \\ \min\limits_i z_{ij}, & Z_{ij} \quad - \quad destimulant \end{cases}, \quad z_{-0j} = \begin{cases} \min\limits_i z_{ij}, & Z_{ij} \quad - \quad stimulant \\ \max\limits_i z_{ij}, & Z_{ij} \quad - \quad destimulat \end{cases} \tag{1}$$

We have calculated the distance between the object and the pattern. In the article we have used the distance:

$$d_{i0} = \sqrt{\sum_{j=1}^{m} \left(z_{ij} - z_{0j}\right)^2} \tag{2}$$

The obtained variable is not normalized. It has been transformed using a formula $m_i = 1 - \left(d_{i0} / d_0\right)$, where m_i is the taxonomic measure of the development of object *i* (EU country), d_0 is the distance between the pattern and the anti-pattern stated by formula $d_0 = \sqrt{\sum_{j=1}^{m} \left(z_{0i} - z_{-0j}\right)^2}$. The higher the level of the phenomenon, the higher the value of the measure, $m_i \in [0,1]$.

In a classic perspective the values of the economic and financial indicators are averaged in time. However, taking into consideration the dynamic character of the variables, the values of taxonomic measures should be analysed in time [6, 7].

2.2 Index HDI construction

The HDI (Human Development Index) indicator is used to determine the level of development of the countries in the world. HDI was constructed in the 90s of the 20th century by Amarty Sen and Mahbuba ul Haqa.

In the evaluation process, the synthetic indicator HDI takes into account three criteria: a long and healthy life, the level of education and life standard. The US programme concerning development recommends calculating the HDI indicator with the use of basic characteristics which are:

- **Health Index**, for LE_i - average lifespan in country *i* according to the formula:

$$H.Ind_i = \frac{LE_i - 25}{65} \tag{3}$$

- **Education Index**, for $LIT.Ind_i$ - illiteracy indicator and $ENR.Ind_i$ - schooling indicator according to the formula:

$$E.Ind_i = \tfrac{2}{3}\left(LIT.Ind_i\right) + \tfrac{1}{3}\left(ENR.Ind_i\right) \tag{4}$$

- **Welfare Index**, where y_i is the income per one inhabitant in a given country:

$$\left(Y.Ind_i\right) = \frac{\log\left(y_i\right) - \log(\$100)}{\log(\$40000) - \log(\$100)}. \tag{5}$$

- **The social development index** per one inhabitant for a given country is calculated according to the formula:

$$HDI_i = \frac{H.Ind_i + E.Ind_i + Y.Ind_i}{3} \tag{6}$$

It is accepted that the value of HDI on a given level indicates a country which is: highly developed (0.801-1), medium developed (0.501-0.8), poorly developed (0-0.5).

3 Selected empirical study results

The empirical analysis was conducted on the base of selected economic indicators for the countries of the EU, a temporal-spatial analysis was conducted. The period of study was 1980 – 2015, however, due to data absence, the consideration was carried out in a narrowed period of time. The source of the data were data bases: UNDP Human Development Report, Eurostat, UNESCO Institute of Statistics.

3.1 Demographic conditions of the EU countries, HDI indicator dynamics

Analysing the data referring to the EU countries, in many of them we can observe two negative demographic phenomena: decreasing number of population and aging of the society.

The following demographic characteristics for the EU countries were used in the analyses (Table 1):

- average and anticipated lifespan,
- the average age in 2015,
- natural growth in 2000 – 2015,
- the percentage of people in age groups: 0-14: 0-14, 15-64, 65+ in 2015.

Country	HDI 1990	HDI 2014	HDI average change rate 1990-2014 (%)	Average natural growth		Average age 2015	Percentage of people in age groups		
				2000/2005	2010/2015		0-14	15-64	65+
							2015		
Netherlands	0.829 (1)	0.922 (2)	0.463	0.550	0.274	42.4	25.8	46.4	27.8
Germany	0.801 (4)	0.916 (3)	0.585	0.077	-0.110	46.3	19.7	47.6	32.7
Denmark	0.799 (5)	0.923 (1)	0.629	0.295	0.395	41.5	27.0	43.9	29.1
Ireland	0.770 (11)	0.913 (4)	0.743	1.781	1.128	35.9	32.9	47.9	19.2
Sweden	0.815 (2)	0.907 (5)	0.466	0.353	0.653	41.2	27.6	40.5	31.8
United Kingdom	0.773 (10)	0.907 (5)	0.697	0.450	0.565	40.5	27.4	44.5	28.1
France	0.779 (8)	0.888 (8)	0.571	0.740	0.547	41.0	28.6	41.7	29.6
Austria	0.794 (6)	0.885 (9)	0.473	0.537	0.368	43.3	21.6	50.6	27.9
Belgium	0.806 (3)	0.890 (7)	0.432	0.462	0.438	41.9	26.7	44.3	29.0
Luxembourg	0.779 (8)	0.902 (6)	0.639	0.973	1.347	39.1	25.4	53.4	21.2
Finland	0.783 (7)	0.883 (10)	0.524	0.268	0.343	42.6	26.1	41.6	32.3
Slovenia	0.766 (9)	0.883 (10)	0.620	0.108	0.241	43.0	21.4	52.2	26.4
Italy	0.766 (9)	0.873 (12)	0.570	0.538	0.208	45.0	21.8	44.3	33.8
Spain	0.756 (14)	0.876 (11)	0.643	1.485	0.436	42.2	23.4	49.0	27.6
Czech Republic	0.761 (12)	0.870 (13)	0.584	-0.039	0.419	40.9	23.0	50.8	26.3
Greece	0.759 (13)	0.865 (14)	0.570	0.099	0.028	43.5	22.6	46.4	31.1
Cyprus	0.733 (16)	0.850 (16)	0.646	1.809	1.076	35.9	23.5	58.4	18.1
Estonia	0.726 (19)	0.861 (15)	0.744	-0.610	-0.284	41.3	24.7	47.2	28.2
Lithuania	0.730 (17)	0.839 (19)	0.607	-1.248	-0.458	39.7	22.4	54.7	22.8
Poland	0.713 (20)	0.843 (18)	0.731	-0.076	0.012	39.4	21.7	56.2	22.0
Slovakia	0.738 (15)	0.844 (17)	0.585	0.013	0.090	38.9	21.4	59.5	19.1
Malta	0.729 (18)	0.839 (19)	0.613	0.351	0.304	41.4	20.8	53.3	26.0
Portugal	0.710(21)	0.830 (20)	0.681	0.393	0.038	43.0	21.8	48.9	29.3
Hungary	0.703 (22)	0.828 (21)	0.714	-0.252	-0.207	41.0	21.9	52.0	26.1
Croatia	0.670 (25)	0.818 (23)	0.872	-0.390	-0.385	43.1	22.0	49.4	28.6
Latvia	0.692 (24)	0.819 (22)	0.735	-1.252	-0.574	41.7	23.5	48.3	28.2
Romania	0.703 (22)	0.793 (24)	0.529	-0.247	-0.260	40.0	21.8	55.9	22.3
Bulgaria	0.695 (23)	0.780 (25)	0.503	-0.809	-0.763	43.4	21.2	48.7	30.1
World's mean	0.597	0.711	0.763	1.244	1.100	29.6	39.5	47.9	12.5

Table 1 Selected demographic characteristics in the EU countries, HDI indicator dynamics

Analysing the characteristics we can state that:

- the average and anticipated lifespan for all the EU countries was gradually increasing from year to year in 1990 – 2015,
- in many countries the demographic structure of population changed; the percentage of people in the age 65+ was increasing, whilst the one in the age 0-14 was decreasing,
- the EU countries featured a low natural growth as well as a low birth rate; the lowest natural growth is observed in the following countries: Lithuania, Latvia, Estonia, Bulgaria,
- the age structure of Europe's population differs significantly compared with the world average,
- Germany is a EU country with the greatest demographic encumbrance,
- amongst the EU countries, the Netherlands is the country with the highest anticipated lifespan – 81 years, Bulgaria is the lowest with 73.5 years.

It was shown that HDI values for the EU countries depend on demographic structure of these countries [6, 7].

The empirical study was conducted on the base of values referring to three fields of measure of the socio-economic development level (HDI) for the EU countries in years 1990-2014, i.e. health, education and life standard, using the following data: average lifespan, schooling indicator, illiteracy indicator, income per 1 inhabitant.

In years 1990-2014 the values for this measure for EU countries fitted in the section (0.670; 0.922). The HDI values in years 1990, 2014 are shown in table 1. From 28 EU countries as many as 26 could already in 2014 be referred to as highly developed, for 2 countries the indicator pointed to a medium development (Romania, Bulgaria).

For four selected countries: the Netherlands, Czech Republic, Poland and Bulgaria the HDI values in years 1990-2014 have been presented in figure 1. In 2014, in the highest places in ranking in the world were: Norway (0.944), Australia (0.935) and Switzerland (0.930). The last place went to Niger (0.348).

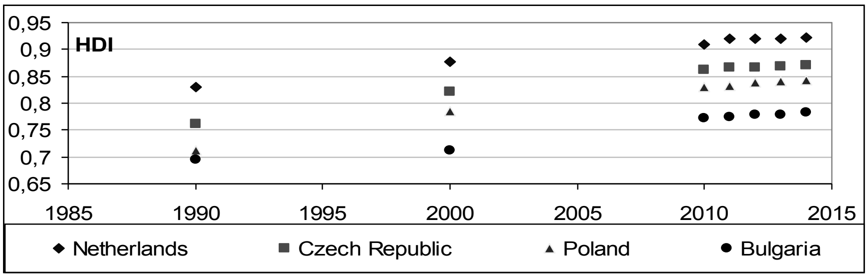


Figure 1 The HDI indicator for selected EU countries in years 1990-2014

From historical data referring to years 1990-2014 we can state that there was a systematic, slow growth of the HDI. On the base of the analysis of regression, analysis of dynamics and the study of the average rate of historic changes of the HDI value [7], we can assume that the values of the indicator in the following years will continue to rise very slowly in the majority of EU countries. In EU member countries, the differentiation between the HDI components is significant, it also clearly differs in comparison with the average in all countries of the world. The difference for selected characteristics, components of the HDI, between the maximum and minimum values is significant: for anticipated lifespan it amounts to 10 years and more, for education-3.8 years. However, the biggest difference is present in the component referring to income per inhabitant, the income of a Luxembourg citizen is almost four times bigger than the income of a citizen of Bulgaria. Nevertheless, taking into account the analysed socio-economic features, all the countries of the EU are situated above the world average.

3.2 Synthetic measure of labour market characteristics

Significant differences in the labour market indicators were observed (Table 2):

- the employment rate differed in 2013 in the EU countries by over 22%; in the Netherlands it was over 60%, whilst in Greece just under 39%,
- the unemployment rate was the lowest in Austria (4.9%), the highest in Greece (27.3%),
- the percentage of not studying and unemployed youth was also changeable in the EU countries: the lowest in France (5%), the highest in Italy (22.2%),
- the GDI significantly differentiated the countries of the EU: the difference for this variable equals as much as 43 115$ per person.

Country	Employment to population ratio (% ages 15 and older), 2013	Labour force participation rate (% ages 15 and older) 2013;	Unemploy-ment total (% of labour force) 2013;	Youth not in school and employment (% ages 15-24); 2013	Gross national income per capita 2013	m_i
Netherlands	60.1	64.1	6.7	5.1	45435	0.845 (1)
Germany	56.7	59.9	5.3	6.3	43919	0.774 (5)
Denmark	58.1	62.5	7.0	6.0	44025	0.813 (3)
Ireland	52.6	60.5	13.0	16.2	39568	0.579 (10)
Sweden	58.9	64.1	8.0	7.4	45636	0.827 (2)
United Kingdom	57.4	62.1	7.7	13.3	39267	0.693 (7)
France	54.2	57.6	5.8	5.0	38056	0.670 (8)
Austria	58.0	61.0	4.9	7.1	43869	0.795 (4)
Belgium	48.8	53.3	8.4	12.7	41187	0.517 (14)
Luxembourg	54.2	57.6	5.8	5.0	58711	0.766 (6)
Finland	54.9	59.8	8.2	9.3	38695	0.691 (7)
Slovenia	51.8	57.7	10.1	9.2	27852	0.540 (12)
Italy	43.1	49.1	12.2	22.2	33030	0.239 (24)
Spain	43.3	59.0	26.1	18.6	32045	0.266 (23)
Czech Republic	55.4	59.5	7.1	9.1	26660	0.585 (9)
Greece	38.7	53.2	27.3	20.4	24524	0.103 (27)
Cyprus	53.6	63.7	15.9	18.7	28633	0.465 (17)
Estonia	56.5	62.0	8.6	11.3	25214	0.575 (11)
Lithuania	53.8	61.0	11.8	11.1	24500	0.534 (13)
Poland	50.7	56.5	10.3	12.2	23177	0.459 (18)
Slovakia	51.1	59.5	14.2	13.7	25845	0.479 (16)
Malta	48.6	52.0	6.4	9.9	27930	0.434 (20)
Portugal	50.4	60.3	16.2	14.1	25757	0.457 (19)
Hungary	46.6	51.9	10.2	15.4	22916	0.336 (22)
Croatia	42.2	51.3	17.3	19.6	19409	0.188 (26)
Latvia	53.8	60.6	11.9	13.0	22281	0.496 (15)
Romania	52.4	56.5	7.3	17.2	18108	0.395 (21)
Bulgaria	46.4	53.3	12.9	21.6	15596	0.231 (25)

Table 2 Selected characteristics of labour market, values of synthetic measure and ranking

Taking into consideration the above mentioned labour market characteristics, a synthetic measure was constructed. Its values as well as the ranking of the EU countries is presented in table 2 and in figure 2. The top most places go to: the Netherlands, Sweden and Denmark, and the lowest places to: Greece, Croatia, and Bulgaria. The Pearson correlation coefficient between the values of the HDI in 2013 and the synthetic measure amounts to 0.7, which suggests a significant interdependence between the studied synthetic measures. Therefore, anticipating a slow growth of HDI measure in the coming years, we can also expect a slow growth of labour market synthetic measure, that is an improvement of the situation on this market in all the EU countries.

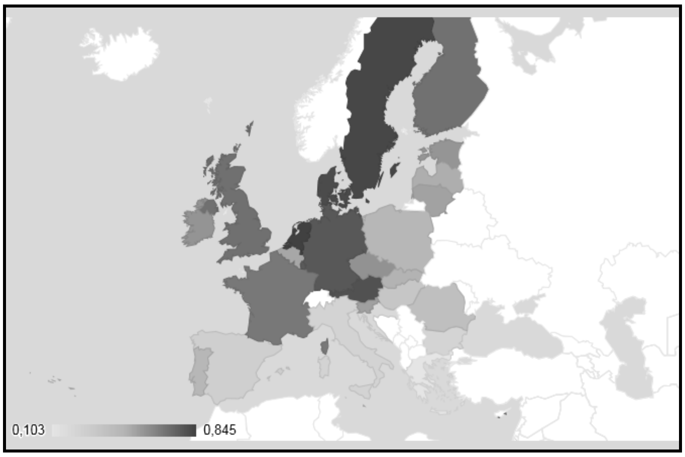


Figure 2 The labour market synthetic measure for the EU countries

4 Conclusion

It can be stated that the changes taking place in the EU countries in years 1990-2015 positively influenced the studied socio-economic indicators as well as the labour market indicators. However, taking into consideration the dynamics of changes, it is necessary to conduct a permanent analysis of indicators and monitor their values in order to prevent unfavourable trends and preserve a balanced development of economic growth in the EU countries.

The observed demographic and economic characteristics as well as labour market characteristics are in connection with each other. The interdependences which took place in the past can be used to anticipate the values of selected characteristics in the future. Unfavourable demographic phenomena, decreasing number of population in many EU countries as well as changes in people's age structure for historical data, in many cases influenced the growth of HDI components and increased the value of synthetic measure for labour market variables. However, we must remember that, during a longer period of time, with a perspective of consecutive decades of the 21st century forthcoming, aging of the population is a negative phenomenon and will most certainly begin to negatively influence the changes in the social and economic field of the EU countries.

A dynamic and stochastic character of the changes in the socio-economic environment should always be included in economic analyses. An element of random can disturb the obtained results (e.g.: an increased migration in a given period). The knowledge of the relation between the variables and of the effects of demographic changes is certainly very important for economists, sociologists, the EU countries' authorities and the Council of Europe. The increased scale of migration to EU countries, which began in 2015, should be studied and its influence in the social and economic field should be analysed in order to prevent negative and unfavourable changes.

References

[1] Biernacki, M.: Kilka uwag o pomiarze dobrobytu społecznego. *Mathematical Economics* **3**, 10 (2006), 115–124.
[2] Duncan, C. J. and Scott, S.: *Human Demography and Disease*. University Press, Cambridge, 1998.
[3] Fihel, A. and Sokólski, M.: *Demografia. Współczesne zjawiska i teorie*. Wydawnictwo Naukowe Scholar, Warszawa, 2012.
[4] Giri, C. N.: *Multivariate Statistical Analysis*. Marcel Dekker, New York, 2004.
[5] Jóźwiak, J.: Demograficzne uwarunkowania rynku pracy w Polsce. In: *Rynek pracy wobec zmian demograficznych w Polsce* (Kiełkowska, M., eds.). Zeszyty Demograficzne, Warszawa, 2013, 8–23.
[6] Mastalerz-Kodzis, A. and Pośpiech, E.: Wielowymiarowa analiza porównawcza w ujęciu dynamicznym na przykładzie wybranych charakterystyk ekonomicznych. *Metody ilościowe w badaniach ekonomicznych* **16**, 4 (2015), 24–33.
[7] Mastalerz-Kodzis, A.: Dynamika przemian społeczno-ekonomicznych krajów Unii Europejskiej, *Studia Ekonomiczne. Zeszyty Uniwersytetu Ekonomiczny w Katowicach* (2016) [in print].
[8] Rencher, A. C.: *Methods of Multivariate Analysis*. John Wiley & Sons, New York, 2002.
[9] Stanton, E. A.: *The Human Development Index: A History*. PERI University of Massachusetts, Amherst, 2007.
[10] Suchecki, B., eds.: *Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych*. C.H. Beck, Warszawa, 2010.
Data: UNDP Human Development Report, Eurostat, UNESCO Institute for Statistics, (16.04.2016).

Mixing Mortality Data Approach Using Czech Cohort Data

Martin Matějka[1], Tomáš Karel[2], Jan Fojtík[3], Pavel Zimmermann[4]

Abstract. Very frequently, more sophisticated mortality models are required for modeling high age mortality where the lack of data is an important and significant problem. The lack of data not only corrupts mortality prediction attributes but may lead to wrong model estimates and inaccurate prediction results. In such situation, additional information of surrounding countries can be incorporated to the mortality model. Shorter prediction intervals as well as better prediction results can be obtained by using credibility mixing of mortality data. This paper presents comparison of two methods that can improve the estimates of old age mortality models. Cohort approach is used. The first improvement is recalculation of the population data with the Extinct cohorts method. Second improvement is incorporation of the data collected from surrounding countries using credibility theory. For all three datasets (original data + 2 improvements), extrapolation of the mortality rate is calculated assuming a parametric function for the increase with age. The estimation error is compared as well as the resulting expected residual lifetime and an annuity value.

Keywords: High-age mortality, Credibility mixing data, Multi-population mortality models

JEL Classification: C10, C11, J10
AMS Classification: 91D20, 62C12

1 Introduction

The mortality modelling for high ages has been an extremely popular and frequently discussed topic for the past years, especially due to the fact that the background for the new pension reform has been discussed. A wrong specification of these models could cause serious problems in the future. Different approaches are usually applied for adult ages as well as for high ages (e.g. more than 80 years). For adult ages, sufficient number of observations allow for specifying models through age specific parameters. For high ages on the other hand, the lack of data is a common issue, thus mortality models are traditionally formulated as one dimensional regression functions of age based on an assumption of e.g. exponential growth [3], [4] or logistic growth [6], [7].

Major problem related to high age mortality modelling is associated with the lack of observations as well as a high occurrence of errors in the data, especially for smaller populations. Two data improvements were used in this article and the impact on the results was tested. The original data for the cohort population size were acquired from the Human Mortality Database (www.mortality.org). The first adjustment applied was replacement of the cohort population size collected for the highest ages by estimates calculated using the Extinct cohorts method [2].

Second adjustment focused on incorporating the information collected for similar populations (Hungary, Germany and Slovakia) into the Czech dataset using a credibility approach. The main advantage of this model is that the confidence intervals in extrapolations are narrower. The credibility approach applied in this article was described in [1]. Originally this methodology was applied on Lee-Carter model for adult data. In this article the methodology was applied for high age data and subsequently the Gompertz–Makeham model to fit the high age mortality in the Czech Republic was used [3].

1.1 Data

The source of all input data is The Human Mortality Database (HMD) [8]. The input data includes cohort information about population size and number of deaths for four countries Czech Republic, Germany, Hungary and Slovakia. These countries were selected with respect to their geographical closeness and similarities. Only

[1] University of Economics, Department of Probability and Statistics, Prague, martin.matejka@vse.cz.
[2] University of Economics, Department of Probability and Statistics, Prague, tomas.karel@vse.cz.
[3] University of Economics, Department of Probability and Statistics, Prague, xfojj00@vse.cz.
[4] University of Economics, Department of Probability and Statistics, Prague, zimmerp@vse.cz.

age above 70 years was considered. The raw data are obviously burdened with relatively high number of errors in comparison the population size. For some cohorts, mortality (specific death rates) are for example decreasing with age, which is contrary to demographic models and data observed for larger populations. For this reason, the population size was for age higher than 81 years replaced by estimates calculated with the Extinct Cohort method. Cohorts with last observed age 110 years were considered as extinct. The extinct cohorts available are from the years 1886 to1901. The extinct cohort method assumes no migration for the highest ages. Due to this reason, population size at some age can be estimated as the sum of deaths observed in all higher ages up to the cohort extinction. Details of the EC methods can be found for example in [9]. The difference between regular mortality data from HMD and Extinct cohort can be seen in

Figure **1**.

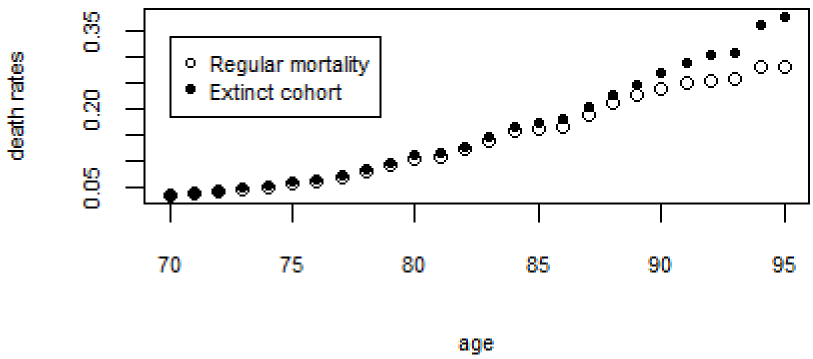


Figure 1 Application of Extinct cohort on 1901 female cohort in the Czech Republic

1.2 Theoretical background

The high age mortality data are very frequently associated with higher heteroscedasticity, missing observations or even outliers. Those can be mitigated by incorporating additional information of surrounding countries. The first step of the mixing mortality data method applied is to calculate a weighted mean of age specific death rates of the additional countries

$$m_{x,c}^{[AVE]} = \sum_{k=1}^{n} w_k m_{x,c}^{[k]}, \tag{1}$$

where $m_{x,c}^{[k]}$ is the age specific death rate in country k for age x in the cohort c. Three additional countries were used in the model (n = 3; Germany, Hungary and Slovakia). In this article, the cohorts c = 1886, 1887, …, 1901 were considered. The optimal values of weights for each country are obtained by minimizing the sum of squared differences between $m_{x,c}^{[AVE]}$ and the Czech age specific death rate $m_{x,c}^{[0]}$:

$$w_k = \arg\left(\min_{w_k} \sum_x \sum_c \left(m_{x,c}^{[0]} - w_k m_{x,c}^{[k]} \right)^2 \right), \tag{2}$$

under the restrictive conditions

$$\begin{aligned} & w_k \geq 0, \quad k = 1, \ldots, n, \\ & \sum_{k=1}^{n} w_k = 1. \end{aligned} \tag{3}$$

The next step of this method is then a calculation of credible specific death rates $m_{x,c}^{[z]}$ calculated as

$$m_{x,c}^{[z]} = m_{x,c}^{[0]} z_{z,c} + m_{x,c}^{[AVE]} \left(1 - z_{x,c}\right), \tag{4}$$

where $z_{x,c}$ is the credible coefficient

$$z_{x,c} = \frac{E_{x,c}^{[0]}}{E_{x,c}^{[0]} + \sum_{k=1}^{n} w_k^{(f)} E_{x,c}^{[k]}}, \tag{5}$$

where $E_{x,c}^{[k]}$ is the population size in k-th country in year x and cohort c and $E_{x,c}^{[0]}$ is the corresponding Czech population size.

Estimation of parameters

The specific death rates in Czech Republic $m_{x,c}^{[0]}$ and credible specific death rates $m_{x,c}^{[z]}$ are then for each cohort fitted with an exponential curve that corresponds with the Gompertz assumptions. Dropping the cohort index c and the dataset index *[0]* or *[z]*, we assume

$$m_x = e^{a+bx}, \tag{6}$$

where a and b are parameters. Parameters are consequently estimated using the maximum likelihood method.

It is assumed that the number of deaths D_x is Poisson distributed with the expected value $m_x E_x$, where m_x is the function specified by the equation (6). The parameters are estimated maximizing the likelihood function

$$l = \sum_x D_x \ln\left(m_x E_x\right) - E_x m_x - \ln\left(D_x!\right). \tag{7}$$

1.3 Results

The following chapter describes a practical application of the presented mortality model (specified by the eq. 6) applied to both Czech and mixed mortality data. The main focus is on the impact of using mixed data on the point estimates of the age specific death rates as well as the corresponding confidence intervals. The confidence intervals were calculated using standard GLM approach.

The following table describes relative difference between point estimates based on Czech and mixed data for each extrapolated age and corresponding cohort of the models based on the Czech and mixed mortality data. The negative values indicate situation when the corresponding point estimate is smaller for the mixed data. The minimum relative difference of the extrapolated area is -7.7 %, the maximum relative difference is 1.3 % and the mean is -3.5 %. Thus it can be concluded that relative difference between point estimates is very small. The point estimates of mixed mortality data are smaller if compared to point estimates of the Czech mortality data.

cohort/age	91	92	93	94	95
1886	-0.0535	-0.0534	-0.0534	-0.0533	-0.0533
1887	-0.0389	-0.0388	-0.0387	-0.0385	-0.0384
1888	-0.0150	-0.0116	-0.0081	-0.0046	-0.0010
1889	-0.0752	-0.0757	-0.0762	-0.0768	-0.0773
1890	-0.0441	-0.0449	-0.0456	-0.0464	-0.0471
1891	-0.0277	-0.0240	-0.0203	-0.0166	-0.0129
1892	-0.0364	-0.0357	-0.0350	-0.0343	-0.0336
1893	-0.0410	-0.0403	-0.0396	-0.0389	-0.0382
1894	-0.0547	-0.0552	-0.0558	-0.0563	-0.0568
1895	-0.0289	-0.0251	-0.0214	-0.0176	-0.0138
1896	-0.0096	-0.0067	-0.0038	-0.0009	0.0020
1897	-0.0061	-0.0014	0.0033	0.0080	0.0127
1898	-0.0378	-0.0366	-0.0354	-0.0342	-0.0330
1899	-0.0493	-0.0508	-0.0523	-0.0538	-0.0554
1900	-0.0374	-0.0346	-0.0318	-0.0290	-0.0261

1901	-0.0506	-0.0504	-0.0501	-0.0499	-0.0497

Table 1 Relative differences between point estimates

The following table describes the relative difference between surfaces of the lower and upper bounds of the confidence intervals for the extrapolated ages. The negative values correspond to narrower confidence intervals of the mixed data model. The minimum relative difference of the extrapolated area is -40.5 %, the maximum relative difference is -7.4 % and the mean is -25.2 %.

Cohort	Czech	mixed	relative difference
1886	0.2535	0.1840	-0.2742
1887	0.2391	0.1704	-0.2875
1888	0.2398	0.1770	-0.2618
1889	0.3056	0.1933	-0.3674
1890	0.2716	0.2039	-0.2492
1891	0.2638	0.1978	-0.2503
1892	0.2502	0.1855	-0.2587
1893	0.2548	0.2091	-0.1794
1894	0.2713	0.1890	-0.3034
1895	0.2505	0.1849	-0.2618
1896	0.2526	0.1749	-0.3075
1897	0.2413	0.1951	-0.1916
1898	0.2716	0.2130	-0.2159
1899	0.3243	0.1929	-0.4054
1900	0.2170	0.2008	-0.0743
1901	0.2314	0.1976	-0.1458

Table 2 Czech and mixed surface of the confidence intervals

Figure 2 illustrates point estimates and confidence intervals of the Czech mortality model for the first cohort, i.e. $c = 1886$.

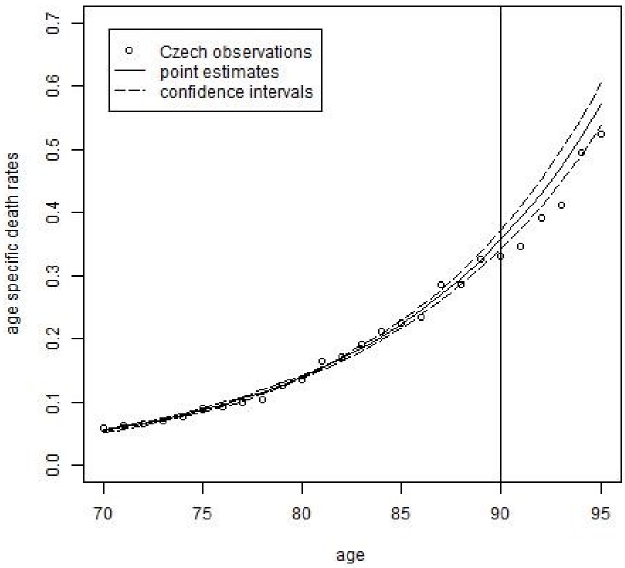


Figure 2 Point estimates and confidence intervals of the Czech mortality model

The following picture represents point estimates and confidence intervals of the Czech mortality model for the particular cohort 1886.

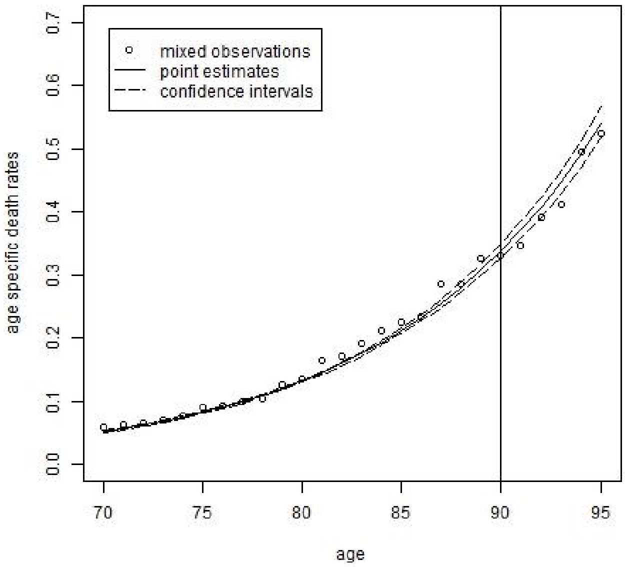


Figure 3 Point estimates and confidence intervals of the mixed mortality model

1.4 Conclusions

Mortality model assuming exponential growth of mortality was used on two mortality datasets. The population size of the first dataset was a combination of the Extinct cohorts method results and publicly available data from the Human Mortality database. The exposure of the second dataset was calculated using credibility mixed mortality death rates and cohort number of deaths. The results of this model were compared measuring the relative difference between point estimates and also surfaces obtained by comparing the lower and upper bounds of the confidence intervals. The overall conclusion is that the point estimates are very similar. Nevertheless the point estimates of the mixed mortality data models are slightly smaller if compared to results received from the Czech mortality model. The comparison of the confidence intervals surface shows that the mortality model based on the mixed data provides significantly narrower confidence intervals.

Acknowledgements

Supported by the grant IG 410025 Využití bayesovských metod pro modelování úmrtnosti of the internal grant agency of the University of economics, Prague.

References

[1] Ahcan, A, D Medved, A Olivieri, and E Pitacco.: *Forecasting Mortality for Small Populations by Mixing Mortality Data*. Insurance Mathematics & Economics, 54, 2014: 12-27.

[2] Burcin, B, K Tesárková, and L Šídlo.: *Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci*. Demografie 52 (2), 2010: 77-89.

[3] Gompertz, Benjamin.: *On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies.* Philosophical Transactions of the Royal Society of London 115, 1825: 513–583.

[4]Koschin, Felix.: *Jak vysoká je intenzita úmrtnosti na konci lidského života. Demografie (2)*, 1999: 105–119.

[5] R CORE TEAM 2013: *R: a language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing, 2012, http://www.r-project.org/.

[6] Thatcher, A *R.: The long-term pattern of adult mortality and the highest attained age.* Journal of the Royal Statistical Society: Series A (Statistics in Society) 162 (1), 1999: 5–43.

[7] Thatcher, A R, V Kannisto, and J W Vaupel.: *The force of mortality at ages 80 to 120.* Odense University Press, 1998.

[8] Wilmonth, J R, V Shkolnikov, and M Barbieri.: *Human mortality database*. 2012.

[9] Wilmoth, J R, K Andreev, D Jdanov, and D A Glei.: *Methods Protocol for the Human Mortality Database*. 2007.

Various Interest Rates Models Used within Real Options Valuation Tools

Kateřina Mičudová[1], Ladislav Lukáš[2]

Abstract. The paper is focused on real options valuation tools in corporate finance. The importance of valuation of real options stems from the fact that it provides more flexible tool for assessment of opportunities within investment projects. Hence, real options represent interesting managerial instrument for investment project evaluation along with traditional procedures based upon cash flows scenes. First, the classical discounted cash flow method is reviewed briefly. Further, we are focused on analytical and numerical methods for real options pricing. Depending on assumptions relating interest rates, there are divided into two groups. The first group of methods is based upon constant interest rate, whilst the second one is based upon stochastic interest rates. We discuss Schwartz-Moon model, Heston model and generalized Black-Scholes model, in particular. We have elaborated numerical implementation of the Heston model and the generalized B-S model in sw Mathematica. Finally, we present some interesting numerical results of both models.

Keywords: real option, interest rate models, pricing problem, stochastic volatility.

JEL Classification: C63, G13
AMS Classification: 91G80

1 Introduction

At present, real options represent interesting managerial instrument for investment project evaluation along with traditional procedures based upon cash flows scenes. The weakness of the discounted cash flow (DCF) method, the most common valuation tool in corporate finance, have become more and more obvious to practitioners and academicians, as well. In general, very volatile economic situation has requested more sophisticated valuation tools to be applied for valuation of projects and firms, too. The importance of valuation of real options stem from the fact it provides more flexible tool for assessment of opportunities within investment projects. The pricing and valuation techniques of real option analysis (ROA) are based upon methods and algorithms broadly used in finance today for pricing financial derivatives, in general. However, financial world is less complex than the real world of corporations with investments, projects, productions, and all kinds of business, see [3], [10]-[12].

Traditional approach to valuing both investment projects and firms, based on net present value (NPV), essentially involves discounting the expected net cash flows. The well-known formulation is following

$$V_0 = \varphi_0 + \sum_{t=1}^{T} \frac{\varphi_t}{\prod_{i=1}^{t}(1+\omega_i)}\,, \tag{1}$$

where φ_t, $t = 0, 1, \ldots, T$ denote corresponding cash flows in discrete times, including the initial cash flow φ_0, T stands for valuation horizon, and set of quantities ω_i represents discounting factors. As usual in case of firm valuation, φ_t and ω_i are called free cash flows to firm (FCFF) and weighted average capital costs (WACC), respectively. Well-known book [4] brings thorough overview of investment valuation methods based upon discounted cash flows and real options, as well. Highly qualified reference book presenting broad field of formulas and terms used in finance in compact form is [2].

There is evident that discounting process defined by $(\omega_i)_{i\,\geq\,1}$ together with risk-free interest rate play the central role not only in (1), but in pricing problems, in general. Sure, it holds for RO pricing, too.

2 Basic stochastic interest rate models

The simplest framework of ROA assumes the risk-free rate used for discounting to be constant over the life of the option. However, much more realistic assumption is to accept a non-constant risk-free rate, which is heavily

[1] University of West Bohemia/Faculty of Economics, Dept. of Economics and Quantitative Methods, Husova 11, 306 14 Pilsen, pitrovak@kem.zcu.cz.
[2] University of West Bohemia/Faculty of Economics, Dept. of Economics and Quantitative Methods, Husova 11, 306 14 Pilsen, lukasl@kem.zcu.cz

supported also by empirical evidence of term structure of government bond yields and other financial instruments analyzed in order to estimate market risk-free rate properly. In general, the problem of risk-free rate determination becomes much more involved, as given in [8], in particular.

In basic stochastic framework, a non-constant risk-free rate is tackled as short rate r_t, which is given by instantaneous forward rate $f(t,T)$, as being well-known and given, e.g. in [2], pp.127-132.

$$r_t = \lim_{K\to t} f(t,T) = -\frac{\partial \ln P(t,t)}{\partial T}, \quad f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T}, \tag{2}$$

where $P(t,T)$ is price of zero bond with maturity time T which, as we know, guarantees payment of unit currency in time T being traded in time $t < T$.

Now, we list the most usual stochastic models for short rate r_t, which can be implemented for generation non-constant risk-free rates. They are defined by the SDE-s, all for $t \geq 0$, as given in [4], p.354, and [11], pp. 96-111, respectively.

- Vasicek model: $$dr_t = \alpha(\gamma - r_t)dt + \sigma dB_t, \quad \alpha > 0, \quad \gamma \in \mathrm{R}, \quad \sigma > 0, \tag{3a}$$
- Hull-White model: $$dr_t = (\phi(t) - \alpha r_t)dt + \sigma dBt, \ \alpha, \sigma \in \mathrm{R}+. \tag{3b}$$
- Ho-Lee model: $$dr_t = \alpha(t)\, dt + \sigma dBt, \quad \alpha, \sigma \in \mathrm{R}+. \tag{3c}$$
- Dothan model: $$dr_t = \alpha\, r_t\, dt + \sigma\, r_t\, dBt, \ \alpha, \sigma \in \mathrm{R}+. \tag{3d}$$
- Black-Derman-Toy model: $$d \ln(r_t) = \phi(t)dt + \sigma dBt, \ \alpha, \sigma \in \mathrm{R}+. \tag{3e}$$
- Black-Karasinski model: $$d \ln(r_t) = (\phi(t) - \alpha \ln(r_t))dt + \sigma dBt, \quad \alpha, \sigma \in \mathrm{R}+. \tag{3f}$$
- Cox-Ingersoll-Ross model: $$dr_t = \alpha(\gamma - r_t)dt + \sigma \sqrt{|r_t|}\, dBt, \quad \alpha, \sigma \in \mathrm{R}+. \tag{3g}$$

where Hull-White model is sometimes called extended Vasicek model. We write all these models with positive volatility σ. Notice, that there exist versions in literature also with time dependent volatility $\sigma(t)$ with non-negative range. The Vasicek model was the first published in that topic, and it is also alternatively called mean-reversion model, as it turns outlying values of r_t from its pre-defined constant level γ. The first four models (3a) – (3d) may produce negative values of r_t, while the last three allow positive values only, as (3e) and (3f) use a lognormal process, and (3g) a square-root one. The time dependent trends, i.e. functions $\alpha(t)$, and $\phi(t)$, are assumed to be linear or exponential with positive range, in general.

We just remark briefly that there is available more general approach to the modeling of interest rate r_t, based upon the HJM (Heath-Jarrow-Morton) framework, in particular. Other promising directions in this research field are volatility approach and pricing kernel approach, respectively. However, these are out of our scope here.

3 Real option pricing framework

In this section, we want to make a short overview of models, which are suitable for solving RO pricing problems in practice either by analytic or numerical methods. Since the topic is rather large, we make just selection under subjective choice of three groups of methods, and refer readers to [1], [4], [5], [7], [9] – [12], for more details.

The methods for real options pricing, depending on assumptions relating interest rates, are divided into two groups. The first group of methods is based upon constant interest rate, whilst the second one is based upon stochastic interest rates. We follow [10] and [11] for brief description of most usual methods from both groups.

The valuation of investment projects with strategic and operating options started in 80-ties of the last century. The main idea coined was to perceive discretionary investment opportunities as so called growth options. Since that time two main directions of real options valuations can be distinguished.

1. Analytical methods. These methods can be divided into closed-form solutions and various methods providing approximate analytical solutions. In such cases, capital budgeting problems are to be written in analytic forms of specific simplified problems with known solutions. In continuous time models, the most important real options considered assume constant risk-free rate, and are following:
   - Option to defer – for the gross project value $(V_t)_{t\geq 0}$, one can usually use a simple diffusion process (4a) given by stochastic differential equation (SDE), with α giving the instantaneous ex-

pected return on the project, σ its instantaneous standard deviation, and B_t denoting standard Brownian motion, as usual

$$dV_t = \alpha V_t\, dt + \sigma V_t\, dB_t, \quad t \geq 0, \quad \alpha, \sigma \in \mathrm{R}^+. \tag{4a}$$

An alternative form providing possibility to include the payout rate D is following

$$dV_t = (\alpha - D)V_t\, dt + \sigma V_t\, dB_t, \quad t \geq 0, \quad \alpha, \sigma \in \mathrm{R}^+. \tag{4b}$$

- Option to shut down or abandon – for the output price of project $(P_t)_{t\geq 0}$, the similar SDE is used just with replacing $(V_t)_{t\geq 0}$, by $(P_t)_{t\geq 0}$, particularly.
- Option to switch – gives a chance to exchange one non-dividend paying risky asset $(V_t)_{t\geq 0}$ for another one $(S_t)_{t\geq 0}$, assuming both processes are governed by SDE (4a), or (4b), but with different constants.
- Simple compound option – calculates the value of an option on a stock which can be seen as a European call option on the value of the firm assets.

2. Numerical methods. These can be divided into methods that approximate the partial differential equations (PDE), and methods that approximate the underlying stochastic processes. The first ones include numerical integration methods including various methods of integral transformations, for example Fourier transformation. The second group of methods consists of various lattice methods, like well-known Cox-Ross-Rubinstein binomial tree method, Monte Carlo simulation methods, and direct methods of numerical solution of SDE-s.

The *first method* we select, there is the Schwartz-Moon model, and we follow [11], pp. 165-170, to formulate it here. Within a framework of ROA, it occupies specific position as being proposed originally for evaluation of research and development projects. However, it seems to be very useful and attractive for firm valuation, as well.

The Schwartz-Moon model is a stochastic model that describes the value of an investment project (the asset) depending on the cost process to complete the project. It also includes an NPV solution which in the case of a deterministic instead of stochastic world can be used as a starting point of the numerical solution. The model itself is based upon three stochastic processes, which also serve as three different sources of uncertainty accepted: i) the investment cost process, ii) the process describing the future payoffs from the investment, and iii) the process that models jump or unexpected catastrophic events.

Its basic approach is to distinguish two stages. The first stage concerns with completing the investment/asset. While the second stage deals with the income generating stream which starts after completion of the investment. In this scope in particular, it stands in line with the well-known deterministic DCF type procedures of valuation of firms.

Changes in the expected cost of project completion are assumed to follow a diffusion process with variance proportional to the level of investment. In general, the model assumes three sources of uncertainty in following forms.

Investment cost process $(K_t)_{t\geq 0}$ is defined as the expected cost to completion $K_t = E(\kappa_t)$, $t\geq 0$, where $(\kappa_t)_{t\geq 0}$ is assumed to be an underlying process which represents the cost to completion of project in theoretical framework, i.e. it exists and under corresponding probability measure yields $E(\kappa_t)$ bounded. The process $(K_t)_{t\geq 0}$ is to take behavior given by following SDE

$$dK_t = -I\, dt + \beta\sqrt{I\, K_t}\, d_1B_t, \quad t \geq 0, \quad I > 0, \; \beta \in \mathrm{R}, \quad K_0 > 0 \text{ const.}, \tag{5}$$

where I gives the rate of investment, β is a parameter that is used for volatility term calibration, and $({}_1B_t)_{t\geq 0}$ is the standard Brownian motion that is assumed to be uncorrelated with the market portfolio and with the aggregated wealth.

Asset value process $(V_t)_{t\geq 0}$ is defined as the expected future net cash flow from the project after its completion in time t. This process $(V_t)_{t\geq 0}$ is assumed to be given by another SDE

$$dV_t = \mu V_t\, dt + \sigma V_t\, d_2B_t, \quad t \geq 0, \quad \mu \in \mathrm{R}, \quad \sigma \in \mathrm{R}^+, \quad V_0 \in \mathrm{R} \text{ const.}, \tag{6}$$

where μ gives instantaneous drift which characterizes the particular investment project, which could be positive or negative, σ is the instantaneous standard deviation of the proportional changes in the value received at completions, and rate of investment, and $({}_2B_t)_{t\geq 0}$ is the standard Brownian motion that is assumed to be uncorrelated with the $({}_1B_t)_{t\geq 0}$ process thus implying the two processes $(K_t)_{t\geq 0}$ and $(V_t)_{t\geq 0}$ to be uncorrelated, too.

Jump or catastrophic process $(N_t)_{t\geq 0}$ is modeled as Poisson process with given rate $\lambda \in]0,1[$ which defines the probability per unit time that the project value will drop to zero prospectively.

The value of the investment project is modeled as a contingent claim with its underlying asset formed by the project value at the completion time involving the expected cost to completion, too. Within continuous time framework, let $F = F(V,K)$ be the function that represents the value of the investment opportunity depending upon variables V and K, where V represents the estimated value of an asset to be obtained at some time in the future, and K is the expected value of the random cost to completion of that asset. Following [11], pp.167-168, the function $F(V,K)$ is formulated as an optimal solution over investment rate I of the following PDE

$$\max_I \left(\frac{1}{2}\sigma^2 V^2 \frac{\partial^2 F}{\partial V^2} + \frac{1}{2}\beta^2 I\, K \frac{\partial^2 F}{\partial K^2} + (\mu - \eta) V \frac{\partial F}{\partial V} - I \frac{\partial F}{\partial K} - (r_f + \lambda) F - I \right) = 0, \quad I \in \mathrm{R}^+, \tag{7}$$

where $(\mu - \eta)$ is the risk-adjusted drift for the asset value, and r_f is the risk-free interest rate being constant in this setting. The investment takes place at the maximum rate I_m or not as follows

$$\begin{aligned} I = I_m, &\text{ for } \frac{1}{2}\beta^2 K \frac{\partial^2 F}{\partial K^2} - \frac{\partial F}{\partial K} - 1 \geq 0, \\ = 0, &\text{ otherwise.} \end{aligned} \tag{8}$$

Equation (7) is an elliptic PDE with free boundary along the line $V^*(K)$ such that $I = I_m$ if $V > V^*(K)$, else $I = 0$, in accordance with (8). The set of critical asset values $V^*(K)$ must be found together with the function $F(V,K)$, and it satisfies the condition being issued from (8)

$$\frac{1}{2}\beta^2 K \frac{\partial^2 F(V^*,K)}{\partial K^2} - \frac{\partial F(V^*,K)}{\partial K} - 1 = 0. \tag{9}$$

Concluding, one can formulate the free boundary problem for investment project in following way

For $V > V^*(K)$, the value of investment opportunity F must satisfy the following PDE being linked to (7)

$$\frac{1}{2}\sigma^2 V^2 \frac{\partial^2 F}{\partial V^2} + \frac{1}{2}\beta^2 I_m K \frac{\partial^2 F}{\partial K^2} + (\mu - \eta) V \frac{\partial F}{\partial V} - I_m \frac{\partial F}{\partial K} - (r_f + \lambda) F - I_m = 0, \tag{10a}$$

otherwise, it must satisfy diffusion equation

$$\frac{1}{2}\sigma^2 V^2 \frac{\partial^2 F}{\partial V^2} + (\mu - \eta) V \frac{\partial F}{\partial V} - (r_f + \lambda) F = 0, \tag{10b}$$

with boundary conditions

$$F(V,0) = V, \quad F(0,K) = 0, \quad \lim_{K\to\infty} F(V,K) = 0. \tag{10c}$$

Numerical solution of free boundary value problem (10a) – (10c) for function F: $\mathrm{R}^+ \times \mathrm{R}^+ \to \mathrm{R}^+$ can be constructed either by finite difference method or finite element method, as well.

The *second method* belongs to stochastic volatility model group. The volatility approach, as mentioned already above in connection with interest rate modeling, is used in field of option valuation problems, too. In general, it provides possibility to construct more realistic models, which are more general than usual Black-Scholes framework of option pricing based upon geometric Brownian motion with volatility σ constant, exclusively.

Following [5], we can write a scalar value asset price process $(S_t)_{t\geq 0}$ as being issued by two-dimensional coupled diffusion process of the general form

$$\begin{aligned} dS_t &= (\alpha_t S_t - D_t)\, dt + \sigma_t S_t\, d_1B_t, \quad \sigma_t^2 = v_t, \quad E(d_1B_t, d_2B_t) = \rho\, dt, \quad \rho \in [-1, 1], \\ dv_t &= b(v_t)\, dt + \xi\, \eta(v_t)\, d_2B_t, \quad t \geq 0, \quad v_t \in \mathrm{R}^+, \quad \sigma \in \mathrm{R}^+, \end{aligned} \tag{11}$$

where ρ is correlation of standard Brownian motion processes d_1B_t, and d_2B_t, as usual. Quantity D_t represents dividends which are paid by S_t. The process $(v_t)_{t\geq 0}$ defines volatility of process $(S_t)_{t\geq 0}$, being expressed by corresponding variance relation $\sigma_t^2 = v_t$, as given in (11). Noting, volatility in finance is expressed usually either by standard deviation or variance, which choice depends mainly upon the context.

In our numerical computations which are presented in **Fig. 1** and **Fig. 2**, we solve option pricing problem with underlying asset following the Heston model, sometimes called square root model, too, which is perhaps the most popular model in field of stochastic volatility. It is a particular case of (11) as follows

$$\begin{aligned} &dS_t = \mu_t\, S_t\, dt + \sigma_t S_t\, d_1B_t, \quad \sigma_t^2 = v_t\,, \quad E(d_1B_t, d_2B_t) = \rho\, dt\,, \quad \rho \in [-1, 1], \\ &dv_t = \alpha(\gamma - \beta\, v_t)\, dt + \xi\, \sqrt{v_t}\, d_2B_t, \quad t \ge 0, \quad \alpha, \gamma, \beta, \xi \in \mathrm{R}^+, \quad 2\alpha\gamma\ \ge\ \xi^{\,2}, \end{aligned} \tag{12}$$

where $2\alpha\gamma \ge \xi^{\,2}$ is the Feller condition ensuring strictly positive instantaneous variance v_t.

For testing numerical procedure we coded in Mathematica, we select a bench-mark problem given in [5], pp. 53-54, 67-70, with following data: call option, asset spot price $S \in [70, 140]$, exercise price $K = 100$, risk-free rate $r_f = 0$, expiration time $T = 0.5$, volatility $\sigma = 0.01$, and further $\alpha\gamma = 0.02$, $\xi = 0.1$, $\alpha\beta = 2$, $\rho = 0.5$, and $\rho = -0.5$. The results are depicted in **Fig. 1**. Further, we compared the results with well-know Black-Scholes model using the same data, and the result is given in **Fig. 2**, left panel. Inspecting these three results, we are hardly able to notice any difference on the first glance. However, in **Fig. 2**, right panel, we present the difference $\Delta = C_{\mathrm{H}}(\rho) - C_{\mathrm{BS}}$, where $C_{\mathrm{H}}(\rho)$ stands for call option price with underlying asset following the Heston model (12) calculated for given data and two values ρ, and C_{BS} denotes the call option price calculated with Black-Scholes model assuming the underlying asset to obey geometric Brownian process.

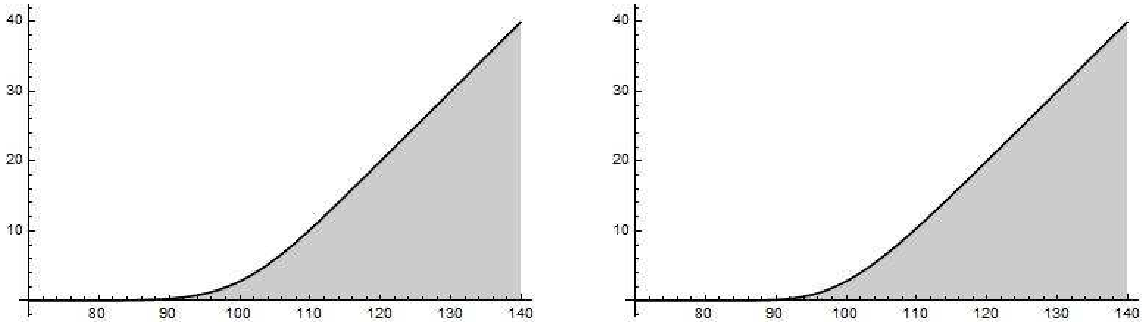


Figure 1 Call option price – Heston model: $\rho = 0.5$ ~ left panel; $\rho = -0.5$ ~ right panel

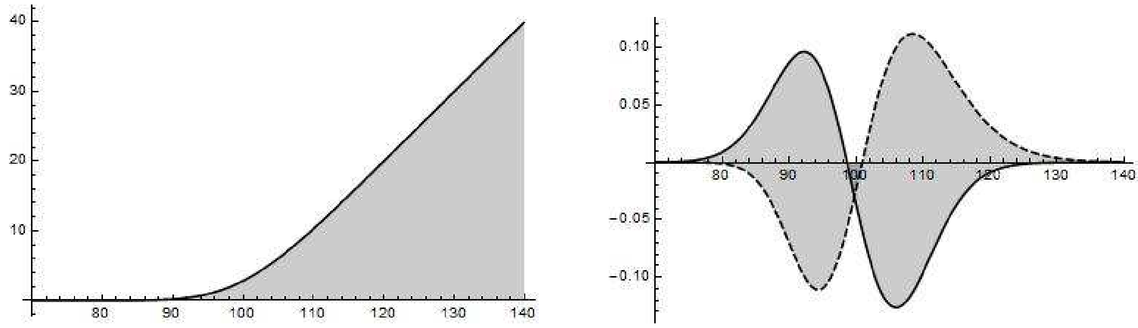


Figure 2 Call option price – B-S model ~ left panel;   difference $\Delta = C_{\mathrm{H}}(\rho) - C_{\mathrm{BS}}$, $\rho = 0.5$ full line, $\rho = -0.5$ dashed line ~ right panel

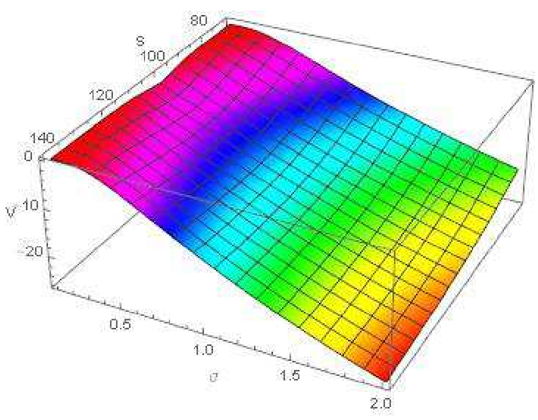


Figure 3 Cubic shape functions $\varphi_n(\xi\,;\,\cdot)$, n=1,…,4, where · stands for unit interpolation parameter

Finally, the *third method* belongs to group of generalized Black-Scholes models. Their common feature is that they assume an underlying asset of option to follow more general process than geometric Brownian one. In **Fig. 3**, we present results gained from option pricing model assuming the underlying asset to follow a subdiffusion geometric mixed Brownian and fractional Brownian process (sg-BfB). The asset price process $(S_t)_{t\geq 0}$ is given by (13), where H is the Hurst exponent of fB process, and α gives the order of $T_\alpha(t)$, which is an inverse α-stable subordinator controlling the time. We refer to [6] for more details and other numerical results.

$$\begin{aligned} S_t &= S_0 \exp(\mu T_\alpha(t) + \sigma M_{\alpha,H}(t)), \quad S_0 > 0, \quad \mu, \sigma \in \mathrm{R}^+, \quad \alpha \in\,]2/3, 1[\\ M_{\alpha,H}(t)) &= aB(T_\alpha(t)) + bB_H(T_\alpha(t)), \quad t \geq 0, \quad a, b \neq 0, \;\; H \in\,]1/2, 1[\; . \end{aligned} \tag{13}$$

The corresponding Black-Scholes PDE for valuation of the call option on underlying asset $V(S_t,t)$ is following

$$\frac{\partial V}{\partial t} + r_f S \frac{\partial V}{\partial S} + \frac{1}{2}\hat{\sigma}^2(t) S^2 \frac{\partial^2 V}{\partial S^2} - r_f V = 0, \qquad \hat{\sigma}^2(t) = \sigma^2 \left[a^2 \frac{t^{\alpha-1}}{\Gamma(\alpha)} + b^2 (\frac{t^{\alpha-1}}{\Gamma(\alpha)})^{2H} (\Delta t)^{2H-1}\right]. \tag{14}$$

In **Fig. 3**, we depict surface representing the difference $\Delta = C_{\text{sgBfB}}(\sigma) - C_{\text{BS}}$, where $C_{\text{sgBfB}}(r_f)$ stands for call option price with underlying asset following the sg-BfB process (13) with variable volatility $\sigma \in [0.1,2]$. The basic data are the same as in previous example, and further, a, b = 0.5, Δt = 0.01. We see that values of all $C_{\text{sgBfB}}(\sigma)$ are less than C_{BS}, for $S \in [70, 140]$, and $\sigma \in [0.1,2]$.

4 Conclusions

Within the paper, we discussed the stochastic interest rate models briefly, first. Further, we present three interesting models which are useful and applicable in ROA, the Schwartz-Moon model, Heston model with stochastic volatility, and generalized B-S model with underlying asset following subdiffusion process, in particular.

Our numerical results confirm that the correlation coefficient relating two standard Brownian motion processes within the Heston model plays very interesting role. Considering the difference between Heston model and classical B-S one, we show reverting influence along whole underlying asset price range given.

Finally, we also show that all values of generalized B-S model with underlying asset following the sg-BfB process are less than corresponding ones calculated by classical B-S model, for given ranges of asset price and volatility.

Acknowledgements

The research project was supported by the grant no. 15-20405S of the Grant Agency, Prague, Czech Republic.

References

[1] Bossu, S.: *Advanced equity derivatives*. Wiley, Hoboken, New Jersey, 2014.

[2] Cipra, T.: *Financial and insurance formulae* (in Czech: *Finanční a pojistné vzorce*). Grada Publishing, Praha, 2006.

[3] Damodaran, A.: *Damodaran on Valuation – Security Analysis for Investment and Corporate Finance*. 2-nd ed., John Wiley and Sons, New York, 2006.

[4] Fries, Ch.: *Mathematical Finance Theory, Modeling, Implementation*. Wiley, Hoboken, New Jersey, 2007.

[5] Lewis, A.A.: *Option Valuation under Stochastic Volatility with Mathematica Code*. 2 ed., Finance Press, Newport Beach, California, 2005.

[6] Lukáš, L.: European Option Pricing Model with Underlying Asset Following Subdiffusion Process. In: *Advanced Methods of Quantitative Finance* (Málek, J., editor). Oeconomica Publ.House, Univ. of Economics, Prague, 2015 Chap.7, 91-106.

[7] Neftci, S.N.: *An Introduction to the Mathematics of Financial Derivatives*. 2.ed, Academic Press, Orlando, Florida, 2000.

[8] Pykhtin, M., Rosen, D.: Pricing Counterparty Risk at the Trade Level and CVA Allocations. Online: https://www.moodys.com/microsites/crc2020/papers/cva_allocations.pdf

[9] Rambeerich, N., et al.: High-order computational methods for option valuation under multifactor models. *EJOR* **224** (2013), 219–226.

[10] Schoene, M.: *Real Options Valuation – The Importance of Stochastic Process Choice in Commodity Price Modelling*. Springer Gabler, Springer Fachmedien, Wiesbaden, 2015.

[11] Schulmerich, M.: *Real Options Valuation – The Importance of Interest Rate Modelling in Theory and Practice*. 2-nd ed., Springer Verlag, Berlin, 2010.

[12] Schwartz, E.S., Trigeorgis, L.: Real Options and Investment under Uncertainty: An Overview. Online: http://mitpress.mit.edu/sites/default/files/titles/content/9780262693189_sch_0001.pdf.

Application of Coalitional Values on Real Voting Data

Elena Mielcová[1]

Abstract. Simple characteristic function form games can be used for description of situation in voting bodies. An influence of coalitions in such types of games are usually measured by so-called power indices. Kojima and Inohara proposed different view on coalition power – they proposed two values – namely viability and blockability values – as a tool to evaluate coalition influence. Thus, the main aim of this article is to apply coalitional values - viability and blockability values - on real voting data. In order to cover all possible outcomes of real voting data, it was necessary to add uncertainty issues into the concept of coalitional values – namely the concept of Atanassov's intuitionistic fuzzy sets approach. Real voting data are in this case represented by a voting data from the Chamber of Deputies of the Parliament of the Czech Republic for two consequent periods. Results of the ex-ante and ex-post coalitional numbers are compared with overall voting outcome. Calculations show an improvement in results when uncertainty issues are considered.

Keywords: blockability value, viability value, parliamentary voting, Chamber of Deputies of the Czech Republic, I-fuzzy sets.

JEL Classification: C71
AMS Classification: 91A80

1 Introduction

A voting game can be described as a simple characteristic function form game. Power indices are considered to be a standard way how to measure a coalition influence in such types of games. However, Kojima and Inohara [5] studied different methods for comparison of coalition influence of characteristic function games; they introduced blockability and viability relations in order to compare influence of players. Moreover, they defined two values – blockability and viability value [6] such that each of the values indicates a coalition influence by a real number. Both coalitional values are bound with solution concepts of a characteristic function game.

The main aim of this article is to apply coalitional values on real voting data. In order to cover all possible outcomes of real voting data, it was necessary to add uncertainty issues into the concept of coalitional values – namely the concept of Atanassov's intuitionistic fuzzy sets approach. The extension of fuzzy sets approach such that membership, nonmembership as well as an uncertainty part of a set is taken into account was introduced by K. T. Atanassov in 1980's as the theory of intuitionistic fuzzy sets [1]. Later Dubois et al. [3] proposed use of different term concerning Atanassov intuitionistic fuzzy sets in order to avoid terminological difficulties in already introduced different approach to intuitionistic fuzzy sets; thus throughout text, the term "I-fuzzy" will be used instead of the term "Atanassov intuitionistic fuzzy".

Real voting data are in this case represented by a voting data from the Chamber of Deputies of the Parliament of the Czech Republic for two consequent periods. Results of the ex-ante and ex-post coalitional numbers are compared with overall voting outcome. Calculations show an improvement in results when uncertainty issues are considered.

2 Preliminaries

This part of the article covers basic definitions of weighted voting game, coalitional values (blockability and viability value), as well as a definition of I-fuzzy weighted coalition game.

2.1 Coalitional values

Any voting of players with different weights can be described as a weighted voting game for n players $N=\{1,2,\dots,n\}$ with weights $w=\{w_1,w_2,\dots,w_n\}$ and quota q. The triple $G=[N,w.q]$ is called a weighted vot-

[1] Silesian University in Opava / School of Business Administration in Karviná, Univerzitní náměstí 1934/3, 733 40 Karviná, Czech Republic, mielcova@opf.slu.cz.

ing game. A coalition T is a subset of players $T \subset N$. A coalition T is winning if its total weight meets or exceeds the quota q, that means $\sum_{i \in T} w_i \geq q.$

Let function v be a characteristic function of a weighted voting game. This characteristic function can obtain only values 0 or 1; $v(T) = 1$ if T wins and $v(T) = 0$ when T loses. The voting game with characteristic function $v: 2^N \to \{0,1\}$ is the simple game that should satisfy these three conditions [4]:

(i) the empty coalition never wins: $v(\varnothing) = 0$,

(ii) the grand coalition always wins: $v(N) = 1$,

(iii) any superset of winning coalition also wins.

In order to compare a coalition influence, Kojima and Inohara [5] proposed to use a blockability value, and a viability value. These two values assign a player with a real number; the bigger the number, the higher a coalition power. The blockability value of the coalition $S \subset N$ is expressed as:

$$\hat{B}_S(N, v) = \frac{\sum_{T \subset N} v(T) - B^*(S)}{\sum_{T \subset N} v(T) - B^*(N)} v(N) \tag{1}$$

where $B^*(S) = \sum_{T \subset N} v(T \setminus S)$. Taking into account that for simple games $v(N) = 1$ and $B^*(N) = 0$, expression (1) can be rearranged as:

$$\hat{B}_S(N, v) = \frac{\sum_{T \subset N} v(T) - B^*(S)}{\sum_{T \subset N} v(T)} \tag{2}$$

Similarly, the viability value of the coalition $S \subset N$ is expressed as:

$$\hat{V}_S(N, v) = \frac{V^*(S)}{V^*(N)} v(N) \tag{3}$$

where $V^*(S) = \sum_{T \subset N} v(S \setminus T)$. In simple games, when $v(N) = 1$, expression (3) can be rearranged as:

$$\hat{V}_S(N, v) = \frac{V^*(S)}{V^*(N)} \tag{4}$$

Example 1. Calculate a blockability value for A, B, C, and D, and a viability value of a coalition ABC for a game

$$G = \left[\acute{N} = \{A, B, C, D\}, w = (80,20,10,90), q = 101\right]$$

Solution:

$$\sum_{T \subset N} v(T) = v(\varnothing) + v(A) + v(B) + v(C) + v(D) + v(AB) + v(AC) + v(AD) + v(BC) + v(BD) + v(CD) + v(ABC) +$$
$$+ v(ABD) + v(ACD) + v(BCD) + v(ABCD) = 0 + 0 + 0 + 0 + 0 + 0 + 0 + 1 + 0 + 1 + 0 + 1 + 1 + 1 + 1 + 1 = 7$$

$$B^*(A) = 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 1 + 0 + 0 + 1 + 0 + 1 + 1 = 1$$

Similarly, $B^*(B) = 4$, $B^*(C) = 6$, $B^*(D) = 2$. Blockability values are:

$$\hat{B}_A(N, v) = \frac{\sum_{T \subset N} v(T) - B^*(A)}{\sum_{T \subset N} v(T)} = \frac{7 - 4}{7} = \frac{3}{7},\ \hat{B}_B(N, v) = \frac{3}{7},\ \hat{B}_C(N, v) = \frac{1}{7},\ \hat{B}_D(N, v) = \frac{5}{7}.$$

In order to find viability values we need to calculate:

$$V^*(N) = \sum_{T\subset N} v(N\setminus T) = v(\varnothing)+v(A)+v(B)+v(C)+v(D)+v(AB)+v(AC)+v(AD)+v(BC)+v(BD)+$$
$$+v(CD)+v(ABC)+v(ABD)+v(ACD)+v(BCD)+v(ABCD)=$$
$$=0+0+0+0+0+0+0+1+0+1+0+1+1+1+1+1=7$$
$$V^*(ABC) = \sum_{T\subset N} v(ABC\setminus T) = v(ABC\setminus\varnothing)+v(ABC\setminus A)+v(ABC\setminus B)+v(ABC\setminus C)+v(ABC\setminus D)+v(ABC\setminus AB)$$
$$+v(ABC\setminus AC)+v(ABC\setminus AD)+v(ABC\setminus BC)+v(ABC\setminus BD)+v(ABC\setminus BD)+v(ABC\setminus CD)+v(ABC\setminus ABC)+$$
$$v(ABC\setminus ABD)+v(ABC\setminus ACD)+v(ABC\setminus BCD)+v(ABC\setminus ABCD)=$$
$$=1+0+0+0+1+0+0+0+0+0+0+0+0+0+0+0=2$$

A viability value for coalition ABC is then

$$\hat{V}_{ABC}(N,v) = \frac{V^*(ABC)}{V^*(N)} = \frac{2}{7}$$

2.2 I-fuzzy weighted coalition game

An intuitionistic fuzzy set is defined as a set of triples $A=\{\langle x_i,\mu_A(x_1),\nu_A(x_1)\rangle\}$ over a fixed set $X=\{x_1,x_2,\ldots,x_n\}$. Functions $\mu_A : x\to L$ and $\nu_A : x\to L$ for L=[0,1] define the degree of membership and the degree of non-membership of the element $x_i \in X$ to $A\subset X$. For an intuitionistic fuzzy set, the condition $0\le \mu_A(x_i)+\nu_A(x_i)\le 1$ holds for all $x_i\in X$ (see [1]).

In this analysis, weights of players will be expressed as intuitionistic fuzzy numbers, defined in [2]. An intuitionistic fuzzy subset $A=\{\langle x_i,\mu_A(x_1),\nu_A(x_1)\rangle\}$ of the real $x_i\in R$ line is called an intuitionistic fuzzy number if:

- A is if-normal;
- A is if-convex (its membership function is fuzzy convex and its non-membership function is fuzzy concave);
- its membership function is upper semi-continuous and its non-membership function is lower semi-continuous;
- A is bounded $A=\{x\in X : \nu_A(x)<1\}$.

In general, any I- fuzzy coalition consists of a group of participating players along with their participation and non-participation level for each player. Any I-fuzzy coalition $\tilde{C}$ is represented as $\tilde{C}=\langle \mu^C,\nu^C\rangle$ with $\mu^C=(\mu_1,\mu_2,\ldots,\mu_n)$ and $\nu^C=(\nu_1,\nu_2,\ldots,\nu_n)$ such that $0\le \mu^C+\nu^C\le 1$ for set of players $N=\{1,2,\ldots,n\}$. Values μ_i and ν_i are expressing level of participation and nonparticipation of player i in coalition $\tilde{C}$, respectively. The empty coalition $\tilde{C}^{\varnothing}=\langle \mu^{\varnothing},\nu^{\varnothing}\rangle$ where $\mu^{\varnothing}=0$ and $\nu^{\varnothing}=1$ for all i. The grand coalition $\tilde{C}^N=\langle \mu^N,\nu^N\rangle$ where $\mu^N=1$ and $\nu^N=0$ for all i. Any crisp coalition $S\subset N$ can be expressed as $\tilde{C}^S=\langle \mu^S,\nu^S\rangle$ where $\mu_i^S=1$ for all $i\in S$ and $\nu_i^S=1$ for all $i\notin S$; $\tilde{C}^i$ denote an I-fuzzy coalition notation for a crisp set $S=\{i\}$.

3 Data description

This analysis is based on the data from the Czech Parliament from the 2006-2013 electoral periods. Basic information on the Czech Parliamentary system as well as the set of all historical votes can be found at the official web site of the Chamber of Deputies of the Czech Parliament URL: www.psp.cz.

The studied time period covered two different electoral periods. Data for the 2006-2010 parliamentary period covered 8740 voting vectors; data for the 2010-2013 parliamentary period covered 5895 voting vectors. The outcome of every vote for every member is one possibility from the set {yes, no, abstain, absent}. Every bill to be passed needs at least as many “yes” votes as quota. Quota is dependent on the number of all present legislators. That means that the possibility “abstain” in fact serves as “no” outcome. Hence, in this analysis the “abstain” outcome is reclassified to “no” outcome.

During the 2006-2010 parliamentary period there were five political parties operating in the Chamber of Deputies of the Czech Parliament:

- Civic Democratic Party (ODS),
- Christian and Democratic Union - Czechoslovak People's Party (KDU-CSL),
- Green Party (SZ),
- Czech Social Democratic Party (CSSD),
- Communist Party of Bohemia and Moravia (KSCM).

Three political parties – ODS, KDU-CSL, and SZ – created governmental coalition; other two political parties – CSSD and KSCM – stayed in opposition. Distribution of seats in the 2006-2010 Chamber of Deputies of the Czech Republic together with party voting success index and blocking value is given in Table 1.

Political party	**ODS**	**CSSD**	**KSCM**	**KDU-CSL**	**SZ**
Seats	81	74	26	13	6
Party Success	0.808	0.812	0.721	0.749	0.678
Blockability value	0.600	0.467	0.467	0.067	0.067

Table 1 Distribution of seats in the 2006-2010 Chamber of Deputies of the Czech Republic

Correlation coefficient between the party success and the blockability value is equal to 0.69, this correlation coefficient is not statistically significant at 5% level of significance (test statistic for zero hypothesis of statistical insignificance of correlation coefficient is t=1.64, p-value p=0.2). The viability value for governmental coalition (ODS. KDU-CSL,SZ) is equal to 0.

During the 2010-2013 parliamentary period there were five political parties operating in the Chamber of Deputies of the Czech Parliament:

- Civic Democratic Party (ODS)
- TOP09
- Veci verejne (VV)
- Czech Social Democratic Party (CSSD)
- Communist Party of Bohemia and Moravia (KSCM)

Three political parties – ODS, TOP09, and VV – created governmental coalition; other two political parties – CSSD and KSCM – stayed in opposition. Distribution of seats in the 2010-2013 Chamber of Deputies of the Czech Republic together with party voting success index and blocking value is given in Table 2.

Political party	**CSSD**	**ODS**	**TOP09**	**KSCM**	**VV**
Seats	56	53	41	26	24
Party Success	0.592	0.943	0.952	0.567	0.831
Blockability value	0.5	0.5	0.25	0.25	0.25

Table 2 Distribution of seats in the 2010-2013 Chamber of Deputies of the Czech Republic

Correlation coefficient between the party success and the blockability value is equal to -0.046, this correlation coefficient is not statistically significant at 5% level of significance (p=0.94). The viability value for governmental coalition (ODS. TOP09,VV) is equal to 0.1875.

4 Uncertainty in Coalitional Values

Let T be a crisp coalition such that $T \subset N$. Let $\tilde{C}$ be an I-fuzzy coalition. Let e^T be a fuzzy coalition created from $\tilde{C}$ such membership and nonmembership function are at basic levels for players not in coalition T, other membership and nonmembership functions are unchanged. The blockability value of the coalition $S \subset N$ for the referral coalition $\tilde{C}$ can be expressed as:

$$\hat{B}_S(N,v) = \frac{\sum_{T \subset N} v(e^T) - B^*(S)}{\sum_{T \subset N} v(e^T)} \tag{5}$$

where $B^*(S) = \sum_{T \subset N} v(e^T \setminus e^S)$. In this case there the referral coalition $\tilde{C}$ is considered to be "a typical coalition" for a played game (for example in voting game it can be a most probable coalition or an announced coalitional partnership). Similarly, the viability value of the coalition $S \subset N$ and values $V^*(S) = \sum_{T \subset N} v(e^S \setminus e^T)$ for the referral coalition $\tilde{C}$ can be expressed as:

$$\hat{V}_S(N,v) = \frac{V^*(S)}{V^*(N)} \tag{6}$$

In the case of real world data, the characteristic function of a game is usually not known. If the number of previous votes is sufficiently high, then the characteristic function values can be obtained directly from data. Another possibility is to express characteristic function using interval numbers representation. Instead of interval numbers, the value of characteristic function for uncertainty interval can be estimated by using reasonable probability density function. In order to calculate a characteristic function in the case of the Chamber of Deputies of the Czech Parliament, a probabilistic approach was used, values $v(e^T)$) were calculated such that:

- if the sum of participation of party members in coalition T (for simplicity called minimal(T)) is greater than 100, then v(T)=1 (in the Chamber of Deputies of the Czech Parliament, if all 200 members are present, then the simple majority rule expects that more than 100 votes ensure passing a bill)
- if the sum of participation of party members in coalition T and all values of uncertainty of participation for all political parties (for simplicity called maximal(T)) is smaller or equal to 100, then v(T)=0.
- if the sum of participation of party members in coalition T is smaller than 100, and the sum of participation of party members in coalition T and all values of uncertainty of participation for all political parties is greater than 100, then the respective value v(T) is calculating as the value of v(T) = 1 - F(100), where F(100) is the value of a cumulative distribution function for the continuous uniform distribution (unif(minimal(T); maximal(T))

Table 3 shows participation and nonparticipation levels of "a typical coalition" calculated as average participation of political parties in winning coalitions and respective participations in losing coalitions.

Political party 2006-2010	**CSSD**	**ODS**	**KSCM**	**KDU-CSL**	**SZ**
Participation level	0.602	0.604	0.605	0.535	0.523
Nonparticipation level	0.205	0.196	0.278	0.196	0.217
Political party 2010-2013	**CSSD**	**ODS**	**TOP09**	**KSCM**	**VV**
Participation level	0.392	0.672	0.761	0.453	0.592
Nonparticipation level	0.354	0.087	0.072	0.407	0.156

Table 3 Typical coalition participation and nonparticipation levels for 2006-2013 Parliaments.

Blocking values for political parties based on participation and nonparticipation levels given in Table 3 are presented in Table 4.

Political party 2006-2010	CSSD	ODS	KSCM	KDU-CSL	SZ
Blocking value	1	1	0.042	0.042	0.014
Political party 2010-2013	**CSSD**	**ODS**	**TOP09**	**KSCM**	**VV**
Blocking value	0.621	0.889	0.759	0.282	0.395

Table 4 Blocking values for political parties present in 2006-2013 Parliaments.

Correlation coefficient between the party success and the blockability value is equal to 0.91 and 0.68, respectively. The first correlation coefficient is statistically significant at 5% level of significance, the second one is not significant, however show improvement comparing to original blockability value (p=0.206).

5 Conclusion

The main aim of this article was to apply coalitional values on real voting data. In order to cover all possible outcomes of real voting data, the I-fuzzy concept was used. Real voting data were represented by a voting data from the Chamber of Deputies of the Parliament of the Czech Republic for two consequent periods: 2006-2010 and 2010-2013 periods.

Results of the ex-ante and ex-post coalitional numbers are compared with overall voting outcome. For the 2006-2010 Parliamentary voting data the correlation coefficient of coefficient of success with ex-ante blockability value was equal to 0.69, this correlation coefficient was not statistically significant at 5% level of significance. The same correlation coefficient for ex-post blockability value was equal to 0.91 and was statistically significant. Similarly, for the 2010-2013 Parliamentary voting data the correlation coefficient of coefficient of success with ex-ante blockability value was equal to -0.046, this correlation coefficient was not statistically significant at 5% level of significance. The same correlation coefficient for ex-post blockability value was equal to 0.68; even though this correlation coefficient was not statistically significant, it showed improvement comparing to original correlation coefficient (p=0.206). Hence, calculations show an improvement in results when uncertainty issues are considered.

Acknowledgements

This research was supported by the GACR project No. 14-02424S.

References

[1] Atanassov, K. T.: Intuitionistic Fuzzy Sets. *Fuzzy Sets and Systems* **20** (1986), 87–96.

[2] Burillo, P., Bustince, H., and Mohedano, V.: Some definitions of intuitionistic fuzzy number. First properties. In: *Proceedings of the 1st Workshop on Fuzzy Based Expert Systems*, 1994, 53–55.

[3] Dubois, D., Gottwals, S., Hajek, P., Kacprzyk, J., and Prade, H.: Terminological difficulties in fuzzy set theory – the case of ‘intuitionistic fuzzy sets’. *Fuzzy Sets and Systems* **156** (2005), 485–491.

[4] Hu, X.: An asymmetric Shapley–Shubik power index. *International Journal of Game Theory* **34** (2006), 229–240.

[5] Kojima, K, and Inohara, T.: Methods for comparison of coalition influence on games in characteristic function form. *Applied Mathematics and Computation* **217** (2010), 4047–4050.

[6] Kojima, K, and Inohara, T.: Coalition values derived from methods for comparison of coalition influence for games in characteristic function form. *Applied Mathematics and Computation* **219** (2012), 1345–1353.

The application of chosen measures of deterministic chaos to building optimal portfolios

Monika Miśkiewicz-Nawrocka[1], Katarzyna Zeug-Żebro[2]

Abstract. Portfolio analysis is one of the most important techniques for investing in the capital market. Its main goal is to diversify the investment risk. In recent years, in addition to classical methods of portfolio analysis have been developed tools that are both modifications of these concept as well as new, alternative diversification techniques of investment portfolio which take into account for example the indicators of fundamental analysis. A new approach proposed in the paper is the use of the measures for identifying chaos, i.e. the largest Lyapunov exponent and the Hurst exponent. Since determinism of chaotic time series indicates on potential possibility of their prediction, it is also expected that has a significant impact on the construction of optimal portfolio.

The paper aims to construct optimal portfolios determined based on the largest Lyapunov exponent, the Hurst exponents, the TMAI measure (Taxonomic Measure of Investment Attractiveness) and the Markowitz portfolio. The test will be conducted on the basis of the financial time series.

Keywords: portfolio analysis, largest Lyapunov exponent, Hurst exponent, measure TMAI.

JEL Classification: C3, C8, G11, E4
AMS Classification: 91B28

1 Introduction

The construction of a optimal portfolio proposed by H. Markowitz [6, 7] started intensive development of scientific field which is a portfolio analysis. Research conducted for many years in various scientific centers have provided new tools and approaches for estimating the shares in the optimal portfolio. One of them is method based on taxonomic measure of investment attractiveness (TMAI) [15, 16]. A new approach proposed by the authors is the use to construct the optimal portfolio of two measures for identifying chaos, i.e. the largest Lyapunov exponent [9, 10] and the Hurst exponent [3, 18].

The aim of the paper will be to an attempt to diversify the risk of the investment portfolio, based on the constructed optimal portfolios determined on the value of the largest Lyapunov exponent, Hurst exponent, the taxonomic measure of investment attractiveness (TMAI) and Markowitz portfolio. In the study we used the financial time series which were the shares price of selected companies listed on the Warsaw Stock Exchange and the indicators describing the economic and financial situation of companies. The data cover the period from 1.01.2005 to 30.09.2013.

2 The largest Lyapunov exponent

Lyapunov exponents are defined as limits [17]

$$\lambda_i(x_0) = \lim_{n\to\infty} \frac{1}{n} \ln\left|\mu_i(n, x_0)\right|, \qquad i = 1. .... m. \quad \text{for } m \geq 1, \tag{1}$$

where $\mu_i(n, x_0)$ are the eigenvalues of the Jacobi matrix of mapping f^n. f^n is an *n*-fold submission of function *f*, and *f* is the function that generates a dynamic system.

[1] University of Economics in Katowice, Department of Mathematics, monika.miskiewicz@ue.katowice.pl.
[2] University of Economics in Katowice, Department of Mathematics, katarzyna.zeug-zebro@ue.katowice.pl.

The Lyapunov exponents measure the rate of divergence or convergence of neighboring trajectories, i.e. the level of chaos in a dynamic system. The largest (maximal) Lyapunov exponent allows to specify the extent of a change (an increase or a decrease) in the distance between the current state x_N of the system and its nearest neighbor x_i in the evolution of the system, and also estimate the distance between the vectors x_{N+1} and x_{i+1}. Based on this distance the value of the forecasts $\hat{x}_{N+1}$ is determined [2, 19].

For real-time series, if you do not know a generator function *f*, the largest (maximal) Lyapunov exponent is estimated based on the relation [17]:

$$\Delta_n = \Delta_0 \cdot e^{n\lambda_{\max}}, \tag{2}$$

as the direction component of the regression equation [4, 5, 12]:

$$\ln \Delta_n = \ln \Delta_0 + \lambda_{\max} n, \tag{3}$$

where Δ_0 is the initial distance between two initially close (in the Euclidean distance sense) points of the reconstructed state space, Δ_n is the distance between these points after *n* iterations and $\lambda_{\max}$ is the largest (maximal) Lyapunov exponent.

Consider a one-dimensional time series, composed of *N* observation $(x_1, x_2, \ldots, x_N)$. Of all the vectors x_t^d of reconstructed state space we choose the vector closest to the vector x_N^d (in terms of Euclidean distance) and it is denoted by $x_{\min}^d$. Let $\Delta_{\min}$ denote the distance between x_N^d and $x_{\min}^d$, and Δ_1 - the distance between x_{N+1}^d and $x_{\min+1}^d$. Assuming that $\Delta_1/\Delta_{\min}$ is a small change in the evolution of the system, the distance between vectors x_{N+1}^d and $x_{\min+1}^d$ is given by [2]:

$$\Delta_1 \approx \Delta_{\min} \cdot e^{\lambda_{\max}}, \tag{4}$$

where $\lambda_{\max}$ is the largest (maximal) Lyapunov exponent.

3 Hurst exponent

The Hurst exponent [3] is another measure that allows for the classification of time series, i.e. to distinguish chaotic time series generated by deterministic dynamic systems from the stochastic time series. The exponent has a value in the range $\langle 0,1 \rangle$. If the time series is generated by a random walk (or a Brownian motion process) it has the value of $H = 0.5$. If $0 \leqslant H < 0.5$ the time series is antipersistent or ergodic. For a series, for which $0.5 < H \leqslant 1$, the series is persistent, i.e. reinforcing trend.

One of the methods of calculating the Hurst exponent is the method of the rescaled range R/S. For time series $\{x_1, x_2, \ldots, x_N\}$ it runs through the following steps[1, 10, 18]:

(1) transform the above time series into $m = N - 1$ logarithmic rates of return:

$$y_k = \log(x_{k+1}/x_k), \quad k = 1.\ 2.\ \ldots.\ N-1. \tag{5}$$

(2) Next, share a series (5) on *T* parts made up of *t* elements: $T = [m/t]$, where [ ] denotes the integer part of the argument. If the quotient m/t is not an integer then $tT < m$ and we use values y_k for $k = 1.\ 2.\ \ldots.\ tT$.

(3) In the next step, define the:

$$z_{ij} = y_{ij} - \bar{y}_j, \tag{6}$$

where: y_{ij} is the *j*-th value in the *i*-th interval and $\bar{y}_j = \frac{1}{t}\sum_{i=1}^{t} y_{ij}$.

(4) A sequence of partial sums z_{lj} for each *i*, is given by:

$$q_{ij} = \sum_{l=1}^{i} z_{lj}, \qquad i = 1.\ 2.\ \ldots.\ t, \quad j = 1.\ 2.\ \ldots.\ T. \tag{7}$$

(5) The range of the *i*-th interval is defined as:

$$R_j = \max(q_{ij}) - \min(q_{ij}). \tag{8}$$

(6) Calculate the rescaled range series (*R*/*S*):

$$\alpha_{jt} = R_j / S_j, \tag{9}$$

where: $S_j = \sqrt{\frac{1}{t}\sum_{i=1}^{t} z_{ij}^2}$.

(7) Next, calculate:

$$(R/S)_t = (1/T)\sum_{j=1}^{T} \alpha_{jt}. \tag{10}$$

(8) The above procedure is carried out for different lengths of time series *t*.

(9) Finally, the value of the Hurst exponent is the slope of the graph of the logarithms $(R/S)_t$ to the axis of logarithms *t*.

4 TMAI – Taxonomic Measure of Investment Attractiveness

One of the methods for the selection of companies that will be included in the optimal portfolio is to determine a taxonomic measure of investment attractiveness (TMAI) [15, 16]. This method allows for a comprehensive evaluation of the companies based on key financial and market indicators and present them in the form of synthetic measure.

Building a taxonomic measure consists of three stages [8, 13, 14]. Having data matrix, we normalize (standardize) the values, following the formula:

$$y_{ij} = \left(x_{ij} - \bar{x}_j\right)/S_j , \qquad i = 1,\ldots,n; \quad j = 1,\ldots,m, \tag{11}$$

where: $\bar{x}_j$ mean of feature *j*, S_j standard deviation for *j*.

Next, the module method is used, and in the normalized matrix of *m* variables, the highest value is taken, module y_{0j}. The Euclidean distance from the module is calculated, using the formula:

$$d_i = \left[\frac{1}{m}\sum_{j=1}^{m}\left(y_{ij} - y_{0j}\right)^2\right]^{1/2}, \quad i = 1,\ldots,n . \tag{12}$$

The shorter the distance of the given object from the module, the lower is the value d_i. The obtained variable is not normalized, which next is transformed into a stimulant using the formula:

$$TMAI_i = 1 - d_i/d_0 , \quad i = 1,\ldots,n , \tag{13}$$

where:

$TMAI_i$ – taxonomic development measure for object *I*,

d_i – distance of *I* object from module,

d_0 – standard to assure that variable $TMAI_i$ will take values ranging from 0 to 1, for example

$$d_0 = \bar{d} + 2S_d ,$$

where: $\bar{d}, S_d$ - mean and standard deviation d_i .

5 The optimal stock portfolio

Contribution of shares in the portfolio were established based on the following optimization problems:

Model I	Model II	Model III	Model IV
$\min S_p^2$.	$\max\left(\sum_{i=1}^{m} TMAI_i x_i\right)$	$\max\left(\sum_{i=1}^{m} TMAI_i x_i\right)$	$\max\left(\sum_{i=1}^{m} \lambda_{\max i} x_i\right)$
$R_p \geq R_0$	$R_p \geq R_0$	$R_p \geq R_0$	$R_p \geq R_0$
$\sum_{i=1}^{m} x_i = 1$	$\sum_{i=1}^{m} S_i x_i \leq S_0$	$\sum_{i=1}^{m} S_i x_i \leq S_0$	$\sum_{i=1}^{m} S_i x_i \leq S_0$
$x_i \geq 0, \quad i = 1,\ldots,m$	$\sum_{i=1}^{m} x_i = 1$	$\sum_{i=1}^{m} A_i x_i \geq A_0$	$\sum_{i=1}^{m} x_i = 1$
	$x_i \geq 0, \quad i = 1,\ldots,m$	$\sum_{i=1}^{m} x_i = 1$	$x_i \geq 0, \quad i = 1,\ldots,m$
		$x_i \geq 0, \quad i = 1,\ldots,m$	

Model V	Model VI	Model VII
$\max\left(\sum_{i=1}^{m} \lambda_{\max i} x_i\right)$	$\max\left(\sum_{i=1}^{m} H_i x_i\right)$	$\max\left(\sum_{i=1}^{m} H_i x_i\right)$
$R_p \geq R_0$	$R_p \geq R_0$	$R_p \geq R_0$
$\sum_{i=1}^{m} S_i x_i \leq S_0$	$\sum_{i=1}^{m} S_i x_i \leq S_0$	$\sum_{i=1}^{m} S_i x_i \leq S_0$
$\sum_{i=1}^{m} A_i x_i \geq A_0$	$\sum_{i=1}^{m} x_i = 1$	$\sum_{i=1}^{m} A_i x_i \geq A_0$
$\sum_{i=1}^{m} x_i = 1$	$x_i \geq 0, \quad i = 1,\ldots,m$	$\sum_{i=1}^{m} x_i = 1$
$x_i \geq 0, \quad i = 1,\ldots,m$		$x_i \geq 0, \quad i = 1,\ldots,m$

Table 1 Optimization problems

where: S_p - the risk of the portfolio m-shares:

$$S_p^2 = \sum_{i=1}^{m} x_i^2 S_i^2 + 2\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} x_i x_j S_i S_j \rho_{ij} , \tag{14}$$

S_i - the standard deviation for *i*-company,

ρ_{ij} - the correlation coefficient of *i*-share with *j*-share,

x_i - contribution of *i*-share in the portfolio,

R_p - the expected rate of return on the portfolio: $R_p = \sum_{i=1}^{m} x_i R_i$,

R_i - the expected rate of return for *i*-company,

R_0 - average rate of return for companies,

S_0 - mean standard deviation,

A_0 - mean asymmetry coefficient,

A_i - the asymmetry coefficient for *i*-company,

$TMAI_i$ - the taxonomic measure of investment attractiveness for *i*-company,

$\lambda_{\max i}$ - the largest (maximal) Lyapunov exponent for *i*-company,

H_i - the Hurst exponent for *i*-company.

6 The purpose and conduct of the study

In the study we used the financial time series which were the shares price of selected companies listed on the Warsaw Stock Exchange: Apator Group JSC (APT), Asseco Poland SA (ACP), Bank Handlowy in Warsaw JSC (BHW), Bank Zachodni WBK JSC (BZW), Dębica SA (DBC), ING Bank Śląski JSC (ING), KGHM Polska Miedź JSC (KGH), LPP JSC (LPP), mBank JSC (MBK), Mostostal Zabrze JSC (MSZ), Orange Poland JSC (OPL), Bank Polska Kasa Opieki (PEO), Powszechna Kasa Oszczędności Bank Polski JSC (PKO), Vistula Group JSC (VST) and Żywiec Group JSC (ZWC). The data cover the period from 1.01.2005 to 30.09.2013.

The largest Lyapunov exponent and the Hurst exponent for the analyzed companies was estimated on the basis of the algorithms described in points 1 and 2. For this were used financial time series which were set up with logarithms of daily returns of closing price indexes of selected companies in period from 1.01.2005 to 30.09.2013[3]. To determine the value of taxonomic measure TMAI used data contained in financial reports of companies [11] for the third quarter of 2013[4].

In the next stape of the studies constructed 56 investment portfolios based on solving optimization problems: model I – model VII. Criteria for selection of companies to a portfolio are shown in Table 1. Due to the value of the synthetic measure TMAI analyzed companies were divided into four groups:

[3] The data comes from stooq.com.
[4] The data come from the author's own calculations based on the financial reports of companies.

- the company's very good: $TMAI_i \geq \overline{TMAI} + S_{TMAI}$,
- the good company: $\overline{TMAI} + S_{TMAI} > TMAI_i \geq \overline{TMAI}$,
- the company's average: $\overline{TMAI} > TMAI_i \geq \overline{TMAI} - S_{TMAI}$,
- the weak company: $\overline{TMAI} - S_{TMAI} > TMAI_i$,

where: $\overline{TMAI}$ - average value of TMAI for the analyzed companies,
S_{TMAI} - standard deviation of TMAI for the analyzed companies.

Based on the above division of companies constructed portfolios 1-8. The value of the largest Lyapunov exponent allowed to choose the companies, whose time series are characterized by chaotic dynamics (portfolio 3) and those for which do not detected chaos (portfolio 4). Due to the value of the Hurst exponent. companies divided into two groups: companies for which $H_i > 0.6$ (portfolio 5) and those for which $H_i < 0.6$ (portfolio 6). In addition, based on the positive rate of return and the value of the coefficient of variation constructed two portfolios (portfolio 8 and 7).

Portfolio 1	**Portfolio 2**	**Portfolio 3**	**Portfolio 4**
All companies	Very good and good companies	Chaos	Absence of chaos
Portfolio 5	**Portfolio 6**	**Portfolio 7**	**Portfolio 8**
Hurst exponent > 0.6	Hurst exponent < 0.6	The coefficient of variation < 10 %	The positive rate of return

Table 2 Conditions for selection of companies for the optimal portfolios

Table 3 shows the value of the expected rate of return and risk of constructed portfolios.

Model		Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5	Portfolio 6	Portfolio 7	Portfolio 8
I	Rate of return	0.00249	0.00297	-0.53913	0.00278	0.00376	-0.08389	-1.24400	0.00287
	Risk	0.00003	0.00020	0.00009	0.00006	0.00016	0.00006	0.00011	0.00004
II	Rate of return	0.00299	0.00299	0.00293	0.00278	0.00372	0.00245	0.00338	0.00299
	Risk	0.00028	0.00029	0.00030	0.00018	0.00025	0.00019	0.00029	0.00028
III	Rate of return	0.00278	0.00302	0.00309	0.00278	0.00372	0.00234	0.00328	0.00287
	Risk	0.00019	0.00029	0.00026	0.00011	0.00020	0.00013	0.00029	0.00018
IV	Rate of return	0.00188	0.00335	-1.37010	0.00278	0.00524	0.00127	-1.24400	0.00287
	Risk	0.00011	0.00028	0.00013	0.00009	0.00021	0.00011	0.00012	0.00011
V	Rate of return	0.00188	0.00333	-1.37010	0.00278	0.00524	0.00127	-1.24400	0.00287
	Risk	0.00011	0.00027	0.00013	0.00009	0.00021	0.00011	0.00012	0.00011
VI	Rate of return	0.00248	0.00363	0.00277	0.00503	0.00491	-1.37086	0.00523	0.00287
	Risk	0.00019	0.00026	0.00019	0.00019	0.00025	0.00019	0.00020	0.00018
VII	Rate of return	0.00338	0.00363	0.00273	0.00483	0.00491	-0.64874	0.00524	0.00347
	Risk	0.00018	0.00026	0.00019	0.00014	0.00025	0.00016	0.00018	0.00018

Table 3 The expected rate of return and risk of the portfolio

On the basis of the data in Table 3 we can conclude that for the optimization models I-V the portfolio 5 which consists of companies for which the Hurst exponent is larger than 0.6 is characterized by the highest expected rate of return. Whereas, for the models VI-VII the highest expected rate of return was obtained for the portfolio 7 (the companies with the coefficient of variation of less than 10%). For optimization problems model II-V and VII, the lowest level of risk was noted for portfolio 4 which included the companies with no evidence of deterministic chaos in the stock prices time series. The highest level of risk was characterized usually for portfolio 2, composed of companies with very good and good TMAI.

Model	Portfolio 1	Portfolio 2	Portfolio 3	Portfolio 4	Portfolio 5	Portfolio 6	Portfolio 7	Portfolio 8
I	0.0954	0.1584	0.1604	0.1399	0.1589	-0.0091	0.1605	0.0865
II	0.1513	0.1701	0.1526	0.1557	0.1572	0.1579	0.1594	0.1510
III	0.1460	0.1679	0.1568	0.1528	0.1582	0.1323	0.1703	0.1450
IV	0.0809	0.1569	0.3020	0.1351	0.1613	-0.0337	0.2760	0.0688
V	0.0809	0.1568	0.3020	0.1351	0.1613	-0.0337	0.2760	0.0688
VI	0.2180	0.1600	0.1715	0.1750	0.1737	0.1042	0.3453	0.2001
VII	0.0955	0.1600	0.1727	0.1355	0.1732	0.0748	0.1998	0.0800

Table 4 The annual rate of return for constructed investment portfolios

Table 4 shows the obtained rates of return for the designated portfolios in the period from 30.09.2013 to 30.09.2014.

Analyzing the obtained rates of return for the designated portfolios (Table 4), it can be seen that the largest profit in the period from 30.09.2013 to 30.09.2014 would give the investment in portfolio 7 in model VI consisted of companies with a coefficient of variation of less than 10%, and then in portfolio 3 in model IV and model V. The lowest rate of return for all optimization models except model II obtained for the portfolio 6 created from the companies for which the Hurst exponent is less than 0.6.

7 Conclusions

In the paper studied attempts to construct the optimal portfolio of shares based on the value of taxonomic measure of investment attractiveness, the largest Lyapunov exponent and the Hurst exponent. The study has revealed that portfolios composed of companies whose time series do not generate chaotic behavior and the companies for whose the Hurst exponent $H_i > 0.6$, they received some of the highest expected rate of returns. In addition, for the portfolio constructed from companies whose time series are not chaotic, reported the lowest level of risk.

References

[1] Chun, S.H., Kim, K.J., Kim, S.H.: *Chaotic analysis of predictability versus knowledge discovery techniques: case study of the Polish stock market*. Expert Systems, Vol. 19, No. 5, 2002, 264 – 272.

[2] Guégan, D., Leroux, J.: *Forecasting chaotic systems: The role of local Lyapunov exponents*. Chaos, Solitons & Fractals, vol. 41, 2009, 2401 – 2404.

[3] Hurst, H. E.:. *Long term storage capacity of reservoirs.* Trans. Am. Soc. Eng. **116**, 1951, 770–799.

[4] Kantz, H.: *A robust method to estimate the maximal Lyapunov exponent of a time series*. Physical Letters A, vol. 185(1), 1994, 77 – 87.

[5] Kantz, H., Schreiber, T.: *Nonlinear time series analysis*. Cambridge University Press, Cambridge, 1997.

[6] Markowitz, H.: *Portfolio Selection*. Journal of Finance, 1952, 77-91.

[7] Markowitz, H.: *Portfolio Selection: Efficient diversification of investments*. 94, Cowles Foundation, New Haven, CT, 1952.

[8] Mastalerz-Kodzis, A., Pośpiech, E.: *Fundamental and Behavioral Methods in Investment Decision Making*. in: Financial Management of Firms and Financial Institutions, Wydawnictwo Technicznego Uniwersytetu w Ostrawie, Ostrawa, 2011, 250-257.

[9] Miśkiewicz-Nawrocka, M., Zeug-Żebro, K.: *The application of Lyapunov exponents to building optimal portfolios*. Studia Ekonomiczne 221, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice, 2015, 61-72.

[10] Miśkiewicz-Nawrocka, M.: *Zastosowanie wykładników Lapunowa do analizy ekonomicznych szeregów czasowych*. Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice, 2012.

[11] Nawrocki, T., Jabłoński, B.: *Inwestowanie na Rynku Akcji. Jak Ocenić Potencjał Rozwojowy Firm Notowanych na GPW w Warszawie*. Wydawnictwo CeDeWu, 2011.

[12] Rosenstein, M. T., Collins, J. J., De Luca, C. J.: *A practical method for calculating largest Lyapunov exponents from small data sets*. Physica D, vol. 65, 1993, 117 – 134.

[13] Tarczyński, W.: *Fundamentalny portfel papierów wartościowych*. Polskie Wydawnictwo Ekonomiczne, Warszawa, 2002.

[14] Tarczyński, W.: *Ocena efektywności metod analizy portfelowej na Giełdzie Papierów Wartościowych w Warszawie za lata 2001-2013*. Zeszyty Naukowe Uniwersytetu Szczecińskiego nr 761, Finanse, rynki finansowe, ubezpieczenia nr 60, Szczecin, 2013, 537-550.

[15] Tarczyński, W.: *A taxonomic measure of the attractiveness of investments in securities*. Przegląd Statystyczny No 3/94, 1994, 275–300.

[16] Tarczyński, W.: *The fundamental attitude to building a stock portfolio*. Argumenta Oeconomica No 1(7), Wrocław, 1999, 153–168.

[17] Zawadzki, H.: *Chaotyczne systemy dynamiczne*. Akademia Ekonomiczna, Katowice, 1996.

[18] Zeug-Żebro, K., Dębicka, J., Kuśmierczyk, P., Łyko, J.: *Wybrane modele matematyczne ekonomii. Decyzje i wybory*. Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław, 2013.

[19] Zhang, J., Lam, K.C., Yan, W.J., Gao, H., Li, Y.: *Time series prediction using Lyapunov exponents in embedding phase space*. Computers and Electrical Engineering 30, 2004, 1-15.

Annealing Based Integer Optimization Heuristic with Lévy Flights

Matej Mojzes [1], Jaromir Kukal [2], Josef Bostik [3], Martin Klimt[4]

Abstract. Novel population based integer optimization heuristic yields from the theory of Mean Field Annealing. Population center and covariance matrix are estimated for given annealing temperature and then used as directional correction of Lévy Flight mutation. The heuristic is of competitive nature like Competitive Differential Evolution. Here, nine Lévy Flight mutations compete and are selected according to their success. Resulting heuristic has four parameters: population size, regularization factor, annealing temperature and Lévy Flight temperature. Depending on the task complexity there is a relationship between searching efficiency and regularization, annealing and heavy-tailed flights. This heuristic is suitable integer optimization tasks with many local extremes. One of them is the clustering problem which can be converted to optimum partition with any penalization function. The clustering can help to classify various states of stock market system according to time series analysis and is used for the demonstration of novel heuristic.

Keywords: Heuristic, Mean Field Annealing, Lévy Flights, Integer Optimization, Clustering.

JEL classification: C44
AMS classification: 90C15

1 Introduction

Stock markets produce very interesting time series data. It may be useful to compare different time periods together and search for patterns in such data. In our paper we are analyzing weekly periods and logarithmic differences of individual daily returns during the week.

We have converted this problem into cluster analysis of weekly stock market data. Such task can be solved for example using well-known k-means clustering algorithm, but this is also known for the fact that is suffers from tendency to end in local minima.

Therefore, more sophisticated heuristics should be employed to ensure better cluster analysis yields in terms of higher cluster quality, time complexity and reliability. In the next sections we are developing novel algorithm for this exact purpose.

2 Meta-Heuristics

There is a variety of optimization methods and meta-heuristics including integer and binary ones. Let $n \in \mathbb{N}$ be task dimension, $\mathbf{x}$, $\mathbf{a}$, $\mathbf{b} \in \mathbb{Z}^n$ be independent variable, its lower and upper bounds satisfying $\mathbf{a} < \mathbf{b}$. The domain of integer optimization is

$$\mathcal{D} = \{\ \mathbf{x} \in \mathbb{Z}^n |\ \mathbf{a} \leq \mathbf{x} \leq \mathbf{b}\} \tag{1}$$

as a frame of objective function

$$\mathrm{f} \colon \mathcal{D} \to \mathbb{R} \tag{2}$$

[1]FNSPE CTU, Dept. of Software Engineering, Trojanova 13, Prague, Czech Republic, mojzemat@fjfi.cvut.cz
[2]FNSPE CTU, Dept. of Software Engineering, Trojanova 13, Prague, Czech Republic, jaromir.kukal@fjfi.cvut.cz
[3]FNSPE CTU, Dept. of Software Engineering, Trojanova 13, Prague, Czech Republic, josef.bostik@fjfi.cvut.cz
[4]FNSPE CTU, Dept. of Software Engineering, Trojanova 13, Prague, Czech Republic, matrli@seznam.cz

minimization. This task can be enriched by threshold

$$f^* \geq \min_{\mathbf{x}\in\mathcal{D}} \mathrm{f}(\mathbf{x}). \tag{3}$$

Then the goal set

$$\mathcal{G} = \{\ \mathbf{x} \in \mathcal{D} \mid \mathrm{f}(\mathbf{x}) \leq f^*\} \tag{4}$$

also contains global optimum of objective function. Finding any $\mathbf{x}_{\mathrm{opt}} \in \mathcal{G}$ is called here as sub-optimization task to be solved. It is useful to define range vector as $\mathbf{d} = \mathbf{b} - \mathbf{a} + \mathbf{1}$ and then denote

$$M^* = \mathrm{card}(\mathcal{G}) \geq 1, \tag{5}$$

$$M = \mathrm{card}(\mathcal{D}) = \prod_{k=1}^{n} d_k \geq 2^n \tag{6}$$

as number of goal states and total number of states. Meta-heuristic approach to sub-optimization task plays main role in the case of NP-hard problems [5] as formula satisfiability [16], set covering [7], knapsack [15], travelling salesman [3], and clustering [2] problems.

Traditional integer optimization meta-heuristics include many variants of Genetic Optimization [6] and discrete variants of Steepest Descent [1], Random Descent [13], Simulated Annealing [9], Fast Simulated Annealing [17], Random Descent with Lévy Flights [8], Particle Swarm Optimization [4], Cuckoo Search [19], [10], Modified Cuckoo Search [18], and many others.

3 Mean Field Integer Flight

The novel integer optimization heuristic is motivated by Evolutionary Search (ES), Mean Field Annealing (MFA), Parzen Estimate (PE) [14] of Probability Density Function (PDF), Lévy Flight Mutation (LFM), and competitive approach.

The *Mean Field Integer Flight* (MFIF) is population based heuristic but the population of $N \in \mathbb{N}$ vectors is unsorted. Denoting $f_k = \mathrm{f}(\mathbf{x}_k)$ we form the population as N-tuple of pairs

$$\mathbf{P} = ((\mathbf{x}_1, f_1), \ldots, (\mathbf{x}_N, f_N)).$$

Traditional MFA is based on the partition function

$$Z = \sum_{k=1}^{N} \exp(-f_k/T_{\mathrm{MFA}}) \tag{7}$$

over all N states for annealing temperature $T_{\mathrm{MFA}} > 0$. Resulting steady state probabilities of MFA are directly

$$p_k = \exp(-f_k/T_{\mathrm{MFA}})/Z \tag{8}$$

for $k = 1, \ldots, N$. The MFA estimate of global minimum is the mean value

$$\mathbf{e} = \sum_{k=1}^{N} p_k \mathbf{x}_k. \tag{9}$$

From the numerical point of view, the probabilities p_k are invariant to any constant shift of function values and therefore we can correct them to

$$f_{k,\mathrm{corr}} = f_k - \min_{j=1,\ldots,N} \{f_j\} \geq 0 \tag{10}$$

to avoid overflow in partition function calculation. After this improvement we obtain $Z_{\mathrm{corr}} \in (1, N]$ and finally $p_k = p_{k,\mathrm{corr}}$. Novel MFIF heuristics also employs Parzen Estimate (PE) of the probability density function (PDF) for given population $\mathbf{P}$ of fixed size N, fixed width $\sigma > 0$, and Gaussian kernel in the form

$$\mathrm{q}(\mathbf{x}) = \frac{1}{N} \cdot \sum_{k=1}^{N} \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}_k\|^2}{2\sigma^2}\right). \tag{11}$$

Steady state probabilities of MFA are directly used for *Weighted Density Estimate* (WDE) as

$$\mathrm{g}(\mathbf{x}) = \sum_{k=1}^{N} \frac{p_k}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}_k\|^2}{2\sigma^2}\right). \tag{12}$$

Very important particular cases are:

- When $T_{\mathrm{MFA}} \to +\infty$ then $\mathrm{g} \to \mathrm{q}$ which is the Parzen Estimate.
- When $\mathbf{x}_{\min}$ is unique minimum from given population $\mathbf{P}$ and $T_{\mathrm{MFA}} \to 0+$ then $\mathrm{g} \to \mathrm{N}(\mathbf{x}_{\min}, \sigma^2\mathbf{I})$ which is Gaussian distribution centered in the best population point.

Using characteristic function

$$\psi(\mathbf{t}) = \mathrm{E}\,\exp(\jmath\mathbf{x}^{\mathrm{T}}\mathbf{t}) \tag{13}$$

we explicitly obtained

$$\psi(\mathbf{t}) = \sum_{k=1}^{N} p_k \exp\left(\jmath\mathbf{x}_k^{\mathrm{T}}\mathbf{t} - \mathbf{t}^{\mathrm{T}}(\sigma^2\mathbf{I})\mathbf{t}/2\right) \tag{14}$$

which is useful for direct calculation of moment characteristics. The first moment is the mean value of sampled population which is well known from the MFA theory as

$$\mathbf{e} = \mathrm{E}\mathbf{x} = \sum_{k=1}^{N} p_k\mathbf{x}_k \tag{15}$$

in formal agreement with (9) meanwhile the covariance matrix is composed from two terms as

$$\mathbf{C} = \mathrm{E}(\mathbf{x}-\mathbf{e})(\mathbf{x}-\mathbf{e}^{\mathrm{T}}) = \mathbf{C}_{\mathrm{raw}} + \sigma^2\mathbf{I} \tag{16}$$

where

$$\mathbf{C}_{\mathrm{raw}} = \sum_{k=1}^{N} p_k(\mathbf{x}_k-\mathbf{e})(\mathbf{x}_k-\mathbf{e})^{\mathrm{T}} \tag{17}$$

is obvious covariance matrix of sampled population and the second term of (16) is resulting effect of Parzen estimation as a kind of statistical regularization.

The main idea behind MFIF method is in directional mutation based on $\mathbf{C}$ of $\mathbf{P}$ but driven by Lévy distribution. Novel *Mean Field Integer Flight Mutation* (MFIFM) consists of several steps:

- Calculate $\mathbf{C}_{\mathrm{raw}}$ for given population $\mathbf{P}$ and temperature $T_{\mathrm{MFA}} > 0$,
- Generate random n-dimensional Gaussian vector $\mathbf{y} \sim \mathrm{N}(\mathbf{0}, \mathbf{I})$,
- Calculate directional vector $\mathbf{z} = (\mathbf{C}_{\mathrm{raw}} + \sigma^2\mathbf{I})^{1/2} \cdot (\mathbf{y}/\|\mathbf{y}\|_2)$,
- Generate $d \sim \mathrm{Levy}(\beta)$ using Lévy distribution with $0 < \beta < 2$,
- Real unlimited mutation of $\mathbf{x} \in \mathcal{D}$ using mutation temperature $T_{\mathrm{mut}} > 0$ produces $\mathbf{r} = \mathbf{x} + T_{\mathrm{mut}} \cdot d \cdot \mathbf{z}$,
- Using component-wise rounding and perturbation via mirroring we calculate $\mathbf{x}_{\mathrm{new}} = \mathrm{P}(\left[\mathbf{r}\right], \mathbf{a}, \mathbf{b})$.

This novel type of mutation operator yields from the properties of given population $\mathbf{P}$ applying MFA theory to obtain the direction of mutation as vector $\mathbf{z}$. Final mutation of $\mathbf{x}$ is realized as perturbed directional integer Lévy flight with dimensionless mutation temperature T_{mut}.

The MFLFM operator has four tuning parameters:

- $T_{\mathrm{MFA}} > 0$ for Mean Field Annealing,
- $T_{\mathrm{mut}} > 0$ for Lévy flights,
- $\beta \in (0, 2)$ for Lévy distribution,
- $\delta > 0$ for Parzen estimate.

4 Basic Frame of MFIF

First we set $N \in \mathbb{N}$, $H \in \mathbb{N}, H \geq 2, n_0 \in \mathbb{N}, \delta > 0$, f^*, $N_{\max}$ as population size, mutation portfolio size, initial counter value, threshold, final value and maximal number of evaluations. Then we introduce mutation family

$$\mathcal{F} = \{\mathrm{MFIFM_1}(\mathbf{x}), \ldots, \mathrm{MFIFM_H}(\mathbf{x})\}\,. \tag{18}$$

The algorithm of *Mean Field Integer Flights* (MFIF) is described in detail in Algorithm 1.

Algorithm 1 MFIF

Set counters $\mathbf{n} = n_0\mathbf{1} \in \mathbb{N}^H$, and mutation portfolio.
Init population $\mathcal{P}$ of size N by uniform sampling from $\mathcal{D}$.
while $f_{\text{best}} > f^*$ and $neval < N_{\max}$: **do**
end while
Using systematic selection strategy find $\mathbf{x}_k \in \mathcal{P}$
for $k = 1, \ldots, N$ **do**
    Generate randomly index j according to mutation probabilities $p_j = n_j/\ ||\mathbf{n}||_1$
    Perform $\mathbf{x}_{\text{new}} = \mathrm{MFIFM}_j(\mathbf{x})$
    Evaluate $f_{\text{new}} = \mathrm{f}(\mathbf{x}_{\text{new}})$.
    if $f_{\text{new}} < f_{(k)}$ **then**
        Update $n_j = n_j + 1$, $\mathbf{x}_k = \mathbf{x}_{\text{new}}$, $f_{(k)} = f_{\text{new}}$
        if $\min_{i=1,\ldots,H} p_i = \frac{\min n_i}{||\mathbf{n}||_1} < \frac{\delta}{H}$ **then**
            Reset counters as $\mathbf{n} = n_0\mathbf{1}$
        end if
    end if
end for

Individual MFIFM_j is described by parameters $(\mathrm{T_{MFA}}, \mathrm{T_{mut}}, \beta, \sigma)_j$ for $i = 1, \ldots, H$. General suggestion is to use fixed $\sigma \in (0, 1)$ for all mutations in the portfolio due to integer nature of searching domain $\mathcal{D}$. The parameter of Lévy distribution can be also set to fixed value $\beta = 1$ for the first experiments. Therefore, the competition of mutations is only about adaptive changing of temperature pairs $(T_{\text{MFA}}, T_{\text{mut}})_j$ for $i = 1, \ldots, H$.

5 Clustering as optimization problem

Traditional k-means algorithm of cluster analysis begins with random partition of data patterns into given number of clusters N_c. Each step of k-means consists of cluster center update and subsequent revision of the pattern partition. The partition quality is expressed as residual sum of squares (SSQ) or standard deviation (s_e) of model error, respectively. Unfortunately, k-means heuristic terminates when any of s_e local minima is reached.

Finally, for any initial partition $\mathbf{x} \in \{1, \ldots, N_c\}^{m^*}$ we obtain adequate final value of $\mathrm{f}(\mathbf{x}) = s_e(\mathbf{x})$ and k-means is not very efficient heuristic beginning with $\mathbf{x_0} \sim \mathrm{U}(\mathcal{D})$ and terminating in any local minimum.

To overcome this effect, we used MFIF heuristics for minimization of $s_e(\mathbf{x})$ on $\mathcal{D}$ and Integer Lévy Flight (ILF) [12] heuristics as referential one.

6 Application to Stock Market State

Let $\{a_k\}_{k=1}^m$, $a_k > 0$ be stock price series. Raw data pre-processing generates return history $\{r_k\}_{k=1}^{m-1}$ where $r_k = \ln(a_{k+1}/a_k)$. The stock data were split into non-overlapping segments of constant length $w \in \mathbb{N}$. Therefore, we obtained statistical samples $\boldsymbol{\xi}_i \in \mathbb{R}^w$ where

$$\boldsymbol{\xi}_i = (r_{wi-w+1}, \ldots, r_{wi}) \tag{19}$$

and $1 \leq i \leq \lfloor (m-1)/w \rfloor = m^*$.

Weekly segments of length $w = 5$ were investigated in this study. Altogether we have analyzed 150 weeks of the *S&P 500 Index* from 5/20/2013 to 3/28/2016.

7 Results

In the first step, we were searching for the best possible cluster quality s_e for $N_c \in \{2, \ldots, 10\}$ and then run each heuristic 100 times with respective $s_e = f^*$ value.

To assess quality of the novel MFIF heuristic, we have used the following parameter setup: $H = 9$, $\beta = 1$, $N = 20$, $\sigma = 10^{-3}$, $T_{\mathrm{MFA}} = 0.1$, $T_{\mathrm{mut}} \in \{0^+, 1/10, 1/5, 1/2, 1, 2, 5, 10, +\infty\}$, $\delta = 1/5$, $n_0 = 1$. Referential ILF heuristic parameters were empirically optimized as well.

Table 1 contains basic heuristic performance measures [11]:

- REL – *reliability*, ratio between the number of successful runs (when f^* was reached) and total number of runs (100),
- MNE as *mean number of objective function evaluations* until optimal solution was found,
- SNE as *standard deviation of the number of evaluations.*

		MFIF			ILF		
N_c	s_e	ENE	REL	SNE	ENE	REL	SNE
2	0.007911	1262	1.00	663	**962**	**1.00**	**950**
3	0.007400	681	1.00	578	**338**	**1.00**	**243**
4	0.007030	1659	1.00	762	**800**	**1.00**	**1076**
5	0.006741	**2038**	**1.00**	**905**	3091	1.00	7609
6	0.006470	**2981**	**1.00**	**1988**	5400	1.00	8138
7	0.006275	**5456**	**1.00**	**3345**	17763	0.85	23255
8	0.006100	**2160**	**1.00**	**1456**	3293	1.00	6941
9	0.005932	**2900**	**1.00**	**1597**	4820	1.00	11415
10	0.005743	**15955**	**1.00**	**9468**	18605	0.74	21334

Table 1 Target cluster quality s_e and heuristic performance measures for MFIF and ILF based on number of clusters N_c

Bold values in Table 1 indicate better performing heuristic for given N_c. Data suggests that MFIF performs better than ILF for $N_c > 4$. However, it is very interesting to see that the SNE/MNE ratio is considerably lower in the case of MFIF compared to ILF and thus MFIF could be considered much more reliable in its search for the optimal stock market data cluster. Still, ILF can be quicker (in terms of MNE) in some cases.

8 Conclusions

Sophisticated heuristics may be very useful when dealing with difficult multimodal optimization problems such as stock market data cluster analysis.

We have shown that the proposed MFIF heuristic can perform more reliably and based on personal preference of the weekly stock market returns also more quickly than the referential ILF method, which is based on reputable Lévy flights as well.

Last, but not least, we have provided suggested parameter setup and based on our experiments it may be also generally useful, but it should be tuned with regard to the given objective function.

Acknowledgements

This paper is supported by grant SGS14/208/OHK4/3T/14 CTU in Prague.

References

[1] Aarts, E. and Lenstra, J. K.: *Local Search in Combinatorial Optimization*, Princeton University Press, Princeton (2003)

[2] Bezdek, J. C. and James, C.: *Pattern Recognition with Fuzzy Objective Function Algorithms*, Kluwer Academic Publisher, Norwell (1981)

[3] Cerny, V.: Thermodynamical Approach to the Travelling Salesman Problem: An Efficient Simulation Algorithm, *Journal of Optimization Theory and Applications* **45**, 1 (1985), 41–51

[4] Eberhart, R. C., Shi, Y. and Kennedy, J.: *Swarm Intelligence*, Morgan Kaufmann, San Francisco (2001)

[5] Garey, M. R. and Johnson D. S.: *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman, New York (1979)

[6] Goldberg, D. E.: *Genetics Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley (1989)

[7] Chvatal, V.: A Greedy Heuristic for the Set-Covering Problem, *Mathematics of Operations Research* **4**, 3 (1979), 233–235

[8] Klimt, M., Kukal, J. and Mojzes, M.: Lévy Flights in Binary Optimization, *Archives of Control Sciences*, **23**, 4 (2013), 447–454

[9] Kirkpatrick, S., Gelatt, C. D. Jr. and Vecchi, M. P.: Optimization by Simulated Annealing, *Science*, **220**, 4598 (1983), 671–680

[10] Kukal, J., Mojzes, M., Tran, Q. V. and Bostik, J.: Integer Cuckoo Search, In: *Proceedings of Mendel 2012 Soft Computing Conference*, Brno (2012), 298–303

[11] Mojzes, M., Kukal, J., Tran, V. Q., Jablonský, J.: Performance Comparison of Heuristic Algorithms via Multi-criteria Decision Analysis. In: *Proceedings of Mendel 2011 Soft Computing Conference*, Brno (2011), 244–251

[12] Mojzes, M., Klimt, M., Kukal, J., Bukovsky, I. and Pitel, J.: Feature Selection via Competitive Lévy Flights To be published in: *Proceedings of 2016 International Joint Conference on Neural Networks*, (2016)

[13] Motwani, R. and Raghavan, P.: *Randomized Algorithms*, Cambridge University Press, Cambridge (1995)

[14] Parzen, E.: On Estimation of a Probability Density Function and Mode, *The Annals of Mathematical Statistics*, **33**, 1065 (1962)

[15] Plateau, G. and Elkihel, M.: A Hybrid Algorithm for the 0-1 Knapsack Problem, *Methods of Oper. Res.*, **49** (1985), 277–293

[16] Selman, B., Mitchell D. and Levesque, H.: Generating Hard Satisfiability Problems, *Artificial Intelligence*, **81** (1996), 17–29

[17] Szu, H. and Hartley, R.: Fast Simulated Annealing, *Physics Letters A*, **122**, 3-4 (1987), 157–162

[18] Walton, S., Hassan, O., Morgan, K. and Brown, M. R.: Modified Cuckoo Search: A New Gradient Free Optimisation Algorithm, *Chaos, Solitons and Fractals*, **44** (2011), 710–718

[19] Yang, X.-S. and Deb, S.: Engineering Optimisation by Cuckoo Search, *International Journal of Mathematical Modelling and Numerical Optimisation*, **1**, 4 (2010), 330–343

Possible and universal robustness of Monge fuzzy matrices

M. Molnárová[1]

Abstract. Robustness of Monge matrices over max-min algebra is studied. The max-min algebra (fuzzy algebra) is an extremal algebra with operations maximum and minimum. An interval matrix A over fuzzy algebra is defined by a lower bound matrix and an upper bound matrix. A is possibly robust, if there is at least one robust matrix in A. If all matrices from A are robust, A is universally robust. Equivalent conditions for robustness of Monge matrices are presented. Equivalent conditions for possible robustness and universal robustness of interval Monge matrices in binary case are presented. Necessary condition for an interval Monge matrix to be possibly robust in general case of max-min algebra was proved.

Keywords: (max, min) algebra, robustness, Monge matrix

JEL classification: C02
AMS classification: 08A72, 90B35, 90C47

1 Introduction

The so-called extremal algebras (at least one operation creates no new elements) are predestined to model applications of discrete dynamic systems. In most frequented cases are operations of addition and multiplication replaced by operations of maximum and addition (max-plus algebra) or maximum and minimum (max-min algebra). The max-min algebra called also fuzzy algebra is used in diverse areas (graph theory, knowledge engineering, managing traffic or production) where the considered systems or devices can be represented by a matrix. Properties of fuzzy matrices were described in [3], [10]. The Monge matrices and their applications were studied in [1], [4]. Robustness of Monge fuzzy matrices was presented in [6], [7], [8]. Robustness of interval fuzzy matrices was studied in [9].

2 Background of the problem

The fuzzy algebra $\mathcal{B}$ is a triple $(B, \oplus, \otimes)$, where $(B, \leq)$ is a bounded linearly ordered set with binary operations *maximum* and *minimum*, denoted by $\oplus$, $\otimes$. The least element in B will be denoted by O, the greatest one by I. By $\mathbb{N}$ we denote the set of all natural numbers. The greatest common divisor of a set $S \subseteq \mathbb{N}$ is denoted by $\gcd S$, the least common multiple of the set S is denoted by $\operatorname{lcm} S$. For a given natural $n \in \mathbb{N}$, we use the notation N for the set of all smaller or equal positive natural numbers, i.e., $N = \{1, 2, \ldots, n\}$.

For any m, $n \in \mathbb{N}$, $B(m, n)$ denotes the set of all matrices of type $m \times n$ and $B(n)$ the set of all n-dimensional column vectors over $\mathcal{B}$. The matrix operations over $\mathcal{B}$ are defined formally in the same manner (with respect to $\oplus$, $\otimes$) as matrix operations over any field. The rth power of a matrix $A \in B(n, n)$ is denoted by A^r, with elements a^r_{ij}. For A, $C \in B(n, n)$ we write $A \leq C$ if $a_{ij} \leq c_{ij}$ holds for all i, $j \in N$.

A *digraph* is a pair $G = (V, E)$, where V, the so-called vertex set, is a finite set, and E, the so-called edge set, is a subset of $V \times V$. A digraph $G' = (V', E')$ is a subdigraph of the digraph G (for brevity $G' \subseteq G$), if $V' \subseteq V$ and $E' \subseteq E$. A path in the digraph $G = (V, E)$ is a sequence of vertices $p = (i_1, \ldots, i_{k+1})$ such that $(i_j, i_{j+1}) \in E$ for $j = 1, \ldots, k$. The number k is the length of the path p and is denoted by $\ell(p)$. If $i_1 = i_{k+1}$, then p is called a cycle. For a given matrix $A \in B(n, n)$ the symbol $G(A) = (N, E)$ stands for the complete, edge–weighted digraph associated

[1]Technical University of Košice, Department of Mathematics and Theoretical Informatics, B. Němcovej 32, 04200 Košice, Slovakia, e-mail: Monika.Molnarova@tuke.sk

with A, i.e. the vertex set of $G(A)$ is N, and the capacity of any edge $(i,j) \in E$ is a_{ij}. In addition, for given $h \in B$, the *threshold digraph* $G(A,h)$ is the digraph $G = (N, E')$ with the vertex set N and the edge set $E' = \{(i,j);\ i,j \in N,\ a_{ij} \geq h\}$.

The following lemma describes the relation between matrices and corresponding threshold digraphs.

Lemma 1. [9] Let A, $C \in B(n,n)$. Let h, h_1, $h_2 \in B$.

(i) If $A \leq C$ then $G(A,h) \subseteq G(C,h)$,

(ii) if $h_1 < h_2$ then $G(A,h_2) \subseteq G(A,h_1)$.

By a *strongly connected component* of a digraph $G(A,h) = (N,E)$ we mean a subdigraph $\mathcal{K} = (N_{\mathcal{K}}, E_{\mathcal{K}})$ generated by a non-empty subset $N_{\mathcal{K}} \subseteq N$ such that any two distinct vertices $i,j \in N_{\mathcal{K}}$ are contained in a common cycle, $E_{\mathcal{K}} = E \cap (N_{\mathcal{K}} \times N_{\mathcal{K}})$ and $N_{\mathcal{K}}$ is the maximal subset with this property. A strongly connected component $\mathcal{K}$ of a digraph is called non-trivial, if there is a cycle of positive length in $\mathcal{K}$. For any non-trivial strongly connected component $\mathcal{K}$ is the *period* of $\mathcal{K}$ defined as $\operatorname{per}\mathcal{K} = \gcd\{\,\ell(c);\ c \text{ is a cycle in } \mathcal{K},\ \ell(c) > 0\,\}$. If $\mathcal{K}$ is trivial, then $\operatorname{per}\mathcal{K} = 1$. By $\mathrm{SCC}^{\star}(G)$ we denote the set of all non-trivial strongly connected components of G.

Let $A \in B(n,n)$ and $x \in B(n)$. The sequence $O(A,x) = \{x^{(0)}, x^{(1)}, x^{(2)}, \ldots, x^{(n)}, \ldots\}$ is the orbit of $x = x^{(0)}$ generated by A, where $x^{(r)} = A^r \otimes x^{(0)}$ for each $r \in \mathbb{N}$.
For a given matrix $A \in B(n,n)$, the number $\lambda \in B$ and the n–tuple $x \in B(n)$ are the so-called *eigenvalue* of A and *eigenvector* of A, respectively, if they are the solution of the *eigenproblem* for matrix A, i.e. they satisfy the equation $A \otimes x = \lambda \otimes x$. The corresponding *eigenspace* $V(A,\lambda)$ is defined as the set of all eigenvectors of A with associated eigenvalue λ, i.e. $V(A,\lambda) = \{x \in B(n);\ A \otimes x = \lambda \otimes x\}$.
Let $\lambda \in B$. A matrix $A \in B(n,n)$ is *ultimately λ–periodic* if there are natural numbers p and R such that the following holds: $A^{k+p} = \lambda \otimes A^k$ for all $k \geq R$. The smallest natural number p with above property is called the period of A, denoted by $\operatorname{per}(A,\lambda)$. In case $\lambda = I$ we denote $\operatorname{per}(A,I)$ by abbreviation $\operatorname{per}A$.

Definition 1. Let $A = (a_{ij}) \in B(n,n)$, $\lambda \in B$. Let $T(A,\lambda) = \{x \in B(n);\ O(A,x) \cap V(A,\lambda) \neq \emptyset\}$. A is called *λ–robust* if $T(A,\lambda) = B(n)$. A λ–robust matrix with $\lambda = I$ is called *a robust matrix.*

In our considerations we will use the following result (adapted for $\lambda = I$) proved in [10] to study robustness of a matrix.

Lemma 2. [10] Let $A = (a_{ij}) \in B(n,n)$. Then A is robust if and only if $\operatorname{per}A = 1$.

3 Robustness of Monge matrices

Definition 2. We say, that a matrix $A = (a_{ij}) \in B(m,n)$ is a convex Monge matrix (concave Monge matrix) if and only if

$$a_{ij} \otimes a_{kl} \leq a_{il} \otimes a_{kj} \qquad \text{for all} \quad i < k,\ j < l$$

$$\big(\, a_{ij} \otimes a_{kl} \geq a_{il} \otimes a_{kj} \qquad \text{for all} \quad i < k,\ j < l\,\big).$$

In this paper, we assume that the considered matrices are convex.

First, we have considered binary case of Monge fuzzy matrices over the set $B = \{0,1\}$. Restricting ourselves to matrices which satisfy the condition $A \geq I_{ad}$ (only the weight of arcs $(1,n)$, $(2,n-1)$, $\ldots$, $(n,1)$ is equal to 1 in I_{ad}) the following necessary and sufficient condition for a binary fuzzy matrix was proved.

Theorem 1. [6] *Let $A = (a_{ij}) \in B(n,n)$ be a binary Monge matrix with $A \geq I_{ad}$. Then A is robust if and only if $G(A,1)$ is strongly connected and contains a loop.*

Afterwards, equivalent conditions for robustness of Monge matrices in general case of max-min algebra was proved. The existence of cycles of odd length, especially loops, in the corresponding digraphs is crucial for robustness of matrices. Hence bellow lemmas were the basis for the proof of the obtained necessary and sufficient condition for a Monge matrix to be robust. Obviously it is enaugh to consider only threshold digraphs $G(A,h)$ for $h \in H = \{a_{ij};\ i,j \in N\}$.

Lemma 3. [7] Let $A \in B(n,n)$ be a Monge matrix. Let $h \in H$. Let $\mathcal{K} \in \mathrm{SCC}^\star(G(A,h))$. Let c be a cycle of odd length $\ell(c) \geq 3$ in $\mathcal{K}$. Then there is a node in c with a loop.

Lemma 4. [7] Let $A \in B(n,n)$ be a Monge matrix. Let $h \in H$. Let for $i, k \in N$ be the loops (i,i) and (k,k) in the digraph $G(A,h)$. Then the nodes i and k are in the same non-trivial strongly connected component $\mathcal{K}$ of $G(A,h)$.

Theorem 2. [7] *Let $A \in B(n,n)$ be a Monge matrix. Then A is robust if and only if for each $h \in H$ the digraph $G(A,h)$ contains at most one non-trivial strongly connected component and this has a loop.*

In bellow two examples we consider fuzzy matrices over the set $B = [0,10]$. Both cases of answer are illustrated. The given Monge matrix is not robust in first example, while the slightly modified matrix in second example does.

Example 1. Let us check the robustness of the given Monge matrix $A \in B(6,6)$ for $B = [0,10]$.

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 & 2 \\ 0 & 0 & 3 & 3 & 3 & 0 \\ 0 & 3 & 3 & 3 & 0 & 0 \\ 0 & 3 & 2 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We shall verify due to Theorem 2 that $G(A,h)$ contains at most one non-trivial strongly connected component and this with a loop for each $h \in H = \{0,1,2,3\}$. Since $G(A,3)$ contains two non-trivial strongly connected components, namely component $\mathcal{K}_1$ generated by the node set $N_{\mathcal{K}_1} = \{2,5\}$ and $\mathcal{K}_2$ generated by the node set $N_{\mathcal{K}_2} = \{3,4\}$ (see Figure 1), the considered matrix is not robust.

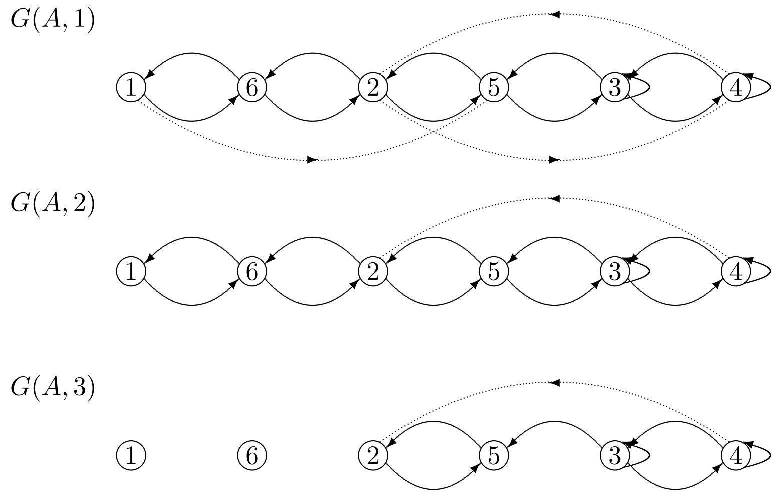


Figure 1: Threshold digraphs in non-robust case

Example 2. Let us check the robustness of the given Monge matrix $C \in B(6,6)$ which was derived from the above matrix A for $B = [0,10]$. Modified element is highlighted by bold character.

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & \mathbf{2} & 2 \\ 0 & 0 & 3 & 3 & 3 & 0 \\ 0 & 3 & 3 & 3 & 0 & 0 \\ 0 & 3 & 2 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

In contrast to the previous example is the considered matrix due to Theorem 2 robust. The threshold digraphs differ only for $h = 3$ (see Figure 2), namely $G(C,3)$ contains exactly one non-trivial strongly connected component and this with a loop.

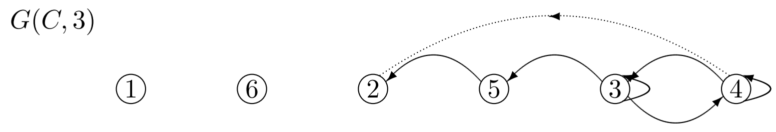


Figure 2: Threshold digraph in robust case

Using the definition of a minimal component and results concerning periodicity of fuzzy matrices proved in [3]

$$\begin{aligned} \mathrm{SCC}^{\star}(A) &= \cup\{\,\mathrm{SCC}^{\star}(G(A,h));\ h \in \{a_{ij}; i,j \in N\}\,\}, \\ \mathrm{SCC}^{\min}(A) &= \{\,\mathcal{K} \in \mathrm{SCC}^{\star}(A);\ \mathcal{K} \text{ is minimal in } \mathrm{SCC}^{\star}(A), \text{ ordered by inclusion}\,\} \end{aligned}$$

we can reformulate the equivalent conditions for a fuzzy matrix to be robust. Since $G(A,h)$ for h corresponding to the smallest element in the given matrix is a complete digraph the set of minimal components is never empty. According to this fact we present a slightly modified theorem from [7].

Theorem 3. [7] *Let $A \in B(n,n)$ be a Monge matrix. Then A is robust if and only if* $\mathrm{SCC}^{\min}(A) = \{\mathcal{K}\}$ *and $\mathcal{K}$ contains a loop.*

Corresponding algorithm checks the robustness in $O(n^3)$ time. If $\mathrm{SCC}^{\min}(A)$ is given, then the robustness can be verified in $O(n)$ time.

4 Robustness of interval Monge matrices

In this section we shall deal with matrices which elements are inexact, namely they are given by intervals, i.e. we will consider interval fuzzy matrices. Similarly to [2], [5], we define an interval matrix $\mathbf{A}$.

Definition 3. Let $\underline{A}, \overline{A} \in B(n,n)$, $\underline{A} \leq \overline{A}$. An interval matrix $\mathbf{A}$ with bounds $\underline{A}$ and $\overline{A}$ is defined as follows

$$\mathbf{A} = [\underline{A}, \overline{A}] = \left\{\, A \in B(n,n);\ \underline{A} \leq A \leq \overline{A} \,\right\}.$$

Since an interval matrix $\mathbf{A}$ is in fact a set of matrices instead of the question about robustness of the matrix we can set two sorts of questions. First, we can ask whether at least one of the matrices in $\mathbf{A}$ is robust or second, whether all of the matrices in $\mathbf{A}$ are robust.

Definition 4. An interval matrix $\mathbf{A}$ is called

- *possibly robust* if there exists a matrix $A \in \mathbf{A}$ such that A is robust,
- *universally robust* if each matrix $A \in \mathbf{A}$ is robust.

Definition 5. An interval matrix $\mathbf{A}^M = [\underline{A}, \overline{A}]$ is called interval Monge, if $\underline{A}$, $\overline{A} \in B(n,n)$ are Monge matrices and $\mathbf{A}^M = \{A \in \mathbf{A};\ A \text{ is Monge}\}$.

Since $\underline{A}$, $\overline{A} \in \mathbf{A}^M$, the set $\mathbf{A}^M$ is non-empty.

Equivalent conditions for a fuzzy interval matrix to be possibly robust were proved in [9]. However, the resulting matrix of the described polynomial algorithm need not to have the monge property ([7]). Moreover, there is no polynomial algorithm for checking the universal robustness of interval matrices in fuzzy algebra.

First, we have considered binary case of Monge fuzzy matrices over the set $B = \{0,1\}$ again. Following necessary and sufficient conditions with corresponding $O(n^3)$ algorithms were proved for interval matrices under restriction $\underline{A} \geq I_{ad}$.

Theorem 4. [8] *An interval Monge matrix $\mathbf{A}^M$ with $\underline{A} \geq I_{ad}$ is possibly robust if and only if $\overline{A}$ is robust.*

Theorem 5. [8] *An interval Monge matrix $\mathbf{A}^M$ with $\underline{A} \geq I_{ad}$ is universally robust if and only if $\underline{A}$ is robust.*

Algorithm **Possible (Universal) Robustness for binary case**

Input. $\mathbf{A}^M = [\underline{A}, \overline{A}]$, I_{ad}.
Output. 'non-Monge matrix' in variable *prbin* if $\mathbf{A}^M$ is not an interval Monge matrix; 'non-proper matrix' in variable *prbin* if $\mathbf{A}^M$ does not satisfied the condition $\underline{A} \geq I_{ad}$; 'yes' in variable *prbin* if $\mathbf{A}^M$ is possibly (universally) robust; 'no' in *binpr* if $\mathbf{A}^M$ is not possibly (universally) robust.
begin

(i) **If** $\underline{A}$ or $\overline{A}$ is not Monge **then** *prbin* :='non-Monge matrix'; **go to** end;

(ii) **If** the condition $\underline{A} \geq I_{ad}$ is not satisfied **then** *prbin* :='non-proper matrix'; **go to** end;

(iii) **If** the digraph $G(\overline{A}, 1)$ ($G(\underline{A}, 1)$) is not strongly connected **then** *prbin* :='no'; **go to** end;

(iv) **If** $G(\overline{A}, 1)$ ($G(\underline{A}, 1)$) contains no loop **then** *prbin* :='no', else *prbin* :='yes';

end

Example 3. Let us consider an interval matrix $\mathbf{A} = [\underline{A}, \overline{A}]$ with bounds $\underline{A}$, $\overline{A} \in B(5, 5)$

$$\underline{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}, \ \overline{A} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Since $\underline{A}$ and $\overline{A}$ are Monge matrices the corresponding interval matrix $\mathbf{A}^M$ is an interval Monge matrix. The digraph $G(\underline{A}, h)$ for threshold $h = 1$ contains two strongly connected components, while the digraph $G(\overline{A}, h)$ for threshold $h = 1$ is strongly connected and contains a loop (see Figure 3). Consequently the considered interval Monge matrix $\mathbf{A}^M$ is possibly but not universally robust.

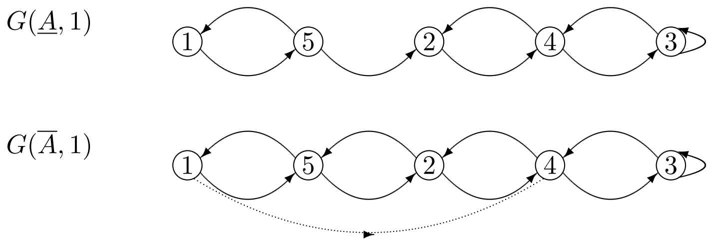


Figure 3: Possible robustness

Afterwards we have considered general case of max-min algebra. We have proved a necessary condition for an interval Monge matrix to be possibly robust.

Theorem 6. *If an interval Monge matrix $\mathbf{A}^M$ is possibly robust then for each $h \in H$ holds: if the digraph $G(\overline{A}, h)$ contains a strongly connected component $\mathcal{K}^\star$ with a loop then this is unique and all non-trivial strongly connected components of the digraph $G(\underline{A}, h)$ are subdigraphs of $\mathcal{K}^\star$.*

Proof. Let $\mathbf{A}^M$ be possibly robust. There exists a robust matrix $A \in \mathbf{A}^M$. Let $h \in H = \{\overline{a}_{ij};\, i, j \in N\} \cup \{\underline{a}_{ij};\, i, j \in N\}$ be arbitrary but fixed. By Theorem 2 the digraph $G(A, h)$ contains at most one non-trivial strongly connected component and this has a loop. If $G(\overline{A}, h)$ contains a non-trivial strongly connected component with a loop then by Lemma 4 this is unique in $G(\overline{A}, h)$. We shall consider two cases.

Case 1. The digraph $G(A, h)$ contains no non-trivial strongly connected component. Since $G(\underline{A}, h) \subseteq G(A, h)$ neither the digraph $G(\underline{A}, h)$ does and the assertion follows.

Case 2. The digraph $G(A, h)$ contains exactly one non-trivial strongly connected component $\mathcal{K}$ and this has a loop. Since $G(A, h) \subseteq G(\overline{A}, h)$ and $\overline{A}$ is a Monge matrix by Lemma 4 the digraph $G(\overline{A}, h)$ contains exactly one non-trivial strongly connected component $\mathcal{K}^\star$ and this has a loop. Moreover, $\mathcal{K} \subseteq \mathcal{K}^\star$. The digraph $G(\underline{A}, h)$ contains eventually non-trivial strongly connected components $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_t$. Since $G(\underline{A}, h) \subseteq G(A, h)$ the components $\mathcal{K}_i \subseteq \mathcal{K}$, for $i = 1, 2, \ldots, t$. Consequently $\mathcal{K}_i \subseteq \mathcal{K}^\star$, for $i = 1, 2, \ldots, t$ and the assertion follows. $\square$

Example 4. Let us consider an interval matrix $\mathbf{A} = [\underline{A}, \overline{A}]$ with bounds $\underline{A}$, $\overline{A} \in B(5,5)$

$$\underline{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \ \overline{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The Monge property of both matrices $\underline{A}$ and $\overline{A}$ guarantees that the corresponding interval matrix $\mathbf{A}^M$ is an interval Monge matrix. The digraph $G(\overline{A}, h)$ for threshold $h = 1$ contains one non-trivial strongly connected component $\mathcal{K}^\star$ with a loop generated by the node set $N_{\mathcal{K}^\star} = \{2, 3, 4\}$. In spite of this fact there exists a non-trivial strongly connected component $\mathcal{K}$ in $G(\underline{A}, h)$ generated by the node set $N_{\mathcal{K}} = \{1, 5\}$ which is not a subdigraph of $\mathcal{K}^\star$ and indeed the interval matrix $\mathbf{A}^M$ is not possibly robust (see Figure 4).

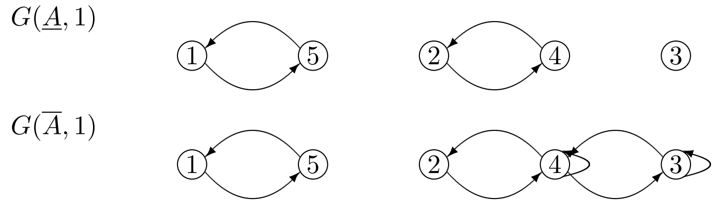


Figure 4: Necessary condition for possible robustness in non-robust case

5 Conclusion

The aim of this paper is to present results concerning robustness for special class of fuzzy matrices, namely Monge matrices with exact and inexact data as well and corresponding polynomial algorithms.

The theory of matrix robustness can reflect the following economic background. The fuzzy matrix $A = (a_{ij}) \in B(n,n)$ can represent the preferences of a customer due to different properties of a commodity on the market. The resulting eigenvector $y = A^k \otimes x$ is the maximum stable vector, i.e. the final vector of preferences.

References

[1] Burkard, R. E., Klinz, B. and Rudolf, R.: Perspectives of Monge properties in optimization, *DAM*, Volume **70** (1996), 95–161.

[2] Fiedler, M., Nedoma, J., Ramík, J., Rohn, J. and Zimmermann, K.: Linear Optimization Problems with Inexact Data. Springer–Verlag, Berlin 2006.

[3] Gavalec, M.: *Periodicity in extremal algebras.* GAUDEAMUS, Hradec Králové, 2004.

[4] Gavalec, M. and Plavka, J.: An $O(n^2)$ algorithm for maximum cycle mean of Monge matrices in max algebra, *DAM*, Volume **127** (2003), 651–656.

[5] Gavalec, M. and Zimmermann, K.: Classification of solutions to systems of two-sided equations with interval coefficients, *Inter. J. of Pure and Applied Math.* Volume **45** (2008), 533–542.

[6] Molnárová, M.: Robustness of (0–1) Monge fuzzy matrices, In: *Proceedings of* 31^{st} *Int. Conference Mathematical Methods in Economics 2013*, Jihlava, 2013, 636–641.

[7] Molnárová, M.: Robustness of Monge matrices in fuzzy algebra, In: *Proceedings of* 32^{nd} *Int. Conference Mathematical Methods in Economics 2014*, Olomouc, 2014, 679–684.

[8] Molnárová, M.: Robustness of Monge fuzzy matrices with inexact data, In: *Proceedings of* 33^{rd} *Int. Conference Mathematical Methods in Economics 2015*, Cheb, 2015, 566–571.

[9] Molnárová, M., Myšková, H. and Plavka, J.: The robustness of interval fuzzy matrices, *DAM*, Volume **438** (2013), 3350–3364.

[10] Plavka, J. and Szabó, P.: On the λ-robustness of matrices over fuzzy algebra, *DAM*, Volume **159** Issue 5 (2011), 381–388.

Interval Fuzzy Matrix Equations

Helena Myšková[1]

Abstract. This paper is dealing with solvability of interval matrix equations in fuzzy algebra. Fuzzy algebra is an algebraic structure, in which classical addition is replaced by maximum and the second operation, multiplication, is replaced by minimum. Fuzzy equations have found a broad area of applications in causal models which emphasize relationships between input and output variables. They are used in diagnosis models or models of non-deterministic systems. We shall deal with the both-sided matrix equations. If we replace elements in matrices with intervals of possible values, we obtain an interval matrix equation. We can define several types of solvability of interval matrix equations. In this paper, we shall deal with four of them. We define the tolerance, weak tolerance, left-weak tolerance, and right-weak tolerance solvability and provide polynomial algorithms for checking of them. We prove, that in fuzzy algebra, weak tolerance, left-weak tolerance, and right-weak tolerance solvability are equivalent to each other.

Keywords: fuzzy algebra, interval matrix equation, tolerance solvability, weak tolerance solvability

JEL classification: C02
AMS classification: 15A18; 15A80; 65G30

1 Motivation

Fuzzy equations have found a broad area of applications in causal models which emphasize relationships between input and output variables. They are used for instance in diagnosis models [9], [11], [12]. The solution of the fuzzy relational equation of the form $A \otimes x = b$, where A is a matrix, b and x are vectors of suitable dimensions and classical addition and multiplication operations are replaced by maximum and minimum, provides a maximal set of symptoms that produce the given fault.

The solvability of the systems of fuzzy linear equations is well reviewed. In this paper, we shall deal with the solvability of fuzzy matrix equations of the form $A \otimes X \otimes C = B$, where A, B, and C are given matrices of suitable sizes and X is an unknown matrix. In the following example we will show one of possible applications.

Example 1. Let us consider a situation, in which there is a supplier of components for four types of computers T_1, T_2, T_3, and T_4. Computers shall be completed in companies C_1, C_2, C_3 and delivered through retail stores R_1, R_2 to the final consumers F_1, F_2 and F_3. Relations between them are expressed in Figure 1.

There is arrow $(T_i\, C_j)$ if computer T_i can be completed in C_j and there is arrow $(R_l\, F_k)$ if consumer F_k purchases computers from retail store R_l. ($i = 1, 2, 3, 4,\ j = 1, 2, 3,\ k = 1, 2, 3,\ l = 1, 2$). Denote by a_{ij} the amount of computers of type T_i, components for which the supplier is able to deliver to C_j. The numbers c_{lk} denote the amount of computers that the retail store R_l is able to deliver to F_k. Suppose that final consumer F_k needs b_{ik} computers of type T_i. We are looking for numbers x_{jl}, which denote numbers of computers completed by C_j for R_l. To meet the requirements of all consumers for computers of type T_1 the following equations must be satisfied:

$$\max\Big\{\min\{a_{11}, x_{11}, c_{11}\}, \min\{a_{12}, x_{21}, c_{11}\}\Big\} = b_{11},$$
$$\max\Big\{\min\{a_{11}, x_{11}, c_{12}\}, \min\{a_{12}, x_{21}, c_{12}\}, \min\{a_{12}, x_{22}, c_{22}\}\Big\} = b_{12},$$
$$\max\Big\{\min\{a_{11}, x_{12}, c_{23}\}, \min\{a_{11}, x_{11}, c_{13}\}, \min\{a_{12}, x_{21}, c_{13}\}, \min\{a_{12}, x_{22}, c_{23}\}\Big\} = b_{13}.$$

[1]Technical University in Košice, Faculty of Electrical Engineering and Informatics, Department of Mathematics and Theoretical Informatics, Němcovej 32, 042 00 Košice, Slovak Republic, helena.myskova@tuke.sk

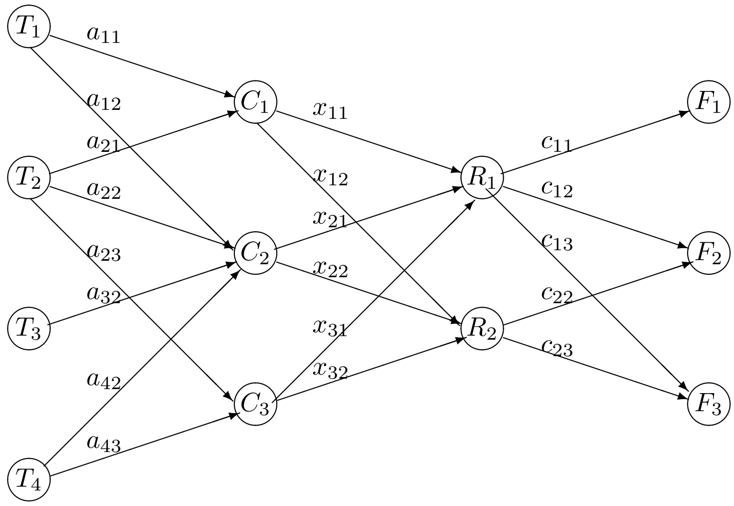


Figure 1 Transportation system

In the same way we can write equations which must be satisfied to ensure requirements of all consumers for computers of types T_2 and T_3.

In general, suppose that there are m computers $T_1, T_2, \ldots, T_m$, n companies $C_1, C_2, \ldots, C_n$, s retail stores $R_1, R_2, \ldots, R_s$ and r final consumers $F_1, F_2, \ldots, F_r$. If there is no connection from T_i to C_j (from R_l to F_k), we put $a_{ij} = O$ ($c_{lk} = O$). to ensure requirements of all consumers for computers of types T_2 and T_3. Our task is to choose the appropriate capacities $x_{jl},\ j \in N = \{1,\ 2, \ldots,\ n\}$, $l \in S = \{1, 2, \ldots, s\}$ such that we ensure requirements of all consumers for all types of computers. We obtain

$$\max_{j \in N,\, l \in S} \min\{a_{ij}, x_{jl}, c_{lk}\} = \ b_{ik} \tag{1}$$

for each $i \in M = \{1, 2, \ldots, m\}$ and for all $k \in R = \{1, 2, \ldots, r\}$.

A certain disadvantage of any necessary and sufficient condition for the solvability of (1) stems from the fact that it only indicates the existence or nonexistence of the solution but does not indicate any action to be taken to increase the degree of solvability. However, it happens quite often in modeling real situations that the obtained system turns out to be unsolvable. One of possible methods of restoring the solvability is to replace the exact input values by intervals of possible values. The result of the substitution is so-called interval fuzzy matrix equation. In this paper, we shall deal with the solvability of interval fuzzy matrix equations. We define the tolerance, right-weak tolerance, left-weak tolerance and weak tolerance solvability and provide polynomial algorithms for checking them.

2 Preliminaries

Fuzzy algebra is the triple $(\mathcal{I}, \oplus, \otimes)$, where $\mathcal{I} = [O, I]$ is a linear ordered set with the least element O, the greatest element I, and two binary operations $a \oplus b = \max\{a, b\}$ and $a \otimes b = \min\{a, b\}$.

Denote by M, N, R, and S the index sets $\{1, 2, \ldots, m\}$, $\{1, 2, \ldots, n\}$, $\{1, 2, \ldots, r\}$, and $\{1, 2, \ldots, s\}$, respectively. The set of all $m \times n$ matrices over $\mathcal{I}$ is denoted by $\mathcal{I}(m, n)$ and the set of all column n vectors over $\mathcal{I}$ we denote by $\mathcal{I}(n)$.
Operations $\oplus$ and $\otimes$ are extended to matrices and vectors in the same way as in the classical algebra. We will consider the *ordering* $\leq$ on the sets $\mathcal{I}(m, n)$ and $\mathcal{I}(n)$ defined as follows:

- for $A, C \in \mathcal{I}(m, n):\ A \leq C$ if $a_{ij} \leq c_{ij}$ for all $i \in M,\ j \in N$,
- for $x, y \in \mathcal{I}(n):\ x \leq y$ if $x_j \leq y_j$ for all $j \in N$.

The *monotonicity of* $\otimes$ means that for each $A, C \in \mathcal{I}(m, n),\ A \leq C$ and for each $B,\ D \in \mathcal{I}(n, s),\ B \leq D$ the inequality $A \otimes B \leq C \otimes D$ holds true.

Let $A \in \mathcal{I}(m,n)$ and $b \in \mathcal{I}(m)$. In fuzzy algebra, we can write the system of equations in the matrix form

$$A \otimes x = b. \tag{2}$$

The crucial role for the solvability of system (2) in fuzzy algebra is played by the *principal solution* of system (2), defined by

$$x_j^*(A,b) = \min_{i \in M}\{b_i : a_{ij} > b_i\} \tag{3}$$

for each $j \in N$, with $\min \emptyset = I$.

The following theorem describes the importance of the principal solution for the solvability of (2).

Theorem 1. [2, 13] *Let $A \in \mathcal{I}(m,n)$ and $b \in \mathcal{I}(m)$ be given.*

(i) *If $A \otimes x = b$ for $x \in \mathcal{I}(n)$, then $x \leq x^*(A,b)$.*

(ii) $A \otimes x^*(A,b) \leq b$.

(iii) *The system $A \otimes x = b$ is solvable if and only if $x^*(A,b)$ is its solution.*

The properties of a principal solution are expressed in the following assertions.

Lemma 2. [1] Let $A \in \mathcal{I}(m,n)$ and $b, d \in \mathcal{I}(m)$ be such that $b \leq d$. Then $x^*(A,b) \leq x^*(A,d)$.

Lemma 3. [3] Let $b \in \mathcal{I}(m)$ and $C, D \in \mathcal{I}(m,n)$ be such that $D \leq C$. Then $x^*(C,b) \leq x^*(D,b)$.

Lemma 4. Let $A \in \mathcal{I}(m,n)$, $b \in \mathcal{I}(m)$ and $c \in \mathcal{I}$. Then for each $j \in N$ the following equality holds:

$$\min\{x_j^*(A \otimes c, b), c\} = \min\{x_j^*(A,b), c\}. \tag{4}$$

Proof. In the case that $x_j^*(A,b) \geq c$ we have $x_j^*(A \otimes c, b) \geq c$, according to Lemma 3, so both minimas are equal to c. In the second case $x_j^*(A,b) = b_i < c$ for some $i \in M$, which follows that $a_{ij} > b_i$. Then $a_{ij} \otimes c > b_i$ and consequently $x_j^*(A \otimes c, b) \leq b_i = x_j^*(A,b)$. Together with $x_j^*(A \otimes c, b) \geq x_j^*(A,b)$ we obtain $x_j^*(A \otimes c, b) = x_j^*(A,b)$ and consequently (4). $\square$

3 Matrix Equations

Let $A \in \mathcal{I}(m,n)$, $B \in \mathcal{I}(m,r)$, and $C \in \mathcal{I}(m,n)$ be matrices with elements a_{ij}, b_{ik}, and c_{lk}, respectively. We shall write system of equalities (1) from Example 1 in the matrix form

$$A \otimes X \otimes C = B. \tag{5}$$

Denote by $X^*(A,B,C) = \big(x_{jl}^*(A,B,C)\big)$ the matrix defined as follows

$$x_{jl}^*(A,B,C) = \min_{k \in R}\{x_j^*(A \otimes c_{lk}, B_k)\}. \tag{6}$$

We shall call the matrix $X^*(A,B,C)$ a *principal matrix solution* of (5). The following theorem expresses the properties of $X^*(A,B,C)$ and gives the necessary and sufficient condition for the solvability of (5).

Theorem 5. [8] *Let $A \in \mathcal{I}(m,n)$, $B \in \mathcal{I}(m,r)$ and $C \in \mathcal{I}(m,n)$.*

(i) *If $A \otimes X \otimes C = B$ for $X \in \mathcal{I}(n,s)$, then $X \leq X^*(A,B,C)$.*

(ii) $A \otimes X^*(A,B,C) \otimes C \leq B$.

(iii) *The matrix equation $A \otimes X \otimes C = B$ is solvable if and only if $X^*(A,B,C)$ is its solution.*

Remark 1. Equality (6) can be written in the form

$$X^*(A,B,C) = (X_1^*(A,B,C), X_2^*(A,B,C), \ldots, X_s^*(A,B,C)),$$

where

$$X_l^*(A,B,C) = \min_{k \in R} x^*(A \otimes c_{lk}, B_k). \tag{7}$$

4 Interval matrix equations

Similarly to [5, 6, 7, 10], we define *interval matrices* $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ as follows:

$$\boldsymbol{A} = [\underline{A}, \overline{A}] = \left\{ A \in \mathcal{I}(m,n);\ \underline{A} \leq A \leq \overline{A} \right\},$$

$$\boldsymbol{B} = [\underline{B}, \overline{B}] = \left\{ B \in \mathcal{I}(m,r);\ \underline{B} \leq B \leq \overline{B} \right\},$$

$$\boldsymbol{C} = [\underline{C}, \overline{C}] = \left\{ C \in \mathcal{I}(s,r);\ \underline{C} \leq C \leq \overline{C} \right\}.$$

Denote by

$$\boldsymbol{A} \otimes X \otimes \boldsymbol{C} = \boldsymbol{B} \tag{8}$$

the set of all matrix equations of the form (5) such that $A \in \boldsymbol{A}$, $B \in \boldsymbol{B}$, and $C \in \boldsymbol{C}$. We call (8) an *interval fuzzy matrix equation*.

We shall think over the solvability of interval fuzzy matrix equation on the ground of the solvability of matrix equations of the form (5) such that $A \in \boldsymbol{A}$, $B \in \boldsymbol{B}$, and $C \in \boldsymbol{C}$. We can define several types of solvability of an interval fuzzy matrix equation. Similarly to the max-plus algebra (see [8]), we define four types of solvability:

Definition 1. Interval fuzzy matrix equation of the form (8) is

- *tolerance solvable* if there exist $X \in \mathcal{I}(n,s)$ such that for each $A \in \boldsymbol{A}$ and for each $C \in \boldsymbol{C}$ the product $A \otimes X \otimes C$ lies in $\boldsymbol{B}$,
- *right-weakly tolerance solvable* if for each $C \in \boldsymbol{C}$ there exist $X \in \mathcal{I}(n,s)$ such that for each $A \in \boldsymbol{A}$ the product $A \otimes X \otimes C$ lies in $\boldsymbol{B}$,
- *left-weakly tolerance solvable* if for each $A \in \boldsymbol{A}$ there exist $X \in \mathcal{I}(n,s)$ such that for each $C \in \boldsymbol{C}$ the product $A \otimes X \otimes C$ lies in $\boldsymbol{B}$,
- *weakly tolerance solvable* if for each $A \in \boldsymbol{A}$ and for each $C \in \boldsymbol{C}$ there exist $X \in \mathcal{I}(n,s)$ such that $A \otimes X \otimes C \in \boldsymbol{B}$.

4.1 Tolerance Solvability

Theorem 6. *Interval fuzzy matrix equation of the form* (8) *is tolerance solvable if and only if*

$$\underline{A} \otimes X^*(\overline{A}, \overline{B}, \overline{C}) \otimes \underline{C} \geq \underline{B}. \tag{9}$$

Proof. A proof is based on the same idea as those of paper [8]. $\square$

Suppose that $m = r = s = n$, i.e., all matrices are square of order n. Checking the tolerance solvability is based on the verifying of inequality (9). Computing $X_l^*(\overline{A}, \overline{B}, \overline{C})$ by (7) for fixed l requires $n \cdot O(n^2) = O(n^3)$ arithmetic operation. Hence, computing the matrix $X^*(\overline{A}, \overline{B}, \overline{C})$ requires $n \cdot O(n^3) = O(n^4)$ operations.

4.2 Right-weak Tolerance Solvability

Lemma 7. Interval fuzzy matrix equation of the form (8) is right-weakly tolerance solvable if and only if the inequality

$$\underline{A} \otimes X^*(\overline{A}, \overline{B}, C) \otimes C \geq \underline{B} \tag{10}$$

holds for each $C \in \boldsymbol{C}$.

Proof. Let $C \in \boldsymbol{C}$ be arbitrary but fixed. The existence of $X \in \mathcal{I}(n,s)$ such that $A \otimes X \otimes C \in [\underline{B}, \overline{B}]$ for each $A \in \boldsymbol{A}$ is equivalent to the tolerance solvability of the fuzzy matrix equation with constant matrix $C = \underline{C} = \overline{C}$, which is, according to Theorem 6, equivalent to (10). Therefore, interval fuzzy matrix equation (8) is right-weakly tolerance solvable if and only if inequality (10) is fulfilled for each matrix $C \in \boldsymbol{C}$. $\square$

Lemma 7 does not give an algorithm for checking the right-weak tolerance solvability. It is easy to see that the tolerance solvability implies the right-weak tolerance solvability. The converse implication may not be valid.

Theorem 8. *An interval fuzzy matrix equation of the form* (8) *is right-weakly tolerance solvable if and only if it is tolerance solvable.*

Proof. Suppose that an interval fuzzy matrix equation is not tolerance solvable, i. e., there exist $i \in M$, $p \in R$ such that $\left[\underline{A} \otimes X^*(\overline{A}, \overline{B}, \overline{C}) \otimes \underline{C}\right]_{ip} < \underline{b}_{ip}$.

Denote by $C^{(p)}$ the matrix with the following entries

$$c_{lk}^{(p)} = \begin{cases} \underline{c}_{lk} & \text{for } k = p,\ l \in S, \\ \overline{c}_{lk} & \text{for } k \in R,\ k \neq p,\ l \in S. \end{cases} \tag{11}$$

We will prove that

$$\left[\underline{A} \otimes X^*(\overline{A}, \overline{B}, C^{(p)}) \otimes C^{(p)}\right]_{ip} = \left[\underline{A} \otimes X^*(\overline{A}, \overline{B}, \overline{C}) \otimes \underline{C}\right]_{ip}. \tag{12}$$

The both sides of (12) we can write as

$$\left[\underline{A} \otimes X^*(\overline{A}, \overline{B}, C^{(p)}) \otimes C^{(p)}\right]_{ip} = \max_{j \in N,\, l \in S} \min\{\underline{a}_{ij}, x^*_{jl}(\overline{A}, \overline{B}, C^{(p)}), c^{(p)}_{lp}\}$$

and

$$\left[\underline{A} \otimes X^*(\overline{A}, \overline{B}, \overline{C}) \otimes \underline{C}\right]_{ip} = \max_{j \in N,\, l \in S} \min\{\underline{a}_{ij}, x^*_{jl}(\overline{A}, \overline{B}, \overline{C}), \underline{c}_{lp}\}$$

It is sufficient to prove that

$$\min\{\underline{a}_{ij}, x^*_{jl}(\overline{A}, \overline{B}, C^{(p)}), c^{(p)}_{lp}\} = \min\{\underline{a}_{ij}, x^*_{jl}(\overline{A}, \overline{B}, \overline{C}), \underline{c}_{lp}\} \tag{13}$$

for each $j \in N$, $l \in S$. The left-hand side of (13) is equal to

$$\min\{\underline{a}_{ij}, x^*_{jl}(\overline{A}, \overline{B}, C^{(p)}), c^{(p)}_{lp}\} = \min\left\{\underline{a}_{ij}, \min_{k \neq p} x^*_j(\overline{A} \otimes \overline{c}_{lk}, \overline{B}_k), x^*_j(\overline{A} \otimes \underline{c}_{lp}, \overline{B}_p), \underline{c}_{lp}\right\}$$

and the right-hand side is equal to

$$\min\{\underline{a}_{ij}, x^*_{jl}(\overline{A}, \overline{B}, \overline{C}), \underline{c}_{lp}\} = \min\left\{\underline{a}_{ij}, \min_{k \neq p} x^*_j(\overline{A} \otimes \overline{c}_{lk}, \overline{B}_k), x^*_j(\overline{A} \otimes \overline{c}_{lp}, \overline{B}_p), \underline{c}_{lp}\right\}.$$

We shall prove that

$$\min\{x^*_j(\overline{A} \otimes \underline{c}_{lp}, \overline{B}_p), \underline{c}_{lp}\} = \min\{x^*_j(\overline{A} \otimes \overline{c}_{lp}, \overline{B}_p), \underline{c}_{lp}\}. \tag{14}$$

According to Lemma 4 we obtain

$$\min\{x^*_j(\overline{A} \otimes \underline{c}_{lp}, \overline{B}_p), \underline{c}_{lp}\} = \min\{x^*_j(\overline{A}, \overline{B}_p), \underline{c}_{lp}\},$$

$$\min\{x^*_j(\overline{A} \otimes \overline{c}_{lp}, \overline{B}_p), \underline{c}_{lp}\} = \min\{x^*_j(\overline{A} \otimes \overline{c}_{lp}, \overline{B}_p), \underline{c}_{lp}, \overline{c}_{lp}\} = \min\{x^*_j(\overline{A}, \overline{B}_p), \underline{c}_{lp}\}.$$

Since (14) is satisfied, from (12) we obtain $\left[\underline{A} \otimes X^*(\overline{A}, \overline{B}, C^{(p)}) \otimes C^{(p)}\right]_{ip} < \underline{b}_{ip}$. According to Lemma 7 an interval fuzzy matrix equation is not right-weakly tolerance solvable.

The converse implication is trivial. □

4.3 Left-weak Tolerance Solvability

Lemma 9. An interval fuzzy matrix equation of the form (8) is left-weakly tolerance solvable if and only if an interval fuzzy matrix equation of the form

$$\boldsymbol{C}^\top \otimes X^\top \otimes \boldsymbol{A}^\top = \boldsymbol{B}^\top \tag{15}$$

is right-weakly tolerance solvable.

Proof. A proof is similar as those of paper [8]. □

Theorem 10. *An interval fuzzy matrix equation of the form* (8) *is left-weakly tolerance solvable if and only if it is tolerance solvable.*

Proof. A proof follows from Theorem 8 and Lemma 9. □

4.4 Weak Tolerance Solvability

Theorem 11. *Interval fuzzy matrix equation of the form* (8) *is weakly tolerance solvable if and only if it is tolerance solvable.*

Proof. Suppose that (8) is weakly tolerance solvable. We obtain the following sequence of implications

$$(\forall A \in \boldsymbol{A})(\forall C \in \boldsymbol{C})(\exists X \in \mathcal{I}(n,s)) A \otimes X \otimes C \in \boldsymbol{B} \overset{\text{Th}\,8}{\Rightarrow} (\forall A \in \boldsymbol{A})(\exists X \in \mathcal{I}(n,s))(\forall C \in \boldsymbol{C}) A \otimes X \otimes C \in \boldsymbol{B}$$
$$\overset{\text{Th}\,10}{\Rightarrow} (\exists X \in \mathcal{I}(n,s))(\forall A \in \boldsymbol{A})(\forall C \in \boldsymbol{C})\, A \otimes X \otimes C \in \boldsymbol{B},$$

hence (8) is tolerance solvable. The converse implication is trivial. □

4.5 Application

Let us return to Example 1. Suppose that the amounts of computers given by the elements of matrices A and C are not exact numbers, but the are given as intervals of possible values. Further, request in provision of computers given by the numbers b_{ik} are from given intervals, too. The tolerance solvability answers the question whether there exist the numbers of computers completed by C_j for R_l such that for each values a_{ij} and c_{lk} from given intervals of possible values the requirements of all consumers are fulfilled. Moreover, in the positive case we obtain the solution given by matrix $X^*(\overline{A}, \overline{B}, \overline{C})$.

References

[1] Cechlárová, K.: Solutions of interval systems in fuzzy algebra. In: *Proceedings of SOR* 2001, (V. Rupnik, L. Zadnik-Stirn, S. Drobne, eds.) Preddvor, Slovenia, 321–326.

[2] Cuninghame-Green, R. A.: *Minimax Algebra.* Springer, Berlin, 1979.

[3] Myšková, H.: Interval systems of max-separable linear equations, *Lin. Algebra Appl.* **403** (2005), 263–272.

[4] Myšková, H.: Fuzzy matrix equations, In: *Proceedings of 33-th International Conference on Mathematical Methods in Economics* (D. Martinčík, J. Ircingová, P. Janeček, eds.), University of West Bohemia, Plzeň, 2015, 572–577.

[5] Myšková, H.: Robustness of interval Toeplitz matrices in fuzzy algebra, *Acta Electrotechnica et Informatica* **12** (4) (2012), 56–60.

[6] Myšková, H.: Interval eigenvectors of circulant matrices in fuzzy algebra, *Acta Electrotechnica et Informatica* **12** (3) (2012), 57–61.

[7] Myšková, H.: Weak stability of interval orbits of circulant matrices in fuzzy algebra, *Acta Electrotechnica et Informatica* **12** (3) (2012), 51–56.

[8] Myšková, H.: Interval max-plus matrix equations, *Lin. Algebra Appl.* **492** (2016),111–127.

[9] Nola, A. Di, Salvatore, S., Pedrycz, W. and Sanchez, E.: *Fuzzy Relation Equations and Their Applications to Knowledge Engineering*, Kluwer Academic Publishers, Dordrecht, 1989.

[10] Plavka, J: On the $O(n^3)$ algorithm for checking the strong robustness of interval fuzzy matrices, Discrete Appl. Math. **160** (2012), 640–647.

[11] Sanchez, E: Medical diagnosis and composite relations, In: *Advances in Fuzzy Set Theory and Applications* (M. M. Gupta, R. K. Ragade, R. R. Yager, eds.) North-Holland, Amsterdam, The Netherlands, 1979, 437–444.

[12] T. Terano, Y. Tsukamoto, Failure diagnosis by using fuzzy logic. In: Proc. IEEE Conference on Decision Control (New Orleans, LA, 1977), 1390–1395.

[13] Zimmermann, K.: *Extremální algebra*, Ekonomicko-matematická laboratoř Ekonomického ústavu ČSAV, Praha, 1976.

Sensitivity analysis of MACD indicator on selection of input periods combination

Martin Navrátil[1], Pavla Koťátková Stránská[2], Jana Heckenbergerová[3]

Abstract. The trend indicator MACD (Moving Average Convergence/Divergence) is commonly used powerful tool of technical analysis. It allows to predict many profitable trades (buying and selling signal), especially when convergence and divergence periods are correctly recognized. The MACD value is dependent on three input parameters, length of short, long and signal period respectively. Generally applied combination of periods 12/26/9 was introduced by Gerald Appel in 1979 for trading with silver. However there is no standard or methodology for selection of optimal input periods and many other combinations are utilized.
The aim of presented contribution is the sensitivity analysis of MACD values on selection of short, long and signal period lengths. In the first part MACD indicator is formally defined using weighted exponential averages, from this definition the dependence of input parameters and resulting MACD value is more understandable. Case study involves three types of shares with growing, falling and changing trend respectively. From results we can conclude that prolonging of signal frame increases sensitivity of MACD indicator. Moreover combination 12/26/9 is not optimal for selected shares and lengths of short and long period are in mutual relation.

Keywords: MACD indicator, weighted exponential average, technical analysis, profitable trade, short period, long period, signal period.

JEL Classification: G17, C61
AMS Classification: 97M30

1 Introduction to Technical Analysis and Indicators

Predicting the future value of securities, especially stock prices, is constantly developing area. Discipline that describes methods of analysis and prediction of stock prices is called technical analysis. Main algorithms are based on statistical, numerical and visual analysis of share prices and sales volumes respectively. There are two essential groups of methods in technical analysis - the technical indicators and graphical formations known as chart patterns [4, 6, 7].

Technical indicators work with recent or close history data with goal to predict future trend of price. They generate so called "buying and selling signals" based on the expected rise or fall of share price. These signals help speculators to decide whether certain trade will be profitable [8].

Most popular methods of technical analysis are based on moving averages (simple, exponential or weighted) with given period. Moreover, for example price oscillators, are constructed using combinations of moving averages with different periods. Famous trend indicator MACD (Moving Average Convergence/Divergence) uses combination of three exponential averages with short, long and signal period respectively. Commonly applied combination of periods 12/26/9 was introduced by Gerald Appel in 1979 for trading with silver. However there is no standard or methodology for selection of optimal input periods and many other combinations are utilized [2, 5, 9].

The aim of presented contribution is sensitivity analysis of MACD values on selection of short, long and signal period lengths. Following section contains definition of MACD indicator and its input parameters. In section Case Study three types of shares with different type of trend are analyzed for selected combinations of input periods using trade success rates. Major results and directions of future work are summarized in the last section.

---

[1] University of Pardubice, Faculty of Electrical Engineering and Informatics, Department of Mathematics and Physics, Studentska 95, 532 10 Pardubice, Czech Republic, e-mail: martin.navratil1@student.upce.cz
[2] Masaryk Institute of Advanced Studies (MIAS CTU), Department of management and quantitative methods, Kolejni 2637/2a, 160 00 Prague 6, Czech Republic, email: stranska.pavla@seznam.cz
[3] University of Pardubice, Faculty of Electrical Engineering and Informatics, Department of Mathematics and Physics, Studentska 95, 532 10 Pardubice, Czech Republic, e-mail: jana.heckenbergerova@upce.cz

2 Trend Indicator MACD (Moving Average Convergence/Divergence)

Let us denote:

i	time index
P_i	stock price in time i
EMA	Exponential Moving Average dataset
PO	Price oscillator dataset
n	period of EMA
ep	exponential percentage.

Exponential moving average of given stock price dataset P is defined as:

$$EMA(P,n)_i = P_i * ep + EMA(P,n)_{i-1} * (1 - ep), \tag{1}$$

where

$$ep = \frac{2}{n+1}. \tag{2}$$

Now it is important to denote input parameters of MACD indicator:

SP	short period
LP	long period
SgP	signal period,

and define MACD value in given time index as follows:

$$MACD_i = EMA(PO,\ SgP)_i = PO_i * \left(\frac{2}{SgP+1}\right) + MACD_{i-1} * \left(1 - \frac{2}{SgP+1}\right) \tag{3}$$

where

$$PO = EMA(P, SP) - EMA(P, LP). \tag{4}$$

In words, from subtraction of first two exponential moving averages with short and long period we are getting price oscillator dataset that moves around zero and indicates price general trend. Third exponential moving average with signal period is applied to price oscillator and it gives us MACD values. Buying and selling signals for selected share are detected as intersections of price oscillator and MACD datasets [10].

Resulting value of MACD indicator is dependent on three input parameters, length of short, long and signal period respectively. Their combination is commonly marked as triplet *SP/LP/SgP*. However there is no standard or methodology for selection of optimal input periods and many other combinations are utilized as summarized in following Table 1.

Used	Short period	Long period	Signal period
In practice	12	26	9
	13	26	9
In FX market	7	28	7
	3	11	17
	5	35	9
	8	17	9
	15	35	5
Automatic Trading System (ATS)	30	55	8
	30	55	20

Table 1 Selected combination of MACD periods [3, 7, 8]

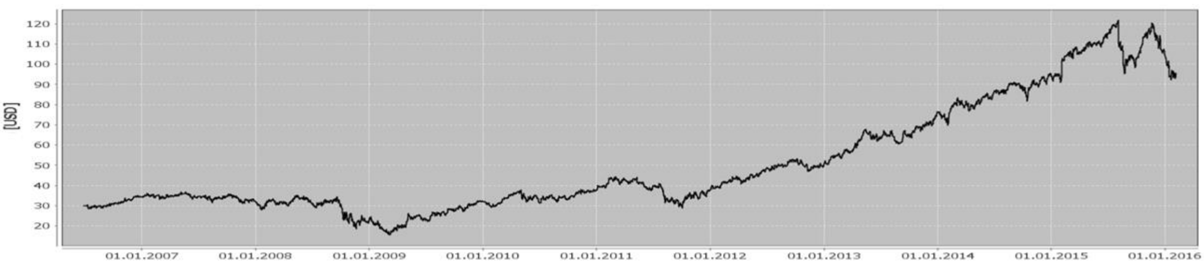


Figure 1 Shares of Walt Disney

3 Sensitivity Analysis – Case Study

For our sensitivity analysis we decided to choose three daily 10years long datasets of shares with different types of trend. Growing trend is represented by shares of Walt Disney illustrated in Fig. 1, falling trend represents shares of Telecom Italia shown in Fig. 2 and variable, frequently changing trend is analyzed for shares of Big Lots inc. illustrated in Fig. 3.

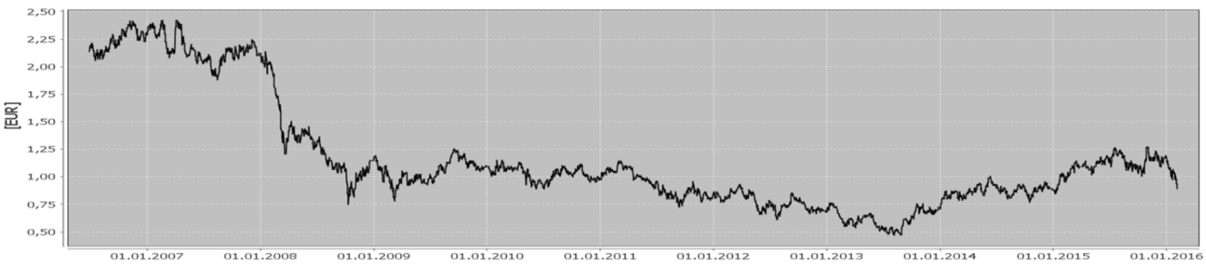


Figure 2 Shares of Telecom Italia

Figure 3 Shares of Big Lots inc.

Boundaries of input parameters were set as:

$$\begin{aligned} SP &= \langle 1; 20 \rangle \\ LP &= \langle SP+1; 40 \rangle . \\ SgP &= \langle 2; 9 \rangle \end{aligned} \tag{5}$$

All possible combinations of input periods with step equal to one were analyzed in our software, which was developed by main author and it is not publically accessible. Each result (combination) was augmented with success rate based on holding duration of selected share. Simple success rate (SSR) is defined as number of profiting trades divided by number of all trades (6), to show how market speculations based on MACD indicator perform.

$$SSR = \frac{w}{n}, \tag{6}$$

where w is number of profitable trades and n is number of all trades based on MACD indicator buying signal. The SSR ranges from 0 (none profitable trade) to 1 (all trades were profitable).

Resulting 5D dependencies are hard to visualize, this is the reason why we decided to specially cumulate information. In following figures Fig. 4, 5 and 6 , we depict these cumulated dependencies as ranges of transaction successfulness based on holding time (y-axis) for combinations of parameters (x-axis) that are sorted first by the short period, then by the long period and finally by the signal period.

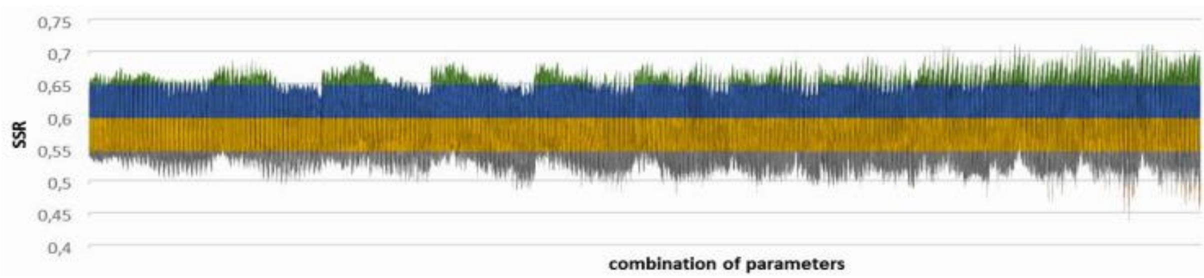


Figure 4 Ranges of MACD success rates for Walt Disney shares with increasing trend

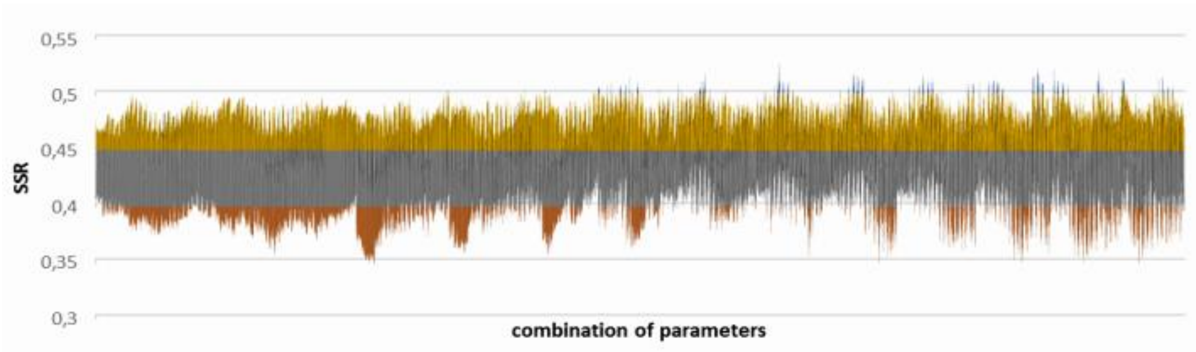


Figure 5 Ranges of MACD success rates for Telecom Italia shares with decreasing trend

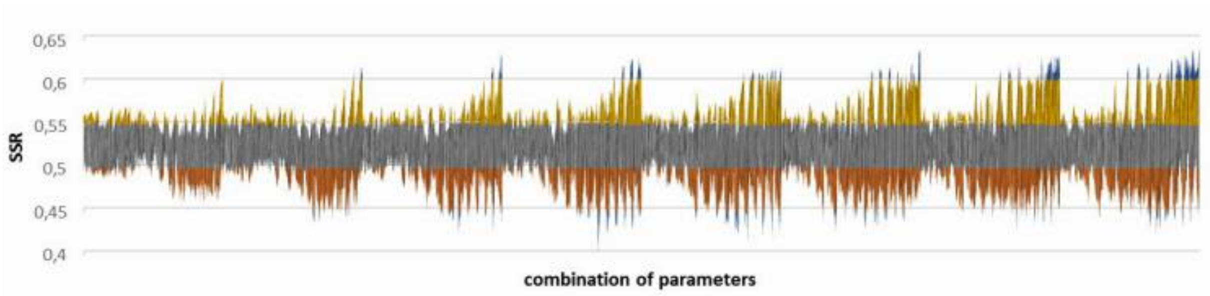


Figure 6 Ranges of MACD success rates for shares of Big Lots inc. with variable trend

Figure 4 illustrates MACD success rates for share of Walt Disney with rising trend. It is obvious that most of the trades with this share are profitable (success rate is higher than 0.5). Moreover 50percent threshold is broken only for a few combinations and only by the bottom boundary of profitability ranges. Opposite statements result from analysis of falling trend of Telecom Italia shares captured in Fig. 5. Only few profitable trades could be found, generally trades with this share are profitless for arbitrary selection of input parameter of MACD indicator. Most interesting result is illustrated in Fig. 6 for changing trend of Big Lots inc. shares. Selection of combination periods highly influences trade profitability. Trades are always profitable for low values of short period (SP), however with rising number of short and long period variation of success rates increases. Proper combination of MACD periods has to be utilized for guarantee of profitable trade.

3.1 Analysis of signal period length influence

Following analysis concentrates on selection of signal period. Figure 7 illustrates dependency between holding duration and SgP value for shares of Big Lots inc. Signal period affects mainly sensitivity of MACD indicator. For low values of SgP more buying and selling signals are generated and MACD sensitivity increases. Moreover Fig. 7 shows that shape of the graph is almost constant plane and therefore dependency between holding duration and SgP value is weak.

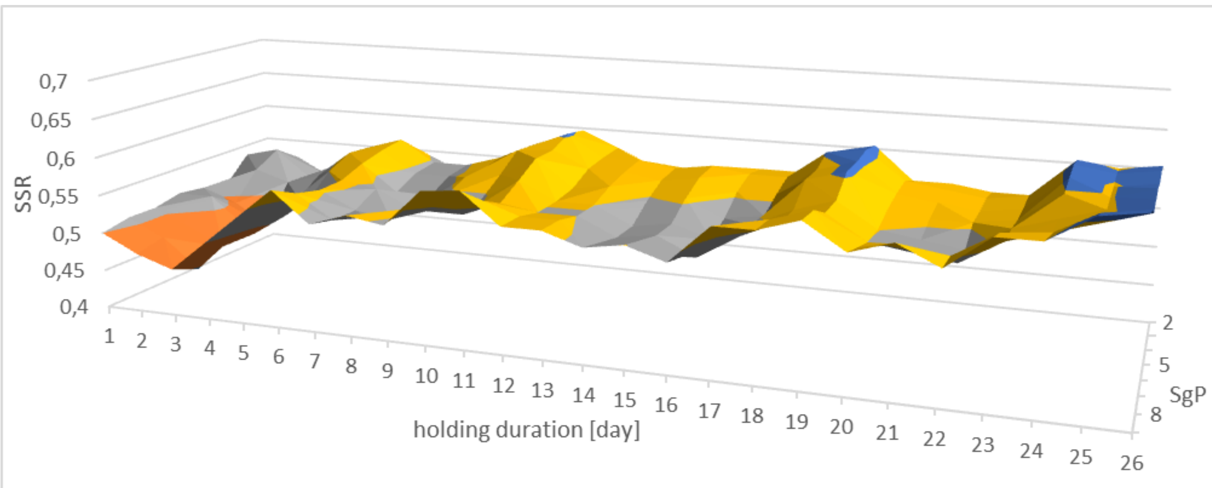


Figure 7 Sensitivity of MACD success rate on selection of signal period and holding duration

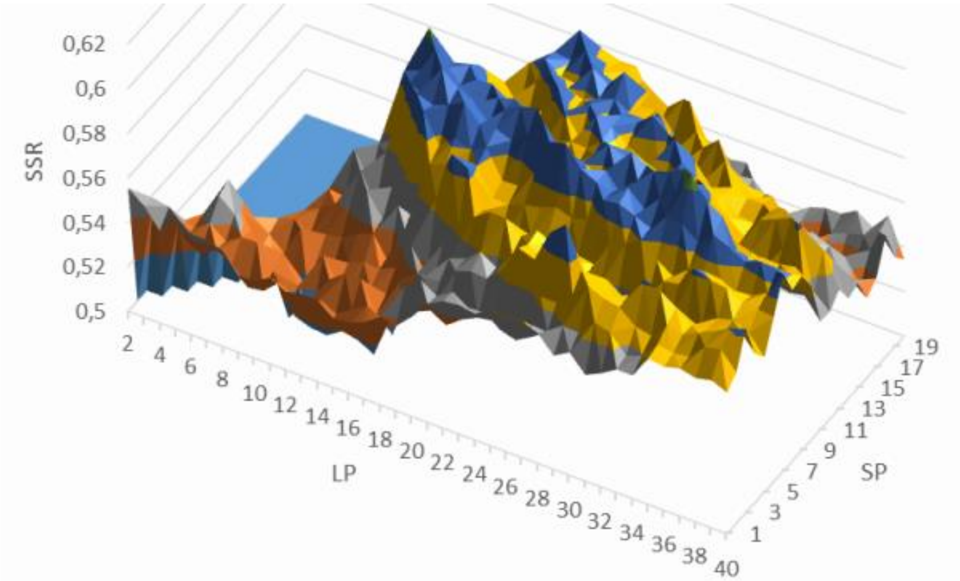


Figure 8 Sensitivity of MACD success rate on selection of signal period and holding duration

3.2 Sensitivity analysis of short and long period selection

Last but not least analysis of short and long period selection was performed for Big Lots inc. shares. 3D graph of success rates for given values of short period and long period is illustrated in Figure 8. Ridge of maximal profitability (around 0.6) is blue with two significant peaks for SP = 9, LP = 33 and SP = 13, LP = 14 respectively. Commonly applied combination of periods 12/26/9 gives resulting value of success rate equal to 0.58. All remarkable peaks augmented with profitability are summarized in following Table 2.

Short period	Long period	Success rate
9	33	0.605
13	14	0.603
10	32	0.601
12	15	0.6
12	26	0.58

Table 2 Selection of MACD parameters combination resulting in high profitability trades with shares of Big Lots inc. augmented with result of 12/26/9 combination

Other interesting result is illustrated in Figure 9, where maximal success rates for different holding periods are depicted. This confirms that there is almost no dependency between duration of holding selected share and setting of MACD parameters. Trend of profitability curve is slightly increasing, however variation around it is very high and it is capturing periodical trading effect with length of one business week (5days).

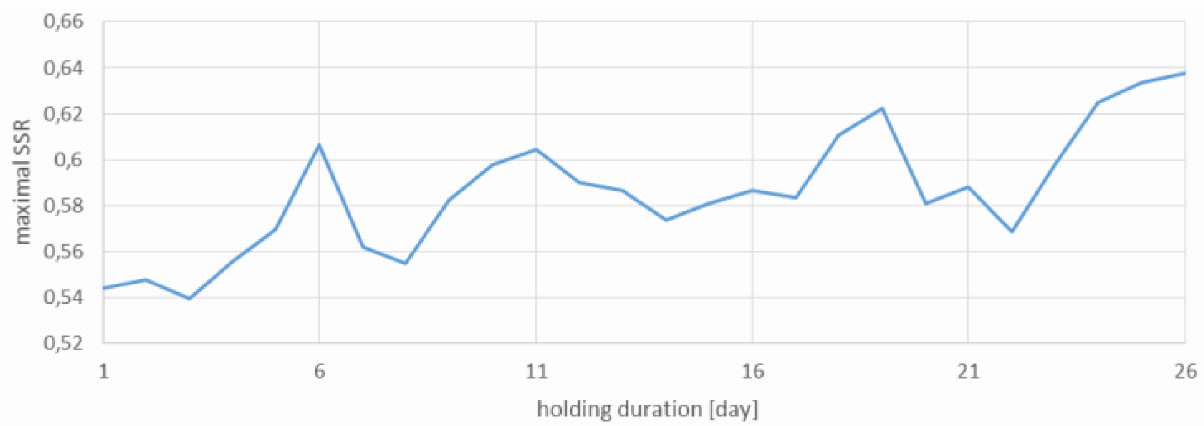


Figure 9 Maximal success rates for different holding durations of Big Lots inc. shares

4 Conclusion

The aim of presented contribution is the sensitivity analysis of MACD values on selection of short, long and signal period lengths. Case study involves three types of shares with growing, falling and changing trend respectively. To show how market speculations based on MACD indicator perform, simple success rate was calculated for each combination of input parameters and duration of holding of selected share. From results we can conclude that prolonging of signal period increases sensitivity of MACD indicator. Lengths of short and long period significantly influence trade profitability. Moreover combination 12/26/9 is not optimal for selected share and lengths of short and long period are in mutual relation. Surprisingly, there is almost no dependency between duration of holding selected share and setting of MACD parameters.

Acknowledgements

This work has been mainly supported by the University of Pardubice via SGS Project (reg. no. SG660026) and by the institutional support.

References

[1] Bender, J. C., Osler, C. L, and Simon, D.: Trading and Illusory Correlations in US Equity Markets. *Review of Finance* **2**, 17 (2012).
[2] Bodas-Sagi, D. J., Fernandez-Blanco, P., Hidalgo, J. I., and Soltero-Domingo, F. J.: A parallel evolutionary algorithm for technical market indicators optimization. *Journal of Nature Computing* **12**, 2 (2013), 195–207.
[3] Devi, M. S., and Singh, K. R.: Study on Mutual Funds Trading Strategy Using TPSO and MACD. *Journal of computer Science and Information Technologies* **5**, 1 (2014), 884–891.
[4] Friesen, G. C., Weller, P. A., and Dunham, L. M.: Price trends and patterns in technical analysis: A theoretical and empirical examination. *Journal of Banking and Finance* **6**, 33 (2009), 1089–1100.
[5] Hanousek, J., and Kopřiva, F.: Do broker/analyst conflicts matter? Detecting evidence from internet trading platforms. *Proceedings of 30th International Conference Mathematical Methods in Economics*, 2012.
[6] Lo, A. W., Mamaysky, H., and Wang, J.: Foundation of Technical Analysis: computational algorithms, statistical inference, and empirical implementation. National Bureau of Economic Research, 2000.
[7] Neely, Ch., Weller, P., and Dittmar, R.: Is Technical Analysis in the Foreign Exchange Market Profitable?: A Genetic Programming Approach. *Journal of Financial and Quantitative Analysis* **4**, 32 (1997), 405–426.
[8] Pring, M. J.: Technical analysis explained: The successful investor's guide to spotting investment trends and turning points. McGraw-Hill Professional, 2002.
[9] Rosillo, R., Fuenten, de al D., and Brugos, J. A. L.: Technical analysis and the Spanish stock exchange: testing the RSI, MACD, momentum and stochastic rules using Spanish market companies. *Journal of Applied Economics* **45**, 12 (2013), 1541–1550.
[10] Tučník, P.: Automatizované obchodní systémy – optimalizace interakce s prostředím trhu při obchodování s komoditami. In *Kognice a umelý život X* (Kelemen, J., and Kvasnicka, V., eds.). Slezská univerzita v Opavě, Opava, 2010, 381–384.

Comparison of an enterprise's financial situation assessment methods. Fuzzy approach versus discriminatory models.

Tomasz L. Nawrocki[1]

Abstract. The main aim of the paper is to compare practical application of different enterprise's financial situation assessment methods and identify their advantages and disadvantages. As subjects of this comparison have been adopted several popular discriminatory models on the one hand, and original model based on the fuzzy approach, on the other. Calculations were performed on the basis of financial data from annual reports for past ten years (2006-2015), published by selected companies listed on the Warsaw Stock Exchange and classified into different sectors of economy: services and trade, industry, finance. Conducted in the paper comparative analysis shows some advantages and disadvantages of considered approaches. For discriminatory models big advantage is the simplicity of use, but on the other hand, fuzzy approach offers far greater flexibility, which inevitably influences the greater objectivity generated in time results.

Keywords: discriminatory models, financial situation assessment, fuzzy logic, solvency.

JEL Classification: C69, G33
AMS Classification: 94D05, 26E50

1 Introduction

A feature of the market economy is operating in conditions of uncertainty and risk, which implies the need to control the corporate financial situation. Evaluation of this situation, as a part of the economic analysis, is an important tool for business management, since the information obtained due to it are the foundation for investment and financing decisions. Because of the imperfections of traditional ratio analysis (increasing the number of indicators rather than lead to a better orientation, results in information noise) for many years has been sought a synthetic measure allowing to evaluate enterprise's condition with the greatest possible accuracy.

In the literature term „financial situation" in context of enterprises has usually two meanings – narrow and wider. In the first it is associated with creditworthiness of entity and concentrate on its financial liquidity and indebtedness analysis. I the second one it is treated as a general financial condition of an enterprise including also issues of profitability, wildly understood efficiency as well as competitive advantages [3]. For purposes of this paper, the first – narrow – meaning of financial situation term was used.

Although the first quantitative assessment models of creditworthiness date back to the 1930s (P.J. Fritz Patrick), a significant development acceleration in this field occurred in 1960s, along with the wider use of statistics and econometrics (W.H. Beaver, E.I. Altman), especially discriminant analysis [4]. At the same time, however, with the increasingly common use of computers and perceiving certain drawbacks and limitations of statistical and econometrical methods [4], [5], it can be seen further develop broadly defined methods of enterprises financial condition assessment, which focus on the neural networks or fuzzy modeling use [5].

Therefore, the main goal of this article is to present a comparative analysis results of corporate financial situation assessment with the use of several popular in the literature discriminatory models and Author's original fuzzy model. Calculations were performed on the basis of financial data from annual reports for past ten years, published by selected companies listed on the Warsaw Stock Exchange and classified into different sectors of economy: services, industry and finance.

2 Research methodology

Among the many discriminatory models, for the purpose of realization the main article goal it was decided to use five of them, which had no certain industry restrictions [4] (in brackets were given the years of model creation): Mączyńska (1994), Gajdka & Stos (2006), Hołda (2001), Prusak (2005) and Hamrol – „*poznański model*" (1999-2002). The reason that probably the most popular in the world Altman's discriminatory model [1] was ignored is this, that the Altman's Z-score model has industry limitation (manufacturers). Discriminatory functions with

[1] Silesian University of Technology, Faculty of Organization and Management, Institute of Economics and Informatics, Roosevelta 26, 41-800 Zabrze, Poland; e-mail: tomasz.nawrocki@polsl.pl.

characterization of used variables are given below. In order to harmonize the limit values in considered models, with regard to the Gajdka & Stos model it was adopted subtracting from the final outcome 0.44, and in the case of the Prusak model adding 0.13. In result, for each considered model the limit value separating enterprises with bad financial situation from the others was 0.

Mączyńska: $$W_M = 1{,}5 \cdot X_1 + 0{,}08 \cdot X_2 + 10 \cdot X_3 + 5 \cdot X_4 + 0{,}3 \cdot X_5 + 0{,}1 \cdot X_6 \quad (1)$$

X_1 – (Profit before tax + Depreciation and amortization) / Reserves and liabilities
X_2 – Total assets / Reserves and liabilities
X_3 – Profit from operations (EBIT) / Total assets
X_4 – Profit from operations (EBIT) / Revenues from sales
X_5 – Inventories / Revenues from sales
X_6 – Total assets / Revenues from sales

Gajdka & Stos: $$W_{GS} = 0{,}20099 \cdot X_1 + 0{,}001303 \cdot X_2 + 0{,}760975 \cdot X_3 + 0{,}965963 \cdot X_4 - 0{,}341096 \cdot X_5 \quad (2)$$

X_1 – Revenues from sales / Total assets
X_2 – Current liabilities x 365 / Cost of goods sold, general administration, selling and distribution
X_3 – Net profit / Total assets
X_4 – Profit before tax / Revenues from sales
X_5 – Reserves and liabilities / Total assets

Hołda: $$W_H = 0{,}605 + 0{,}681 \cdot X_1 - 0{,}0196 \cdot X_2 + 0{,}00969 \cdot X_3 + 0{,}000672 \cdot X_4 + 0{,}157 \cdot X_5 \quad (3)$$

X_1 – Current assets / Current liabilities
X_2 – Reserves and liabilities / Total assets
X_3 – Net profit / Total assets
X_4 – Current liabilities x 360 / Cost of goods sold, general administration, selling and distribution
X_5 – Revenues from sales / Total assets

Prusak: $$W_P = -1{,}5685 + 6{,}5245 \cdot X_1 + 0{,}148 \cdot X_2 + 0{,}4061 \cdot X_3 + 2{,}1754 \cdot X_4 \quad (4)$$

X_1 – Profit from operations (EBIT) / Total assets
X_2 – Cost of goods sold, general administration, selling and distribution / Trade and other payables
X_3 – Current assets / Current liabilities
X_4 – Profit from operations (EBIT) / Revenues from sales

Hamrol - *„poznański model”*: $$W_{H-poz} = -2{,}368 + 3{,}562 \cdot X_1 + 1{,}588 \cdot X_2 + 4{,}288 \cdot X_3 + 6{,}719 \cdot X_4 \quad (5)$$

X_1 – Net profit / Total assets
X_2 – (Current assets – Inventories) / Current liabilities
X_3 – (Equity capital + Non-current liabilities) / Total assets
X_4 – Net profit on sales / Revenues from sales

In contrast to the approach used in mentioned above statistical and econometrical methods [9], assessment criteria selection in proposed by the Author original solution based on the fuzzy logic was carried out arbitrarily, on the basis of author's knowledge and practical experience in the field of financial analysis. Proposed model generally bases on an earlier concept [6], but for purposes of this research more efforts were made to limit number of variables and, at the same time, take into account all the dimensions of enterprises financial condition assessment, that are crucial from the viewpoint of bankruptcy risk, or creditworthiness.

Therefore, in the proposed model it was assumed that the financial situation of an enterprises can be analysed and assessed in two basic dimensions considered in the literature on financial analysis [8]: its financial liquidity and indebtedness. In the assessment of financial liquidity there were three basic dimensions included: static, income (cash flows) and structural (net working capital), and in the case of indebtedness two: debt level and ability to service the debt.

The structure of the suggested enterprise's financial situation assessment model, along with the most detailed assessment criteria within the particular modules, is presented in Figure 1.

In the proposed model firstly it is intended to obtain partial assessments within the distinguished basic assessment criteria of financial situation. These assessments will result from the ratios calculated on the basis of data from financial statements. Next aggregated assessment results may be obtained in the areas of financial liquidity in static, income and structural dimension, debt level and ability to service the debt. Furthermore, these results constitute foundations for calculating general situation measures in the areas of financial liquidity and indebtedness, so that in the final stage, on their basis, it is possible to achieve overall financial situation assessment for the analysed enterprise.

The calculation tool in the suggested solution is based on the fuzzy set theory, which is one of the approximate reasoning methods [7], [10].

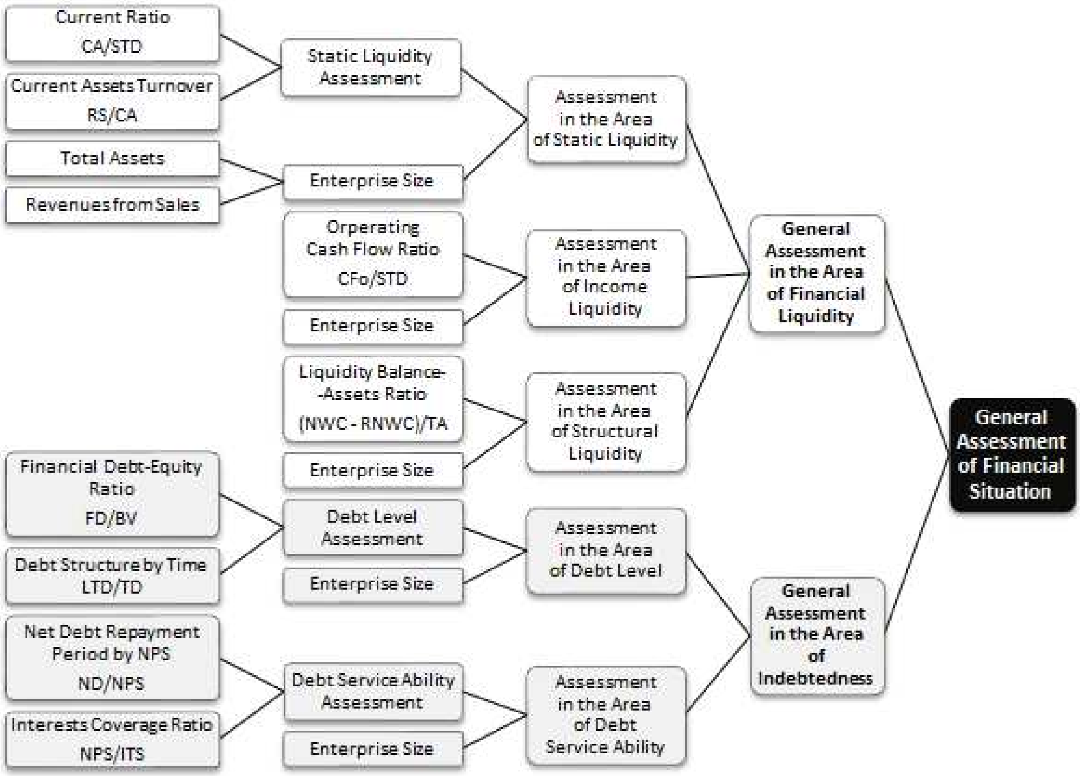


Figure 1 General structure of an enterprise's financial situation assessment model

CA – current assets, STD – short-term debt, CFo – cash flows from operations, NWC – net working capital (current assets – short-term liabilities), RNWC – request for net working capital (non-financial current assets – non-financial short-term liabilities), TA – total assets, FD – financial debt, BV – book value (equity capital), TD – total debt (liabilities and reserves), LTD – long-term debt, ND – net debt (FD – cash), NPS – net profit from sales, ITS – interests.

It was assumed, that the financial data for calculation purposes (regarding both discriminatory models and original fuzzy model) will be obtained from financial reports of analyzed companies. In all cases of variables based on balance sheet (financial situation statement) values it is also assumed to use their average values from two-year period.

Detailed assumptions of fuzzy model construction

In relation to the construction of proposed fuzzy model, based on the Mamdani approach [7], the following assumptions were made:

- for all input variables of the model, the same dictionary of linguistic values was used, and their value space was divided into three fuzzy sets named {low, medium, high};
- for output variables of the model, in order to obtain more accurate intermediate assessments, the space of linguistic values was divided into five fuzzy sets named {low, mid-low, medium, mid-high, high};
- in case of all membership functions to the particular fuzzy sets, a triangular shape was decided for them (Figure 2 and Figure 3);
- the values of fuzzy sets characteristic points (x_1, x_2, x_3) for the particular input variables of the model were determined partly basing on the literature on enterprises financial analysis and partly arbitrarily, basing on the distribution of analysed variables values and on the author years of experience in the area of companies financial situation analysis (Table 1);
- for fuzzification of input variables, the method of *simple linear interpolation* was used [2];
- fuzzy reasoning in the particular knowledge bases of the model was conducted using *PROD* operator (fuzzy implication) and *SUM* operator (final accumulation of the conclusion functions received within the particular rule bases into one output set for each base) [7];
- for defuzzification of fuzzy reasoning results within the particular rule bases simplified *Center of Sums* method was used [7].

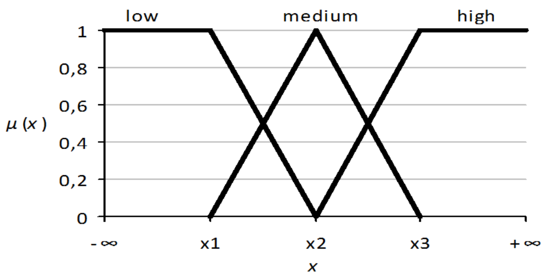


Figure 2 The general form of input variables membership function to distinguished fuzzy sets.

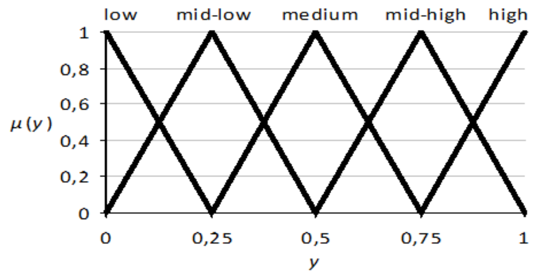


Figure 3 The output variables membership function to distinguished fuzzy sets

	x1 $\mu(x)$ **= 1/low**	**x2** $\mu(x)$ **= 1/medium**	**x3** $\mu(x)$ **= 1/high**
Current Ratio (-)	0	1	2
Current Assets Turnover (-)	0	2	4
Operating Cash Flow Ratio (-)	0	0,5	1
Liquidity Balance-Assets Ratio (-)	-0,1	0,0	0,1
Financial Debt-Equity Ratio (-)	0	0,5	1
Debt Structure by Time (-)	0	0,5	1
Net Debt Repayment Period by EBIT (years)	0	3	6
Interests Coverage Ratio (-)	0	2	4
Total Assets ('000 PLN)	0	172.000	344.000
Revenues from Sales ('000 PLN)	0	200.000	400.000

Table 1 The values of fuzzy sets characteristic points for particular input variables of the enterprise's financial situation assessment fuzzy model

Next, taking into consideration the general structure of the financial situation assessment model presented in Figure 1, author, basing on his knowledge and experience in the area of analysed issue, designed 12 rules bases in the form of „IF – THEN" (11 bases with 9 rules and 1 base with 27 rules), achieving this way a „ready to use" form of the financial situation assessment fuzzy model.

The intermediate and final assessments generated by the model take values in the range between 0 and 1, where from the viewpoint of analysed issue, values closer to 1 mean a very favourable result (better financial situation, lower risk of financial problems and bankruptcy), while values closer to 0 indicate a result less favourable (worse financial situation, higher risk of financial problems and bankruptcy). It also needs to be noted, that due to the fuzzy model assumptions, since the outputs of the fuzzy rule-based model are defuzzified, the information about the uncertainty of an assessment is to some extent lost. However, on the other hand this procedure provides an information about narrower areas of financial situation assessment.

All calculations related to the presented fuzzy model were performed in MS Excel.

3 Research results

In order to compare results generated by considered discriminatory models and proposed original fuzzy model, the financial situation assessment was conducted for 12 companies from different sectors of the economy, which shares are listed on the WSE – Orange PL (telecommunications), MW Trade (finance), Neuca (wholesale trade), Sfinks (restaurants), Polnord (real estate developer), Marvipol (real estate developer/retail trade), PBG (construction), Budimex (construction), JSW (coalmine), KGHM (copper mine), Żywiec Group (brewery), PKM Duda (meat industry). According to the adopted methodology, the basis for the financial situation assessment of mentioned above entities were data acquired from published by them in the years 2006-2016 annual reports.

Despite methodological differences, both discriminatory models and proposed original fuzzy model inform about the same category – financial situation of enterprises. This informational (and not strictly numerical) surface was adopted as the basis for comparative analysis of investigated companies. The main issue was to compare a general perception of the enterprises' financial situation through the prism of general results, obtained with the use of different methods. For this reason it was considered, that the graphical form of results presentation and their comparison will be adequate and sufficient at the same time. The results obtained during the research are presented in Figures 4-9. Analyzed companies were grouped according to similarity of business activity – service and trade, real estate developers and construction, manufacturers.

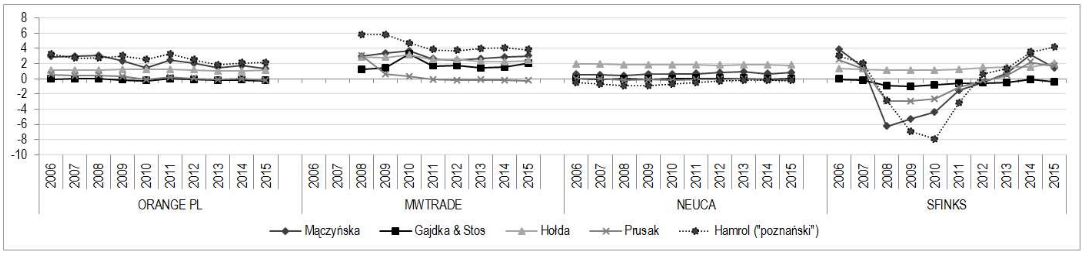


Figure 4 Financial situation of servicing and trading companies based on discriminatory models

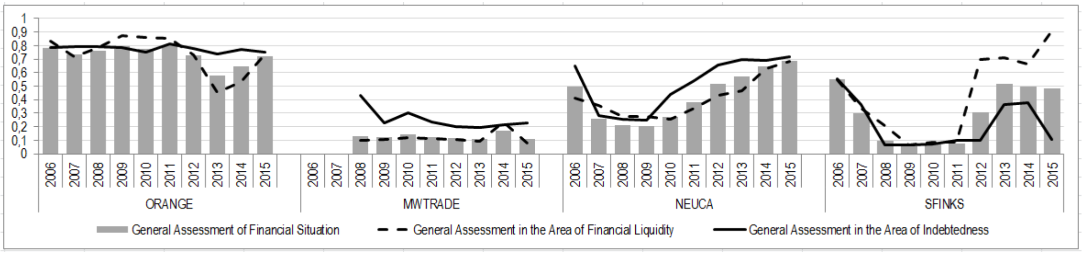


Figure 5 Financial situation of servicing and trading companies based on original fuzzy model

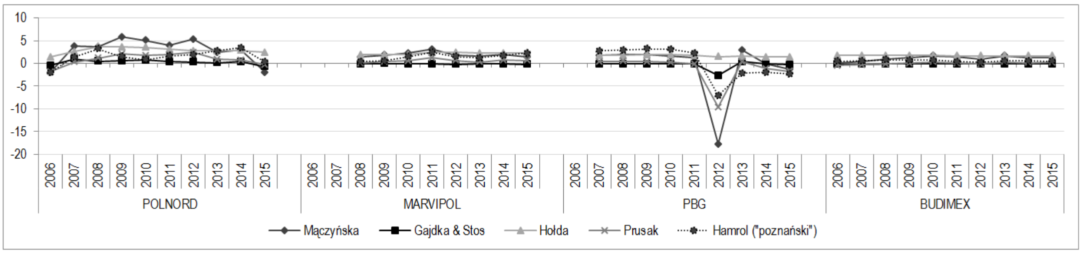


Figure 6 Financial situation of re developers and construction companies based on discriminatory models

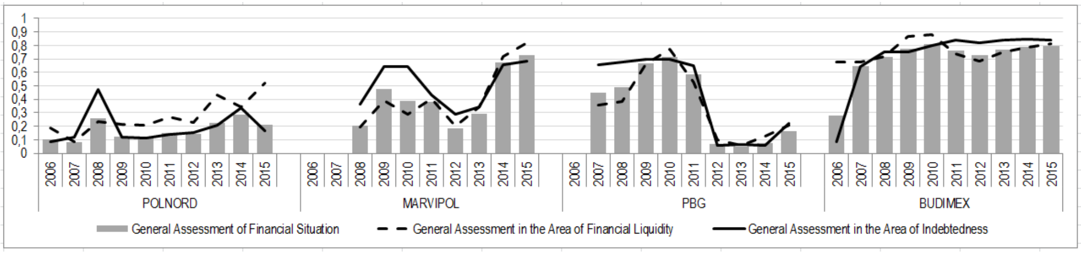


Figure 7 Financial situation of re developers and construction companies based on original fuzzy model

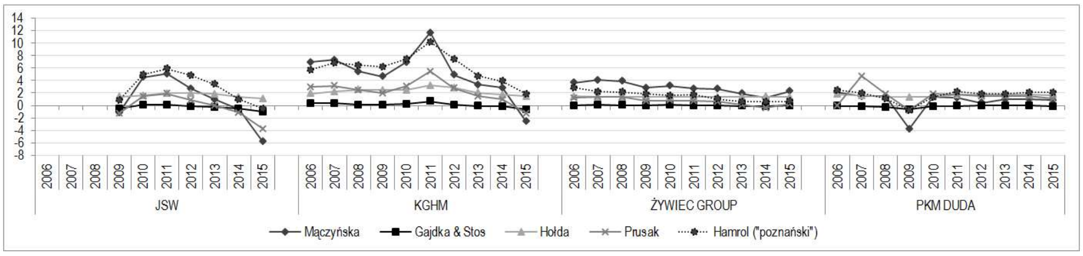


Figure 8 Financial situation of manufacturing companies based on discriminatory models

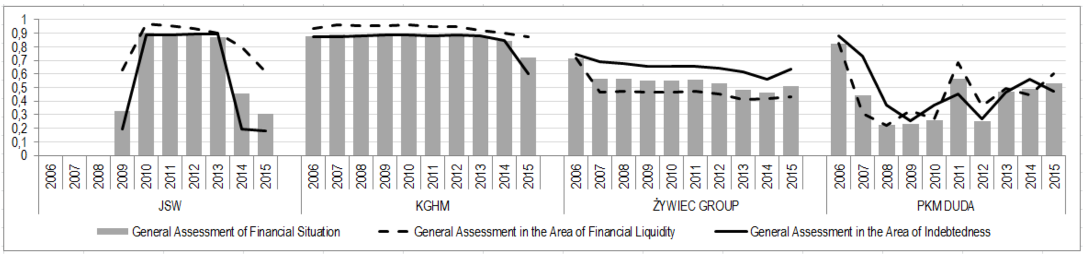


Figure 9 Financial situation of manufacturing companies based on original fuzzy model

Taking into account carried out research the attention should be paid to two issues. The first is significant differences in financial situation assessments of companies analyzed over the concerned period, which were obtained using the discriminatory models. Most often the lowest results, including signalling the possibility of bankruptcy, indicated the Gajdka & Stos model, which in most cases was not however covered in the future. On the other hand, the results obtained on the basis of the Hołda model were less susceptible to changes in time and remained at relatively high levels despite rather turbulent changes in the financial situation of investigated companies, with the interim systemic bankruptcy inclusive (Sphinx, PBG, PKM Duda).

The second issue concerns whereas the main purpose of the article, which is a comparison of investigated companies financial situation assessment results, obtained using discriminatory models and original fuzzy model. As it can be seen in Figures 4-9, in several cases, these results differ quite substantially (Orange PL, MW Trade, Neuca, Polnord, Budimex, JSW, KGHM, Żywiec Group). The reasons for this situation can be traced to excessive referencing of discriminatory models on values or indicators, which in recent years are often subject to distortion by one-off events (profit from operations, profit before tax, net profit), or specific characteristics of the business activity (current assets, current liabilities), which also affects in minus objectivity of generated assessments. On the other hand, when it comes to the directions of financial situation changes in analyzed companies, both in the case of discriminatory models and the original fuzzy model they were generally similar.

4 Conclusions

Comparative analysis of corporate financial situation assessment, that was conducted on the purpose of this paper with the use of selected discriminatory models and original fuzzy model, shows some advantages and disadvantages of each of these approaches. For discriminatory models big advantage is the simplicity of use, which boils down to a calculation of several indicators based on published financial statements and discriminant function of a specific formula. The issue of application is in turn a significant barrier to the wider use of an approach based on fuzzy logic, where there is no simple formula, but a complex fuzzy model, in which the final result is obtained usually through the use of specific software. On the other hand, an approach based on the fuzzy logic is characterized by a far greater flexibility (eg. the possibility of selection a universal from sectoral viewpoint variables or replacing the variable distorting the final results, burdened by one-off events), which inevitably influences the greater objectivity generated in time results than in the case of discriminatory models (here a considerable importance for the results quality has quantitatively-sectoral scope of the research sample, based on which the discriminant function is developed).

Although discriminatory models are, and probably will still continue to be, popular in the quick testing of enterprise's financial situation assessment, it is undoubtedly the use of fuzzy logic provides ample opportunities for the development in this research area.

Acknowledgements

This research was financed from BK 13/010/BK_16/0018.

References

[1] Altman, E.I.: An emerging market credit scoring system for corporate bonds, *Emerging Markets Review*, 6/2005, p. 311-323.

[2] Bartkiewicz, W.: *Zbiory rozmyte* [in:] *Inteligentne systemy w zarządzaniu*, Zieliński, J.S., (Ed.), PWN, Warszawa, 2000, p. 72-140.

[3] Bombiak, E.: Modele dyskryminacyjne jako metoda oceny sytuacji finansowej przedsiębiorstwa, *Zeszyty Naukowe Akademii Podlaskiej w Siedlcach*. *Seria*: *Administracja i Zarządzanie*, nr86/2010, s. 141-152.

[4] Jagiełło, R.: *Analiza dyskryminacyjna i regresja logistyczna w procesie oceny zdolności kredytowej przedsiębiorstw*, Materiały i Studia NBP, nr 286/2013. Available from: http://www.nbp.pl/publikacje/materialy_i_studia/ms286.pdf.

[5] Korol, T.: Modele prognozowania upadłości przedsiębiorstw – analiza porównawcza wyników sztucznych sieci neuronowych z tradycyjną analizą dyskryminacyjną, *Bank i Kredyt*, 6/2005, p. 10-17.

[6] Nawrocki, T.: The concept of an enterprise's financial situation assessment fuzzy model, *Proceedings of the International Conference on Mathematical Methods in Economics* 2015, p. 584-589.

[7] Piegat, A.: *Modelowanie i sterowanie rozmyte*. EXIT, Warszawa, 2003.

[8] Wędzki, D.: *Analiza wskaźnikowa sprawozdania finansowego* t.2. Wolters Kluwer, Kraków, 2009.

[9] Wiatr, M.S.: *Zarządzanie indywidualnym ryzykiem kredytowym*, SGH, Warszawa 2011.

[10] Zadeh, L.A.: Fuzzy sets, *Information and Control* 8 (1965), p. 338-353.

Prediction of Daily Traffic Accident Counts and Related Economic Damage in the Czech Republic

Dana Nejedlová[1]

Abstract. Free datasets describing weather and traffic accidents in the Czech Republic have been used for the training of neural network that would predict the number of traffic accidents and the level of economic damage for a given day. The aim of the research is to find out whether there are enough statistical dependencies in the available data so that a practically usable predictor could be trained from them. The Pearson's chi-squared test was used to select input attributes for the neural network. The selected attributes are month, day of week, temperature in two selected preceding days and in the current day, precipitation, and snow. The neural network has been trained on the daily data of the years 2009 till 2014 divided into training and development test sets. The accuracy of the network after this training on more recent days is higher than majority voting, which can motivate a future research.

Keywords: Pearson's chi-squared test, feed-forward neural network, traffic accidents.

JEL Classification: C32, C45, R41
AMS Classification: 68T, 62H

1 Introduction

Predicting traffic accidents is a challenging task because they are not planned and apart from e.g. insurance frauds everyone wishes to avoid them. Scientific studies of traffic accidents focus on identifying factors that influence their incidence. Article [6] models by logistic regression the influence of rainfall on single-vehicle crashes using the information about crash severity, roadway geometries, driver demographics, collision types, vehicle types, pavement conditions, and temporal and weather conditions from databases available in Wisconsin, United States of America. Article [8] models by linear regression traffic volume in Melbourne, Australia. It identifies weather as a significant factor influencing the traffic volume as well as daytime and nighttime periods, day of the week, month, and holidays. The traffic volume is then used to predict the count of accidents by a nonlinear model that takes into account the observation that the number of accidents is directly proportional to the traffic volume and to the severity of weather conditions, but at the same time the inclement weather deters motorists from venturing onto the road and thus reduces the traffic volume, which justifies the used nonlinear regression model. Article [2] presents autoregressive time-series model of daily crash counts for three large cities in the Netherlands. This study uses data about daily crash counts, accurate daily weather conditions (wind, temperature, sunshine, precipitation, air pressure, and visibility), and daily vehicle counts for each studied area. This article also shows that if information on daily traffic exposure is not available, the information about day of the week is a good substitute. Article [15] uses autoregressive time-series model of annual road traffic fatalities in Great Britain and the monthly car casualties within the London congestion charging (CC) zone to investigate the impact of various road safety measures and legislations (seat-belt safety law, penalty points for careless driving, driving with insurance, and seat-belt wearing for child passengers). This study uses also the data about the traffic volume in the form of annual vehicle-kilometers travelled in Great Britain. Article [1] suggests a negative binomial model as the most appropriate for predicting traffic accidents using various details about studied location of traffic and about the drivers (age and gender). Available databases allowed assigning accidents to particular segments of State Road 50 in Central Florida, United States of America. Each segment was characterized by roadway geometries, traffic volumes, and speed limits. Weather was not included in the variables that contribute to accident occurrence. Results of this research identify combinations of qualities of drivers and roads which are most likely to result in an accident, which can be used as a source for arranging preventive interventions.

This contribution identifies attributes significantly influencing traffic accident counts in the Czech Republic. It draws these attributes from free publicly available datasets about daily traffic accident counts, related economic damage, and weather. These datasets do not contain as much details as those that were available to the above-mentioned studies. It is not surprising that the precision of prediction of traffic accidents using these datasets is low but when this precision is higher than the precision of a mere guessing then it indicates the existence of

[1] Technical University of Liberec, Faculty of Economics, Department of Informatics, Studentská 1402/2, 461 17 Liberec 1, Czech Republic, e-mail: dana.nejedlova@tul.cz.

statistical dependencies that hold true in the whole studied time period. The existence of prevailing statistical dependencies can motivate future research using longer data series, additional data sources that could possibly become available, and various data models, the result of which could be a practically usable predictor of the number of accidents and the economic damage caused by these accidents for a given time period of the future. The data modeling technique used in this study is Pearson’s chi-squared test to identify the input data for the feed-forward neural network.

2 Data sources

Datasets used in this study are sources of information about traffic accidents, weather, and geomagnetic activity as some studies, e.g. [17, 18], indicate a possible effect of solar and geomagnetic activity on traffic accidents.

2.1 Traffic accidents

Dataset of traffic accidents is published by the Police of the Czech Republic on their web page [14]. The earliest record in this source is for February 10, 2009. The dataset is updated on a daily basis, and the data can be downloaded separately for each day. The data are structured by date and by 14 Czech regions, each pair of day and region having the following attributes: number of accidents, fatalities, severe injuries, minor injuries, and economic damage in thousands of Czech crowns (CZK). The accidents are classified also by their causes (speeding, not giving way, inappropriate overtaking, inappropriate driving, other cause, alcohol) in such a way that the sum of the number of accidents having a certain cause is equal or higher than the number of accidents, suggesting that some accidents have been assigned to more than one cause. There is an error in this database: on August 26, 2012 there are no data.

2.2 Weather

Weather dataset [11, 12] was acquired from the National Centers for Environmental Information. The earliest record in this source is for January 1, 1775. The dataset is updated on a daily basis with a several day delay. Data for the Czech Republic can be ordered from [13] by clicking the “Add to Cart” button, “View All Items” button, selecting the “Custom GHCN-Daily CSV” option, selecting the Date Range, clicking the “Continue” button, checking all Station Detail & Data Flag Options flags, selecting Metric Units, clicking the “Continue” button, entering email address, and clicking the “Submit Order” button. The data are structured by meteorological station and date of observation. For each of this pair the following attributes are available: precipitation in mm, snow depth in mm, and maximum and minimum temperature in degrees C. There are some errors in this dataset: records for March 18, 2007 and September 9, 2013 are missing. Some other days exhibited missing measurements in all meteorological stations. For these days the missing values have been added to the dataset from internet weather service [4] that presents measurements of a number of Czech meteorological stations for a selected day. Of all available meteorological stations the data from station Ruzyně were used.

2.3 Geomagnetic activity

Dataset of geomagnetic activity can be downloaded for a submitted time period from web page [5]. The earliest available date of observation is January 1, 2000. The data are structured by days.

3 Data preprocessing

Available time series of days have been divided into 3 parts defined by Table 1. The first two parts called Training Set and Development Test Set have been used for selecting useful attributes for the feed-forward neural network that would predict the number of accidents and related economic damage for a given day and for the estimation of parameters of this neural network. The third part called Test Set is used for the presentation of results in Section 5. This methodology of dividing the data into three sets is recommended e.g. in [7] and also in [3] where the Development Test Set is called Validation Set.

Set	From Date	To Date	Number of Days
Training	2009.02.10	2014.02.09	1826
Development Test	2014.02.10	2015.02.09	365
Test	2015.02.10	2016.04.19	435

Table 1 Division of available data

In order to obtain meaningful results of analysis the acquired data have been transformed into values as described below.

3.1 Traffic accidents

Of all available information in [14] the total number of accidents and the total economic damage per day for all Czech regions have been selected as the values that will be predicted.

The number of accidents

There is a rising tendency in this time series that demanded transforming the original values into their differences from quadratic trend of the period spanning the Training and Development Test sets. The values of these differences have been transformed into three approximately equifrequent categories defined by Table 2. Without applying the quadratic trend the count of the "Low" category would be inadequately low in Development Test and Test sets containing late days, see Table 1, which would make the prediction of the Test Set from the Training and Development Test sets impossible.

Economic damage

This time series has been processed the same way as the number of accidents, see Table 2, but before this step the original values have been deprived of outliers. The outliers have been defined as the values above 50,000 thousand CZK. The outliers have been substituted by the mean value of the joint Training and Development Test sets equal to 16,876 thousand CZK, which resulted in classifying them into the "High" category.

Time Series Quadratic Trend	**The Number of Accidents $y = x - 6.37E\text{-}06 \cdot t^2 - 0.00471 \cdot t - 196$**			**Economic Damage $y = x - 6.69E\text{-}04 \cdot t^2 + 1.5025 \cdot t - 13297$**		
Border Values	**Lower than –20**	**>= –20 and < 22**	**>= 22**	**Lower than –2000**	**>= –2000 and < 1300**	**>= 1300**
Set / Category	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
Training	602	628	596	611	611	604
Development Test	132	100	133	119	126	120
Test	141	123	171	131	139	165

Table 2 Division of daily values in the traffic accidents dataset into categories

In Table 2 the x symbol stands for the original number of accidents and the number of thousand CZK with outliers substituted by the mean value, t stands for the order of the day in the time series where the date 2009.02.10 has t equal to 1, and y is the resulting value which is assigned the category defined by border values.

3.2 Weather

From weather dataset [11, 12], where each row is an observation of a single meteorological station for a single day, all available information has been utilized. To predict traffic accidents for a given day, weather for this day had been transformed into single values of attributes listed in Section 2.2.

Precipitation and snow depth

Two columns for a binary value of either precipitation or snow have been added to the weather dataset [11, 12] with value equal to zero when the observed value of precipitation or snow in a meteorological station in this row was either zero or missing. When the observed value was above zero, its binary value was equal to one. A single value of either precipitation or snow depth for a single day has been computed as arithmetic mean of these binary values for all meteorological stations respectively for precipitation and snow depth.

Temperature

For each meteorological station a single temperature for a single day has been computed as arithmetic mean of its maximum and minimum temperature. Missing values have been ignored. Missing values of both maximum and minimum temperature resulted in a missing value of arithmetic mean. A single value of temperature for a single day has been computed as arithmetic mean of these mean values for all meteorological stations ignoring missing values. The resulting temperature has been classified into categories Low, Medium, and High using border values 5 and 14 degrees C in a similar way to the border values –20 and 22 or –2000 and 1300 in Table 2.

3.3 Geomagnetic activity

Each row of the dataset obtainable from web page [5] contains values of 8 K-indices measured in a three-hour interval in a given day characterizing geomagnetic activity of this day. Missing measurements denoted as "N" have been replaced by a preceding measurement. The underlying theory is explained in [16]. Each day has been classified into one of possible categories Quiet, Unsettled, Active, Minor storm, Major storm, and Severe storm according to rules stated in [9].

4 Pearson's chi-squared tests

Pearson's chi-squared tests [19] employ categorical values obtained in data preprocessing described in Section 3. Data involved in their computation come from merged Training and Development Test sets. The results are summarized in Table 3.

Row	Test	Degrees of Freedom	0.995 Quantile	X^2
1	Temperature and the Number of Accidents	4	14.8643	48.567
2	Temperature and the Level of Economic Damage	4	14.8643	40.686
3	Geomagnetic Activity and the Number of Accidents	6	18.5490	6.0727
4	Precipitation and the Number of Accidents	2	10.5987	0.2950
5	Precipitation and the Level of Economic Damage	2	10.5987	6.9586
6	Snow and the Number of Accidents	2	10.5987	41.0788
7	Snow and the Level of Economic Damage	2	10.5987	44.3082
8	Month and the Number of Accidents	22	42.7958	229.361
9	Month and the Level of Economic Damage	22	42.7958	130.746
10	Day of the Week and the Number of Accidents	12	28.2995	1094.896
11	Day of the Week and the Level of Economic Damage	12	28.2995	745.122
12	Week after the Transition to or from Daylight Saving Time (DST) and the Number of Accidents	4	14.8643	14.263

Table 3 Tested dependencies

Remarks to the rows of Table 3 are in Table 4.

Row	Commentary to the Statistical Tests for a Given Row in Table 3
1, 2	There is a low number of accidents and a low level of economic damage on days with low temperature and vice versa.
3	The influence of geomagnetic activity on the number of accidents is not significant. Only 4 categories of the lowest levels of geomagnetic activity listed in Section 3.3 have been observed, that is why there are 6 degrees of freedom: $(4 - 1) \cdot (3 - 1)$, 3 standing for Low, Medium, and High number of accidents.
4, 5	The influence of precipitation on and the number of accidents and on the level of economic damage is not significant. Precipitation was treated as a binary value. The first category was for days with zero precipitation and the second category was for days with the precipitation above zero. The computation of precipitation is described in Section 3.2.
6, 7	There is a low number of accidents and a low level of economic damage on days with non-zero snow depth and vice versa. Snow depth was treated the same way as precipitation in a contingency table.
8-11	Month and day of the week are strongly correlated with the number of accidents and the level of economic damage, suggesting that the primary cause of accidents and economic damage is the level of traffic.
12	Week after the transition to or from daylight saving time and the number of accidents might be somewhat correlated. Days have been classified into three categories: days of the week after the transition to the DST, days of the week after the transition from DST, and the other days. The week begins on Sunday.

Table 4 Details about tested dependencies in **Table 3**

Tested dependencies with X^2 very close to a 0.995 Quantile (i.e. 1 – 0.995 is the probability that the tested pairs of attributes are independent) have been examined more closely in a series of contingency tables differing in a shift of the attributes Temperature, Week after the Transition to or from DST, Precipitation, and Geomagnetic Activity relatively to the other attributes The Number of Accidents and The Level of Economic Damage. The result can be seen in Figure 1. The horizontal axis shows the number of days by which the attributes have been shifted. Had there been a peak around the zero shift, a significant dependency could have been detected in this way. Peaks in the negative part of the horizontal axis mean that the shifted attributes have a delayed effect on accidents and related economic damage. Peaks in the positive part of the horizontal axis mean that the categories of accidents and related economic damage have counts different from random distribution on days preceding the change in the shifted attributes.

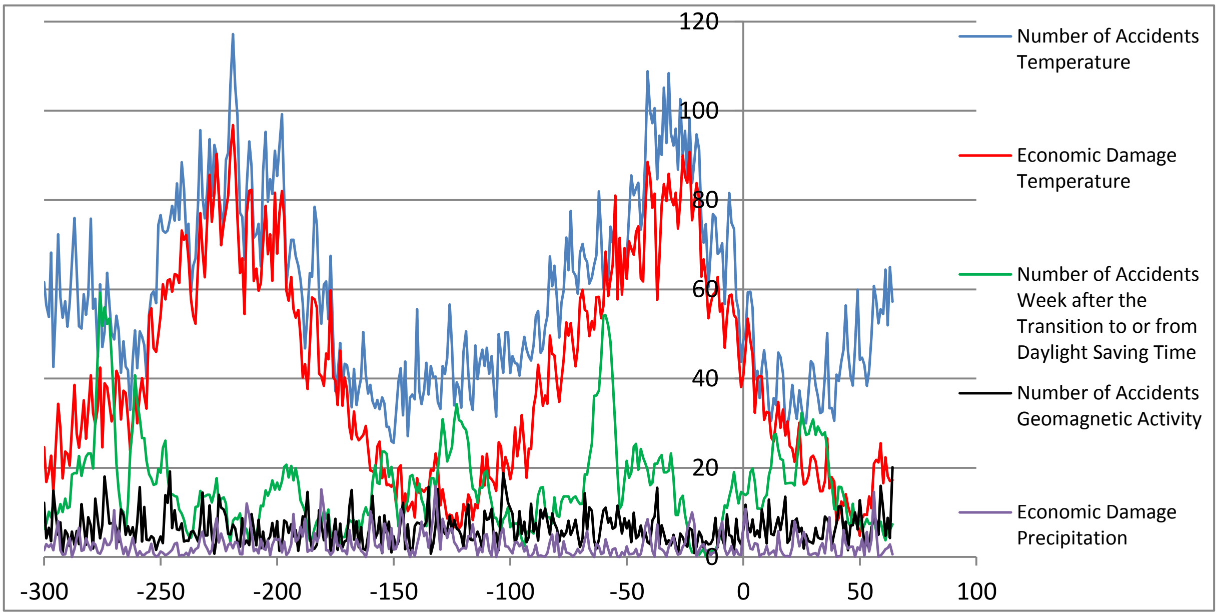


Figure 1 Pearson's chi-squared tests with a time shift in days

It can be seen from Figure 1 that Precipitation and Geomagnetic Activity do not have a significant influence at least in the used categorical representation.

The Transition to or from DST has a significant peak distant approximately a month from the zero shift in the positive part of the horizontal axis. Approximately 2 weeks after the time shift to or from DST the Number of Accidents becomes randomly distributed. Examining the related contingency table reveals that there is a relatively low number of accidents before the spring time shift to DST and a relatively high number of accidents before the autumn time shift from DST. Two biggest peaks at –60 and –276 are probably a manifestation of a yearly temperature and traffic cycle described in the next paragraph. It is probable that all significant peaks on the Transition to or from DST dataset in Figure 1 are caused by this cycle.

The data series involving Temperature in Figure 1 suggest that accidents and related economic damage have a cyclic character and that they react to the change in temperature with approximately a month's delay. This delay suggests that the primary cause of accidents and economic damage is the level of traffic. There is a low number of accidents and low level of economic damage in days with low temperature and vice versa in the contingency table for a minus-32-day shift. And there is a high number of accidents and high level of economic damage in days with low temperature and vice versa in the contingency table for a minus-219-day shift. It means that these two peaks in Figure 1 are anticorrelated. Although having correlated (or anticorrelated) attributes in the input of a predictor is not recommended, see e.g. [3], the neural network described in Section 5 has been usually more accurate when it used both of them.

5 Prediction using the feed-forward neural network

A feed-forward neural network has been reported in paper [10] dealing with traffic crashes occurred at intersections as a better solution than the Negative Binominal Regression employed also in articles [1, 2, 15].

A feed-forward neural network has been programmed in C language according to the algorithm described in [20] and used to predict separately the category of the number of accidents and the level of economic damage from month, day of week, temperature 219 days ago, temperature 32 days ago, temperature for the current day, precipitation and snow depth. All categorical attributes have been encoded in the form of the so-called *dummy variables* described in [3]. Precipitation and snow depth have been represented respectively as a single value described in Section 3.2. The network has been trained according to the Training Set and the training has been stopped when its accuracy of prediction of the Development Test Set was maximal. Neural network with final weights after this training has predicted correctly the category of the Number of Accidents in the range of 53% to

64% of days and the category of the Level of Economic Damage in the range of 43% to 50% of days in the Test Set, while guessing the most frequent category has accuracy of 39% and 38% respectively.

6 Conclusion

Artificial neural network can predict the number of accidents and the level of economic damage with accuracy slightly above majority voting. Its low accuracy is most likely caused by imprecise data. It would be possible, however, to create training data for a specific region and use the neural network trained with the similar methodology as described in this contribution for e.g. assessing the impact of local preventive police interventions.

References

[1] Abdel-Aty, M. A., and Radwan, A. S.: Modeling traffic accident occurrence and involvement. *Accident Analysis and Prevention*. **32** (2000), 633–642.

[2] Brijs, T., Karlis, D., and Wets, G.: Studying the effect of weather conditions on daily crash counts using a discrete time-series model. *Accident Analysis and Prevention*. **40** (2008), 1180–1190. DOI: 10.1016/j.aap.2008.01.001.

[3] Hastie, T., Tibshirani, R., and Friedman, J.: *The Elements of Statistical Learning*. 2nd edition. Springer, New York, 2009. [accessed 2016-03-01]. Available from: http://statweb.stanford.edu/~tibs/ElemStatLearn/.

[4] In-počasí: Archiv. [online]. (2016). [accessed 2016-04-04]. Available from WWW: http://www.in-pocasi.cz/archiv/.

[5] Institute of Geophysics of the Czech Academy of Sciences, v. v. i.: Data archive - K indices. [online]. (2016). [accessed 2016-04-04]. Available from WWW: http://www.ig.cas.cz/en/structure/observatories/geomagnetic-observatory-budkov/archive.

[6] Junga, S., Qinb, X., and Noyce, D. A.: Rainfall effect on single-vehicle crash severities using polychotomous response models. *Accident Analysis and Prevention*. **42** (2010), 213–224. DOI: 10.1016/j.aap.2009.07.020.

[7] Jurafsky, D, and Martin, J. H.: *Speech and Language Processing*. Prentice Hall, Inc., Englewood Cliffs, New Jersey 07632, 2000.

[8] Keay, K., and Simmonds, I.: The association of rainfall and other weather variables with road traffic volume in Melbourne, Australia. *Accident Analysis and Prevention*. **37** (2005), 109–124. DOI: 10.1016/j.aap.2004.07.005.

[9] Kubašta, P.: *Výpočet geomagnetické aktivity*. [online]. (2011). [accessed 2016-01-28]. Available from WWW: https://www.ig.cas.cz/kubasta/PersonalPages/zpravy/noaaVypocet.pdf.

[10] Liu, P., Chen, S.-H., and Yang, M.-D.: Study of Signalized Intersection Crashes Using Artificial Intelligence Methods. In: *Proceedings of the 7th Mexican International Conference on Artificial Intelligence* (Gelbukh, A., Morales, E. F., eds.). Springer-Verlag, Berlin Heidelberg, 2008, 987–997.

[11] Menne, M. J., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., Anthony, S., Ray, R., Vose, R. S., Gleason, B. E., and Houston, T. G.: *Global Historical Climatology Network - Daily (GHCN-Daily), Version 3*. [FIPS:EZ - Czech Republic - 2000-01-01--2016-03-31]. NOAA National Climatic Data Center. (2012), DOI: 10.7289/V5D21VHZ. [accessed 2016-04-14].

[12] Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., and Houston, T. G.: An Overview of the Global Historical Climatology Network-Daily Database. *J. Atmos. Oceanic Technol.* **29** (2012), 897-910. DOI: 10.1175/JTECH-D-11-00103.1.

[13] National Centers for Environmental Information: Daily Summaries Location Details. [online]. 2016. [accessed 2016-04-04]. Available from WWW: http://www.ncdc.noaa.gov/cdo-web/datasets/GHCND/locations/FIPS:EZ/detail.

[14] Policie České republiky (Police of the Czech Republic): *Statistiky dopravních nehod* (*Traffic Accidents Statistics*). [online]. (2015). [accessed 2016-04-14]. Available from WWW: http://aplikace.policie.cz/statistiky-dopravnich-nehod/default.aspx.

[15] Quddus, M. A.: Time series count data models: An empirical application to traffic accidents. *Accident Analysis and Prevention*. **40** (2008), 1732–1741. DOI: 10.1016/j.aap.2008.06.011.

[16] Reeve, W. D.: *Geomagnetism Tutorial*. [online]. (2010). [accessed 2016-01-28]. Available from WWW: http://www.reeve.com/Documents/SAM/GeomagnetismTutorial.pdf.

[17] Střeštík, J., and Prigancová, A.: On the possible effect of environmental factors on the occurrence of traffic accidents. *Acta geodaet., geophys. et montanist. Acad. Sci. Hung*. **21** (1986), 155-166.

[18] Verma, P. L.: Traffic accident in India in relation with solar and geomagnetic activity parameters and cosmic ray intensity (1989 to 2010). *International Journal of Physical Sciences*. **8**.10 (2013), 388-394. DOI: 10.5897/IJPS12.733.

[19] Wilcox, R. R.: *Basic Statistics*. Oxford University Press, Inc., New York, 2009.

[20] Winston, P. H.: *Artificial intelligence*. 3rd edition. Addison Wesley, Reading, MA, 1992.

Wage rigidities and labour market performance in the Czech Republic

Daniel Němec[1], Martin Macíček[2]

Abstract. The main goal of our contribution is to evaluate the impact of structural characteristics of the Czech labour market on the business cycle and to quantify the role of wage rigidities in the labour market performance. The Czech economy is represented by a small dynamic stochastic general equilibrium model of an open economy. The model economy consists of three types of agents: firms, households and central bank. Model parameters are estimated using the quarterly data of the Czech Republic for the period from the first quarter of 2000 to the last quarter of 2014. The comparison of two alternative wage setting model schemes shows that the Nash bargaining wage setting with the real wage persistence fits the data better than the Calvo type of wage rigidities modelling approach. The existence of hiring costs in the Czech economy was proved as significant and the estimated hiring costs account for 0.77% of the gross domestic product.

Keywords: wage rigidities, labour market frictions, DSGE model, hiring costs, Czech labour market

JEL classification: E32, J60
AMS classification: 91B40

1 Introduction

The aim of this paper is to evaluate the impact of structural characteristics of the Czech labour market on the business cycle and to quantify the role of wage rigidities in the labour market performance in the period of growth and prosperity on one hand and the deep fall into the worldwide recession, that initiated in 2008 in the USA by the severe problems of a mortgage sector and spread rapidly throughout the whole world, on the other hand. To achieve this goal we make use of a Dynamic Stochastic General Equilibrium (DSGE) model of the small open economy. Our paper extends the previous investigations of the efficiency and flexibility of the Czech labour market carried out by Němec [5] and Němec [6]. We incorporate two types of wage rigidities into the model in order to take into consideration the complicated process of wage setting and permit involuntary unemployment in the model. Since firms are nowadays changing their hiring strategies and are far more prone to find and train their own talents, rather then relying on a short time working relationships, the model considers hiring costs which could be considerably high.

2 The Model

We use a medium-scale small open economy (SOE) DSGE model presented in Sheen and Wang [7]. The authors build the model by further developing the works of Adolfson et al. [1] and Blanchard and Galí [2]. The model economy consists of three types of agents: firms, households and central bank. We assume that government distributes all taxes to households as a lump sum payment, so there is no need to introduce government in the model explicitly. Firms are divided into four categories. Domestic intermediate goods-producing firms employ labour and capital services from households, depending on the relative wage and rental price of capital, to produce intermediate goods. These goods are further sold to domestic final producers, who combine intermediate goods and transform them into a homogeneous goods. These

[1]Masaryk University, Faculty of Economics and Administration, Department of Economics, Lipová 41a, 602 00 Brno, Czech Republic, nemecd@econ.muni.cz
[2]Masaryk University, Faculty of Economics and Administration, Department of Economics, Lipová 41a, 602 00 Brno, Czech Republic, 369753@mail.muni.cz

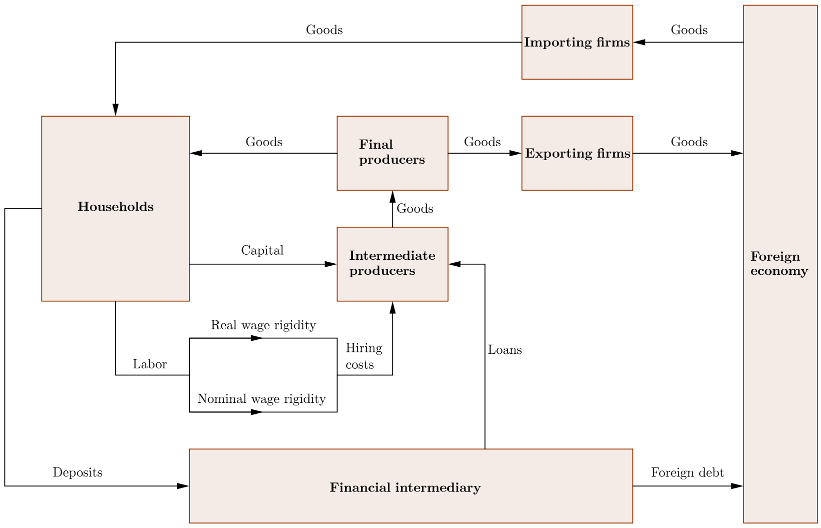


Figure 1 Simplified diagram of the model flows

goods, that can be used either for consumption and investment, are sold both to domestic households and abroad. Trading with foreign economies is ensured by the existence of importing and exporting firms. The New Keynesian framework assumes monopolistic competition, thus the firms themselves set their prices optimally subject to Calvo [3] type pricing mechanism. Since inflation is assumed to be well maintained by the central bank, firms which do not optimize their prices in current period simply index their prices to the inflation. Firms have to finance all their wage bills by borrowing at the nominal cash rate. Infinitely long lived optimizing households gain utility from leisure, real balances and consumption, that is subject to the habit formation. They offer capital and labour to intermediate goods producing firms. These employers face hiring costs when employing a new worker. These costs can take a form of initial trainings for a new employee, time until she starts to do her job properly, time of her colleagues when giving her advices, the processes of choosing the right candidates, the costs of assessment centers etc., firms internalize these costs when making their price decisions, making inflation dependent on lagged, current and future expected unemployment rate. Further two alternative schemes of wage setting are considered. Nominal wage rigidities are arising from assumption that the labour is differentiated, thus giving workers power to set wages according to the Calvo [3] type of nominal rigidities. Real wage rigidities alternative is introduced by making real wages depend on the lagged real wage and the current Nash bargaining wage. A central bank conducts monetary policy by applying a Taylor type interest rate rule. Further insight into the model functioning provides simplified model diagram displayed in Figure 1.

3 Data and methodology

To estimate the model we use Czech quarterly time series from 2000Q1 to 2014Q4. The data are taken from databases of Czech National Bank, Czech Statistical Office, Ministry of labour and Social Affairs and Eurostat. Economic and Monetary Union of the European Union is perceived as a foreign economy. More detailed description of original time series is offered in Table 1. Kalman filter is employed to evaluate the likelihood function of the model and two Metropolis-Hastings chains are run to estimate the posterior density. One milion draws are taken and 70% are discarded in order to minimize the impact of initial values. We also try to control the variance of the candidate distribution to achieve the acceptance ratio of draws around 0.3. Foreign economy is modeled independently as a VAR(1) process. Many of the model parameters are calibrated. Discount factor β and steady state–level of unemployment $\overline{U}$ are set to match the observed interest rate and unemployment rate of the data sample. Parameter ϑ is set to 1, according to Blanchard and Gali [2], implying the unit elasticity of hiring cost to the current labour market conditions. Following Jääskelä and Nimark [4], steady–state level of domestic goods inflation $\overline{\pi}^d$ is set to 1.005 and the shares of consumption and investments in imports are set to $\omega_c = 0.2$, $\omega_i = 0.5$ respectively. The curvature of the money demand function $\sigma_q = 10.62$, constant determining the level of utility the households gain from real balances $A_q = 0.38$, elasticity of labour supply $\sigma_L = 1$ and

Time series	Description	Detail
CPI	Consumer price index	Seasonally adjusted, 2014 = 100
Wnom	Nominal wages	
C	Real consumption	
I	Real investment	Gross fixed capital formation
X	Real exports	
M	Real imports	
Y	Real GDP	
Ynom	Nominal GDP	
R	Nominal interest rate	3M Interbank Rate - PRIBOR
U	Unemployment rate	
E	Nominal exchange rate	Exchange Rate CZK/EUR
Yst	Foreign real GDP	GDP EMU
CPIst	Foreign CPI	CPI EMU, 2014 = 100
Rst	Foreign nominal interest rate	3M EURIBOR
RER	Real exchange rate	= E × (CPI/CPIst)
DI	GDP deflator	=100 × (Ynom/Y)
W	Real wages	=100 × (Wnom/DI)
L	labour force	Seasonally adjusted
Lgr	labour force growth	= $\log(L_t) - \log(L_{t-1})$

Table 1 Original time series description

utilization cost parameter $\sigma_a \equiv \frac{a''(1)}{a'(1)} = 0.049$ are set according to Adolfson et al. (2007) [1]. Capital share α, separation rate δ, depreciation rate δ_k are set to match the Czech macroeconomic conditions.

Parameter	Description	Value
β	discount factor	0.99
δ	separation rate	0.1
ϑ	elasticity of hiring cost to labour market condition	1
δ_k	depreciation rate	0.02
α	capital share	0.35
ω_c	import share of consumption good	0.2
ω_i	import share of investment good	0.5
σ_q	money demand curvature parameter	10.62
σ_L	labour supply elasticity	1
A_L	labour dis-utility constant	1
A_q	money utility constant	0.38
σ_a	utilization cost parameter	0.049
$\overline{\pi}^d$	steady-state level of domestic goods inflation	1.005
$\overline{U}$	steady-state level of unemployment	0.071

Table 2 Calibrated parameters

All model calibrated parameters are summarized in Table 2. In prior definition we use the setting of Adolfson et al. [1] and Jääskelä and Nimark [4]. We set this priors as symmetrical as possible to prevent a potential distortion of the posterior estimates.

4 Estimation results and model assessment

In this paper two rival models are considered. RWHC model containing real wage rigidities and hiring costs and the alternative NWHC model that assumes nominal wage rigidities with the assumption of hiring costs. We compare the posterior log-likelihoods of the models using the Bayes factor. The RWHC model provides a better data fit compared to its rival. The logarithm of Bayes factor under the null hypothesis of nominal wage rigidities extends to 15, meaning the model with real wage rigidities is strongly favored by the data. The results of the benchmark model estimation are thoroughly depicted in Table 3 and Table 4. The posterior estimate of a habit formation parameter b is 0.78. This indicates substantially high level of households preference of a smooth consumption paths, i.e., households are willing to save and borrow money to keep their level of consumption relatively stable in time. The key parameter in labour market setup is the constant B, which determines the steady state level of hiring costs. The point estimate of this parameter is 0.19 with 90% confidence interval ranging from 0.12 to 0.26. This result indicates the significant support of hiring costs by the data. Although the hiring costs are not included in the theoretical model's output we are able to derive the hiring costs to GDP ratio using some algebraic manipulations. The resulting share is 0.77% with the 90% confidence interval between 0.48 to

1.08 percent. The estimated posterior mean of parameter B is a bit smaller under nominal wage rigidity, with the value of $B = 0.15$ and the resulting hiring cost to GDP ratio of 0.60%.

			Prior		Posterior			
Parameter	Description	Distribution	Mean	Std	Mean	Std	5%	95%
b	habit formation	Beta	0.65	0.10	0.78	0.05	0.71	0.86
B	tightness coefficient	Normal	0.12	0.05	0.19	0.04	0.12	0.26
$\tilde{S}''$	curvature of investment adjustment cost	Normal	7.694	4.00	11.31	2.68	6.83	15.57
$\tilde{\phi}_a$	risk premium	InvGamma	0.01	0.01	0.52	0.36	0.09	1.00
f	real wage AR(1) persistence	Normal	0.50	0.20	0.42	0.09	0.27	0.57
η_c	imported consumption elasticity	InvGamma	1.20	5.00	4.44	1.22	2.51	6.27
η_i	imported investment elasticity	InvGamma	1.20	5.00	1.18	0.14	1.05	1.36
η_f	export elasticity	InvGamma	1.20	5.00	6.01	1.20	4.03	7.90
$\overline{\mu}^z$	steady-state growth rate	Normal	1.0025	0.001	1.004	0.001	1.002	1.005
$\overline{\lambda}^d$	s–s. markup: domestic	InvGamma	1.20	3.00	1.91	0.26	1.47	2.32
$\overline{\lambda}^{mc}$	s–s. markup: imported-consumption	InvGamma	1.20	3.00	1.09	0.03	1.04	1.13
$\overline{\lambda}^{mi}$	s–s. markup: imported-investment	InvGamma	1.20	3.00	1.11	0.04	1.05	1.17
Calvo lottery								
ξ_d	domestic firm	Beta	0.60	0.15	0.35	0.07	0.23	0.45
ξ_{mc}	consumption import firm	Beta	0.60	0.15	0.36	0.06	0.25	0.46
ξ_{mi}	investment import firm	Beta	0.60	0.15	0.56	0.15	0.31	0.80
ξ_x	exporter	Beta	0.60	0.15	0.57	0.06	0.47	0.68
Monetary policy parameters								
ρ_R	interest rate smoothing	Beta	0.85	0.05	0.78	0.03	0.73	0.83
r_π	inflation response	Normal	1.70	0.30	2.46	0.22	2.10	2.82
r_y	output response	Normal	0.125	0.05	0.04	0.03	-0.02	0.10
r_s	real exchange rate response	Normal	0.00	0.05	0.004	0.04	-0.07	0.08
$r_{\Delta\pi}$	inflation change response	Normal	0.00	0.10	0.15	0.08	0.02	0.27
$r_{\Delta y}$	output change response	Normal	0.00	0.10	0.16	0.05	0.07	0.24
Persistence parameters								
ρ_{ζ^c}	consumption preference	Beta	0.50	0.15	0.55	0.11	0.37	0.73
ρ_{ζ^N}	labour preference	Beta	0.50	0.15	0.57	0.14	0.34	0.79
ρ_{μ^z}	permanent technology	Beta	0.50	0.15	0.66	0.08	0.55	0.79
ρ_ϵ	temporary technology	Beta	0.50	0.15	0.89	0.05	0.82	0.97
ρ_ϕ	risk premium	Beta	0.50	0.15	0.59	0.16	0.33	0.85
ρ_Γ	investment-specific technology	Beta	0.50	0.15	0.59	0.08	0.45	0.72
$\rho_{\tilde{z}^*}$	asymmetric foreign technology	Beta	0.50	0.15	0.52	0.15	0.27	0.77
ρ_{λ^d}	markup: domestic	Beta	0.50	0.15	0.83	0.09	0.70	0.95
$\rho_{\lambda^{mc}}$	markup: imported-consumption	Beta	0.50	0.15	0.88	0.05	0.81	0.95
$\rho_{\lambda^{mi}}$	markup: imported-investment	Beta	0.50	0.15	0.50	0.15	0.26	0.75
ρ_{λ^x}	markup: export	Beta	0.50	0.15	0.87	0.07	0.78	0.96

Table 3 Prior and posterior

The elasticity of investment with respect to price of an existing capital can be derived from the first order condition of households choosing the level of investment maximizing their utility function and linearizing this condition around the steady state. Solving the resulting equation forward and inserting back the original model variables we obtain the expression from which the elasticity can be easily derived. We use the equation to determine the elasticity of investment to a temporary rise in the price of current installed capital as $\frac{1}{(\overline{\mu}^z)^2\tilde{S}''}$ and the elasticity of investment to a permanent rise in the price of installed capital as $\frac{1}{(\overline{\mu}^z)^2\tilde{S}''(1-\beta)}$. The posterior estimate of an investment adjustment cost $\tilde{S}''$ is 11.31. Evaluating these expressions at our point estimates delivers the elasticities of 0.088 for a temporary rise of a price and 8.78 for a permanent rise respectively. The posterior estimate of the risk premium parameter $\tilde{\phi}_a = 0.52$ implies (using the uncovered interest rate parity condition), that a 1 percent increase in net foreign assets reduces the domestic interest rate by 0.52 percent. The point estimate of real wage persistence f reached the value 0.42 implying a significant degree of real wage rigidity in the labour market. The comparison of values of consumption η_c and investment η_i imported elasticity parameters reveals, that the demand for a consumption of a foreign goods is significantly more sensitive to the changes of relative prices than the demand for investment goods. The export elasticity is with its value $\eta_f = 6.01$ most sensitive. The point estimate of $\overline{\mu}^z$ at 1.004 suggests, that the economy is growing only slowly in the steady state, this could well be caused by the time period chosen, where the economic crisis covers the substantial part of a data range. From the estimates of steady–state markups we can derive the information, that the domestic firms have the most market power compared to other types of firms, since they can afford to set a highest markup. The estimates of the Calvo lottery parameters vary greatly across the different firm types. The lowest estimate of Calvo parameter belongs to the domestic firm. The value $\xi_d = 0.35$ implies the average price fix duration of slightly more than 1.5 quarters. The second least rigid price setting mechanism goes to consumption importing firms with average price fix duration of 1.6 quarters. The results indicate that the investment importing firms together with exporters are unable to adjust their prices most frequently with average price fix durations of 2.27 and 2.32 quarters respectively.

		Prior			Posterior			
Parameter	Description	Distribution	Mean	Std	Mean	Std	5%	95%
σ_{ζ^c}	preference shock: consumption	InvGamma	0.50	3.00	3.29	0.83	2.05	4.46
σ_{ζ^N}	preference shock: labour supply	InvGamma	0.50	3.00	2.70	0.58	1.77	3.61
σ_{μ^z}	permanent technology shock	InvGamma	0.50	3.00	0.50	0.10	0.35	0.66
σ_{ϵ}	temporary technology shock	InvGamma	0.50	3.00	0.43	0.06	0.32	0.53
σ_{ϕ}	risk premium shock	InvGamma	0.50	3.00	0.41	0.18	0.15	0.66
σ_{Γ}	investment specific technology shock	InvGamma	0.50	3.00	11.21	3.16	6.28	16.13
$\sigma_{\tilde{z}^*}$	asymmetric foreign technology shock	InvGamma	0.50	3.00	0.22	0.07	0.12	0.32
σ_R	monetary policy shock	InvGamma	0.50	3.00	0.28	0.06	0.18	0.37
σ_{λ^d}	markup shock: domestic production	InvGamma	0.50	3.00	1.22	0.38	0.70	1.71
$\sigma_{\lambda^{mc}}$	markup shock: imported-consumption	InvGamma	0.50	3.00	3.89	0.93	2.58	5.21
$\sigma_{\lambda^{mi}}$	markup shock: imported-investment	InvGamma	0.50	3.00	0.45	0.33	0.12	0.90
σ_{λ^x}	markup shock: export	InvGamma	0.50	3.00	4.96	1.03	3.43	6.41

Table 4 Standard errors of exogenous shocks

The point estimates of monetary policy parameters reveal that the central bank makes an effort to keep the interest rate path relatively smooth, $\rho_R = 0.78$, and responses more than proportionately to inflation when setting the nominal interest rate, $r_\pi = 2.46$. The parameter of response to the real exchange rate is not significant, therefore it seems that the Czech national bank is not influenced by its changes directly when targeting interest rates. The direct response to the output turned out to be insignificant too. Although the values of last two parameters are small, they are significant. Thus the central bank puts a little weight on direct responses to the inflation change and output change.

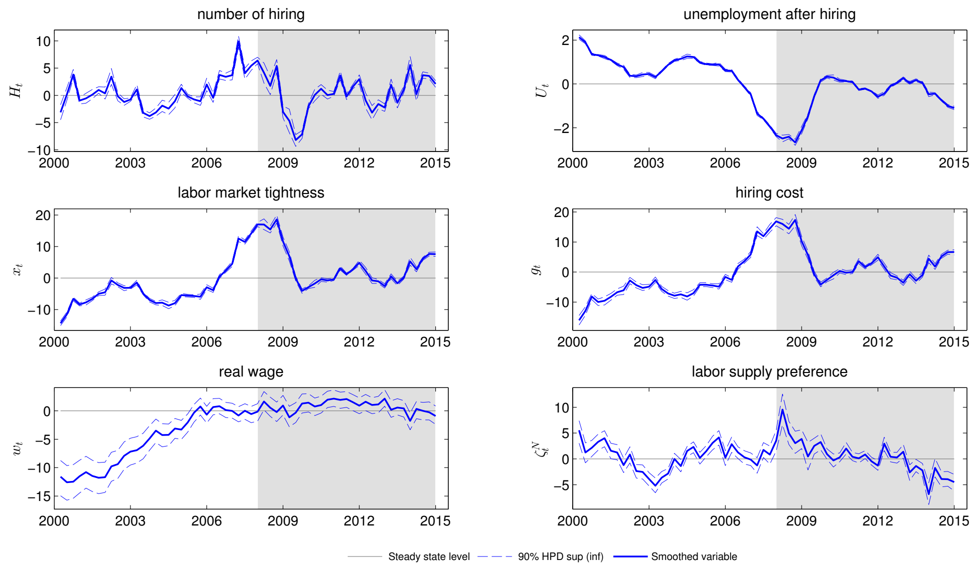


Figure 2 Smoothed paths of labour market variables

We proceed with the evaluation of the Czech labour market and its structural changes throughout the last 15 years, using our benchmark model. To analyze the development of the Czech labour market we focus on the following variables: H_t (number of hiring – number of hires is the difference between current and lagged employment, after accounting for separation), U_t (unemployment after hiring – due to the assumption of normalized labour force, we can write $U_{t-1} = 1 - N_{t-1}$, where N_t is the employment after hiring ends in time t), x_t (labour market tightness – can be interpreted as the job-finding rate from the perspective of the unemployed), g_t (stationary hiring cost), w_t (real wage – compiled as a weighted sum of a Nash bargaining wage and its own lagged value), ζ_t^N (labour supply preference – the time varying weight that households assign to the labour dis-utility relatively to the utility gained from consumption and real balances). In Figure 2 the smoothed paths of these variables, together with 90% HPD interval, are depicted. The shaded area denotes the period of the latest recession starting in 2008. This moment is a clear break point for the Czech labour market and the worldwide economy in general. The number of hiring oscillated until 2006 when the short period of growth started. In 2008 it reached its peak and then collapsed during the most severe period of financial crisis. From 2011 until now this variable gained again the stability and returned on its original path. The unemployment was declining continuously till 2008, where the unemployment rate hit the bottom. After a short period of a gradual rise the variable returned to the steady state level. In the last year the unemployment seems to be at the beginning of the another period of a decline. The labour market tightness and hiring cost are tightly bounded together. Both variables were growing steadily and hit the peak in 2008 corresponding to the lowest value of the unemployment rate. When the number of unemployed is low, the labour market is tight and therefore it is more expensive for firms to hire new workers. Following a sudden dip both labour market tightness and hiring costs fluctuated slightly and in 2014 took a growth path. The real wage experienced a long period of a remarkable growth. In 2006 the level of real wages had stabilized and since then fluctuated around the steady state level. During the last periods it seems that the real wages are slightly decreasing. The last variable depicted in the Figure 2 is the labour supply preference. We can see that from 2002 to 2005 the households were giving a relatively small weight to the dis-utility from working. In 2008 this variable shot up and returned to the steady state level in 2012. During this period the labour supply preference was very high relatively to the consumption and real balances preference, meaning agents were hesitating a lot, whether to actively participate in the working process. During the last periods the variable has been decreasing significantly.

5 Conclusion

The comparison of two alternative wage setting model schemes revealed that the Nash bargaining wage setting with the real wage persistence, reflecting the cautious approach of agents, fits the data better than the Calvo type of wage rigidities modeling approach. During the conjuncture preceding the crisis, number of hiring were exceptionally high, corresponding with the period of the lowest unemployment throughout the last fifteen years. This boom was accompanied with the increased labour market tightness which further resulted in costly hiring for firms hiring new employees. The outbreak of a sharp economic slump that followed decreased the new hires rapidly, the unemployment shot up and households started to consider their labour supply preferences more seriously. The latest improvement in the unemployment rate, tendencies in the labour market related variables and the optimistic economic atmosphere indicate the hints of hope to the Czech labour market and the economy in general.

Acknowledgements

This work was supported by funding of specific research at Faculty of Economics and Administration, project MUNI/A/1040/2015. This support is gratefully acknowledged.

References

[1] Adolfson, M., Lasén, S., Lindé, J., and and Villant, M.: Bayesian estimation of an open economy DSGE model with incomplete pass-through. *Journal of International Economics* **72** (2007), 481–511.

[2] Blanchard, O., and Galí, J.: labour Markets and Monetary Policy: A New Keynesian Model with Unemployment. *American Economic Journal: Macroeconomics* **2** (2010), 1–30.

[3] Calvo, G. A.: Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* **12** (1983), 383–398.

[4] Jääskelä, J. P., and Nimark, K.: A Medium-Scale New Keynesian Open Economy Model of Australia. *Economic Record* **87** (2011), 11–36.

[5] Němec, D.: Investigating Differences Between the Czech and Slovak Labour Market Using a Small DSGE Model with Search and Matching Frictions. *The Czech Economic Review* **7** (2013), 21–41.

[6] Němec, D.: Measuring Inefficiency of the Czech Labour Market. *Review of Economic Perspectives - Národohospodářský obzor* **15** (2015), 197–220.

[7] Sheen, J., and Wang, B. Z.: An Estimated Small Open Economy Model with Labour Market Frictions. *SSRN Electronic Journal* (2012).

Parametric rules for stochastic comparisons

Sergio Ortobelli[1], Tommaso Lando[2], Carla Nardelli[3]

Abstract. In the financial literature, the second-degree stochastic dominance (2-SD) and the increasing and convex order (icx) are frequently used to rank distributions in terms of expectation and risk aversion. These dominance relations may be applied to the empirical distribution functions, obtained by the historical observations of the financial variables, or to parametric distributions, used to model such variables. In this regard, it is well documented that empirical distributions of financial returns are often skewed and exhibit fat tails. For this reason, we focus on the stable Paretian distribution, that has been shown to be especially suitable for modeling financial data. We analyze these two different approaches, namely parametric and non-parametric, and compare them by performing an empirical study. The analyzed dataset consists of the returns of a set of financial assets among the components of the S&P500.

Keywords: stochastic dominance, heavy tails, risk aversion, stable distribution.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Stochastic dominance relations are aimed at ranking probability distributions according to an order of preference. Their usefulness in many fields of study, such as economics and finance is well known. In this regard, in a financial framework, investors are generally interested in ranking distributions of financial random variables, such as asset returns, according to their gain/loss expectations and their attitude towards risk. Several dominance rules have been proposed in the financial literature that conform to different risk profiles. These rules may be applied to the empirical distribution functions or, sometimes, to parametric models used for fitting financial distributions. In the latter case, dominance rules can be re-formulated as a comparison between parametric values, that is clearly an advantage in terms of computation as well as ease of interpretation.

With regard to parametric models, in the financial literature it is well known that asset returns are not normally distributed. Several studies by Mandelbrot (see e.g. [3,4]) and Fama (see e.g. [1]) recognized an excess of kurtosis and skewness in the empirical distributions of financial assets, that lead to the rejection of the assumption of normality. It has been widely documented that the distribution of financial assets is generally heavy tailed, so that the main alternative models that have been proposed in the literature, among which we may cite the Student's t distribution, the generalized hyperbolic distribution [8] and the stable Paretian distribution [9] do not necessarily have finite variance. In particular, the stable Paretian distribution is mainly justified by a generalized version of the central limit theorem, ascribable to Gnedenko and Kolmogorov [2], which basically states that the sum of a number of i.i.d. random variables with heavy tailed distributions (the classical assumption of finite variance is relaxed) can be approximated by a stable Paretian model.

This paper is concerned with the issue of analyzing the stochastic dominance rules that are the most suitable for dealing with heavy tailed distributions, such as the stable Paretian one. Generally, one of the concepts most frequently employed for comparison of financial random variables is the second-degree stochastic dominance (2-SD) which conforms to the idea that distributions with higher expectation and an inferior degree of dispersion (to be understood as the risk) should be preferred. Complementarily to the 2-SD, we may also recall the increasing and convex order (icx), which conforms to the idea that distributions with higher expectation and higher degree of dispersion should be preferred. However, it may often happen that the 2-SD is not verified, so that some finer ranking criteria need to be introduced. We may cite the i-th-degree stochastic dominance (i-SD) with i>2, such as the 3-SD, that obeys to the principle of downside risk aversion, although i-SD requires finite moment of order i-1, in order to be defined. Hence, it is sometimes more suitable to deal with stochastic dominance relations that are based on i-th degree integration of the quantile distribution, that is, the i-th degree inverse stochastic dominance

[1] University of Bergamo, department SAEMQ, via dei Caniana 2 Bergamo (Italy);
and VSB-TU Ostrava, department of Finance, Sokolskà trida 33 Ostrava (Czech republic), sergio.ortobelli@unibg.it.
[2] University of Bergamo, department SAEMQ, via dei Caniana 2 Bergamo (Italy);
and VSB-TU Ostrava, department of Finance, Sokolskà trida 33 Ostrava (Czech republic), tommaso.lando@unibg.it
[3] University of Bergamo, department SAEMQ, via dei Caniana 2 Bergamo (Italy), carla.nardelli@unibg.it

(i-ISD Muliere and Scarsini 1989). Both i-SD and i-ISD attach more weighting to the low tail of the distribution, emphasizing the degree of risk aversion towards the worst case scenario. Differently, some alternative approaches have been recently proposed in order to emphasize both tails of the distribution, such as the moment dispersion order [5], that is especially suitable for dealing with stable Paretian distributions.

Recently, Ortobelli et al. [5] studied the parametric relations that allow one to verify the 2-SD and the icx order, under the assumption that the random variables are stable distributed. In an empirical context, these parametric rules can serve as an approximation to 2-SD or to icx. In this regard, we argue that it would be useful to understand if these approximations are close to what observed, that is, 2-SD or icx between empirical distributions.

In section 2, we study the different approaches (non-parametric and parametric) from a theoretical point of view and propose some parametric ranking criteria, based on the results in [5], namely the asymptotic 2-SD and the asymptotic icx. These ordering rules serve as proxy to the 2-SD and the icx, respectively. Then, in section 3 we perform an empirical analysis of a set of financial assets (among the components of the S&P500) with the aim of identifying the couples of distributions that cannot be ranked according to the different orderings. In particular, we compare the results obtained with the parametric and the non-parametric approaches.

2 Methods

Let F_X be the distribution function of the random variable X. Stochastic dominance relations generally determine preorders in the set of random variables, or, equivalently, distribution functions. We recall that a preorder is a binary relation $\leq$ over a set S that is reflexive and transitive. In particular, observe that a preorder $\leq$ does not generally satisfy the antisymmetry property (that is, $a \leq b$ and $b \leq a$ does not necessarily imply $a = b$) and it is generally not total (that is, each pair a, b in S is not necessarily related by $\leq$).
The main stochastic dominance rules can be defined as follows.

Definition 1. *We say that X first-degree dominates (1-SD) Y, $X \geq_{SD} Y$, iff*

$$F_X(t) \leq F_Y(t), \forall t \in \mathbb{R}$$

or, equivalently, $E(g(X)) \geq E(g(Y))$ for every increasing function g such that the expectations exist.

The 1-SD relation is quite a strong condition, often referred to as the concept of one random variable being "stochastically larger" than another. The 1-SD implies the following two weaker criteria, that conform to the idea that distribution with higher dispersion and inferior (or greater) dispersion, or "risk", should be preferable for any investor who is risk averse (or risk lover, respectively).

Definition 2. *We say that X second-degree dominates (2-SD) Y, $X \geq_{SD}^2 Y$, iff*

$$\int_{-\infty}^{t} F_X(z)dz \leq \int_{-\infty}^{t} F_Y(z)dz\,, \forall t \in \mathbb{R}$$

or, equivalently, $E(g(X)) \geq E\big(g(Y)\big)$ for every increasing and concave function g such that the expectations exist.

Definition 3. *We say that X increasing-convex dominates (icx) Y, $X \geq_{icx} Y$, iff*

$$\int_{t}^{\infty} \big(1 - F_X(z)\big)dz \geq \int_{t}^{\infty} \big(1 - F_Y(z)\big)dz\,, \forall t \in \mathbb{R}$$

or, equivalently, $E(g(X)) \geq E\big(g(Y)\big)$ for every increasing and convex function g such that the expectations exist.

When 2-SD and/or icx cannot be verified, some weaker orders need to be introduced. In particular, we shall need the following order of dispersion with respect to a center, that can also be interpreted as an order of tailweightness.

Definition 4. *Let X and Y be random variables such that $X \in L^p$, i.e. $E(|X|^p) < \infty$ for $p \geq 1$. We say that X dominates Y with respect to the central moment dispersion (cmd) order and write $X \geq_{cmd} Y$ if and only if $\varphi_{X-E(X)}(p) \leq \varphi_{Y-E(Y)}(p)$, $\forall p \geq 1$, where $\varphi_Z(p) = \operatorname{sign}(p)E(|Z|^p)$.*

Stable distributions depend on location and scale parameters, moreover they are identified by a parameter which specifies the shape of the distribution in terms of skewness (to be intended as the disproportion between the left and the right tails) and, more importantly, a parameter that describes the asymptotic behavior of the tails. In this paper we stress the critical role of the tail parameter (also known as the stability index).

Let X be a random variable in the domain of attraction of a stable law with finite mean $E(X) = \mu$ and infinite variance. Let $\{X_i\}_{i\in\mathbb{N}}$ be independent and identically distributed observations of X. Thus, we know there exist a sequence of positive real values $\{d_i\}_{i\in\mathbb{N}}$ and a sequence of real values $\{a_i\}_{i\in\mathbb{N}}$, such that, as $n\to+\infty$:

$$\frac{1}{d_n}\sum_{i=1}^{n} X_i + a_n \xrightarrow{d} X'$$

where $X' \sim S_\alpha(\sigma, \beta, \mu)$ is an α-stable Paretian random variable, where $0 < \alpha \leq 2$ is the so-called stability index, which specifies the asymptotic behavior of the tails, $\sigma > 0$ is the dispersion parameter, $\beta \in [-1,1]$ is the skewness parameter and $\mu \in \mathbb{R}$ is the location parameter. Observe that, in this paper, we consider the parameterization for stable distributions proposed by Samorodnitsky and Taqqu [10]. We recall that, if < 2 , then $E(|X|^p) < \infty$ for any $p < \alpha$ and $E(|X|^p) = \infty$ for any $p \geq \alpha$. Therefore, stable distributions do not generally have finite variance, this happens only when $\alpha = 2$ (i.e. the Gaussian distribution, $E(|X|^p) < \infty$ for any p). In this paper, we shall focus on the case $\alpha > 1$.

Unfortunately, except in few cases, we do not have a closed form expression for the density (and thereby the distribution function) of stable Paretian distribution, which is identified by its characteristic function. Therefore, it is not possible to verify 1-SD, 2-SD and icx by applying directly definitions 1,2,3. However, in a recent paper, Ortobelli et al. [5] have determined the parametric conditions under which it is possible to verify the 2-SD and the icx, under stable Paretian assumptions. We briefly summarize them in what follows.

Theorem 1. *Let $X_1 \sim S_{\alpha_1}(\sigma_1, \beta_1, \mu_1)$ and $X_2 \sim S_{\alpha_2}(\sigma_2, \beta_2, \mu_2)$.*

1. *If $\alpha_1 > \alpha_2 > 1$, $\beta_1 = \beta_2$, $\sigma_1 \leq \sigma_2$ and $\mu_1 \geq \mu_2$, then $X_1 \geq_{SD}^2 X_2$. If $\mu_2 \geq \mu_1$ then $X_2 \geq_{icx} X_1$.*
2. *If $\alpha_1 = \alpha_2 > 1$, $\sigma_1 = \sigma_2$, $\mu_1 = \mu_2$ and $|\beta_1| < |\beta_2|$, then $X_1 \geq_{cmd} X_2$.*

Theorem 1 states that i) the stability index is crucial for establishing a dominance (2-SD or icx), whilst ii) the skewness parameter may yield the cmd order, that is, a weaker order of dispersion (or risk) around the mean value. Put otherwise, an inferior degree of skewness generally corresponds to a less spread out distribution. Therefore, we argue that a risk averse investor would generally prefer $X_1$to X_2 if $\alpha_1 \geq \alpha_2 > 1$, $|\beta_1| \leq |\beta_2|$, $\sigma_1 \leq \sigma_2$ and $\mu_1 \geq \mu_2$, whilst a risk lover would generally prefer X_2 to X_1 if $\alpha_1 \geq \alpha_2 > 1$, $|\beta_1| \leq |\beta_2|$, $\sigma_1 \leq \sigma_2$ and $\mu_2 \geq \mu_1$. This can be formalized by the following two definitions.

Definition 5. *Let $X_1 \sim S_{\alpha_1}(\sigma_1, \beta_1, \mu_1)$ and $X_2 \sim S_{\alpha_2}(\sigma_2, \beta_2, \mu_2)$ stable distributed random variables. We say that X_1 dominates X_2 with respect to the second degree asymptotic dominance (2-ASD) and write $X_1 \geq_{ASD}^2 X_2$ iff $\alpha_1 \geq \alpha_2 > 1$, $|\beta_1| \leq |\beta_2|$, $\sigma_1 \leq \sigma_2$ and $\mu_1 \geq \mu_2$, with at least one strict inequality.*

Definition 6. *Let $X_1 \sim S_{\alpha_1}(\sigma_1, \beta_1, \mu_1)$ and $X_2 \sim S_{\alpha_2}(\sigma_2, \beta_2, \mu_2)$ stable distributed random variables. We say that X_2 dominates X_1 with respect to the asymptotic increasing and convex order (A-icx) and write $X_2 \geq_{A-icx} X_1$ iff $\alpha_1 \geq \alpha_2 > 1$, $|\beta_1| \leq |\beta_2|$, $\sigma_1 \leq \sigma_2$ and $\mu_2 \geq \mu_1$, with at least one strict inequality.*

Note that the value of the location parameter plays a critical role in both definitions 5 and 6.
We expect that 2-SD and 2-ASD, as well as icx and A-icx, tend to be equivalent when the stable distribution is well fitting to the empirical distribution (see the recent studies of [5] and [6]). The following empirical analysis is aimed at the verification and the analysis of the dominance rules represented by definitions 2,3,5,6.

3 Empirical analysis

The results of section 3 have several applications in different areas of study, because of the fundamental role of the stable distribution, discussed in the introduction. It is well known that the stable Paretian model is especially suitable for approximating the empirical distribution of the returns of financial assets. In this section, we apply the stochastic dominance rules stated in the above to real financial data, in order to verify empirically the conformity between the parametric and the non-parametric approaches.

The dataset consists of the components of S&P500, starting from 7/1/1998 until 7/4/2016. We assume that the number of trading days per year is 250 and, accordingly, we consider two different timeframes for collecting data: 125 days (6 months) or 500 days (2 years). Moreover, we update (recalibrate) the datasets within a moving window of historical observations (i.e. 125 or 500 days) every 20 days (monthly) or 125 days (6 months), thus we obtain 2×2 different sequences of empirical distribution functions.

We are mainly concerned with the differences between the parametric and non-parametric rules. The non parametric approach consists in applying definitions 2 and 3 (2-SD and icx) to the empirical distribution function.

This is done for each couple of distributions, that is, for each couple of assets, within a given timeframe. If we assume that the S&P500 at time t consists of $n(t)$ assets (generally 500), we approximately have $n(t)\times(n(t)\text{-}1)$ couples of distributions (yielded by 125 or 500 observations, according to the different moving window) at each recalibration time.

Differently, the parametric approach is performed as follows. For each empirical distribution, we estimate the unknown parameters $(\alpha, \sigma, \beta, \mu)$ of the stable Paretian distribution with the maximum likelihood (ML) method (see [7]), then we apply definitions 5 and 6 in order to verify whether the conditions for 2-ASD or A-icx are satisfied, for each couple of distributions.

In our analysis, we especially focus on the number (of percentage) of assets that are not dominated by any of the other assets. Non-dominated distributions may consist of one distribution that dominates all the others and especially a set of distributions among which is not possible to establish a ranking. The results are summarized in Table 1.

	6 months		*2 years*	
criteria	*rec 20*	*rec 125*	*rec 20*	*rec 125*
2-SD	0.058	0.059	0.074	0.078
2-ASD	0.127	0.117	0.151	0.142
icx	0.018	0.017	0.030	0.031
A-icx	0.079	0.078	0.090	0.087

Table 1 Percentages of non-dominated distributions according to different timeframes and different recalibration times

Surprisingly, the percentage of couples that cannot be ranked, according to any preorder, is always very low, in that the ranking between the assets is "almost" complete in some cases. In particular, icx and A-icx obtain the inferior percentages of non-dominated couples, thus risk seekers could almost reach complete unambiguous rankings. We observe that, by increasing the timeframe of historical observations, the percentage of non-dominated couples generally decrease (slightly). With regard to the 2-SD and the icx, this means that, by increasing the number of observations, the number of cases when the curves $\int_{-\infty}^{t} F_X(z)dz$ or $\int_{t}^{\infty}(1 - F_X(z))dz$ intersect, generally increase, as the shape of the empirical distribution may change with more freedom, in that it exhibits more flexibility. Similarly, with regard to the parametric approach, we argue that, by increasing the number of observations, the probability of obtaining couples of parametric values that do not comply to the criteria of definitions 5,6 generally increase. Moreover, it should be stressed, that the parametric rules (def.'s 5 and 6) represent a good indicator of the 2-SD and the icx, although they generally overestimate the percentage of non-dominated couples. In fact, this conversely means that the set of ranked couples is generally larger, so that if we verify the dominance relation of def. 5 (or 6) w.r.t. the estimated parameters, it is likely that conditions of def. 2 (or 3) are verified as well. Put otherwise, the set of ranked couples, according to def.'s 5 and 6, is approximately contained in the set of ranked couples according to def.'s 2 and 3.

4 Conclusion

We have examined the impact of different types of preoders on a real dataset. The aim was to stress the differences and the links between the parametric and non-parametric approaches by computing and investigating the number (percentages) of couples of distributions (i.e. empirical or fitted distributions) that can (or cannot) be dominated by others. The parametric approach may be useful for approximating the number of non-dominated empirical distributions. Moreover, we stress that the main advantage of the parametric method is its computational speed, compared to the non-parametric approach, that require the verification of an integral condition on the whole support.

Acknowledgements

This paper has been supported by Italian funds ex MURST 60% 2013 and 2014 and MIUR PRIN MISURA Project, 2013-2015. The research was also supported through the Czech Science Foundation (GACR) under projects 13-13142S and 15-23699S. Moreover, the research was done through SP2014/16, and SGS research project of VSB-TU Ostrava, and furthermore by the European Social Fund in the framework of CZ.1.07/2.3.00/20.0296.

References

[1] Fama, E.: Mandelbrot and the stable paretian hypothesis. *Journal of Business*, 36 (1963), 420-429.

[2] Gnedenko, B.V., Kolmogorov A.N.*: Limit distributions for sums of independent random variables,* Addison-Wesley Mathematics Series, Addison-Wesley, Cambridge, 1954.

[3] Mandelbrot, B.: New methods in statistical economics. *Journal of Political Economy* 71 (1963), 421-440.

[4] Mandelbrot, B.: The variation of some other speculative prices. *Journal of Business* 40 (1967), 393-413.

[5] Ortobelli, S., Lando, T., Petronio, F., Tichy, T.: Asymptotic stochastic dominance rules for sums of iid random variables. *Journal of computational and applied mathematics* 300 (2016), 432-448.

[6] Ortobelli, S., Lando, T., Petronio, F., Tichy, T.: Asymptotic multivariate dominance: a financial application. *Methodology and computation in applied probability (*2016), 1-19. DOI 10.1007/s11009-016-9502-y.

[7] Nolan, J. P.: Maximum likelihood estimation and diagnostics for stable distributions. In *Lévy processes* (2001), 379-400, Birkhäuser Boston.

[8] Platen, E., Rendek, R.: Empirical evidence on student-t log-returns of diversified world stock indices. *Journal of statistical theory and practice* 2(2) (2008), 233-251.

[9] Rachev, S.T., Mittnik, S.: *Stable paretian models in finance*, Wiley, New York, 2000.

[10] Samorodnitsky, G., Taqqu M.,S.: *Stable Non Gaussian Random Processes*, Chapman & Hall, 1994.

Vehicle Scheduling with Roundabout Routes to Depot

Stanislav Palúch[1], Tomáš Majer[2]

Abstract. Minimum fleet size vehicle scheduling problem is to arrange a given set of trips into the minimum number of running boards. This problem can be modeled as an assignment problem or maximum flow problem both with polynomial complexity. Slovak and Czech bus providers often require that vehicles visit depot during working day in order to deliver money or to make a vehicle maintenance or refueling. Just mentioned procedure is called a roundabout route to depot. This paper studies ways how to manage that every running board contains a feasible roundabout route to depot.

The proposed procedure is like that: First calculate bus schedule with the minimum number of vehicles. Then create a large set of fictive dummy trips starting and finishing in depot with time duration equal to the required time length of stay in depot. Afterwards calculate bus schedule with the minimum number of vehicles that contains exactly so many fictive trips as the number of vehicles.

A maximum group matching formulation of this problem is proposed and some computer experiments with Gurobi solver for corresponding MILP model are discussed.

Keywords: vehicle scheduling, mixed linear programming, group matching.

JEL classification: C44
AMS classification: 90C15

1 Introduction

A bus trip is a quadruple (d_k, a_k, u_k, v_k) where d_k is departure time, a_k is arrival time, u_k is departure bus stop and v_k is arrival bus stop of the trip k.

Let $\mathbf{M} = \big\{m(u,v)\big\}$ be a time distance matrix determining the travel time $m(u,v)$ from bus stop u to bus stop v.

Trip j can be carried immediately after trip i by the same bus if

$$d_j \geq a_i + m(v_i, u_j), \tag{1}$$

i.e. if a bus after arriving to the arrival bus stop v_i of trip i can pull to the departure bus stop u_j of the trip j sufficiently early. In this case we will say, that the trip j can be linked after trip i and we will write $i \prec j$.

A running board of a bus is sequence of trips $i_1, i_2, \dots, i_r$ such that for very k, $1 \leq k < r$ it holds $i_k \prec i_{k+1}$. By other words, a running board is a sequence of trips which can be carried by the same bus in one day. It represents a day schedule of work for one bus.

The goal of bus scheduling with minimum number of buses is the following: Given the set $\mathcal{S}$ of trips to arrange all trips from $\mathcal{S}$ into minimum running boards. Resulting set of running boards is called a bus schedule.

Relation $\prec$ on the set $\mathcal{S}$ can be modeled by a digraph $G = (\mathcal{S}, E)$ where $E = \{(i,j) | i \in \mathcal{S}, j \in \mathcal{S}, i \prec j\}$.

[1]University of Žilina, Department of mathematical methods and operations analysis, Univerzitná 8215/1, 010 26 Žilina, e-mail: stanislav.paluch@fri.uniza.sk.

[2]University of Žilina, Department of mathematical methods and operations analysis, Univerzitná 8215/1, 010 26 Žilina, e-mail: tomas.majer@fri.uniza.sk.

Suppose that $\mathcal{S} = 1, 2, \ldots, n$. Denote by x_{ij} a decision binary variable with the following meaning

$$x_{ij} = \begin{cases} 1 & \text{if the trip } j \text{ is linked immediately after trip } i \text{ in a running board} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Imagine that we have a bus schedule with $n = |\mathcal{S}|$ buses – every bus makes only one trip. This situation corresponds to $x_{ij} = 0$ for all $(i, j) \in E$. When we realized one linkage for $(i, j) \in E$ (what is indicated by setting $x_{ij} = 1$) we have saved one bus. Therefore the more variables are equal to 1 the less vehicles are used in corresponding bus schedule. In order to obtain a feasible bus schedule, at most one trip j can be linked after arbitrary trip i and at most one trip i can be linked before arbitrary trip j. Just mentioned facts lead to the following mathematical model.

$$\text{Maximize} \quad \sum_{\substack{ij,\\(i,j)\in E}} x_{ij} \tag{3}$$

$$\text{subject to:} \quad \sum_{\substack{i,\\(i,j)\in E}} x_{ij} \leq 1 \quad \text{for } j = 1, 2, \ldots, n \tag{4}$$

$$\sum_{\substack{j,\\(i,j)\in E}} x_{ij} \leq 1 \quad \text{for } i = 1, 2, \ldots, n \tag{5}$$

$$x_{ij} \in \{0, 1\} \tag{6}$$

Mathematical model (3) – (6) is an instance of assignment problem, therefore condition (6) can be replaced by

$$x_{ij} \geq 0 \tag{7}$$

and linear program (3) – (5), (7) has still integer solution.

2 Model with depot visiting

Slovak and Czech bus providers often require that vehicles visit depot during working day in order to deliver money or to make a vehicle maintenance or refueling.

A visit of the depot can be modeled as a special trip with both arrival place and departure place equal to depot and arrival and departure times determined by time interval necessary for all procedures required. In most cases this time interval is at least 30 minutes long and can be considered as a safety break for a driver.

In practice, bus staying in depot do not have assigned fixed time positions, they are scheduled in weak traffic hours.

Suppose, we have calculated a bus schedule with the minimum number of vehicles r till now without roundabout routes to depot.

We propose following procedure in order to manage desired visits of all r buses in depot:

Propose the set D of dummy trips representing bus stays in depot. Define digraph

$$\overline{G} = (S \cup D, A) \tag{8}$$

where $S \cup D$ is the union of the set of original set of trips together with the set of all dummy trips – visits of depot. The arc set A of $\overline{G}$ is the set of all ordered pairs (i, j) of elements of $S \cup D$ such that $i \prec j$.

The goal is to arrange all trips from S and r trips from D into minimum number of running boards. We introduce two decision variables x_{ij} and z_k with following meaning:

$$x_{ij} = \begin{cases} 1 & \text{if the trip } j \in S \cup D \text{ is linked immediately after trip } i \in S \cup D \text{ in the same running board} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

$$z_i = \begin{cases} 0 & \text{if the dummy trip } i \in D \text{ is chosen into a running board} \\ 1 & \text{if the dummy trip } i \in D \text{ is not used in any running board} \end{cases} \tag{10}$$

The number of chosen depot stays should be equal exactly to r (the minimum number of running boards without roundabout routes) therefore $|D| - \sum_{i \in D} z_i = r$

If the dummy trip $i \in D$ was not used then all variables x_{ij} have to be equal to zero for all $j \in V$. This is ensured by the constraint

$$z_i + \sum_{\substack{j; \\ (i,j) \in A}} x_{ij} \leq 1 \tag{11}$$

since $z_i = 1$ for any unused trip $i \in D$.

Similarly if dummy trip $j \in D$ is not used in any running board then zero values of all x_{ij} for all $i \in V$ are guaranteed by

$$z_j + \sum_{\substack{i; \\ (i,j) \in A}} x_{ij} \leq 1 \tag{12}$$

On the other hand, if dummy trip $i \in D$ is used what is indicated by $z_i = 0$ then the constraint (11) guarantees that at most for one $j \in S$ is $x_{ij} = 1$. Similarly, if trip $j \in D$ is used what is indicated by $z_j = 0$ then the constraint (12) guarantees that at most for one $i \in V$ is $x_{ij} = 1$.

Hence, now we deal with the following optimization problem:

$$\text{Maximize} \quad \sum_{\substack{ij \\ (i,j) \in A}} x_{ij} \tag{13}$$

$$\text{subject to:} \quad \sum_{\substack{i, \\ (i,j) \in A}} x_{ij} \leq 1 \quad \text{for } j \in S \tag{14}$$

$$z_j + \sum_{\substack{i, \\ (i,j) \in A}} x_{ij} \leq 1 \quad \text{for } j \in D \tag{15}$$

$$\sum_{\substack{j, \\ (i,j) \in A}} x_{ij} \leq 1 \quad \text{for } i \in S \tag{16}$$

$$z_i + \sum_{\substack{j, \\ (i,j) \in A}} x_{ij} \leq 1 \quad \text{for } i \in D \tag{17}$$

$$|D| - \sum_{i \in D} z_i = r \tag{18}$$

$$x_{ij} \in \{0,1\} \quad \text{for all } i,j \text{ such that } (i,j) \in A \tag{19}$$

$$z_i \in \{0,1\} \quad \text{for all } i \in D \tag{20}$$

2.1 Design of the set D of dummy trips

A deficit function $f(x)$ is a function with domain $\langle 0, 1440)$ defined as follows

$$f(x) = |\ \{\ i \mid x \in \langle d_i, a_i \rangle, i \in V\}\ | \tag{21}$$

Every $x \in \langle 0, 1440)$ represents time in a day in minutes, value $f(x)$ is the number of trips running in time moment x. A deficit function can help us to determine rush and weak traffic time interval. Typical deficit function has its rush hours in the morning and after noon and weak hours between 10:00 and 12:00 o'clock before noon. Roundabout routes should be placed into weak hour in order not to raise the number of vehicles.

Remember that a trip is a quadruple (d_k, a_k, u_k, v_k) where d_k is departure time, a_k is arrival time, u_k is departure bus stop and v_k is arrival bus stop of the trip k.

Every dummy trip representing stay in depot should have departure and arrival place equal to depot. Let $W \subset V$ be a set of trips running in week hours. If $(d_i, a_i, \mathit{depot}, \mathit{depot})$ is a dummy trip then its latest time position is such that there exist a trip (d_j, a_j, u_j, v_j) such that

$$a_i = d_j - m(\mathit{depot}, u_j) \tag{22}$$

otherwise this dummy trip could be shifted later. Therefore the set D of possible dummy trips can be defined as follows:

$$D = \{\ (a - 30, a, depot, depot) \quad | \quad a = d_j - m(depot, u_j), j \in W\} \tag{23}$$

When creating digraph $\overline{G} = (V \cup D, A)$, exclude all arcs (i, j) such that $i \in D$ and $j \in D$ from arc set A in order to get away of possibility that two roundabout routes will be assigned to the same vehicle.

Another possibility how to reduce the possibility of assigning two dummy trips to the same running board is following. If there are two dummy trip $i \in D$, $j \in D$ such that there exist a trip $k \in V$ such that $i \prec k \prec j$ then remove one of dummy trips from D.

3 Computational experiment

We used Gurobi MILP solver on Intel Xeon with 4 CPU Cores to solve linear model (13)–(20).

We had available real world data for municipal bus public transport in Slovak town Martin–Vrútky. The public transport system of this town consisted of 725 trips organized in 18 lines. Corresponding deficit function is presented on Figure 1.

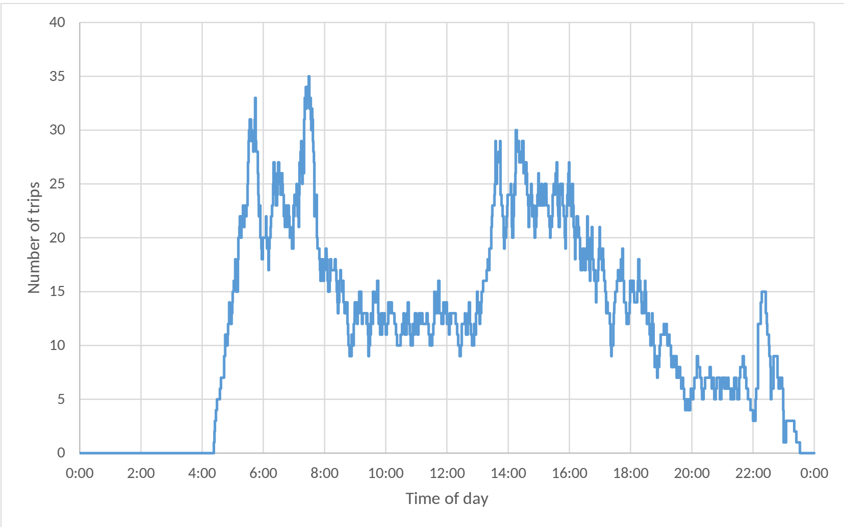


Figure 1 Deficit function of typical working day

Depot is close to the bus stop 59. Using deficit function we have planed time positions of roundabout dummy trips to depot between 8:00 and 12:00 – in hours with weak traffic.

First calculation of the minimum number of running boards (without roundabout routes to depot) solved by model (3)–(7) yielded 39 buses. Therefore we tried first to use $r = 39$ dummy trips – $r = 39$ in equation (18). The result was not satisfactory since sometimes two dummy trips were placed into one running board and so another running boards remained without roundabout trip.

So we incremented the number of dummy trips to depot (constant r in equation (18)) and re–solved model.

We repeated this procedure until every running board contained at least one dummy trip. A feasible solution without need for more buses appeared for $r = 58$ wheres the number of running boards remained equal to 39. Computational steps are shown on Table 1.

Number of dummy trips taken (r)	39	40	41	42	43	44	45	46	47	48
Number of running boards	39	39	39	39	39	39	39	39	39	39
Number of boards without depot	11	10	9	9	8	7	6	6	5	6
Computational time [s]	6.40	7.16	6.18	6.82	7.51	6.25	6.13	6.58	6.24	6.41
Number of dummy trips taken (r)	49	50	51	52	53	54	55	56	57	58
Number of running boards	39	39	39	39	39	39	39	39	39	39
Number of boards without depot	3	3	2	2	2	2	2	2	2	0
Computational time [s]	7.09	6.22	6.75	6.79	7.01	6.93	7.34	6.52	6.84	7.44

Table 1 Dependence of the number of running boards without roundabout route on the number r of included dummy trips.

It can be easy seen that the solution with 39 running boards and 58 roundabout trips has to contain many running boards with more that one roundabout trip. An example of a running board with more than one dummy trip is shown on Table 2.

Line	Trip	Dep.stop	Dep.time	Arriv.stop	Arriv.time
10	1	31	04:35	91	04:55
10	2	91	05:00	31	05:20
13	3	59	05:29	43	06:06
30	12	10	06:21	59	06:42
27	13	59	06:47	70	07:20
27	16	70	07:25	59	07:58
70	27	81	08:20	91	08:37
11	16	91	08:53	31	09:20
***	***	59	09:23	59	10:23
70	37	59	10:32	91	10:52
***	***	59	11:12	59	12:12
1	10	10	13:25	102	13:45
2	7	59	13:50	26	14:07
52	18	44	14:10	56	14:25
12	15	18	14:43	24	15:05
30	74	10	15:26	59	15:47
30	81	59	15:48	10	16:11
30	84	10	16:26	59	16:47
30	90	10	17:11	59	17:32
30	101	59	18:13	10	18:36
30	106	10	19:21	56	19:37
10	129	31	20:40	91	21:00
30	112	10	22:36	59	22:57

Table 2 Example of a running board with more than two dummy trips

Solution containing exactly one dummy trip in every running board can be simply obtained by discarding abundant dummy trips.

4 Conclusion

We have defined problem of vehicle scheduling with roundabout routes to depot. We have formulated MILP mathematical model and computed running boards for public transport system in Slovak town Martin–Vrútky. Computation experiment showed that this model can be applied for real world instances of vehicle scheduling problem to find time positions of roundabout routes to depot. However, the presented model does not guarantee that chosen dummy trips are assigned to every running board – therefore some experiments are needed.

Therefore, the further research will be focused on ways how to prohibit assignment of more than one dummy trip to one running board.

Authors of this paper have taken part in bus public transport of more than twenty Czech and Slovak towns, cooperation with everyone of them was only temporary.

Acknowledgements

The research was supported by the Slovak Research and Development Agency grant APVV-14-0658 "Optimization of urban and regional public personal transport".

References

[1] Ceder, A.: Public-transit vehicle schedules using a minimum crew-cost approach. *Total Logistic Management* **4** (2011), 21–42.

[2] Gurobi Optimization Inc.: *Gurobi Optimizer Reference Manual*, http://www.gurobi.com, 2015.

[3] Palúch, S.: Two approaches to vehicle and crew scheduling in urban and regional bus transport. In: *Proceeding of the Quantitative Methods in Economics (Multiple Criteria Decision Making XIV)*, Iura Edition, Bratislava, 2008, 212–218.

[4] Palúch, S.: Vehicle and crew scheduling problem in regular personal bus transport - the state of art in Slovakia. In: *Proceeding of the International Conference on Transport Science*, Slovene Association of Transport Sciences, Portorož, Slovenija, 2013, 297–304.

[5] Peško, Š.: Flexible bus scheduling with overlapping trips. In: *Proceeding of the Quantitative Methods in Economics (Multiple Criteria Decision Making XIV)*, Iura Edition, Bratislava, 2008, 225–230.

[6] Peško, Š.: The minimum fleet size problem for flexible bus scheduling. *Studies of the Faculty of Management Science and Informatics* **9** (2001), 75–81.

[7] Wren, A., Rousseau, J.M.: Bus driver scheduling. In: *Proceeding of the Computer-Aided Transit Scheduling*, Springer Verlag, 1995, 173–187.

Monetary Policy in an Economy with Frictional Labor Market - Case of Poland

Adam Pápai[1]

Abstract. The aim of this paper is to investigate the impacts and effectiveness of the monetary policy in the Polish economy during the last 14 years. Furthermore, we are interested in the effects of the Great Recession, which was caused by the financial crisis in 2007, on the behavior of key macroeconomic variables. To achieve our goals, we choose and estimate a dynamic stochastic general equilibrium model specified for a small open economy. The model is characterized by an explicitly defined labor market. The model version of this sector of the economy contains several frictions in form of a search and matching function, wage adjustment costs, hiring costs and wage bargaining processes. We implement a few different Taylor-type equations, which represent the decision making rule of the monetary authority. This approach helps us identify the main relationships, which the monetary policy maker takes into consideration while setting the interest rates. We then examine the impacts of monetary policy shock on the output and labor market.

Keywords: DSGE model, monetary policy, labor market frictions, Polish economy.

JEL classification: E32, J60
AMS classification: 91B40

1 Introduction

In this paper, we focus on the influence of the monetary policy maker's decisions on labor market. It might not be the principal goal of the central banks to interfere in the movements on the labor market, however the decisions of the monetary authority may affect the preferences of employers and workers nonetheless. It is therefore desirable to investigate the connections between the monetary authority and the variables describing the labor market. To be able to assess these relationships, we follow up on our previous research (Pápai and Němec [5] and Pápai and Němec [6]), and estimate a dynamic stochastic general equilibrium (DSGE) model for a small open economy. In this paper we choose to focus more on the monetary policy rather than the labor market. We selected the economy of Poland, although it hardly classifies as small. Therefore, our goal is also to test, whether such model can replicate the low dependence of Poland on the foreign sector.

DSGE models are widely used tools for forecasting and decision making in the macroeconomic sector. The great degree of modifiability increased their popularity in the last two decades. The literature regarding DSGE models is vast. Some, like Christiano et al. [3] incorporated financial frictions and labor market rigidities into a otherwise standard small open economy model. Their estimation results on Swedish data suggest, that introducing such limitations improve the accuracy of the model's behavior. They conclude, that "labor market tightness (measured as vacancies divided by unemployment) is unimportant for the cost of expanding the workforce. In other words, there are costs of hiring, but no significant costs of vacancy postings."

Antosiewicz and Lewandowski [2] analyze the factors behind the financial crisis and their transmission into labor markets of selected European countries, including Poland and the Czech Republic. They find a significantly higher effect of productivity shock to the wages in the Czech Republic than in Poland. Also, in all analyzed countries except of the Czech Republic, job destruction shocks influenced consumption significantly. For Poland, this suggests a decrease of consumption in times of high job loss probability.

[1]Masaryk University, Faculty of Economics and Administration, Department of Economics, Lipová 41a, Brno, Czech Republic, papai@mail.muni.cz

In both countries, the unemployment rate is mostly influenced by foreign demand and job destruction shocks. However, the model of these authors does not contain monetary policy in any form.

Tonner et al. [7] expand the core model of the Czech national bank by various implementations of labor market. They provide results of multiple predictions to assess the ability of the different model specifications to match the development of the selected macroeconomic variables. Their findings indicate, that altering the baseline model within acceptable bounds, the linking of unemployment with the labor market markup gives more accurate forecasting results.

2 Model

We selected a DSGE model originally developed by Albertini et al. [1], who used New Zealand data for their estimation. This model focuses on the analysis of the driving forces on the labor market and hence uses a detailed description of relationships among labor market variables. The model contains several types of rigidities. First, search and matching mechanism, created by Mortensen and Pissarides [4] is present to describe the flows on the labor market. This function matches unfilled and therefore unproductive job positions with unemployed workers. The search and matching introduces a rigidity to the model, because not all people who search for job and not all firms who want to fill their vacancies can do so. Another friction in the model has the form of vacancy posting costs. Firms cannot create vacant job positions freely, but must decide, whether the potential gains outweigh the induced costs. Finally, there are price and wage adjustment costs, which similarly to the previous frictions, induce actual costs to the firms.

The model does not contain capital and government and consists of households, firms and the monetary authority. The representative household decides between consumption and labor. Current consumption depends on habit to achieve a smoother level over time. Labor brings disutility to the households. However, if the people decide to work, they get a compensation in form of hourly wages. There are three types of firms, each facing some kind of rigidity. Producers create intermediate goods and sell them on a perfectly competitive market. They hire employees and negotiate their wages and the hours using a Nash bargaining process. Retailers buy the unfinished products and combine them to sell final goods on a monopolistically competitive market to the households. This gives them the ability to adjust the prices. Finally, like the domestic retailers, importers sell foreign finished goods on the domestic market for consumption. They can also adjust their prices, while facing price adjustment costs. The last agent in the model is the monetary authority.

Monetary authority

The monetary policy maker has a key role in the macroeconomic realm. It is therefore desirable to incorporate the monetary authority into the DSGE model. The most common way is to define it using the Taylor rule.

$$i_t = \rho_r i_{t-1} + (1-\rho_r)(\rho_\pi \pi_{t+1} + \rho_y y_t + \rho_{\Delta y}\Delta y_t + \rho_{\Delta e}\Delta e_t) + \epsilon_m \tag{1}$$

$$i_t = \rho_r i_{t-1} + (1-\rho_r)(\rho_\pi \pi_{t+1} + \rho_y y_t + \rho_{\Delta y}\Delta y_t) + \epsilon_m \tag{2}$$

$$i_t = \rho_r i_{t-1} + (1-\rho_r)(\rho_\pi \pi_{t+1} + \rho_y y_t) + \epsilon_m \tag{3}$$

$$i_t = \rho_r i_{t-1} + (1-\rho_r)(\rho_\pi \pi_{t+1} + \rho_{\Delta y}\Delta y_t) + \epsilon_m \tag{4}$$

Equations (1)–(4) show the slight modifications to the Taylor rule (presented in log-linearized forms), which were used for our estimations. Equation (1) represents the original variation taken from Albertini et al. [1], where the central bank sets the current nominal interest rate (i_t) based on its previous value, the expected inflation gap (π_{t+1}), output gap (y_t), difference between the current and previous output (Δy_t) and nominal exchange rate (Δe_t) gaps. Parameter ρ_r is the interest rate smoothing parameter, while the other ρ-s represent the dependence of the interest rate on the development of the previously mentioned variables. The other three specifications of the Taylor rule leave out, among other things, the nominal exchange rate (e_t). We leave out this variable to investigate this way, whether the monetary policy maker reacts to the impulses arising from the interactions between Poland and other economies. The omission of output shows us, whether the central bank makes a decision based on the current output gap or also takes into consideration the difference of the output gaps and vice versa. We evaluate the plausibility of different model specifications using impulse response functions.

3 Data, methodology and calibration

Our sample covers the period between the first quarter of 2002 and the last quarter of 2015, which is 56 periods in total. We chose our observed variables as in Albertini et al. [1] with one exception, given the unavailability of quarterly time series for hours worked. Seven time series were selected to characterize the domestic economy. The harmonized unemployment rate (u_t), number of vacancies (v_t), hourly earnings (w_t), CPI (for calculation of inflation (π_t)), gross domestic product (y_t) and three month interbank interest rate (i_t) were acquired from the OECD database. The exchange rate between Polish zloty and euro was downloaded from the sites of the European central bank. The foreign sector (y_t^*, π_t^*, i_t^*), represented by the Eurozone is present in the model in the form of AR(1) processes. The variables are entered into the model in per capita terms. Each of the time series is seasonally adjusted. We use demean (for π_t, i_t, π_t^* and i_t^*) and Hodrick-Prescott filter (for the rest of the variables), with smoothing parameter $\lambda = 1600$ to detrend the time series and get the cyclical components of the data.

We use Bayesian techniques to estimate models (1)–(4). This allows us to impose additional information for the model in the form of prior densities. Furthermore, the Kalman filter allows us to investigate the trajectories of unobserved variables. For each estimation, we generate two chains of Metropolis-Hastings algorithm, while targeting acceptance ratio of 30%. Matlab and its toolbox Dynare is used to acquire our estimation results.

Standard parameter calibration methods are used to help pinpoint the steady state. Some parameter values are taken form the literature, while others are calculated from the data. The discount factor was set to a widely used value of 0.99. The share of labor in production was set to 2/3. The parameter of home bias in consumption was calculated from the data as the import share on GDP (0.2843). This relatively small number supports our presumption of Poland not being a small open economy. The steady state of unemployment (0.1227) was calculated as a sample mean of unemployment rate. The prior densities of estimated parameters are presented in Table 1.

4 Estimation results

Description		Prior density	Model 1	Model 2	Model 3	Model 4
Habit	ϑ	$\beta(0.5, 0.15)$	0.6766	0.6833	0.6035	0.6750
El. of substitution (dom. & for.)	η	$\Gamma(1, 0.2)$	0.4454	0.4875	0.3926	0.5137
Bargaining power of firms	ξ	$\beta(0.5, 0.2)$	0.6790	0.5847	0.7292	0.5627
Elasticity of matching	ν	$\beta(0.5, 0.2)$	0.8591	0.8825	0.7307	0.8862
Price and wage setting						
Back. looking price (dom. good)	γ_H	$\beta(0.75, 0.1)$	0.9108	0.9109	0.8875	0.9044
Back. looking price (for. good)	γ_F	$\beta(0.75, 0.1)$	0.6457	0.6922	0.6974	0.6880
Back. looking wage parameter	γ_W	$\beta(0.75, 0.1)$	0.3603	0.3626	0.4484	0.3865
Price adj. cost (dom. good)	ψ_H	$\Gamma(50, 15)$	83.1461	83.0849	80.6558	81.4520
Price adj. cost (for. good)	ψ_F	$\Gamma(50, 15)$	54.5994	53.3332	59.9069	52.5602
Wage adj. cost	ψ_W	$\Gamma(50, 15)$	19.5984	21.2508	13.8994	20.1162
Monetary policy						
Interest rate smooth.	ρ_r	$\beta(0.5, 0.15)$	0.6666	0.6768	0.6081	0.6570
Inflation	ρ_π	$\Gamma(1.5, 0.25)$	2.9685	3.0574	2.8051	3.0736
Output gap	ρ_y	$\mathcal{N}(0.25, 0.1)$	0.2693	0.2180	0.2930	–
Difference of output	$\rho_{\Delta y}$	$\mathcal{N}(0.25, 0.1)$	0.1154	0.1418	–	0.1310
Difference of exchange rate	$\rho_{\Delta e}$	$\mathcal{N}(0.25, 0.1)$	-0.0702	–	–	–

Table 1 Estimation results - parameter means ('–' means not estimated)

Table 1 shows the results of parameter estimates for all four models. Our first finding is, that the differences in the results of parameter estimates for these models are small and mostly insignificant. This could be given by the fact, that the changes made are really small. Therefore, we first describe the results of parameter estimations and not the differences between the models. Second, most of the posterior values are different from our priors. The deep habit parameter (ϑ) is relatively high, meaning the households try to smooth their consumption. The firm's bargaining power parameter (ξ) varies across the models, but is not extraordinarily high. This could suggest, that there are some relatively powerful trade unions in Poland (see [8]). On the other hand, the elasticity of matching function with respect to the number of job seekers (ν) is estimated to be around 0.88. This means, that the number of matches depends more on the unemployed than on the vacancies.

The price and wage setting parameters show us the degree of rigidities in the Polish economy. The backward looking parameters (γ) indicate how much the current value depends on its previous level. Our estimation results suggest, that the prices of the domestically produced goods are the stickiest, followed by the prices of the imported products. Wages are the least rigid among these three frictions. The price and wage adjustment cost parameters (ψ) indicate the size of costs induced by the changing of the product's price or worker's wage. Again, the value for the prices of domestic goods is the highest and the wage setting parameter is the lowest. Meaning, it is cheaper for producers to change the wage than for retailers to adjust the price. These results show low degrees of rigidities on the labor market of Poland.

Lastly, the parameters important for the monetary policy maker are also presented in table 1. The main reason for leaving out the dependency of interest rates in the Taylor rule on the nominal exchange rate was the estimation result of the first model. Here, the posterior mean resulted in a negative number, while the 90% posterior density interval was between -0.1567 and 0.0071, so we could not rule out statistical insignificancy. The interest rate smoothing parameter was estimated in all cases above 0.6 which suggests relatively stable values. The parameter describing the dependency on inflation is estimated to be around 3, which is in line with the current literature. The monetary policy maker takes into consideration the output gap more than the differences of this variable.

If we look at the differences among the four models, we find that the parameter estimation results are similar in models 1, 2 and 4. Model 3 sticks out the most. This is even more visible in the two following figures. Figures 1 and 2 show the results of impulse responses to monetary and technology shocks respectively. Models 2 and 4 behave in both cases almost identically and are at the very least at the boundaries of the 90% highest posterior density interval (HPDI) of model 1. These three models behave in accordance with the literature and our expectations. Model 3 however exhibits different behavior, which in several cases contradicts with the actual behavior of the real economy. Therefore, we will focus our further interpretation only on the three similar models.

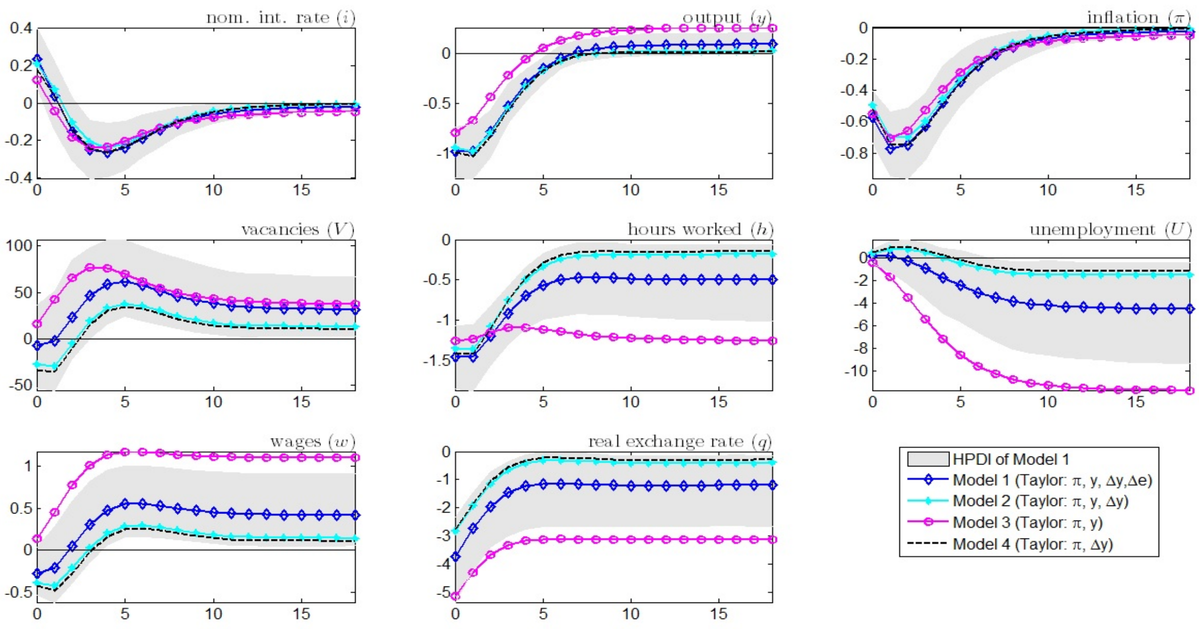


Figure 1 Impulse responses to a monetary shock

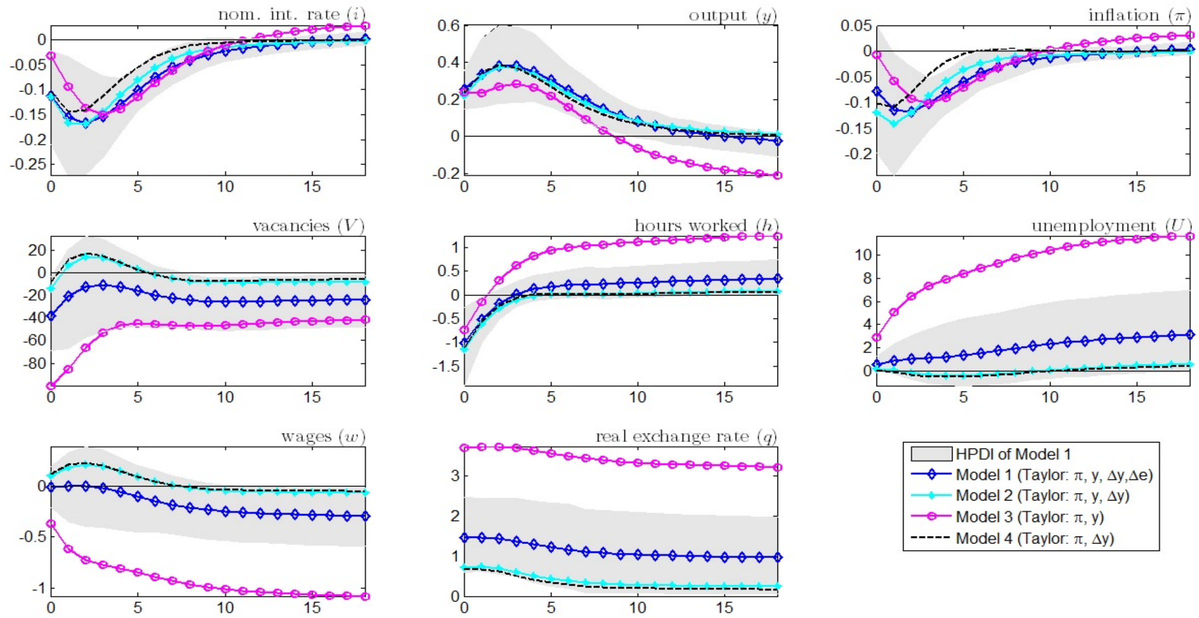


Figure 2 Impulse responses to a technology shock

At this point we would also like to note, that the variables of vacancy and unemployment are divided by economically active population and their volatility is smaller than the volatility of the rest of the variables. High deviations from steady state therefore do not mean high changes of the actual number of vacancies and number of unemployed. Furthermore, the impulse responses of vacancies, unemployment and wages contain zero in the 90% HPDI, so we cannot rule out statistical insignificancy.

The transmission mechanism of monetary policy is presented as a monetary shock in figure 1. It causes an initial increase of interest rates. This restrictive monetary policy causes a drop of output and decrease of inflation. Because the unemployment does not change substantially, the firms lower the hours of workers to compensate for the lower production. Wages and vacancies also decrease, as is expected, when the output drops. Finally, this monetary policy creates an appreciation of exchange rate.

Figure 2 shows the reactions of selected variables to a positive technology shock, that increases the output. The reactions of labor market variables are limited. Only models 2 and 4 indicate a lagged increase of vacancies created by the firms during expansion periods. The initial decrease of hours worked, when the labor is improved by the better technology, fades away during $4-5$ periods. These two models also show a slight increase of wages, while the unemployment remains without change and the exchange rate depreciates.

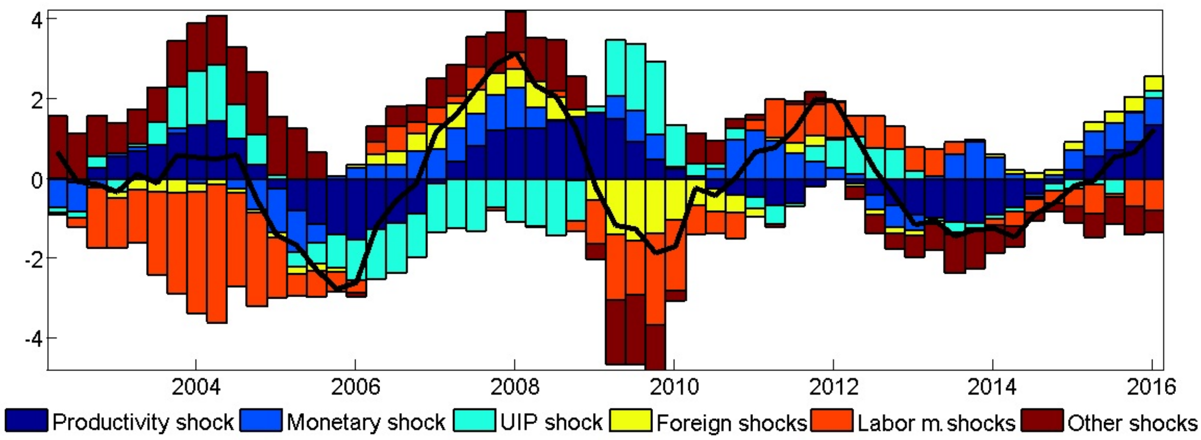


Figure 3 Historical shock decomposition of output

The models are not able to capture the counter cyclical changes in the variable of unemployment. This could be given by the unavailability of the time series of hours and the model therefore interprets the changes in the workforce as intensive (hours) and not extensive (unemployed).

Finally, in figure 3 we present the historical shock decomposition of output from the results of model 2. There were some substantial negative labor market shocks at the beginning of the observed period. The most interesting conclusion of this figure is, that the foreign shocks influenced the Polish output mainly during the Great Recession. Their presence was very little in the other periods. This could suggest, that the Polish economy is, in fact, not so small and open.

5 Conclusion

In this paper we estimated four slightly different model specifications for a small open economy. We used quarterly data of Poland for the last 14 years. Our findings suggest that there are significant price rigidities in the Polish economy, while the frictions on the labor market are not so grave. The impulse response functions indicate, that the monetary policy is able to affect the real economy, though its influence on labor market is marginal. Also, the decision making by the monetary authority does not depend on the development of the exchange rate, however omission of output difference from the Taylor rule causes serious distortion in the estimated results. Finally, given that we indicated the size of the Polish economy, it would be more suitable to estimate models that capture its size better.

Acknowledgements

This work was supported by funding of specific research at Faculty of Economics and Administration, project MUNI/A/1040/2015. This support is gratefully acknowledged.

References

[1] Albertini, J., Kamber, G., and Kirker, M.: Estimated small open economy model with frictional unemployment. *Pacific Economic Review*, **17**, 2 (2012), 326–353.

[2] Antosiewicz, M., and Lewandowski, P.: *What if you were German? - DSGE approach to the Great Recession on labour markets.* IBS Working Papers. Instytut Badan Strukturalnych (2014).

[3] Christiano, L. J., Trabandt, M., and Walentin, K.: Introducing Financial Frictions and Unemployment into a Small Open Economy Model. *Journal of Economic Dynamics and Control*, **35**, 12 (2011), 1999–2041.

[4] Mortensen, D. T., and Pissarides, C. A.: Job creation and job destruction in the theory of unemployment. *Review of Economic Studies*, **61**, 3 (1994), 397–415.

[5] Pápai, A., and Němec, D.: Labour market rigidities: A DSGE approach. *Proceedings of 32nd International Conference Mathematical Methods in Economics* (Talašová, J., Stoklasa, J., Talášek, T., eds.), Olomouc: Palacký University, 2014, 748–753.

[6] Pápai, A., and Němec, D.: Labor Market Frictions in the Czech Republic and Hungary. *Proceedings of 33nd International Conference Mathematical Methods in Economics* (Martinčík, D., Ircingová, J., Janeček, P., eds.), Plzeň: University of West Bohemia, 2015, 606–611.

[7] Tonner, J., Tvrz, S., and Vašíček, O.: *Labour Market Modelling within a DSGE Approach.* CNB Working paper series. 6/2015.

[8] Trade unions. Link: http://www.worker-participation.eu/National-Industrial-Relations/Countries/Poland/Trade-Unions

Adaptive parameter estimations of Markowitz model for portfolio optimization

Josef Pavelec[1], Blanka Šedivá [2]

Abstract. This article is focused on a stock market portfolio optimization. The used method is a modification of traditional Markowitz model which extends the original one for adaptive approaches of parameter estimations. One of the basic factors which significantly influence optimal portfolio is the method of estimations of return on assets, risk and covariance between them. Since the stock market processes tend to be not stationary we can expect that prioritization of recent information will lead to improvement of these parameter estimations and thus to better results of the entire model. For this purpose a modified algorithm was design to estimate expected return and correlation matrix more stable. For implementation and verification of this algorithm we needed to build a program which was able to download historical stock market data from the internet and compute optimal portfolio using either traditional Markowitz model and its modified approach. Obtained results will be compared to the traditional Markowitz model provided real data.

Keywords: Markowitz model, estimation of parameters, adaptive method.

JEL classification: G11
AMS classification: 91G10

1 Introduction

There are many articles about optimal portfolio in the science of mathematics. The reason is that investing on a stock market with potential of profit is interesting for a large amount of people in the world. Thanks to this fact, there is also many approaches how mathematicians try to model the stock market. Some of them tend to believe that there is no relationship between the history and the future. This group use methods of random walk and tries to simulate large number of scenarios to forecast the future. The other group of scientists believe that there is a strong relation between historical prices and the future ones. In this article will focus on this approach and try to treat one of the biggest milestones of these methods - stability of the model provided the parameter estimation. We will introduce two different ways to improve the stability of optimization - matrix cleaning and data weighting.

2 Portfolio theory: basic results

Suppose we have a set of N financial assets characterized by their random return in chosen time period, so the random vector is the vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_N)$ of random returns on the individual assets.

The distribution of vector $\mathbf{X}$ is characterized by vector of expected value with elements $\mathbb{E}X_i = r_i$ and by covariance matrix $\boldsymbol{V}$ whose i, j^{th} element is the covariance between the X_i^{th} and the X_j^{th} random variables. The elements on diagonal σ_i^2 of matrix $\boldsymbol{V}$ represent variances of asset i.

The Markowitz's theory of optimal portfolio is focused on the problem to find optimal weight of each assets such that overall portfolio provides the best return for a fixed level of risk, or conversely the smallest risk for a given overall return [5]. More precisely, the average return R_p of a portfolio P of N assets is defined as $R_p = \sum_{i=1}^N w_i r_i$ where w_i is the amount of capital invested in the assets i and r_i are expected returns of the individual assets. Similarly the risk of a portfolio P can by associated with the total variance $\sigma_P^2 = \sum_{i,j=1}^N w_i V_{ij} w_j$ or in alternative form $\sigma_P^2 = \sum_{i,j=1}^N w_i \sigma_i C_{ij} \sigma_j w_j$ where σ_i^2 is the

[1]University of West Bohemia, Plzeň, Czech Republic, Department of Mathematics, jpavelec@kma.zcu.cz
[2]University of West Bohemia, Plzeň, Czech Republic, Department of Mathematics, sediva@kma.zcu.cz

variance of asset i and C is the correlation matrix. The optimal portfolio which minimizes σ_P^2 for a given value of R_p can be easily found introducing a Lagrange multiplier and leads to a linear problem where the matrix $\boldsymbol{C}$ has to be invertible [3],[2].

The resulting from a Markowitz optimization scheme, which gives the portfolio with the minimum risk for a given return $R_P = \sum w_i r_i$

$$w_i \sigma_i = R_p \frac{\sum_j C_{ij}^{-1} r_j/\sigma_j}{\sum_{i,j} r_i/\sigma_i C_{ij}^{-1} r_j/\sigma_j} \tag{1}$$

By redefining w_i as $w_i \sigma_i$ the σ_i is absorbed in r_i and w_i and the equations (3) can be write in matrix notation

$$\boldsymbol{w}_C = R_p \frac{\boldsymbol{C}^{-1}\boldsymbol{r}}{\boldsymbol{r}^T \boldsymbol{C}^{-1} \boldsymbol{r}} \tag{2}$$

and the corresponding risk of the portfolio over the period using this construction is

$$\sigma_P^2 = \frac{R_p^2}{\boldsymbol{r}^T \boldsymbol{C}^{-1} \boldsymbol{r}} \tag{3}$$

From mathematical equation (2) is obvious that usability of Markowitz model strongly depends on input data which are used for asset mean return estimations and the dominant role for stability is given by quality of estimation of the covariance matrix.

3 Parameter estimations - stability of the model

3.1 Empirical correlation matrix

Suppose we have N stock return series with T elements each. If we want to measure and optimize the risk of this portfolio, it is necessary to use a reliable estimate of the covariance matrix $\boldsymbol{V}$ or correlation matrix $\boldsymbol{C}$.

If r_i^t is the daily return of stock i at time t, the empirical variance of each stock is given by

$$\sigma_i^2 = \frac{1}{T} \sum_t \left(r_i^t - \overline{r}_i \right)^2 \tag{4}$$

and can be assumed for simplicity to be perfectly known. We also suppose, as usual, the daily return of stock r_i^t is demeaned ($\overline{r}_i = 0$). The empirical correlation matrix is obtained as

$$E_{ij} = \frac{1}{T} \sum_t x_i^t x_j^t, \quad \text{where } x_i^t = r_i^t/\sigma_i \tag{5}$$

or in matrix form $\boldsymbol{E} = (1/T)\, \boldsymbol{X}^T \boldsymbol{X}$, where $\boldsymbol{X}$ is the normalization $T \times N$ matrix of return $X_{it} = r_i^t/\sigma_i$.

3.2 Random matrix theory and matrix cleaning

For a set of N different assets, the correlation matrix contains $N(N-1)/2$ entries, which must be determined from N time series of length T. If T is not very large compared to N, we can expect that the determination of covariances is noisy and therefore that the empirical correlation matrix is to a large extent random. Because a covariance matrix is positive semidefinite, that the structure of it can by describe by real eigenvalues and corresponding eigenvectors. Eigenvalues of the covariance matrix that are small (or even zero) correspond to portfolios of stocks that have non-zero returns, but extremely low or vanishing risk; such portfolios are invariably related to estimation errors resulting from insufficient data. One of the approaches used to eliminate the problem of small eigenvalues in the estimated covariance matrix is the so-called random matrix technique. Random matrix theory (RMT), first developed by authors such as Dyson [4] and Mehta [9] for physical application, but there are also many results of interest in a financial context [7], [1], [11].

The spectral properties of $\boldsymbol{C}$ may be compared to those of random correlation matrix. As described by [7], [11] and others, if $\boldsymbol{R}$ is any matrix defined by $\boldsymbol{R} = (1/T)\,\boldsymbol{A}^T\boldsymbol{A}$, where $\boldsymbol{A}$ is an $N \times T$ matrix whose elements are i.i.d random variables with mean zero and fixed variance σ^2, than in the limit $T, N \to \infty$ keeping ratio $Q = T/N \geq 1$ constant, the density of eigenvalues of $\boldsymbol{R}$ is given by

$$P(\lambda) = \frac{Q}{2\pi\sigma^2}\frac{\sqrt{(\lambda_{max} - \lambda)(\lambda - \lambda_{min})}}{\lambda}, \quad \lambda_{min} \leq \lambda \leq \lambda_{max}, \tag{6}$$

where the maximum and minimum eigenvalues are given by

$$\lambda_{max/min} = \sigma^2\left(1 \pm \sqrt{\frac{1}{Q}}\right)^2. \tag{7}$$

The distribution $P(\lambda)$ are known as the Marčenko-Pastur density [8] and the theoretical maximum and minimum eigenvalues determined the bounds for random matrix. If the eigenvalues of matrix are beyond these bounds, it is said that they deviate from random. If we apply this theoretical background of RMT to the correlation matrix we can separate the noise and non-noise parts of $\boldsymbol{E}$. We cleaned the matrix by following procedure: 1. to construct the empirical correlation matrix as (5), 2. separate the noisy eigenvalues from non-noisy eigenvalues as (6), 3. to keep the non-noisy eigenvalues the same and to take average of the noisy eigenvalues, 4. to replace each eigenvalue associated with the noisy part by average of the eigenvalues, 5. to reconstruct correlation matrix. The simple repair mechanism, based on the spectral decomposition of the correlation matrix, is described for example in [6].

3.3 Exponential weights - parameter estimation

Another method how we can minimize the non-stability of the model is weighting. Since we suppose that the most recent data are the most relevant and the older ones influence the future less, we designed model, which weights the data exponentially to the history. This idea was introduced in [10], where we can also find more details. The parameters we need to estimate are:

- Return - estimate of expected return on asset X_i is in this case calculated by weighted mean:

$$\widehat{r}_i = \left(\sum_{t=1}^{T} r_i^t \cdot \delta^t\right) / \left(\sum_{t=1}^{T} \delta^t\right) \tag{8}$$

- Risk - the estimated expected return of asset X_i is in this case calculated by sample weighted variance:

$$\widehat{\sigma}_i = \sqrt{\frac{\sum\limits_{t=1}^{T} \delta^t \cdot (r_i^t - \overline{r}_i)^2}{\sum\limits_{t=1}^{T} \delta^t} \cdot \frac{T}{T-1}}, \tag{9}$$

- Empirical correlation matrix is computed using exponentially weighted by

$$E_{ij} = \left(\sum_{t=1}^{T} \delta^t x_i^t x_j^t\right) / \sum_{t=1}^{T} \delta^t, \quad \text{where } x_i^t = r_i^t / \widehat{\sigma}_i \tag{10}$$

  where $\delta \in (0, 1\rangle$ is weighting parameter. This parameter δ is sometimes called a "forgetting coefficient" and should be close to 1. The smaller it is, the faster older data get unsignificant.

This approach is implemented and tested in created program StockMaTT, where we can either analyse our data using traditional Markowitz model (uses linear data weights) or using this adaptive approach with exponential weighting. The results of these two methods vary and depend on its parameters - on data history length for linear model and on λ for the adaptive model.

4 Data analysed

4.1 Program StockMaTT

So that we can verify the model and figure out it's results we needed to build a programme tool. A satisfying environment needs to have sufficient mathematical and statistical background and also needs to be fast enough to implement quite a complicated algorithm on complex data.

A tool that fitted our needs was the programming language MATLAB® in combination with its GUI environment that makes the user's controlability comfortable. The main screen where we update data and compute the optimal portfolio can be found in the Figure 1.

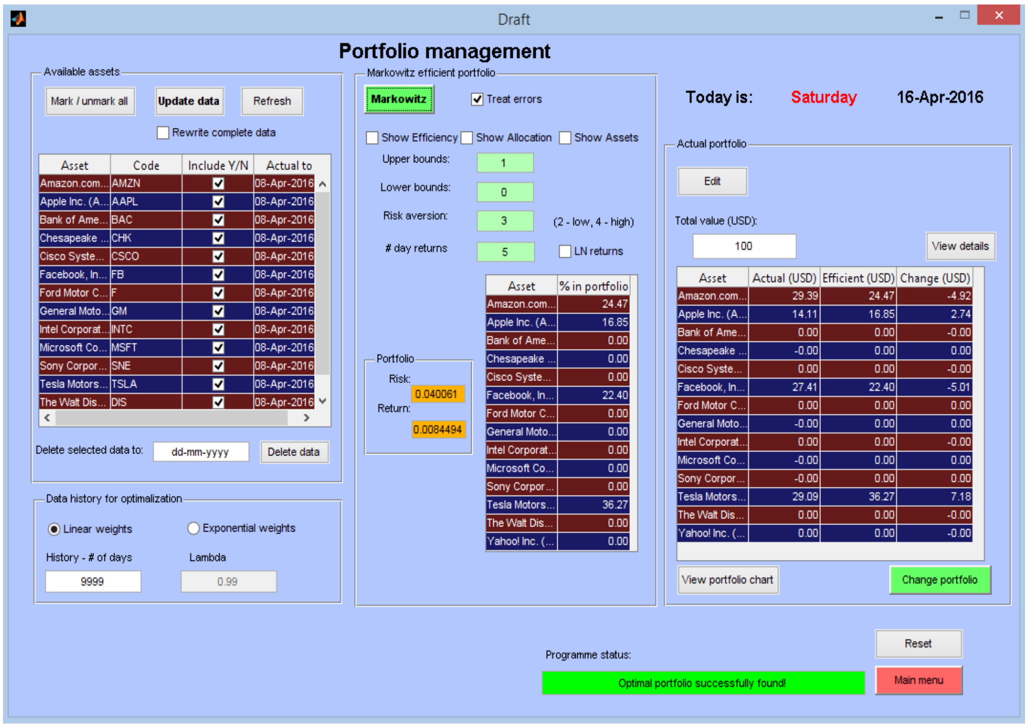


Figure 1 Window for data update and portfolio optimization

4.2 Real data

The model was also tested on real data. Since the focus of this article is on portfolio optimization we chose to use the data from the stock market. As a prefered server was chosen server *yahoo.finance.com*.

For our analysis we used the time series of daily closing prices of stocks available mainly on *NASDAQ*, which is the biggest stock market in the USA with over 3900 assets from about 39 countries. The financial instruments that can be traded here besides stocks are also options and futures.

Asset split and dividends

Downloaded data also needed to be "smoothened" for unexpected jumps caused by splitting the stocks and also for dividends. The stock spilt is a phenomenon that happens usually for expensive assets when the stakeholders want to support the stock liquidity. Usually they decide to split the stock in the rate of X:1 which means that suddenly all stock holders have X times more assets with $\frac{1}{X}$ of its value. This phenomenon causes this obvious jumps that can bee seen for example on the Apple Inc. stock in the Figure 2.

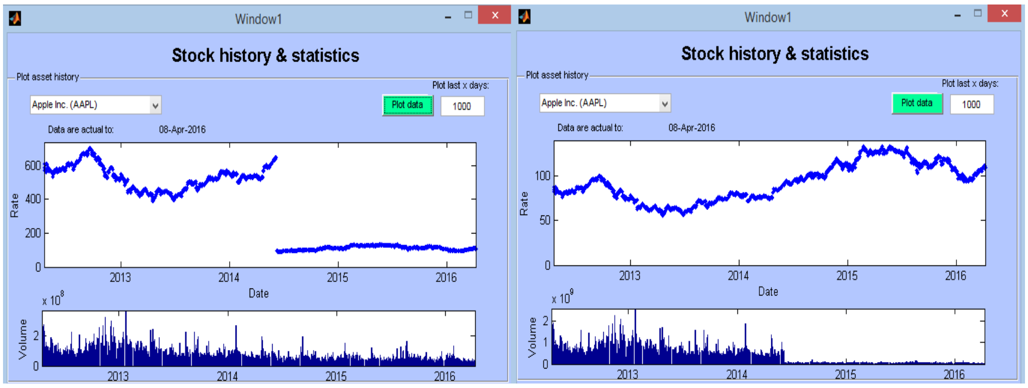


Figure 2 Sample asset split - Apple Inc. - before and after smoothening

4.3 Matrix cleaning based on random matrix theory

For better understanding of the data we applied the method described in subsection 3.2. Firstly we constructed the empirical measured correlation matrix $\boldsymbol{E}$ by using several different numbers of observations (T) and we analysed the distribution of the eigenvalues. We compared the empirical distribution of the eigenvalues of the correlation matrix with the theoretical prediction given by (6) based on assumption that the correlation matrix is purely random. The results are summarized in Table 1.

# observation	% of $\lambda < \lambda_{min}$	% of $\lambda_{min} < \lambda < \lambda_{max}$	% of $\lambda_{max} < \lambda$
15	7%	91%	3%
30	20%	75%	5%
50	33%	60%	7%
100	54%	38%	8%
200	74%	19%	8%

Table 1 Comparing eigenvalues of empirical correlation matrix and Marčenko- Pastur bounders.

From these results we can see that the important information about asset mutual connections is carried by 3 to 7% of eigenvalues of the correlation matrix. By increasing the number of observations on which is based the correlation matrix estimation, slightly increases the number of non-random correlations, but also increases the instability of the correlation matrix since its eigenvalues are very small. As an optimal number of observations in this case seems to be number between 30 and 50.

Figure 3 shows the results of our experiments on the data with 50 observations used for estimation of the correlation matrix. There is histogram of eigenvalues of the empirical correlation matrix and for comparison we plotted the density of (6) for $Q = 3.8462$ and $\sigma = 0.660$. A better fit can be obtained with a smaller value of $\sigma = 0.462$ (dashed blue line).

This result is in accordance with similar articles, for example [7] or [11] and shows problems related to correct market risk estimation.

5 Conclusion

The goal of this article was to develop an advance approach for stability treatment of portfolio optimization. We have developed two methods to minimize the effects of unstability and tested these methods on real data. For purposes of comparison of traditional Markowitz model and the weighting modification we created a SW solution StockMaTT, where we could try to simulate investments with both methods and different parameters. As we expected the traditional Markowitz model is very sensitive on input data and using the weighting we obtained different results. To sum up, as a possible treatment of the unstability can be recommended both methods described in this article. A potential topic for the following studies could be the correct estimation of weighting parameter.

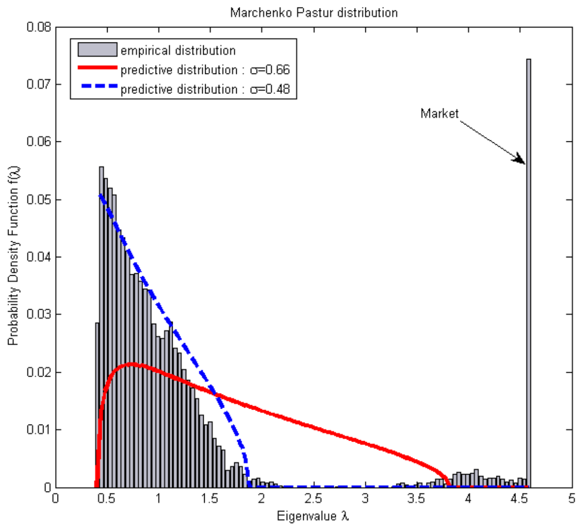


Figure 3 Empirical and prediction distribution of eigenvalues for $T = 50$.

Acknowledgements

This article was supported by the European Regional Development Fund (ERDF), project "NTIS - New Technologies for the Information Society", European Centre of Excellence, CZ.1.05/1.1.00/02.0090.

References

[1] Daly, J., Crane, M. and Ruskin, H. J.: Random matrix theory filters in portfolio optimisation: a stability and risk assessment. *Physica A: Statistical Mechanics and its Applications*, 387.16 (2008), pp. 4248–4260.

[2] Danielsson, J.: *Financial risk forecasting: the theory and practice of forecasting market risk with implementation in R and Matlab* (Vol. 588). John Wiley & Sons, 2011.

[3] Dupačová, J.: Markowitzův model optimální volby portfolia - předpoklady, data, alternativy *(n.d.). Retrieved April 25, 2016, from http://msekce.karlin.mff.cuni.cz/ dupacova/downloads/Markowitz.pdf*

[4] Dyson, Freeman J.: Correlations between eigenvalues of a random matrix. *Comm. Math. Phys.* 19, no. 3 (1970), pp. 235–250.

[5] Elton, E. J., Gruber, M. J., Brown, S. J. and Goetzmann, W. N.: *Modern portfolio theory and investment analysis.* John Wiley & Sons, 2009.

[6] Gilli, M., Maringer, D. and Schumann, E.: *Numerical methods and optimization in finance.* Academic Press, 2011.

[7] Laloux, L., Cizeau, P., Potters, M. and Bouchaud, J. P.: Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*, 3.03 (2000), pp. 391–397.

[8] Marčenko, V. A. and Pastur, L. A.: Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1.4 (1967), 457.

[9] Mehta, M. L.: *Random matrices* (Vol. 142). Academic Press, 2004.

[10] Pavelec, J.: *Programový nástroj pro volbu optimálního portfolia.* Diplomová práce. Západočeská univerzita v Plzni. Fakulta aplikovaných věd, 2013.

[11] Potters, M., Bouchaud, J. P. and Laloux, L.: Financial applications of random matrix theory: Old laces and new pieces. *arXiv preprint physics/0507111* (2005).

Fuzzy Decision Matrices Viewed as Fuzzy Rule-Based Systems

Ondřej Pavlačka[1], Pavla Rotterová[2]

Abstract. A decision matrix represents a particular problem of decision making under risk. Elements of the matrix express the consequences if a decision-maker chooses a particular alternative and a particular state of the world occurs. Assuming the probabilities of the states of the world are known, alternatives are compared on the basis of the expected values and variances of their consequences.

In practice, the states of the world are often described only vaguely; they are expressed by fuzzy sets on the universal set on which the probability distribution is given. In literature, the common approach in such a case consists in computing crisp probabilities of the fuzzy states of the world by formula proposed by Zadeh. In the paper, we first discuss the problems connected with the common approach. Consequently, we introduce a new approach, in which a decision matrix with fuzzy states of the world does not describe discrete random variables but fuzzy rule-based systems. We illustrate the problem by examples.

Keywords: decision matrices, decision making under risk, fuzzy states of the world, fuzzy rule-based systems.

JEL classification: C44
AMS classification: 90B50

1 Introduction

In decision making under risk, decision matrices (see Tab. 1) are often used as a tool of risk analysis (see e.g. [1, 3, 8]). They describe how consequences of alternatives $x_1, \dots, x_n$ depend on the fact which of possible and mutually disjoint states of the world $S_1, \dots, S_m$ will occur. The probabilities of occurrences of the states of the world are given by $p_1, \dots, p_m$. Thus, the consequence of choosing an alternative x_i, $i \in \{1, \dots, n\}$, is a discrete random variable H_i that takes on the values $h_{i,1}, \dots, h_{i,m}$ with probabilities $p_1, \dots, p_m$. The alternatives are usually compared on the basis of the expected values $EH_1, \dots, EH_n$ and variances $var\, H_1, \dots, var\, H_n$ of their consequences.

	S_1	S_2	$\cdots$	S_m		
	p_1	p_2	$\cdots$	p_m		
x_1	$h_{1,1}$	$h_{1,2}$	$\cdots$	$h_{1,m}$	EH_1	$var\, H_1$
x_2	$h_{2,1}$	$h_{2,2}$	$\cdots$	$h_{2,m}$	EH_2	$var\, H_2$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_n	$h_{n,1}$	$h_{n,2}$	$\cdots$	$h_{n,m}$	EH_n	$var\, H_n$

Table 1 A decision matrix.

In practice, the states of the world are often specified only vaguely, like e.g. "inflation is low", etc. Talašová and Pavlačka [7] showed that in such a case it is more appropriate to model the states of the world by fuzzy sets on the universal set on which the probability distribution is given. They proposed to

[1]Palacký University Olomouc, Faculty of Science, Department of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic, ondrej.pavlacka@upol.cz

[2]Palacký University Olomouc, Faculty of Science, Department of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic, pavla.rotterova01@upol.cz

proceed in the same way as in the case of crisp (i.e. exactly described) states of the world; they set the probabilities of the fuzzy states of the world applying the formula proposed by Zadeh [9] and treated the consequences of alternatives as discrete random variables.

However, Pavlačka and Rotterová [4] showed that Zadeh's probabilities of fuzzy events lack the common interpretation of a probability measure. Another problem is what does it mean to say that the particular fuzzy state of the world will occur (see the discussion in Section 2). Therefore, in Section 3, a new approach to this problem will be introduced, in which a decision matrix with fuzzy states of the world does not describe discrete random variables but fuzzy rule-based systems.

2 A decision matrix with fuzzy states of the world according to [7]

First, let us describe and analyse the approach to the extension of a decision matrix to the case of fuzzy states of the world that was considered by Talašová and Pavlačka [7].

Let us assume that a probability space $(\Omega, \mathcal{A}, P)$ is given, where Ω denotes a non-empty set of all elementary events, $\mathcal{A}$ represents the set of all considered random events ($\mathcal{A}$ forms a σ-algebra of subsets of Ω), and $P : \mathcal{A} \to [0,1]$ is a probability measure that assigns to each random event $A \in \mathcal{A}$ its probability $P(A) \in [0,1]$.

A vaguely defined state of the world can be appropriately expressed by a fuzzy set S on Ω. A *fuzzy set S on Ω* is determined by the membership function $\mu_S : \Omega \to [0,1]$. The interpretation of membership degrees $\mu_S(\omega)$, $\omega \in \Omega$, is explained in the following example.

Example 1. Let us consider a vaguely defined state of the economy "inflation is low", denoted by S. Let Ω be a set of all inflation rates and let an inflation rate $\omega \in \Omega$ occur. If $\mu_S(\omega) = 1$, then we would definitely say that inflation is low. If $\mu_S(\omega) \in (0,1)$, then we would say that inflation is low only partly. Finally, if $\mu_S(\omega) = 0$, then we would say that inflation is not low.

As the probability space $(\Omega, \mathcal{A}, P)$ is given, Talašová and Pavlačka [7] assumed that fuzzy states of the world are expressed by fuzzy sets on Ω that are called fuzzy random events. A *fuzzy random event* S is a fuzzy set on Ω whose membership function μ_S is $\mathcal{A}$-measurable (see Zadeh [9]). This assumption means that the α-cuts $S_\alpha := \{\omega \in \Omega \mid \mu_S(\omega) \geq \alpha\}$, $\alpha \in (0,1]$, are random events, i.e. $S_\alpha \in \mathcal{A}$ for any $\alpha \in (0,1]$.

For the case of fuzzy random events, Zadeh [9] extended the given probability measure P in the following way: A probability $P_Z(S)$ of a fuzzy random event S is defined as

$$P_Z(S) := E(\mu_S) = \int_\Omega \mu_S(\omega) dP. \tag{1}$$

The existence of the above Lebesgue-Stieltjes integral follows directly from the assumption that μ_S is $\mathcal{A}$-measurable. It can be easily shown (see e.g. [2, 4, 9]) that the mapping P_Z possesses analogous mathematical properties as a probability measure P, and thus, it can be called a probability measure.

Remark 1. Let us note that any crisp set $S \subseteq \Omega$ can be viewed as a fuzzy set of a special kind; the membership function μ_S coincides in such a case with the characteristic function χ_S of S. In such a case, $S_\alpha = S$ for all $\alpha \in (0,1]$. This convention allows us to consider crisp states of the world as a special kind of fuzzy states of the world. It can be easily seen that for any crisp random event $S \in \mathcal{A}$, $P_Z(S) = P(S)$.

Talašová and Pavlačka [7] considered the following extension of the decision matrix given by Tab. 2[1]: The fuzzy states of the world are fuzzy random events $S_1, \ldots, S_m$ that form a *fuzzy partition of* Ω, i.e. $\sum_{j=1}^m \mu_{S_j}(\omega) = 1$ for any $\omega \in \Omega$. The probabilities $p_{Z1}, \ldots, p_{Zm}$ of the fuzzy states of the world are given by $p_{Zj} := P_Z(S_j)$, $j = 1, \ldots, m$. Since $S_1, \ldots, S_m$ form a fuzzy partition of Ω, $\sum_{j=1}^m p_{Zj} = 1$. Therefore, Talašová and Pavlačka [7] treated the consequence of choosing an alternative x_i, $i = 1, \ldots, n$, as a discrete random variable H_i^Z that takes on the values $h_{i,1}, \ldots, h_{i,m}$ with probabilities $p_{Z1}, \ldots, p_{Zm}$.

[1] Besides the fuzzy states of the world, Talašová and Pavlačka [7] considered also fuzzy consequences of alternatives $H_{i,j}$ instead of the crisp consequences $h_{i,j}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. As we are focused strictly on the fuzzification of the states of the world here, only crisp consequences will be considered further.

They compared the alternatives $x_1, \ldots, x_n$ on the basis of the expected values $EH_1^Z, \ldots, EH_n^Z$ and variances $var\ H_1^Z, \ldots, var\ H_n^Z$ of their consequences, computed for $i = 1, \ldots, n$ as follows:

$$EH_i^Z = \sum_{j=1}^{m} p_{Zj} \cdot h_{i,j}, \tag{2}$$

and

$$var\ H_i^Z = \sum_{j=1}^{m} p_{Zj} \cdot (h_{i,j} - EH_i^Z)^2. \tag{3}$$

	S_1	S_2	$\cdots$	S_m		
	p_{Z1}	p_{Z2}	$\cdots$	p_{Zm}		
x_1	$h_{1,1}$	$h_{1,2}$	$\cdots$	$h_{1,m}$	EH_1^Z	$var\ H_1^Z$
x_2	$h_{2,1}$	$h_{2,2}$	$\cdots$	$h_{2,m}$	EH_2^Z	$var\ H_2^Z$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_n	$h_{n,1}$	$h_{n,2}$	$\cdots$	$h_{n,m}$	EH_n^Z	$var\ H_n^Z$

Table 2 A fuzzy decision matrix considered in [7].

Thus, we can see that by applying the Zadeh's crisp probabilities of fuzzy states of the world, the extension of the decision matrix is straightforward. Now, let us discuss the problems that are connected with this way of extension of a decision matrix tool.

As it is written in Introduction, the element $h_{i,j}$ of the matrix given by Tab. 1 describes the consequence of choosing the alternative x_i if the state of the world S_j occurs. If we consider fuzzy states of the world instead of crisp ones, a natural question arises: What does it mean to say "if the fuzzy state of the world S_j occurs"? Let us suppose that some $\omega \in \Omega$ occurred. If $\mu_{S_j}(\omega) = 1$, then it is clear that the consequence of choosing the alternative x_i is exactly $h_{i,j}$. However, what is the consequence of choosing x_i if $0 < \mu_{S_j}(\omega) < 1$ (which also means that $0 < \mu_{S_k}(\omega) < 1$ for some $k \neq j$)? Thus, perhaps it is not appropriate in case of a decision matrix with fuzzy states of the world to treat the consequence of choosing x_i as a discrete random variable H_i^Z that takes on the values $h_{i,1}, \ldots, h_{i,m}$.

Moreover, it was pointed out by Rotterová and Pavlačka [5] that the Zadeh's probabilities $p_{Z1}, \ldots, p_{Zm}$ of fuzzy states of the world express the expected membership degrees, in which the particular states of the world will occur. Thus, they do not have in general the common probabilistic interpretation - a measure of a chance that a given event will occur in the future, which is desirable in case of a decision matrix.

Therefore, we cannot say that the values $EH_1^Z, \ldots, EH_n^Z$, given by (2), and $var H_1^Z, \ldots, var H_n^Z$, given by (3), express the expected values and variances of consequences of the alternatives. Hence, ordering of the alternatives based on these characteristics is questionable.

3 Fuzzy rule-based systems derived from a decision matrix with fuzzy states of the world

In this section, let us introduce a different approach to the model of decision making under risk described by the decision matrix with fuzzy states of the world presented in Tab. 2. Taking into account the problems discussed in the previous section, we suggest not to treat the consequence of choosing an alternative x_i, $i \in \{1, \ldots, n\}$, as a discrete random variable H_i^Z taking on the values $h_{i,1}, \ldots, h_{i,m}$ with the probabilities $p_{Z1}, \ldots, p_{Zm}$. Instead of this, we propose to treat the information about the consequence of choosing x_i as the following basis of If-Then rules:

$$\begin{aligned}
&\text{If the state of the world is } S_1, \text{ then the consequence of } x_i \text{ is } h_{i,1}.\\
&\text{If the state of the world is } S_2, \text{ then the consequence of } x_i \text{ is } h_{i,2}.\\
&\ldots\\
&\text{If the state of the world is } S_m, \text{ then the consequence of } x_i \text{ is } h_{i,m}.
\end{aligned} \tag{4}$$

For computing the output of the fuzzy rule-based system (4), it is appropriate to apply the so-called *Sugeno method of fuzzy inference*, introduced by Sugeno [6]. For any $\omega \in \Omega$, the consequence of choosing x_i is given by

$$H_i^S(\omega) = \frac{\sum_{j=1}^m \mu_{S_j}(\omega) \cdot h_{i,j}}{\sum_{j=1}^m \mu_{S_j}(\omega)} = \sum_{j=1}^m \mu_{S_j}(\omega) \cdot h_{i,j}, \tag{5}$$

where we used the assumption $\sum_{j=1}^m \mu_{S_j}(\omega) = 1$ for any $\omega \in \Omega$. Let us note that the assumption that the fuzzy states of the world $S_1, \ldots, S_m$ form a fuzzy partition of Ω can be omitted in this approach.

Since we operate within the given probability space $(\Omega, \mathcal{A}, P)$, H_i^S is a random variable. A distribution function of H_i^S is given for any $h \in \mathbb{R}$ as follows:

$$F_{H_i^S}(h) = P(H_i^S \leq h) = \int_{\{\omega \in \Omega | H_i^S(\omega) \leq h\}} dP.$$

Remark 2. It can be easily seen from (5) that in the case of crisp states of the world $S_1, \ldots, S_m$, the random variables $H_1^S, \ldots, H_n^S$ coincide with discrete random variables $H_1, \ldots, H_n$ taking on the values $h_{i,1}, \ldots, h_{i,m}$, $i = 1, \ldots, n$, with probabilities $p_j = P(S_j)$, $j = 1, \ldots, m$. Hence, this new approach can be also considered as an extension of a decision matrix to the case of fuzzy states of the world.

Analogously as in the previous approach, the alternatives $x_1, \ldots, x_n$ can be compared on the basis of the expected values and variances of $H_1^S, \ldots, H_n^S$. Let us derive now the formulas for their computation. In addition to that, we will compare these characteristics with the characteristics of $H_1^Z, \ldots, H_n^Z$.

First, let us show now that the expected values $EH_1^S, \ldots, EH_n^S$ coincide with $EH_1^Z, \ldots, EH_n^Z$. For $i = 1, \ldots, n$, we get

$$EH_i^S = \int_\Omega H_i^S(\omega) dP = \int_\Omega \sum_{j=1}^m \mu_{S_j}(\omega) \cdot h_{i,j}\, dP = \sum_{j=1}^m \int_\Omega \mu_{S_j}(\omega)\, dP \cdot h_{i,j} = \sum_{j=1}^m p_{Zj} \cdot h_{i,j} = EH_i^Z. \tag{6}$$

Hence, from the point of view of the expected values of consequences that are taken into account in comparison of alternatives, both the approaches are the same. Nevertheless, as it will be shown next, the variances of $H_1^S, \ldots, H_n^S$ are generally different from $var\ H_1^Z, \ldots, var\ H_n^Z$. For $i = 1, \ldots, n$, $var\ H_i^S$ is given as follows:

$$var\ H_i^S = E\left(H_i^S\right)^2 - \left(EH_i^S\right)^2 = \int_\Omega H_i^S(\omega)^2\, dP - \left(EH_i^S\right)^2 = \int_\Omega \left(\sum_{j=1}^m \mu_{S_j}(\omega) \cdot h_{i,j}\right)^2 dP - \left(EH_i^S\right)^2.$$

For $var\ H_i^Z$, $i = 1, \ldots, n$, it holds that

$$var\ H_i^Z = E\left(H_i^Z\right)^2 - \left(EH_i^Z\right)^2 = \sum_{j=1}^m p_{Zj} \cdot h_{i,j}^2 - \left(EH_i^Z\right)^2 = \int_\Omega \sum_{j=1}^m \mu_{S_j}(\omega) \cdot h_{i,j}^2\, dP - \left(EH_i^Z\right)^2.$$

Hence, employing $\left(EH_i^S\right)^2 = \left(EH_i^Z\right)^2$,

$$\begin{aligned} var\ H_i^Z - var\ H_i^S &= \int_\Omega \sum_{j=1}^m \mu_{S_j}(\omega) \cdot h_{i,j}^2\, dP - \int_\Omega \left(\sum_{j=1}^m \mu_{S_j}(\omega) \cdot h_{i,j}\right)^2 dP \\ &= \int_\Omega \left[\sum_{j=1}^m \mu_{S_j}(\omega) \cdot h_{i,j}^2 - \left(\sum_{j=1}^m \mu_{S_j}(\omega) \cdot h_{i,j}\right)^2\right] dP \geq 0, \end{aligned}$$

since the integrand is clearly non-negative (it represents the variance of a discrete random variable that takes the values $h_{i,1}, \ldots, h_{i,m}$ with "probabilities" $\mu_{S_1}(\omega), \ldots, \mu_{S_m}(\omega)$). Moreover, it is obvious that the difference of variances is equal to zero, if and only if $h_{i,j} = h_{i,k}$ for any $j \neq k$ such that both $E\mu_{S_j}$ and $E\mu_{S_k}$ are positive.

Thus, whereas the expected values $EH_1^S, \ldots, EH_n^S$ and $EH_1^Z, \ldots, EH_n^Z$ are the same, the variances $var\, H_1^S, \ldots, var\, H_n^S$ and $var\, H_1^Z, \ldots, var\, H_n^Z$ generally differ. This means that ordering of the alternatives that is based on the expected values and variances of the consequences can be different for both the above described approaches. Let us illustrate this fact by the following numerical example.

Example 2. Let us consider the following situation: We can realize one of the two possible projects, denoted by x_1 and x_2. The result depends solely on the fact what kind of a government coalition will be established after the parliamentary election. Let us assume that only the six possible coalitions, denoted by $\omega_1, \ldots, \omega_6$, can be established after the election. For the sake of simplicity, let the probabilities of establishing each coalition be the same. Thus, the probability space $(\Omega, \mathcal{A}, P)$ is given, where $\Omega = \{\omega_1, \ldots, \omega_6\}$, $\mathcal{A}$ is the family of all subsets of Ω, and P is the probability measure such that $P(\{\omega_k\}) = 1/6$, $k = 1, \ldots, 6$, and $P(A) = \sum_{\{k|\omega_k \in A\}} P(\{\omega_k\})$, for any $A \subseteq \Omega$.

We distinguish three possibilities (states of the world) - a right coalition (S_1), a centre coalition (S_2), and a left coalition (S_3). Let the above mentioned three states of the world be expressed by the following fuzzy sets defined on Ω:

$$\begin{aligned}
S_1 &= \left\{ {}^{1}|_{\omega_1}, {}^{0.8}|_{\omega_2}, {}^{0.2}|_{\omega_3}, {}^{0}|_{\omega_4}, {}^{0}|_{\omega_5}, {}^{0}|_{\omega_6} \right\}, \\
S_2 &= \left\{ {}^{0}|_{\omega_1}, {}^{0.2}|_{\omega_2}, {}^{0.8}|_{\omega_3}, {}^{1}|_{\omega_4}, {}^{0.5}|_{\omega_5}, {}^{0}|_{\omega_6} \right\}, \\
S_3 &= \left\{ {}^{0}|_{\omega_1}, {}^{0}|_{\omega_2}, {}^{0}|_{\omega_3}, |_{\omega_4}, {}^{0.5}|_{\omega_5}, {}^{1}|_{\omega_6} \right\},
\end{aligned}$$

where elements of the sets are in the form ${}^{\mu_{S_j}(\omega_k)}|_{\omega_k}$, $j = 1, 2, 3$, and $k = 1, \ldots, 6$. How the future yield (in %) depends on the fact which of these three types of a coalition will occur in the future is described in Tab. 3.

	S_1	S_2	S_3
x_1	15	0	-4
x_2	15	-3	0

Table 3 The future yields (in %) of x_1 and x_2 based on the type of a coalition.

First, let us compute the expected values and variances of the random variables H_1^Z and H_2^Z. The probabilities of the fuzzy states of the world S_1, S_2, and S_3, computed according to (1), are given as follows:

$$P_Z(S_1) = 0.333,\ P_Z(S_2) = 0.417,\ \text{and}\ P_Z(S_3) = 0.25.$$

Hence, we get

$$\begin{aligned}
EH_1^Z &= 0.333 \cdot 15 + 0.417 \cdot 0 + 0.25 \cdot (-4) = 4\ \%, \\
EH_2^Z &= 0.333 \cdot 15 + 0.417 \cdot (-3) + 0.25 \cdot 0 = 3.75\ \%,
\end{aligned}$$

and

$$\begin{aligned}
var\ H_1^Z &= 0.333 \cdot (15-4)^2 + 0.417 \cdot (0-4)^2 + 0.25 \cdot (-4-4)^2 = 63, \\
var\ H_2^Z &= 0.333 \cdot (15-3.75)^2 + 0.417 \cdot (-3-3.75)^2 + 0.25 \cdot (0-3.75)^2 = 64.69.
\end{aligned}$$

By applying the rule of maximization of the expected value and minimization of the variance, we would choose the alternative x_1 over x_2. However, as it was stated in Section 2, the interpretation of the characteristics EH_i^Z and $var\ H_i^Z$, $i = 1, 2$, is questionable. Thus, the selection of x_1 is not trustworthy.

Now, let us construct the random variables H_1^S and H_2^S. Since the universal set Ω is a discrete set, H_1^S and H_2^S are discrete random variables. According to (5), we obtain the following:

$$H_1^S(\omega_1) = 15\ \%,\ H_1^S(\omega_2) = 12\ \%,\ H_1^S(\omega_3) = 3\ \%,\ H_1^S(\omega_4) = 0\ \%,\ H_1^S(\omega_5) = -2\ \%,\ \text{and}\ H_1^S(\omega_6) = -4\ \%,$$

$$H_2^S(\omega_1) = 15\ \%,\ H_2^S(\omega_2) = 11.4\ \%,\ H_2^S(\omega_3) = 0.6\ \%,\ H_2^S(\omega_4) = -3\ \%,\ H_2^S(\omega_5) = -1.5\ \%,\ \text{and}\ H_2^S(\omega_6) = 0\ \%.$$

Both the random variables H_1^S and H_2^S taking on these values with the probabilities equal to 1/6.

According to (6), the expected values EH_1^S and EH_2^S coincide with EH_1^Z and EH_2^Z, i.e. $EH_1^S = 4\ \%$ and $EH_2^S = 3.75\ \%$. The variances of H_1^S and H_2^S are given as follows:

$$\begin{aligned}
var\ H_1^S &= \frac{\sum_{k=1}^{6} (H_1^S(\omega_k) - 4)^2}{6} = 50.33, \\
var\ H_2^S &= \frac{\sum_{k=1}^{6} (H_2^S(\omega_k) - 3.75)^2}{6} = 47.03.
\end{aligned}$$

Thus, we can see that under this approach, we would not be able to decide between x_1 and x_2 according to the rule of maximization of the expected value and minimization of the variance, since $EH_1^S > EH_2^S$ and $var\ H_1^S > var\ H_2^S$.

4 Conclusion

We have dealt with the problem of extension a decision matrix to the case of fuzzy states of the world. We have analysed the approach to this problem proposed by Talašová and Pavlačka [7] that is based on applying the Zadeh's probabilities of the fuzzy states of the world. We have found out that the meaning of obtained characteristics of the consequences of alternatives, namely the expected values and variances, is questionable. Therefore, we have introduced a new approach that consists in deriving fuzzy rule-based systems from a decision matrix with fuzzy states of the world. In such a case, the obtained characteristics of consequences, based on which the alternatives are compared, are clearly interpretable. We have proved that the resulting expected values of consequences are for both the approaches the same, whereas the variances generally differ. In numerical example, we have shown that the final ordering of the alternatives according to the both approaches can be different.

Next research in this field could be focused on the cases, where the consequences of alternatives are given by fuzzy numbers, and/or where the underlying probability measure is fuzzy.

Acknowledgements

Supported by the project No. GA 14-02424S of the Grant Agency of the Czech Republic and by the grant IGA_PrF_2016_025 *Mathematical Models* of the *Internal Grant Agency of Palacký University Olomouc*. The support is greatly acknowledged.

References

[1] Chankong, V., Haimes, Y.Y.: *Multiobjective decision making: theory and methodology.* Elsevier Science B.V., New York, 1983.

[2] Negoita, C.V., Ralescu, D.: *Applications of Fuzzy Sets to System Analysis.* Birkhuser Verlag–Edituria Technica, Stuttgart, 1975.

[3] Özkan, I., Türken, IB.: Uncertainty and fuzzy decisions. In: *Chaos Theory in Politics*, Springer Netherlands, 2014, 17–27.

[4] Pavlačka, O., Rotterová, P.: Probability of fuzzy events. In: *Proceedings of the 32nd International Conference Mathematical Methods in Economics* (J. Talašová, J. Stoklasa, T. Talášek, eds.), Palacký University Olomouc, Olomouc, 2014, 760–765.

[5] Rotterová, P., Pavlačka, O.: Probabilities of Fuzzy Events and Their Use in Decision Matrices. *International Journal of Mathematics in Operational Research.* To Appear.

[6] Sugeno, M.: *Industrial applications of fuzzy control.* Elsevier Science Pub. Co., New York, 1985.

[7] Talašová, J., Pavlačka, O.: Fuzzy Probability Spaces and Their Applications in Decision Making. *Austrian Journal of Statistics* **35**, 2&3 (2006), 347–356.

[8] Yoon, K.P., Hwang, Ch.: *Multiple Attribute Decision Making: An Introduction.* SAGE Publications, California, 1995.

[9] Zadeh, L.A.: Probability Measures of Fuzzy Events. *Journal of Mathematical Analysis and Applications* **23**, 2 (1968), 421–427.

Valuation of the externalities from biogas stations

Marie Pechrová[1], Václav Lohr[2]

Abstract. The aim of the paper is to valuate the externalities caused by projects realized with financial support of Rural Development Program for the Czech Republic for years 2014-2020, particularly the externalities from biogas stations. We applied hedonic model valuation as we suppose that it can give the most objective price of externalities related to the biogas stations as the price is derived from the value of real estates.
A case study was done on eight biogas stations in Jihomoravsky region. The data were obtained from real estate server at the beginning of 2.Q of 2016. Price of flats for sale was explained in the regression model by number of rooms, acreage and distance to the biogas station. Linear and log-linear relation was taken into account. As expected, the effect of the distance from biogas station on the flat prices is negative as its presence lowers the price by 0.15% according to linear regression model, or by 0.40% based on log-linear form of the model.

Keywords: biogas station, rural development program, externalities

JEL Classification: D61, C21
AMS Classification: 62J05

1 Introduction

The aim of the paper is to evaluate the externalities caused by projects realized with financial support of Rural Development Program for the Czech Republic for years 2014-2020 (RDP), particularly the externalities from biogas stations (BGS). "Externalities are defined as benefits or costs, generated as by-products of an economic activity that do not accrue to the parties involved in the activity." [1] Externalities are generated as side effects of an activity and affect (usually negatively) the production or consumption possibilities of other economic agents without permission or compensation. Besides, the market prices of the externalities often do not exist. Therefore, they have to be monetary valuated. "Monetary valuation is the practice of converting measures of social and biophysical impacts into monetary units and is used to determine the economic value of non-market goods, i.e. goods for which no market exists." [8]

There is a complex of externalities related to the BGS: noise from operation of the BGS or from transport of the feed, odor, esthetic point of view etc.; are negatively perceived by local inhabitants. In this sense, the presence of the BGS lowers their welfare. Therefore, when selecting the projects to be financed from RDP, not only economic, but also social point of view should be applied. Cost benefit analysis (CBA) is recommended as the suitable method for project evaluation. CBA measures inputs and outputs in monetary units. "It understands the benefits as every increase of the utility and costs as every decrease of the utility, what is sometimes hard to be determined in monetary terms." [7]

There are various methods to monetary valuate externalities related to certain type of projects. Particularly, hedonic price method (HPM) was selected and applied in this article. First, there are presented the results or researches which used HPM. Next section describes the method in detail together with used data and models. Then the results of the analysis are presented and discussed. Last section concludes.

1.1 Results of previous researches

To the best authors' knowledge, there is no study which would assess the externalities related to the BGS. Therefore, we present results of researches aimed on evaluation of the odor from animal sites or landfills as we suppose that odor nuisance is one of the main externalities from BGS despite the fact that the researchers and experts are not agreed on whether or how much the biogas station smells. [7]

Van Broeck, Bogaert and de Meyer [13] examined the externalities related to odor from animal waste treatment facility based on 1200 real estate prices in Belgium. They compared those houses situate within and outside odor affected areas at the same statistical sector. Log-linear model showed that an extra sniffing unit corresponds to a decrease of the average net actual value by 650 euros or by 0.4%. Linear regression model estimated de-

[1] Institute of Agricultural Economics and Information, Department of the impacts of agricultural policy, Mánesova 75, Prague, Czech Republic, e-mail: pechrova.marie@uzei.cz
[2] Czech University of Life Sciences Prague, Faculty of Economics and Management, Department of Information Technologies, Kamýcká 129, Prague, Czech Republic, e-mail: lohr@pef.czu.cz

crease by 1526 euros (0.8%). Also Milla, Thomas and Ansine [4] found negative and significant influence of proximity to swine facilities on the selling price of residential properties. Similar results were achieved also by Palmquist, Roka and Vukina [6] who analyzed the effect of large-scale hog operations on residential property values. “An index of hog manure production at different distances from the houses was developed. It was found that proximity caused statistically significant reduction in house prices of up to 9 percent depending on the number of hogs and their distance from the house.” [6]

Bouvier et al. [2] examined six landfills, which differed in size, operating status, and history of contamination, and estimated by multiple regression model the effect of each landfill on rural residential property values. “In five of the landfills, no statistically significant evidence of an effect was found. In the remaining case, evidence of an effect was found, indicating that houses in close proximity to this landfill suffered an average loss of about six percent in value.” [2] The proximity to the site indeed plays important role as the effect is visible only to the certain distance. Nelson, Genereux and Genereux [5] examined the price effects of one landfills on 708 houses’ values in Minnesota, USA in 1980s. They came to the conclusion that the influence is visible up to 2-2.5 miles. At the landfill boundary, the prices decrease by 12% and about 1 mile away by 6% at about 1 mile.

Beside the distance, also the character of the area is important determinant of the effect. Reichert, Small and Mohanty [10] analyzed the effect of five municipal landfills on residential property values in a major metropolitan area (Cleveland, Ohio, USA). They concluded that landfills would likely have an adverse impact upon housing values when the landfill was located within several blocks of an expensive housing area (the negative impact was between 5.5%-7.3% of market value depending upon the actual distance from the landfill), but in case of less expensive and older areas the landfill effect was considerably less pronounced (3%-4% of market value), and for predominantly rural areas there was no effect measured at all. The specifics of rural areas are analyzed for example by Šimpach and Pechrová [12]. Also the size of the landfill plays the role. Ready [9] found out that landfills that accept high volumes of waste (≥ 500 tons/day) decrease neighboring residential property values on average by 12.9%. This impact diminishes with distance at a gradient of 5.9% per mile in this case while lower-volume landfills decrease property values on average by 2.5% with a gradient of 1.2% per mile. Besides 20-28% of low-volume landfills have no impact at all on nearby property values. [9] Intensity of the decrease of real estate’s prices in similar scope was found also in California (USA) cities where industrial odors were present. Saphores and Aguilar-Benitez [11] found statistically significant reduction in house prices of up to 3.4% using GLS model with controlled heteroskedasticity and GIS software.

2 Data and methods

2.1 Hedonic pricing method

The essence of hedonic pricing method is the evaluation of the externality’s price based on the decrease of the value of properties located near the site which cause loss of benefits. It belongs to the type of methods which are based on the alternative market (real estate market) when the real one does not exist. This method requires detailed data about the sales transactions and characteristics of the sold estates. When the frequency of property sales is not sufficiently high (for example in rural areas) the usage of HPM might be problematic. The price of the real estate P (or the logarithm of the price, depending on the chosen type of the functional form) is normally explained by three categories of variables (see equation 1): C – physical characteristics that contribute to the price of a house such as the number of rooms, lot size, date of construction, type of ownership, energy demand etc.), N – neighborhood characteristics (distance to the center, supermarket, school, park, etc.), E – characteristics of the surrounding environment (distance to the activity which nagging the local residents that is operation of BGS in our case).

$$P = f(C, N, E), \tag{1}$$

Hedonic price models are commonly estimated by the method of ordinary least squares (OLS) in cases when cross-sectional data are available or by fixed effects models when the data are of panel nature.

2.2 Data

Data of real estates for sale in Jihomoravsky region of the Czech Republic were gathered. There were originally 1480 real estates for sale at the beginning of the second quarter of year 2016 in size categories from 1+kk up to 5+1. The size of the flat was measured by the number of rooms (e.g. 1+1 or 2+kk are 2 rooms) and the acreage in m^2. The proxy for the externalities related to BGS was chosen the distance. The further is the house for sale; the lower is the price of the externality. There were 8 biogas stations operating in the region which used as a feed the organic waste from agricultural production. The characteristics of the BGS are displayed at Table 1.

Location	Power [kW]	Type of feed	Operational since	No. of nearby houses	Avg. distance from BGS
Čejč	1072	pig manure	2007	8	11.08
Hodonín	261	pig manure, bedding, maize silage, corn silage	autumn 2009	97	4.41
Mikulčice	500	silage, pig manure, farmyard manure, corn leftovers, moldings of tomatoes and apples	N/A	36	10.96
Mutěnice	526	pig manure, maize silage	2009	22	4.92
Šebetov	130	pig manure	1997	19	8.70
Suchohrdly u Miroslavi	495	pig manure, maize straw	2008	38	12.17
Švábenice	526	pig manure, maize silage	2008	47	9.47
Velký Karlov	1480	animal by-products, manure, waste	2006	51	11.56

Table 1 Characteristics of the biogas stations in the sample; **Source** [3], own elaboration

The distance to the BGS was measured based on the GPS system. Longitude and latitude of particular real estate and of particular BGS were recalculated to Cartesian coordinate system and based on Pythagoras theorem the shortest (air) distance was calculated. Consequently, only those estates which distance to the nearest biogas station was lower than 15 km were selected and used for consequent analysis. Also the cases with incomplete data were dropped. Finally there were 318 real estates included into the analysis.

Descriptive characteristics of the sample are displayed at Table 2. Average price of the real estate was 1.5 mil. CZK. On average, the flat had 3 rooms (2+1 or 3+kk) and its size was 72 m^2. The closest estate to the BGS was 240 meters, the upper threshold was set on 15 km. On average, the estates were 8 km far from BGS. Majority of flats was close to BGS station in Hodonín run by company M PIG spol. s r.o.; 51 real estates were less than 15 km from BGS in Velký Karlov (Jevišovice) operated by ZEVO spol. s r.o.; 47 were close to Švábenice, etc. Above stated determinants were included into model. Originally, also the distance to the nearest biggest city (center) was calculated, but it showed to be statistically insignificant. Hence, it was dropped from the analysis.

Variable	Mean	Std. dev.	Min	Max
y – price [CZK]	1 468 006	665 768	325 000	3 749 000
x_1 – no. of rooms	3.13	0.97	1	6
x_2 – size [m^2]	71.96	33.77	18	382
x_3 – distance to nearest BGS [km]	8.43	4.34	0.24	14.98

Table 2 Descriptive characteristics of the sample; **Source** own calculation

Data were checked for the presence of multicollinearity. The pair correlation coefficients were not higher than 0.8, hence, there was no multicollinearity in the data detected.

2.3 Models

Hedonic price function has no known explicit form. Therefore, various forms are usually used such as linear, log-linear, log–log, and linear with Box–Cox transformations of the price and continuous regressors. In our case two forms of a model were selected as optimal based on multiple coefficient of determination (R^2). Firstly, the relation between price and its determinants was modelled by linear function. Linear regression model estimated by OLS is defined as (2):

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u_t , \tag{2}$$

where y is explained variable, x are explanatory variables with parameters β, and u_t represents stochastic term. Second form of the model was exponential function (3). Hence, a growth model was constructed.

$$y = e^{\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u_t} \tag{3}$$

In order to estimate (2) by OLS, the model was linearized by natural logarithm to log-linear form (4).

$$\ln y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u_t \tag{4}$$

Another possible form (not applied in the paper) is exponential function linearized as (5).

$$\ln y = \beta_1 \ln x_1 + \beta_2 \ln x_2 + \beta_3 \ln x_3 + u_t \tag{5}$$

All models provided better results when the constant was not included. The calculations were done in econometric software Stata 11.2.

3 Results and Discussion

Results of the linear form of the model with estimated parameters are presented in equation (6) and Table 3.

$$y = 148\,187{,}10 x_1 + 10284{,}55 x_2 + 25934{,}39 x_{3_t}, \tag{6}$$

The parameters have expected sign. Increase of each explanatory variable by a unit cause increase of price by certain amount of money. Increase by 1 room represents average increase in price by 148 187 CZK. Taking into account that average price of 1-room flat was 1.01 mil. CZK, of 2-room flat 1.26 mil. CZK, 3-room flat 1.51 mil. CZK, 4-room flat 1.56 mil CZK, the intensity of increase seems reasonable.

When the flat is larger by 1 m^2, the price increases by 10 285 CZK. When the real estate is 1 km further from the nearest BGS station, its price is higher by 25 934 CZK. All coefficients are statistically significant at 95% level. Average price of the house (when data from Table 2 are put into the equation 5) is 1 421 953 mil. CZK. Based on elasticities, it can be seen that the acreage of the flat influences the price the most. Increase of acreage by 1% means increase in price by 0.52%. On the other hand, when the distance of the real estate to BGS decreases by 1%, the price decreases by 0.15%. Therefore, we may conclude that the presence of the BGS could cause the decrease in flat price by 0.15%.

The model fit the data well, from 85.60% (after adjustment). Also the whole model is statistically significant (F-value enables to reject H_0 that all parameters of explanatory variables are jointly equal to zero).

Source	Sum of squares	Degrees of freedom	Mean sum of squares
Model	7.08e^{14}	3	2.36e^{14}
Residual	1.18e^{14}	315	3.74e^{11}
Total	8.26e^{14}	218	2.60e^{12}

F(3, 315)	630.92
Prob > F	0.00
R^2	85.73%
Adj. R^2	85.60%
Root MSE	6.10e^5

y	Coefficient	Std. error	t-value	p-value	95% confidence interval		Average elasticity
x_1	148187.10	31218.22	4.75	0.000	86764.53	209609.70	0.33%
x_2	10284.55	1161.22	8.86	0.000	7999.82	12569.27	0.52%
x_3	25934.39	6877.98	3.77	0.000	12401.79	39466.98	0.15%

Table 3 Results of linear regression model; **Source** own calculation

Results of log-linear form of a model are displayed in equation (7) and Table 4.

$$\ln y = 2{,}6721 x_1 + 0{,}0211 x_2 + 0{,}3987 x_3 \tag{7}$$

Parameters of log-linear model could be interpreted in percentage terms. As expected when the distance increases by 1 km, price of the real estate increases, on average by 0.3987%. Increase of the number of rooms by 1 cause increase of price by 2.67% and increase by 1 m^2 by 0.02%. Explanatory power of the model is better than in previous case – the changes of real estate's price are explained by changes of explanatory variables from 93.89% (based on the adjusted coefficient of determination). The model as a whole is statistically significant as same as all its parameters.

Source	Sum of squares	Degrees of freedom	Mean sum of squares
Model	$5.95e^4$	3	19822.76
Residual	$3.83e^3$	315	12.17
Total	63301.51	318	199.06

F(3, 315)	1628.96
Prob > F	0.00
R^2	93.94%
Adj. R^2	93.89%
Root MSE	3.49

ln y	Coefficient	Std. error	t-value	p-value	95% confidence interval	
x_1	2.67	0.18	1.16	0.000	2.32	3.02
x_2	0.02	0.01	3.18	0.002	0.01	0.03
x_3	0.40	0.04	10.16	0.000	0.32	0.48

Table 4 Results of log-linear regression model; **Source** own calculation

The results of our research show lower intensity of price decrease than other researches. The closest are to the study of van Broeck, Bogaert and de Meyer [13]. They estimated the same intensity of price decrease based on log-linear model (0.4%), but their research had the advantage that dispersion studies of odor were available for the sites. Therefore, the decrease of price is related to the increase of amount of sniffing units as the real estate is closer to the animal waste treatment facility. In our case, the price decrease is measured by the distance to the BGS. Linear regression models provided different results – van Broeck, Bogaert and de Meyer [13] reported the decrease by 0.8%, but in our case the calculated elasticity for the average value of other variables revealed lower elasticity (0.52%).

Palmquist, Roa and Vukina [6] found statistically significant reduction in house prices of up to 9% in case of hog farms. This is relatively high intensity in comparison with our results taking into account the fact that BGS included in the research were mostly fed by pigs' manure and hence the nature of the odor is similar. Different type of odor (local industrial) can explain different intensity in case of research by Saphores and Aguilar-Benitez [11], where the decrease was 3.4%. Also the distance matters. As the examined houses in our study were the closest to Hodonín BGS, the decrease of price might be steeper. However, running separate model only for those houses near Hodonín BGS did not bring statistically significant results and hence the model is not presented in the article.

4 Conclusion

The aim of the paper was to evaluate the externalities related to the biogas stations (BGS). From a complex set of externalities which are related to the operations of BGS such as odor, noise, esthetical point of view etc., the odor was taken as the most important and the results of our research are compared to studies evaluating the price of smell. A hedonic pricing method (as one valuation methods which utilize substitute markets) was used. The price of the flat was explained by its number of rooms, acreage and distance to BGS. Data were obtained by special software from the internet server. They included 318 flats for sale in Jihomoravsky region from the beginning of second quarter of the year 2016. Linear and log-linear (exponential) types of models were estimated by ordinary least squares method.

The results of linear regression show that when the house is closer to the BGS by 1 km, its price is lower by 25 934 CZK. Using elasticity for average value it was calculated that when the distance to BGS decreases by 1%, the price of the house decreases by 0.15%. Based on log-linear model, we can conclude that when the distance decreases by 1 km, price of the real estate decreases by 0.3987%.

However, despite that the fit of both models is high; they are simplified to some extent as they do not include other price's determinants. It was not possible to use the software to download the data. To manually collect information about other characteristics of the flats which determine their price such as the type of ownership, whether there is cellar, balcony, garage etc. and include them into the model remains a challenge for future research. We will further develop our model and try set the price of externalities related to BGS as objective as possible so it will be possible to use it in CBA analysis of projects financed from RDP. Also the usage of contingent valuation method is planned.

Acknowledgements

The paper was supported by the thematic task No. 4107/2016 (No. 18/2016) of Institute of Agricultural Economics and Information.

References

[1] Birol, E.: Using Economic valuation techniques to inform water resources management: A survey and critical appraisal of available techniques and an application. *Science of the Total Environment* **365** (2006), 105 – 122.

[2] Bouvier, R. A., Halstead, J. M., Conway, K. S., Manalo, A. B.: The Effect of Landfills on Rural Residential Property Values: Some Empirical Evidence. *Journal of the Regional Analysis and Policy* **30** (2000), 23–37.

[3] CZBA (Czech Biogas Association). Mapa bioplynových stanic. [on-line]. Available from http://www.czba.cz/mapa-bioplynovych-stanic [cit. 15-03-2016]

[4] Milla, K., Thomas, M.H., Ansine, W.: Evaluating the Effect of Proximity to Hog Farms on Residential Property Values: A GIS – Based Hedonic Price Model Approach. *URISA Journal* **17** (2005), 27–32.

[5] Nelson, A., Genereux, J., Genereux, M.: Price Effects of Lands Fills on House Values. *Land Economics* **68** (1992), 259–365.

[6] Palmquist, R. B., Roka, F. M., Vukina, T.: Hog Operations, Environmental Effects, and Residential Property Values. *Land Economics* **73** (1997), 114–124.

[7] Pechrová, M.: Valuation of Externalities Related to Investment Projects in Agriculture. In: *Proceedings of the 20th International Conference Current Trends in Public Sector Research*. Masaryk's university, Brno, 2016, 352–360.

[8] Pizzol, M., Weidema, B., Brand, M., Osset, P.: Monetary valuation in Life Cycle Assessment: a review. *Journal of Cleaner Production* **86** (2015), 170–179.

[9] Ready, R. C.: Do Landfills Always Depress Nearby Property Values? *Rural Development Paper* **27** (2005), 1–29.

[10] Reichert, A., Small, M., Mohanty, S.: The Impact of Landfills on Residential Property Values. *Journal of Real Estate Research* **7** (1992), 297–314.

[11] Saphores, J.-D., Aguilar-Benitez, I.: Smelly local polluters and residential property values: A hedonic analysis of four Orange County (California) cities. *Estudios Económicos* **20** (2005), 197–218.

[12] Šimpach, O., Pechrová, M.: Rural Areas in the Czech Republic: How Do They Differ from Urban Areas? In: *Proceedings of the 22nd International Scientific Conference on Agrarian Perspectives - Development Trends in Agribusiness*. Czech University of Life Sciences Prague, Prague, 322–334.

[13] van Broeck, G., Bogaert, S., de Meyer, L.: *Odours and VOCs: Measurement, Regulation and Control Techniques*. Kassel university press, Diagonale 10, Kassel, 2009. ISBN: 978-3-89958-608-4.

Multiple criteria travelling salesman problem

Pekár Juraj[1], Lucia Mieresová[2], Reiff Marian[3]

Abstract: Presented paper address the problem of linking the traveling salesman problem with the multiple criteria iterative - interactive methods. Thus the aim of the paper is to solve of multiple criteria traveling salesman problem, since research publications in this area are less frequent in comparison to the classical traveling salesman problem. In addressing the problem, iterative method is used for finding the compromise solution of multiple criteria decision making problem. In application part of paper, we demonstrate the way how to solve the problem of multiple criteria traveling salesman problem on data from the region of Slovakia. The first part of the paper deals with the definition of classical traveling salesman problem, multiple criteria traveling salesman problem and their mathematical formulations. The second part contains description of multiple criteria iterative method. The specification and solutions of real problem is presented in the third part.

Key words: Multiple criteria problem, traveling salesman problem, iterative method.

JEL Classification: C02, C61
AMS Classification: 90C11, 90B06

1 Introduction

Vehicle routing problems ([2], [3], [4], [6]) are well known mainly for its utilization in many real-world applications. In this paper, we focused on analysis of the traveling salesman problem (TSP) and ways of its compromise solution computation, whereas the analysis is focused on multiple criteria modification of TSP compromise solution identification. Models of multiple criteria decision making offer solving tools for different types of problems, and in the our paper we have chosen to apply ISTM (Interactive Step Trade-off Method) on TSP. Interactive Step Trade-off Method [5], [7] reprezents iterative procedure, where the decision-maker based on on the decision-making process carried out by the two entities (analyst and decision-maker) select the alternatives from the set of effective solutions. In ISTM method, the decision maker, after obtaining the first solution has to decide, whether the target needs to be improved and should be maintained at least at current levels, or whether some certain part of it can be relaxed. Based on this compromise analysis, , ISTM procedure will find new effective solution that can satisfy decision maker preferences.

2 Modification of ISTM for solution of multiple criteria TSP

In the next section we present a modification ISTM for solution of multiple criteria traveling salesman problem. Steps of computation procedure ISTM for obtaining solution of multiple criteria TSP can be summarized as follows:

Step 1 Define multiple criteria optimization problem for solution of the routing problem TSP

$$\begin{aligned} \max f_1(x) &= -\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}^{(1)} x_{ij} \\ \max f_2(x) &= -\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}^{(2)} x_{ij} \\ &\dots \\ \max f_k(x) &= -\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}^{(k)} x_{ij} \end{aligned} \tag{1}$$

[1] University of Economics in Bratislava, juraj.pekar@euba.sk.
[2] University of Economics in Bratislava, mieresova.euba@gmail.com
[3] University of Economics in Bratislava, reiff@euba.sk

$$\sum_{i=1}^{n} x_{ij} = 1, \qquad j = 1, 2, 3, \dots, n \qquad i \neq j$$
$$\sum_{j=1}^{n} x_{ij} = 1, \qquad i = 1, 2, 3, \dots, n \qquad i \neq j$$
$$y_i - y_j + n x_{ij} \leq n - 1, \qquad i, j = 2, 3, \dots, n$$
$$x_{ij} \in \{0,1\}, \qquad i, j = 1, 2, \dots, n$$

Step 2 Utilize the method for generating efficient alternatives based on an ideal alternative $\mathbf{f}^* = (f_1^*, f_2^*, \dots f_k^*)$ [5] for specified values of weights λ_j $(j = 1, \dots, k)$.

$$\begin{array}{ll} \min \; f(\mathbf{X}, \mathbf{y}, \alpha) = \alpha & \\ -\sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}^{(1)} x_{ij} + \frac{1}{\lambda_1} \alpha \geq f_1^* & \\ -\sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}^{(2)} x_{ij} + \frac{1}{\lambda_2} \alpha \geq f_2^* & \\ \vdots & \\ -\sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}^{(k)} x_{ij} + \frac{1}{\lambda_k} \alpha \geq f_k^* & i, j = 1, 2, \dots, k \\ \sum_{i=1}^{n} x_{ij} = 1, \; j = 1, 2, 3, \dots, n \qquad i \neq j & \\ \sum_{j=1}^{n} x_{ij} = 1, \; i = 1, 2, 3, \dots, n \qquad i \neq j & \\ y_i - y_j + n x_{ij} \leq n - 1, \; i, j = 2, 3, \dots, n & \\ x_{ij} \in \{0,1\}, \; i, j = 1, 2, \dots, n & \\ \alpha \geq 0 & \end{array} \tag{2}$$

Step 3 Categorize objective functions into the following subsets:
W – Subset subscript of goal objective functions, which value needs to be improved from current level $f_j(x^{t-1})$
R – Subset subscript of goal objective functions, which value at least should be maintained on current level $f_j(x^{t-1})$
Z – Subset subscript of goal objective functions, which value may be deteriorated from current level $f_j(x^{t-1})$

Let
$$\begin{Bmatrix} W = \{i \mid i = i_1, i_2, \dots, i_w\} \\ R = \{j \mid j = j_1, j_2, \dots, j_r\} \\ Z = \{l \mid l = l_1, l_2, \dots, l_z\} \end{Bmatrix} \tag{3}$$
where $W \cup R \cup Z = \{1, 2, \dots, k\}$ and $W \cap R \cap Z = \emptyset$.

Step 4 Let assume that u_i $(i \in W)$ denotes auxiliary variable, than the problem solved at this stage can be defined as follows:

$$\text{Max} \quad u(\mathbf{u}) = \sum_{i \in W} \lambda_i u_i$$
$$x \in \Omega \tag{4}$$
$$x \in \Omega_u = \begin{Bmatrix} f_i(x) - h_i u_i \geq f_i(x^{t-1}), u_i \geq 0, i \in W \\ f_j(x) \geq f_j(x^{t-1}), j \in R \\ f_l(x) \geq f_l(x^{t-1}) - d f_l(x^{t-1}), l \in Z \end{Bmatrix}$$

where Ω still represents former set of feasible solutions and Ω_u represents the set of additional structural constraints. u_i represents improvements $f_i(x)$, $u(\mathbf{u})$ is utility function representing improvement of objective functions, in which improvement is requested and λ_i is positive weight factor, defined according to relative importance of goal objective function in subset W. We leave value $\lambda_i = 1$ $(i \in W)$.

Step 5 Solve the problem stated in step 4. Optimal solution denoted as $\mathbf{x}^t$, represent new effective solutions of former problem.

Step 6 If decision maker is not satisfied with obtained solution x^t, in $t = t + 1$ procedure returns to step 3. Iterative procedure stops if:
a) there is no requirement to improve any target value, or
b) deterioration of desired goals is not feasible.

3 Application of MTSP utilizing ISTEM on data from Slovakia

In the analysis nodes representing district towns of central Slovakia regions were used: Žilina and Banská Bystricka County. The list nodes on which the analysis was performed from Žilina County is as follows: 1 - Bytča, 2 – Čadca, 3 – Dolný Kubín, 4 – Kysucké nové mesto, 5 - Liptovský Mikuláš, 6 – Martin, 7 – Námestovo, 8 – Ružomberok, 9 – Turčianske teplice, 10 – Tvrdošín, 11 – Žilina; and Banska Bystrica County: 12 – Banská Bystrica, 13 – Banská Štiavnica, 14 – Brezno, 15 – Detva, 16 – Krupina, 17 – Lučenec, 18 – Poltár, 19 – Revúca, 20 – Rimavská sobota, 21 – Veľký Krtíš, 22 – Zvolen, 23 – Žarnovica, 24 – Žiar nad Hronom.
When defining the tasks, set of three objectives were defined. Firs objective is to minimize the driving distance. Second objective in our case is to minimize distance traveled including detours of mountain passes and roads above specific height above sea level. The target mountain passes detour reflects safety requirements especially under the winter conditions on the route. Third objective represents preference of minimization of time duration of route

Parameters:
Matrix $\boldsymbol{D}^{(1)} = \{d^{(1)}{}_{ij}\}$, represents the value of the minimum distances between towns.
Matrix $\boldsymbol{D}^{(2)} = \{d^{(2)}{}_{ij}\}$, represents the value of the minimum distances between towns omitting mountain passes.
Matrix $\boldsymbol{D}^{(3)} = \{d^{(3)}{}_{ij}\}$, represents the time value of the minimum distances between towns.
In the first step, each of the three objectives are optimized, solve model (1) for each individual objective function. Afterwards pay-off table [1] is created from optimization results, the optimal values for each objective function are set on the diagonal in Table 1.

	$f_1(x^1)$	$f_2(x^2)$	$f_3(x^3)$
x^1	–759	–952.4	–784
x^2	–835.3	–869.1	–831
x^3	–763.5	–950.4	–781

Table 1 Pay-off table

Based on pay-off table, the relative weights of tree targets were calculated with following values:

$$[\lambda_1, \lambda_2, \lambda_3] = [0.3943, \qquad 0.3245, \qquad 0.2817]$$

In this stage it is possible to use method for generating effective alternatives, the method is based on ideal vector in order to find initial effective solution, solve model (2).

Based on output of optimization software GAMS, the following initial effective solution was acquired. x^0 represents binary matrix, that depicts the order in which the nodes should be visited. In order to obtain values of $f(x^0)$ it is necessary to make calculations based on the obtained binary matrix and matrices of values of individual criterions.

$$x^0 = \begin{bmatrix}
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0
\end{bmatrix}$$

$$F(\boldsymbol{x}^0) = [784{,}9; \quad 875{,}4; \quad 810]$$

The resulting values can be interpreted as follows: Based on solution, the sequence of locations is determined, i.e. schedule of truck loading is determined. On the day of delivery, according to specific situation on the route, one of the three above listed solutions is implemented. If the decision-maker accepted the solution, in the case of the shortest route 784.9 km are driven. In case, the safety requirements are preferred and mountain passes and roads need to be avoided 875.4 km are driven. In the case of quickest route is preferred, it will take 810 minutes to drive calculated route.

In order to acquire first interaction, the first interactive table is constructed (Table 2).

$f_1(\boldsymbol{x}^{(0)})$	$f_2(\boldsymbol{x}^{(0)})$	$f_3(\boldsymbol{x}^{(0)})$
784,9	875,4	810
W	Z	Z
	$df_2(\boldsymbol{x}^{(0)}) = 60$	$df_3(\boldsymbol{x}^{(0)}) = 30$

Table 2 First interaction of calculation procedure ISTM

In this step, the decision maker and analysts interacts. It is necessary that the decision maker define his satisfaction with the solution that is provided. The decision maker can accept the proposed solution or identify which of these objectives needs to be improved. Compromises as shown in Table 2 mean that the decision maker would like to improve $f_1(\boldsymbol{x})$ at the expense of $f_2(\boldsymbol{x})$ and $f_3(\boldsymbol{x})$ so that $f_2(\boldsymbol{x})$ and $f_3(\boldsymbol{x})$ may be relaxed as follows: second target by 60 units, that means by 60 kilometers and third target by 30 units, that means by 30 minutes.
On the basis of these values, the task which ensure the implementation of the conditions laid down by the decision maker is formulated (4). Variable $u_1$that represents improvement of utility function of first target (increment of the first criterion value) is introduced at this stage. Solving the problem leads to the following effective solution.

$$x^{(1)} = \begin{bmatrix}
0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0
\end{bmatrix}$$

$$u_1 = 0{,}1258, \quad F(\boldsymbol{x}^{(1)}) = [775{,}3; \quad 934{,}9; \quad 797{,}0]$$

Afterwards, the second interactive table is constructed, that represents the value of individual objective functions (Table 3).

$f_1(\boldsymbol{x}^{(1)})$	$f_2(\boldsymbol{x}^{(1)})$	$f_3(\boldsymbol{x}^{(1)})$
775,3	934,9	797

Table 3 Second interaction of calculation procedure ISTM

Assuming the decision maker is satisfied with values $f_1(x)$, $f_2(x)$ and $f_3(x)$, obtained solution is the best compromise solution. After the calculation at this step, the value of the first objective function was improved, therefore, let assume that the decision maker is satisfied with the achieved results and doesn't required further improvement of the goals. Selected compromise solution in this case would constitute the following route: Bytča – Kysucké Nové Mesto – Čadca – Námestovo – Tvrdošín - Dolný Kubín – Ružomberok – Liptovský Mikuláš – Brezno - Revúca - Rimavská Sobota – Poltár – Lučenec – Veľký Krtíš – Detva - Banská Bystrica – Zvolen – Krupina – Banská Štiavnica – Žarnovica – Žiar nad Hronom – Turčianske Teplice – Martin - Žilina - Bytča.

The resulting values represent the following:
If the decision-maker accepted the solution, in the case of the shortest route 775.3 km are driven. In case, the safety requirements are preferred and mountain passes and roads need to be avoided 934.9 km are driven. In the case of quickest route is preferred, it will take 797 minutes to drive obtained route.

Conclusion

In the article the authors outline the procedure to solve multiple criteria traveling salesman problem. They are focused on an interactive method ISTM, which finds a compromise solution based on the interaction between decision-makers and analysts. Multiple criteria TSP solution procedure was proposed in the paper. Three objectives were defined: minimum driving distance, the minimum distance traveled by omitting the mountain passes and minimum time length. The result is a succession of towns of central Slovakia on the route. This procedure can be used to set the parameters of distribution of goods, in cases where the loading of goods is carried out at a time when conditions are not known and the decision on the selection of type of route (shortest, fastest route, under bad weather route conditions bypassing mountain passes) is carried out before actual delivery. The computational experiments were based on regional data of Slovakia. Software implementation was realized in GAMS (CPLEX solver 12.2.0.0).

Acknowledgements

This paper is supported by the Grant Agency of Slovak Republic – VEGA, grant no. 1/0245/15 „Transportation planning focused on greenhouse gases emission reduction“.

References

[1] BENAYOUN, R, MONTGOLFIER, J., TERGNY, J., LARITCHEV, O.: *Linear programing with multiple objective functions: Step Method(STEM),* Mathematical Programming, North-Holland Publishing Company, 1971.

[2] ČIČKOVÁ, Z., BREZINA, I.: An evolutionary approach for solving vehicle routing problem. In: *Quantitative methods in economics multiple criteria decision making XIV*. IURA EDITION, Bratislava, 2008, 40–44.

[3] ČIČKOVÁ, Z., BREZINA, I., PEKÁR, J.: *Solving the Real-life Vehicle Routing Problem with Time Windows Using Self Organizing Migrating Algorithm*. Ekonomický časopis **61**(5/2013), 497–513.

[4] ČIČKOVÁ, Z., BREZINA, I., PEKÁR, J.: *Open vehicle routing problem.* In: *Mathematical Methods in Economics 2014*. Olomouc: Faculty of Science, Palacký University, 2014, 124 – 127.

[5] LIU, G. P., YANG, J. B., WHIDBORNE, J. F.: *Multiobjective Optimisation and Control*, Baldock, Hertfordshire, England, 2003.

[6] MILLER, C.E., TUCKER, A.W., ZEMLIN, R.A.: *Integer programming Formulation of Traveling Salesman Problems*. Journal of the ACM, **7**(4/1960), 326-329.

[7] YANG J.B., CHEN CH., ZHANG Z. J.: *The Interactive Step Trade-off Method (ISTM) for Multiobjective Optimization*. IEEE Transactions on Systems, Man, and Cybernetics, **20**(3/1990), 688-695.

Heuristics for VRP with Private and Common Carriers

Jan Pelikan[1]

Abstract. The topic of the paper is a heuristic for a modification of vehicle routing problem. In the presented problem there are two kinds of carriers. The first carrier is a producer owning a fleet of vehicles with given capacity and costs per km. The second carrier is common carrier whose services are used by producer at fixed price of transported unit not dependant on distance from depot. A set of customers with given demand are known. The total costs that are minimized consists from private carrier costs and common carrier costs. Main question is what quantity of goods is transported by private and by common carrier. The problem is studied with and without split demand and with uniform and heterogenous fleet of vehicles. The mathematical model is shown and a modified insert heuristic is proposed. The model and proposed heuristic are tested on instances from testing dataset, results of the model obtained in limited computational time and results of insert heuristic with sequential and parallel routes creation are compared.

Keywords: vehicle routing problem, insert heuristic, integer programming.

JEL classification: C44
AMS classification: 90C15

1 Introduction

Transportation of goods from producers to consumers presents a substantial part of the costs and is therefore the subject of optimization problems solved by the models and methods of operations research. Typical problem is traveling salesman problem (TSP), or vehicle routing problem (VRP), where goods are transported from depot to custometrs through a communications network (roads, railways etc.). There are many modifications of those problems as for example VRP with time windows, with heterogenous fleet of vehicles, with more depots, with split delivery of demands, the pick and delivery problem, VRP with stochastic demand etc.

The paper deals with modification of VRP, which consists of two kinds of carriers, private and common carriers. Goods from depot to customers can be transported or by private or common carrier, or partially by both carriers. Similarly to the classic VRP customers' demand is given, it can be split (SDVRPPC) or non-split (VRPPC).

The problem aims at goods transportation from depot to customers with minimal costs which are the sum of private carrier costs and common carrier costs. Costs of private carrier depend on the length of vehicles routes multiplied by the costs per km of vehicles. Unlike private carrier costs of common carrier depend only on quantity of transported goods multiplied price of unit quantity transport. Common carrier costs do not depend on transported distance. There is a fleet of vehicles of private carrier located in depot, the capacity of vehicles is limited, there is no limit of capacity of common carrier.

In the problem SDVRPPC is to decide what part of customer's demand is delivered by private and by common carrier. In case of non-split demand problem VRPPC question is only whether the customer's request is delivered by private or common carrier. A further problem is to create optimal routes for private carrier vehicles under condition that the route for each vehicle must carry the quantity of goods that does not exeed the capacity of the vehicle. The problem is NP hard because clasical vehicle routing problem can be reduced to this problem.

[1]University of Economics, W. Churchill sq. 4, Prague, pelikan@vse.cz

Tabu search heuristic is published in [2] for the problem with nonsplit demand. A heuristic proposed in [3] selects customers according the distance from depot, customers which are closed to depot are assigned to private carrier, the others to common carriers. A modification of the Clarke and Wrights savings algorithm is used in [1] to solve so called the truckload and less-than-truckload problem, where private carrier uses less-than-truckload vehicles while an outside carrier trucks.

2 Mathematical model of the problem

The mathematical model of vehicle routing problem with private and common carriers with split demand (SDVRPPC) and without split demand (VRPPC) are shown in this section. The models contain hudge number of binary variables, so the optimal solution cannot be obtained in limited computation time for real problems.

2.1 Split delivery model SDVRPPC

Let $G = \{V, E\}$ is undirected complete graph, $V = \{1, 2, \cdots, n\}$, depot is node 1, nodes $2, 3, \cdots, n$ are customers. A number of vehicles of the private carrier is m.

Parameters of the model are:
$d_{i,j}$ distance between node i and node j,
q_i demand of node i,
w_s capacity of s-th vehicle, $s = 1, 2, \cdots, m$,
p_s costs per km of the s-th vehicle,
c_c, costs of the transport a unit of goods by common carrier.

Variables of the model are:
x^s_{ij} binary, equals 1 if vehicle s travels from node i to node j,
y^s_i an amount of goods (a part of demand of node i), which is transported by s-th vehicle of private carrier,
z_i an amount of goods (a part of demand of node i), which is transported by common carrier (outside carrier),
u^s_i variables in anti-cyclic constraints.

$$\sum_{s,i,j} p_s d^s_{ij} x^s_{ij} + c_c \sum_i z_i \to min \tag{1}$$

$$\sum_i x^s_{ij} = \sum_i x^s_{ji} \qquad \forall s, j \tag{2}$$

$$z_i + \sum_s y^s_i = q_i \qquad \forall i \tag{3}$$

$$\sum_i y^s_i \le w_s \qquad \forall s \tag{4}$$

$$q_i \sum_j x^s_{ij} \ge y^s_i \qquad \forall s, i \tag{5}$$

$$\sum_j x^s_{1j} \le 1 \qquad \forall s \tag{6}$$

$$u^s_i + 1 + n(1 - x^s_{ij}) \le u^s_j \qquad \forall s, i \quad j > 1 \tag{7}$$

$$x_{ij} \quad \text{binary}, \qquad z_i \ge 0, y^s_i \ge 0 \qquad \forall s, i, j. \tag{8}$$

In this formulation the objective function (1) minimizes the sum of travel costs of vehicles of private carrier and common carrier costs. Constraint (2) states that the same vehicle must enter and leave a node. Inequality (3) assures that capacity of s-th vehicle is not exceeded . Equation (4) means that the demand each node will be satisfied. Constraint (5) prevents supply of node by the vehicle s if this vehicle does not enter the node. At least one leave the depot is possible for each vehicle states (6). Anti-cyclic conditions are in (8).

2.2 Nonsplit delivery model VRPPC

Nonsplit delivery model contains binary variable z_i which is equal 1 if common carrier assures whole demand q_i of the node i.

$$\sum_{s,i,j} p_s d_{ij}^s x_{ij}^s + c_c \sum_i z_i q_i \to min \tag{9}$$

$$\sum_i x_{ij}^s = \sum_i x_{ji}^s \qquad \forall s, j \tag{10}$$

$$\sum_{j,s} x_{ij}^s = (1 - z_i) \qquad \forall i \tag{11}$$

$$u_i^s \le w_s \qquad \forall s, i \tag{12}$$

$$\sum_j x_{1j}^s \le 1 \qquad \forall s \tag{13}$$

$$u_i^s + q_j + n(1 - x_{ij}^s) \le u_j^s \qquad \forall s, i \quad j > 1 \tag{14}$$

$$x_{ij}^s, z_i \text{ binary} \qquad \forall s, i, j. \tag{15}$$

The objective function (9) minimizes the sum of travel costs of vehicles of private carrier and common carrier costs. Constraint (10) states that the same vehicle must enter and leave a node. Equation (11) means that node not served by common carrier has to be served by some vehicle of the private carrier. Inequality (12) assures that capacity of s-th vehicle is not exceeded. At least one leave the depot is possible for each vehicle states (13). Anti-cyclic conditions and defining of load u_i^s of the vehicle s entering node i are in (14).

3 Insert heuristic

The following notation is used for the modified insert heuristic: the route of s-th vehicle is $R^s = \{v_1^s, v_1^s, \dots, v_{h_s}^s\}$ where $v_1^s = v_{h_s}^s = 1$ and delivery to node i by vehicle s is denoted as y_i^s.
The symbol $\Delta(i, j, k)$ means $\Delta(i, j, k) = p(s)(d_{i,j} + d_{j,k} - d_{i,k})$.

Step 1: {choice of vehicle s and its initial route} Vehicle s and node k are selected in order to maximize the value $\sigma(s, k, y_k^s) = c_c y_k^s - p_s(d_{1,k} + d_{k,1})$, where y_k^s is maximum uncovered demand of node k which does not exceed available capacity of the vehicle s, ie. $0 < y_k^s \le q_k$ and $y_k^s \le w_s$. Value $\sigma(s, k, y_k^s)$ has to be positive. If no vehicle with initial route is selected then stop, otherwise put $v_1^s := 1; v_2^s := k; v_3^s := 1; q_k := q_k - y_k^s; w_s := w_s - y_k^s; h_s := 3$ and go to Step 2.

Step 2: {insertion} Repeatedly select a node k and a quantity y_k^s with the edge (v_j^s, v_{j+1}^s) to maximize value $\delta(s, k, j, y_k^s) = c_c y_k^s - p_s(d_{v_j^s,k} + d_{k,v_{j+1}^s} - d_{v_j^s,v_{j+1}^s})$ where value y_k^s cannot exceed the available capacity of the vehicle s, ie. $0 < y_k^s \le q_k$ and $y_k^s \le w_s$. Value $\sigma(s, k, y_k^s)$ has to be positive. If no node k is selected then go ro step 1, otherwise put $q_k := q_k - y_k^s; w_s := w_s - y_k^s; h_s := h_s + 1$, the edge (v_j^s, v_{j+1}^s) is deleted from the route and edges (v_j^s, k) and (k, v_{j+1}^s) are added to the route, go to Step 2.

Step 3: {common carriers transport} The remaining demand q_k is assigned to transport by common carrier.

4 Main features of the problem and heuristic

Proposition 1. *If $\{R_s; s = 1, 2, \dots m\}$ is optimal, then inequality*

$$p_s(d_{v_j^s,v_{j+1}^s} + d_{v_{j+1}^s,v_{j+2}^s} - d_{v_j^s,v_{j+2}^s}) \le c_c y_{v_{j+1}^s}^s \tag{16}$$

is valid for all s and $j = 1, \dots, h_s - 2$.

Proof. Proof by contradiction. □

Proposition 2. *If inequality (16) holds for some s, then (17) is valid.*

$$p_s \sum_{j=1}^{h_s-1} d_{v_j^s, v_{j+1}^s} \leq c_c \sum_i y_i^s \tag{17}$$

Left hand side of the inequality (17) represends costs of private carrier to provide transport y_j^s by s-th vehicle to the customers $j = 2, 3, \dots n$, right hand side are costs of common carrier by the same amount of goods.

Proof. Proof by induction over number of nodes of the route. □

Proposition 3. *The solution obtained from insert heuristic satisfies the inequality (16) and (17).*

Proof. When the node v_{j+1}^s was inserted into the edge (v_j^s, v_{j+2}^s), the inequality

$$\Delta(v_j^s, v_{j+1}^s, v_{j+2}^s) \leq c_c y_{j+1}^s \tag{18}$$

is valid for the quantity y_{j+1}^s. We have to prove that the inequality

$$\Delta(v_j^s, v_{j+1}^s, k) \leq c_c y_{j+1}^s \tag{19}$$

is valid after subsequential insertion the node k into the edge (v_{j+1}^s, v_{j+2}^s).

The node v_{j+1}^s was inserted into the edge (v_j^s, v_{j+2}^s) as it is true that

$$c_c y_{j+1}^s - \Delta(v_j^s, v_{j+1}^s, v_{j+2}^s) \geq c_c y_k^s - \Delta(v_j^s, k, v_{j+2}^s) \tag{20}$$

for all nodes k and the value y_k^s as well as $c_c y_{j+1}^s - \Delta(v_j^s, v_{j+1}^s, v_{j+2}^s) \geq 0$.

After insertion v_{j+1}^s into edge (v_j^s, v_{j+2}^s) the node k is inserted into edge (v_{j+1}^s, v_{j+2}^s) with value $y_k^{'s} \leq y_k^s$ so we have

$$c_c y_k^{'s} - \Delta(v_{j+1}^s, k, v_{j+2}^s) \geq 0 \tag{21}$$

From (20) and (21) we get

$$\begin{gathered} c_c y_{j+1}^s \geq c_c y_{j+1}^s - \Delta(v_j^s, k, v_{j+2}^s) + \Delta(v_j^s, v_{j+1}^s, v_{j+2}^s) \geq \\ \geq c_c y_k^{'s} - \Delta(v_j^s, k, v_{j+2}^s) + \Delta(v_j^s, v_{j+1}^s, v_{j+2}^s) \geq \Delta(v_{j+1}^s, k, v_{j+2}^s) - \Delta(v_j^s, k, v_{j+2}^s) + \Delta(v_j^s, v_{j+1}^s, v_{j+2}^s) = \\ = p_s\{(d_{v_{j+1}^s,k} + d_{k,v_{j+2}^s} - d_{v_{j+1}^s,v_{j+2}^s}) + (d_{v_j^s,v_{j+1}^s} + d_{v_{j+1}^s,v_{j+2}^s} - d_{v_j^s,v_{j+2}^s}) - (d_{v_j^s,k} + d_{k,v_{j+2}^s} - d_{v_j^s,v_{j+2}^s})\} = \\ = \Delta(v_j^s, v_{j+1}^s, k)). \end{gathered}$$

□

Remark 1. Non-optimality of the result of insert heuristic exists in two causes:
a) vehicles routes are not optimal, they can be improved by using some heuristic,
b) it is not optimal amount of goods delivered to the nodes by private and common carrier. It is possible to exchange goods between private and common carrier.

5 Numerical example

The model and insert heuristic was tested on example S6A which is published in http://www.uv.es/ belengue/carp.html, results are shown in the Table 1. The parameters of the instance S6A are: n=31, m=5, c_c=70. Insert heuristic is written in VBA language. The model is solved using CPLEX 12.0 on PC (IntelCore2Quad, 2,83GHz). The best value of object function is shown but the result is not optimal (computation was aborted), the gap of object function is placed under the value of total costs in parenthesis (in % of the lower bound). Value NP are costs of private carries, NC are costs of common carrier. Result of the model is better than the result of heuristic in case split delivery problem, for case non-split problem it is opposite.

	insert heuristic			model		
	NP	NC	NP+NC	NP	NC	NP+NC
non-split	3745	700	4445	960	10780	11740 (90%)
split	3650	630	4280	2880	630	3510 (44%)

Table 1 Results of S6A

Conclusion

An interesting modification for vehicle routing problem is studied, mathematical model of the problem is presented and the new heuristic method is proposed. Main features of the heuristic are shown. As to solve the mathematical model is very difficult for real problem, the proposed heuristic method gives us good solution in polynomial time. Additionaly, the heuristic solution meets the necessary condition for the optimal solution.

Acknowledgements

Supported by the grant No. 16-00408S of the Czech Grant Agency.

References

[1] Chu, Ch-W.: A heuristic algorithm for the truckload and less-than-truckload problem., *European Journal for Operations Research* **165** (2005), 657-667.

[2] Cote,JF. and Potvin, JY.: A tabu search heuristics for the vehicle routing problem with private fleet and common carrier, *European Journal for Operations Research* **198**,2 (2009), 464-469.

[3] Plevny, M.: Vehicle routing problem with a private fleet and common carriers - the variety with a possibility of sharing the satisfaction of demand. In: *Mathematical Methods in Economics 2013*, 31st international conference proceedings, College of Polytechnics Jihlava, (2013), s.730-736. ISBN 978-80-87035-76-4.

[4] www.uv.es/ belengue/carp.html

Fuzzy Decision Analysis Module for Excel

Radomír Perzina[1], Jaroslav Ramík[2]

Abstract. There exists wide range of software products to support decision making. Main disadvantage of those software products is that they are commercial and relatively expensive and thus it prevents them to be used by small companies, students or researchers. This paper introduces a new Microsoft Excel add-in FuzzyDAME which is completely free and was developed to support users in multicriteria decision making situations with uncertain data. It can be also used by students to help them understand basic principles of multicriteria decision making, because it doesn't behave as a black box but displays results of all intermediate calculations. The proposed software package is demonstrated on an illustrating example of real life decision problem – selecting the best advertising media for a company.

Keywords: analytic hierarchy process, analytic network process, feedback, fuzzy, multi-criteria decision making, pair-wise comparisons.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Decision making in situations with multiple variants is an important area of research in decision theory and has been widely studied e.g. in [4], [7], [8], [9], [10], [11], [12]. There exists wide range of computer programs that are able to help decision makers to make good decisions, e.g.
Expert Choice (http://www.expertchoice.com),
Decisions Lens (http://www.decisionlens.com),
Mind Decider (http://www.minddecider.com),
MakeItRational (http://makeitrational.com) or
Super Decisions (http://www.superdecisions.com).

Main disadvantage of those programs is that they are commercial and relatively quite expensive and thus it prevents them to be used by small companies or individual entrepreneurs.

Here we introduce a new Microsoft Excel add-in named FuzzyDAME – Fuzzy Decision Analysis Module for Excel which is a successor of another add-in developed by the authors called DAME that was used just for crisp evaluations [9]. The main advantage of the new add-in FuzzyDAME is possibility to work with both crisp and fuzzy evaluations. Comparing to other software products for solving multicriteria decision problems, FuzzyDAME is free, able to work with scenarios or multiple decision makers, allows for easy manipulation with data and utilizes capabilities of widespread spreadsheet Microsoft Excel. Users can structure their decision models into three levels – scenarios/users, criteria and variants. Standard pair-wise comparisons are used for evaluating both criteria and variants. For each pair-wise comparison matrix there is calculated an inconsistency index. There are provided three different methods for the evaluation of the weights of criteria, the variants as well as the scenarios/users – Saaty's Method [11], Geometric Mean Method [1] and Fuller's Triangle Method [4]. Multiplicative and additive syntheses are supported. For ranking fuzzy weights there are supported three methods – center of gravity method, optimistic alfa cut and pessimistic alfa cut method [3]. Using FuzzyDAME pairwise comparisons method with fuzzy evaluations can be also applied on a multicriteria decision making problem of ranking given alternatives, see [10].

2 Software Description

FuzzyDAME works with all current versions of Microsoft Excel from version 97 up to version 2016. The add-in can be downloaded at http://www.opf.slu.cz/kmme/FuzzyDAME. It consists of three individual files:

- FuzzyDAME.xla – main module with user interface, it is written in VBA (Visual Basic for Applications),
- FuzzyDAME.xll – it contains special user defined functions used by the application, it is written in C# and linked with Excel-DNA library (http://exceldna.codeplex.com).

[1] Silesian University in Opava, School of Business Administration in Karviná, Department of Informatics and Mathematics, University Square 1934/3, Karviná, Czech Republic, perzina@opf.slu.cz.
[2] Silesian University in Opava, School of Business Administration in Karviná, Department of Informatics and Mathematics, University Square 1934/3, Karviná, Czech Republic, ramik@opf.slu.cz.

- FuzzyDAME64.xll – the same as FuzzyDAME.xll just for 64-bit version of Microsoft Excel. Correct .xll version is automatically loaded based on the installed Microsoft Excel version.

All files should be placed in the same folder and macros must be enabled before running the module (see Microsoft Excel documentation for details). FuzzyDAME itself can be executed by double clicking on the file FuzzyDAME.xla. After executing the add-in there will appear a new menu item “FuzzyDAME” in the Add-ins ribbon (in older Excel versions the menu item “FuzzyDAME” will appear in the top level menu). A new decision problem can be generated by clicking on “New problem” item in the main FuzzyDAME menu. Then there is shown a form with main decision problem characteristics, see Fig. 2.

In the top panel there are basic settings: Number of scenarios, criteria and variants. In case a user doesn’t want to use scenarios or there is just a single decision maker, the number of scenarios/users should be set to one. In the second panel “Other settings” we can choose multiplicative or additive synthesis model. Here we set the mode how to compare scenarios/users and criteria by using either pairwise comparison matrix or by setting the weights directly. In the last panel the users choose how to evaluate variants according to the individual criteria. There are three options: Pairwise – each pair of variants is compared individually, Values max – indicates maximization criterion where each variant is evaluated by single value, e.g. revenues and Values min – indicates minimization criterion where each variant is evaluated by single value, e.g. costs. Then, after confirmation, a new Excel sheet with forms is created, where user can update the names of all elements and evaluate criteria and variants using pairwise comparison matrices as shown in the example in Fig. 1.

Criteria	Crit 1	Crit 2	Crit 3	0.002	Criteria weights	
Crit 1	1	3	5			0.648329
Crit 2	0.333333	1	2			0.229651
Crit 3	0.2	0.5	1			0.12202

Figure 1 Pairwise comparison matrix example

In the pairwise comparison matrix users enter values only in the upper triangle. The values in the lower triangle are reciprocal and automatically recalculated. If criterion (variant) in the row is more important than the criterion (variant) in the column user enters values from 2 to 9 (the higher the value, the more important is the criterion in the row). If the criterion (variant) in the row is less important than the criterion (variant) in the column the user enters values from 1/2 to 1/9 (the less the value, the less important is the criterion in the row). If criterion (variant) in the row is equally important to the criterion (variant) in the column user enters value 1 or leaves it empty. In the top right corner inconsistency index, which should be less than 0.1, is calculated. If this index is greater than 0.1, we should reconsider our pairwise comparisons, so that they are more consistent. In the very right column the weights of individual criteria (variants) based on the values in the pairwise comparison matrix are calculated and evaluation method is selected. The weights are automatically calculated using eq. (1).

3 Mathematical Background

Here we briefly describe mathematical background that is used for calculations in FuzzyDAME.

3.1 Crisp Calculations

In case of crisp evaluations, i.e. when we are able to compare all pairs of criteria/variants exactly we use geometric mean method for calculation criteria/variants weights w_k from pairwise comparison matrices, see e.g. [10], as follows.

$$w_k = \frac{\left(\prod_{j=1}^{n} a_{kj}\right)^{1/n}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}\right)^{1/n}}, \; k = 1,2,\ldots,n, \tag{1}$$

where w_k is the weight of k-th criterion (variant), a_{ij} are values in the pairwise comparison matrix, and n is number of criteria (variants).

The inconsistency index is calculated using the formula [1]

$$GCI = \frac{2}{(n-1)(n-2)} \sum_{i<j} \log^2\left(a_{ij} \cdot \frac{w_j}{w_i}\right). \tag{2}$$

When we enter some values into individual pairwise comparison matrices all weights are recalculated, so that we obtain an immediate impact of our each individual entry. The matrix of normalized values as well as the graph with total evaluation of variants is then shown at the bottom of the sheet. The resulting vector of weights of the variants **Z** is given by the formula (3).

$$\mathbf{Z} = \mathbf{W}_{32}\mathbf{W}_{21}, \tag{3}$$

where $\mathbf{W}_{21}$ is the $n\times 1$ matrix (weighing vector of the criteria), i.e.

$$\mathbf{W}_{21} = \begin{bmatrix} w(C_1) \\ \vdots \\ w(C_n) \end{bmatrix}, \tag{4}$$

and $\mathbf{W}_{32}$ is the $m\times n$ matrix

$$\mathbf{W}_{32} = \begin{bmatrix} w(C_1,V_1) & \cdots & w(C_n,V_1) \\ \vdots & \cdots & \vdots \\ w(C_1,V_m) & \cdots & w(C_n,V_m) \end{bmatrix}, \tag{5}$$

where $w(C_i)$ is the weight of criterion C_i, $w(V_r,C_i)$ is the weight of variant V_r subject to the criterion C_i.

3.2 Fuzzy Calculations

In practice it is sometimes more convenient for the decision maker to express his/her evaluation in words of natural language saying e.g. “possibly 3”, “approximately 4” or “about 5”, see e.g. [11]. Similarly, he/she could use the evaluations as “A is possibly weak preferable to B”, etc. It is advantageous to express these evaluations by fuzzy sets of the real numbers, e.g. triangular fuzzy numbers. A triangular fuzzy number a is defined by the triple of real numbers, i.e. $a = (a^L; a^M; a^U)$, where a^L is the Lower number, a^M is the Middle number and a^U is the Upper number, $a^L \le a^M \le a^U$. If $a^L = a^M = a^U$, then a is the crisp number (non-fuzzy number). In order to distinguish fuzzy and non-fuzzy numbers we shall denote the fuzzy numbers, vectors and matrices by the tilde, e.g. $\tilde{a} = (a^L; a^M; a^U)$. It is known that the arithmetic operations +,-, * and / can be extended to fuzzy numbers by the *Extension principle*, see e.g. [3].

If all elements of an $m\times n$ matrix **A** are triangular fuzzy numbers then we call **A** the triangular fuzzy matrix and this matrix is composed of the triples of real numbers. Particularly, if **A** is a pair-wise comparison matrix, we assume that it is reciprocal and there are ones on the diagonal.

In order to find the "optimal" variant we have to calculate the triangular fuzzy weights as evaluations of the relative importance of the criteria and evaluations of the variants according to the individual criteria. We assume that there exists vectors of triangular fuzzy weight $\tilde{w}_1, \tilde{w}_2, ..., \tilde{w}_n$, $\tilde{w}_i = (w_i^L; w_i^M; w_i^U)$, i = 1,2,...,n, which can be calculated by the following formula (see [7]):

$$\tilde{w}_k = (w_k^L; w_k^M; w_k^U)\text{, } k = 1,2,...,n, \tag{6}$$

where

$$w_k^S = \frac{\left(\prod_{j=1}^{n} a_{kj}^S\right)^{1/n}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} a_{ij}^M\right)^{1/n}}\text{, } S \in \{L, M, U\}. \tag{7}$$

In [3], the method of calculating triangular fuzzy weights by (7) from the triangular fuzzy pair-wise comparison matrix is called by the *logarithmic least squares method*. This method can be used both for calculating the triangular fuzzy weights as relative importance of the individual criteria and also for eliciting relative triangular fuzzy values of the criteria of the individual variants out of the pair-wise comparison matrices.

Now we calculate the synthesis - the aggregated triangular fuzzy values of the individual variants by formula (8), applied for triangular fuzzy matrices

$$\tilde{\mathbf{Z}} = \tilde{\mathbf{W}}_{32}\tilde{\mathbf{W}}_{21}\,. \tag{8}$$

Here, for addition, subtraction and multiplication of triangular fuzzy numbers we use the fuzzy arithmetic operations mentioned earlier, see e.g. [7].

The simplest method for ranking a set of triangular fuzzy numbers in (8) is *the center of gravity method*. This method is based on computing the *x*-th coordinates of the center of gravity of each triangle given by the corresponding membership functions of , $i = 1,2,...,n$. Evidently, it holds

$$z_i^g = \frac{z_i^L + z_i^M + z_i^U}{3}. \tag{9}$$

By (9) the variants can be ordered from the best to the worst. There exist more sophisticated methods for ranking fuzzy numbers, for a comprehensive review of comparison methods see [3]. In FuzzyDAME we apply two more methods which are based on α-cuts of the fuzzy variants being ranked.

Particularly, let $\tilde{z}$ be a fuzzy alternative, i.e. fuzzy number, $\alpha \in (0,1]$ be a preselected aspiration level. Then the *α-cut of* $\tilde{z}$, $[\,\tilde{z}\,]_\alpha$, is a set of all elements x with the membership value greater or equal to α, i.e. $[\,\tilde{z}\,]_\alpha = \{x|\ \mu_{\tilde{z}}(x) \geq \alpha\}$.

Let $\tilde{z}_1$ and $\tilde{z}_2$ be two fuzzy variants, $\alpha \in (0,1]$. We say that $\tilde{z}_1$ is *R-dominated* by $\tilde{z}_2$ at the level α if $\sup[\,\tilde{z}_1\,]_\alpha \leq \sup[\,\tilde{z}_2\,]_\alpha$. Alternatively, we also say that $\tilde{z}_2$ *R-dominates* $\tilde{z}_1$ at the level α.

If $\tilde{z} = (z^L, z^M, z^U)$ is a triangular fuzzy number, then

$\sup[\,\tilde{z}_1\,]_\alpha = z_1^U - \alpha(z_1^U - z_1^M)$, $\sup[\,\tilde{z}_2\,]_\alpha = z_2^U - \alpha(z_2^U - z_2^M)$,
as it can be easily verified.

We say that $\tilde{z}_1$ is *L-dominated* by $\tilde{z}_2$ at the level α if $\inf[\,\tilde{z}_1\,]_\alpha \leq \inf[\,\tilde{z}_2\,]_\alpha$. Alternatively, we also say that $\tilde{z}_2$ *L-dominates* $\tilde{z}_1$ at the level α.

Again, if $\tilde{z} = (z^L, z^M, z^U)$, then

$\inf[\,\tilde{z}_1\,]_\alpha = z_1^L + \alpha(z_1^M - z_1^L)$, $\inf[\,\tilde{z}_2\,]_\alpha = z_2^L + \alpha(z_2^M - z_2^L)$.

Applying *R* and/or *L* domination we can easily rank all the variants (at the level α).

4 Case Study

Here we demonstrate the proposed add-in FuzzyDAME on a decision making situation selecting the "best" advertising media for a company. Advertising is an important part of the company's marketing strategy. Advertising objectives may be to inform, remind, convince, create prestige, or to increase sales and profits. Different media have varying capacity to meet these objectives. Here we assume a company with objective to increase its profit. The goal of this realistic decision situation is to find the best variant from 4 pre-selected ones (Internet, TV, Radio and Press) according to 3 criteria: Overall Impression (pairwise), Look&Feel (pairwise) and Costs (pairwise). We assume that a decision maker is not able to precisely compare each pair of criteria/variants so it becomes more natural to use fuzzy evaluations, e.g. by triangular fuzzy numbers, see [10]. Setting of parameters can be seen on the Fig. 2.

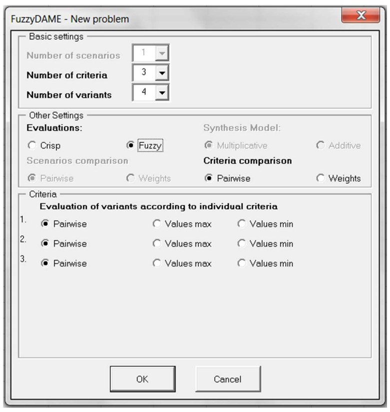


Figure 2 Case study – New problem form

After confirmation by OK a new Excel sheet with forms is created, here the user can set in all required values. First we fill in the importance of the criteria that is given by the fuzzy pair-wise comparisons, see Fig. 3. Note that we use triangular fuzzy numbers, i.e. each pair-wise comparison input is expressed by 3 numbers.

Criteria	Impression			Look&Feel			Costs		
Impression	1	1	1	1/3	1	2	3	3	4
Look&Feel	0.5	1	3	1	1	1	1	2	4
Costs	0.25	0.333333	0.333333	0.25	0.5	1	1	1	1

Figure 3 Case study – criteria fuzzy comparison

Then the corresponding triangular fuzzy weights are calculated, i.e. the relative fuzzy importances of the individual criteria are shown, see Fig. 4.

0.307457	0.443429	0.614913
0.244028	0.387371	0.7039
0.122014	0.1692	0.213178

Figure 4 Case study – criteria fuzzy weights

Next step is to include fuzzy evaluations of the variants according to the individual criteria by three pair-wise comparison matrices, see Fig. 5.

Impressio	Internet			TV			Radio			Press		
Internet	1	1	1	1/2	1/3	1/3	1	2	4	1	1	2
TV	2	3	3	1	1	1	1	3	4	2	3	5
Radio	0.25	0.5	1	0.25	0.333333	1	1	1	1	1/3	1	2
Press	0.5	1	1	0.2	0.333333	0.5	0.5	1	3	1	1	1

Look&Fee	Internet			TV			Radio			Press		
Internet	1	1	1	1	2	2	2	4	6	1	1	3
TV	0.5	0.5	1	1	1	1	1/3	1/2	1/2	1/4	1/3	1
Radio	0.166667	0.25	0.5	2	2	3	1	1	1	1/6	1/4	1/2
Press	0.333333	1	1	1	3	4	2	4	6	1	1	1

Costs	Internet			TV			Radio			Press		
Internet	1	1	1	1/3	1/4	1/4	1/4	1/3	1/2	1/4	1/3	1/3
TV	3	4	4	1	1	1		1/2		1	1	3
Radio	2	3	4	2	2	2	1	1	1	1/2	2	3
Press	3	3	4	0.333333	1	1	0.333333	0.5	2	1	1	1

Figure 5 Case study – evaluation of variants

Then the corresponding fuzzy matrix $\mathbf{W}_{32}$ of fuzzy weights is calculated, see Fig. 6.

0.165835	0.197212	0.308653	0.254382	0.35975	0.523967	0.077443	0.089423	0.106343
0.308653	0.497504	0.607426	0.096644	0.11493	0.179875	0.242409	0.260486	0.342818
0.082917	0.13945	0.259545	0.103851	0.127191	0.199065	0.260486	0.407682	0.484818
0.103204	0.165835	0.241534	0.193288	0.398129	0.473458	0.166435	0.242409	0.368382

Figure 6 Case study – fuzzy weights of variants

Now the final aggregation of results is automatically calculated and we can see the total evaluation of variants as triangular fuzzy numbers, the center of gravity (COG) of each variant and the corresponding Rank in Fig. 7 as well as the graphical representation in Fig. 8.

Zn=	L	M	U	COG	Rank
Internet	0.122512	0.241937	0.581285	0.315245	2
TV	0.148059	0.309202	0.57321	0.34349	1
Radio	0.082619	0.180086	0.403072	0.221926	4
Press	0.099206	0.268775	0.560321	0.309434	3

Figure 7 Case study – total evaluation of variants

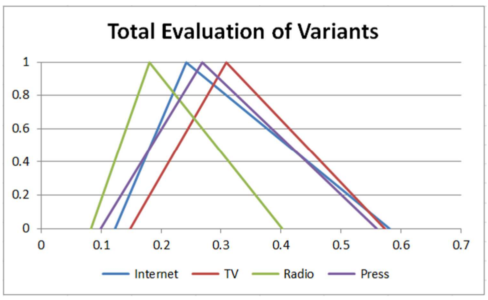


Figure 8 Case study – total evaluation of variants – graph

As we can see the best variant is TV with center of gravity 0.343 followed by Internet with center of gravity 0.315, third one is Press with center of gravity 0.309 and the last one is Radio with center of gravity 0.222.

Apart from center of gravity method the add-in supports also two other variant ranking methods *R*-domination (optimistic) and *L*-domination (pessimistic) as described in section III.B. These methods require alfa level which can be set in the Excel sheet. For the experiments we used alfa level 0.8, see Fig. 9. Final rank for those two methods can be seen in Fig. 10 and Fig. 11.

Parameters for Evaluation of Variants:	
Method	Optimistic
Alfa	0.8

Figure 9 Case study – ranking method selection

Zn=	L	M	U	COG	Rank
Internet	0.122512	0.241937	0.581285	0.315245	1
TV	0.148059	0.309202	0.57321	0.34349	2
Radio	0.082619	0.180086	0.403072	0.221926	3
Press	0.099206	0.268775	0.560321	0.309434	4

Figure 10 Case study – total evaluation of variants – optimistic method

Zn=	L	M	U	COG	Rank
Internet	0.122512	0.241937	0.581285	0.315245	4
TV	0.148059	0.309202	0.57321	0.34349	2
Radio	0.082619	0.180086	0.403072	0.221926	3
Press	0.099206	0.268775	0.560321	0.309434	1

Figure 11 Case study – total evaluation of variants – pessimistic method

5 Conclusions

In this paper we have proposed a new Microsoft Excel add-in FuzzyDAME for solving decision making problems with certain and uncertain data. Comparing to other decision support software products FuzzyDAME is free, able to work with scenarios or multiple decision makers, can process crisp of fuzzy evaluations, allows for easy manipulation with data and utilizes capabilities of widespread spreadsheet Microsoft Excel. On two realistic case studies we have demonstrated its functionality in individual steps. We have also shown that introducing fuzzy comparisons can change rank of alternatives. This add-in is used by students in the course Decision Analysis for Managers at the School of Business Administration in Karvina, Silesian University in Opava. It can be recommended also for other students, researchers or small companies

Acknowledgements

This research was supported by the grant project of GACR No. 14-02424S.

References

[1] Aguaron, J. and Moreno-Jimenez, J.M.: The geometric consistency index: Approximated thresholds. *European Journal of Operational Research* **147** (2003), 137-145.
[2] Buckley, J.J.: Fuzzy hierarchical analysis. *Fuzzy Sets and Systems* **17** (1985), 233-247.
[3] Chen, S.J., Hwang, C.L. and Hwang, F.P.: Fuzzy multiple attribute decision making. *Lecture Notes in Economics and Math. Syst.*, Vol. 375, Springer-Verlag, Berlin – Heidelberg, 1992.
[4] Fishburn, P. C.: A comparative analysis of group decision methods, *Behavioral Science* **16** (1971), 538-544.
[5] Horn, R. A. and Johnson, C. R.: *Matrix Analysis*, Cambridge University Press, 1990.
[6] Gass, S.I. and Rapcsák, T.: Singular value decomposition in AHP. *European Journal of Operational Research* **154** (2004), 573–584.
[7] Ramik, J. and Perzina, R.: Fuzzy ANP – a New Method and Case Study. In: *Proceedings of the 24th International Conference Mathematical Methods in Economics 2006*. University of Western Bohemia, 2006.
[8] Ramik, J. and Perzina, R.: Microsoft Excel Add-In for Solving Multicriteria Decision Problems in Fuzzy Environment. In: *Proceedings of the 26th International Conference Mathematical Methods in Economics 2008*. Technical University of Liberec, 2008.
[9] Ramik, J. and Perzina, R.: Solving Multicriteria Decision Making Problems using Microsoft Excel. In: *Proceedings of the 32nd International Conference Mathematical Methods in Economics 2014*. Palacky University Olomouc, 2014.
[10] Ramik, J.: Strong Consistency in Pairwise Comparisons Matrix with Fuzzy Elements on Alo-Group. In: *Proc. WCCI 2016 Congress*, FUZZ-IEEE. Vancouver, 2016, to appear.
[11] Saaty, T.L.: *Multicriteria decision making - the Analytical Hierarchy Process*. Vol. I., RWS Publications, Pittsburgh, 1991.
[12] Saaty, T.L.: *Decision Making with Dependence and Feedback – The Analytic Network Process*. RWS Publications, Pittsburgh, 2001.
[13] Van Laarhoven and P.J.M., Pedrycz, W.: A fuzzy extension of Saaty's priority theory. *Fuzzy Sets and Systems* **11** (1983), 229-241.

The treatment of uncertainty in uniform workload distribution problems

Štefan Peško[1], Roman Hajtmanek[2]

Abstract. We deal with solving an NP-hard problem which occurs in scheduling of workload distribution with uncertainty. Let us have m vehicles working according to n days schedule. Every day a vehicle performs some work, so for each day we have m working jobs to be done. Let us have the matrix with uncertainty elements giving the profit gained by work of i-th vehicle on j-th day. The total profit produced by i-th vehicle is the sum of elements in row i of matrix. If there is a big difference between the profits gained, the drivers of vehicles will have very different wages and workloads. This can be improved by suitable permutations of elements in the individual columns of matrix so that the row sums are well balanced. It means that we need to split the total workload to given number of approximately equal parts. The level of regularity is expressed by irregularity measure. The lower measure implies the higher regularity. The goal is to minimize the total irregularity of parts of schedule with uncertainty. We study a mixed quadratic programming (MIQP) formulation of the modified problem. Computational experiments with heuristic based on repeated solution of the MIQP problem are presented.

Keywords: schedule with uncertainty, uniform split, irregularity measure, matrix permutation problem

JEL classification: C02, C61, C65, C68, J33
AMS classification: 90C10, 90C15

1 Introduction

In this paper we investigate the solution of an NP-hard problem which occurs in uniform scheduling of workload distribution with uncertainty. We need to split total workload to given number of approximately equal parts. The level of uniformity is expressed by irregularity measure. The lower measure implies the higher regularity. The goal is to minimize the irregularity measure of parts of the concrete schedule.

Our model is motivated by the following practical job scheduling problem:

Let us have m vehicles working according to n days (periodic) schedule. Every day a vehicle performs some work, so for each day we have m working jobs to be done. Let us have the $m \times n$ matrix $\mathbf{A}$ with elements a_{ij} giving the profit gained by work of i-th vehicle on j-th day. The total profit produced by i-th vehicle is the sum of elements in row i of matrix $\mathbf{A}$. If there is a big difference between the profits gained, the drivers of vehicles will have very different wages and workloads. This can be improved by suitable permutations of elements in the individual columns of matrix $\mathbf{A}$ so that the row sums are well balanced.

The first formulation of this problem was mentioned in the 1984 article [2] by Coffman and Yannakakis. Further investigation of *Matrix Permutation Problem* (MPP) can be found in Černý [3] (in Slovak) and Tegze and Vlach [13] (in English) giving the optimality conditions for the case of two-column matrix. The approaches based on graph theory for solving graph version of the MPP are studied by Czimmermann [4, 5]. In the Dell'Olmo et al. [7] 2005 review article, 32 types of uniform k-partition problems are defined and examined from the viewpoint of computational complexity, classified by particular measure of set uniformity to be optimized. Most of the studied problems can be solved by linear time algorithms, some

[1]University of Žilina, Univerzitná 8215/1, 010 026 Žilina, Slovakia, stefan.pesko@fri.uniza.sk
[2]University of Žilina, Univerzitná 8215/1, 010 026 Žilina, Slovakia, roman.hajtmanek@fri.uniza.sk

require more complex algorithms but can still be solved in polynomial time, and three types are proved to be NP-hard.

The paper by Peško and Černý [9] presents different real situations where making managerial decisions can be influenced by methodology of fair assignment. The MPP with interval matrix presented in Peško [10] is motivated by practical need of regularity in job scheduling when inputs are given by real intervals. It is shown that this NP-hard problem is solvable in polynomial time for two-column interval matrix. Stochastic algorithm for general case of problem based on aggregated two-column sub-problems is presented. The possibility of using this algorithm as the alternative to solving the set partitioning model for a fair scheduling of workload distribution can be found in paper Peško and Hajtmanek [11].

Presently, the investigation of MPP problem continues in the working paper [1] which state the column permutation algorithm in a special case and also in the general form used in our article. The computational experiments or comparisons of different approaches to stochastic algorithm based on repeating solving aggregated two-column subproblems are given in the paper Peško and Kaukič [12]. The review of problems from years 1984-2015 can be found in paper Czimmermann, Peško and Černy [6].

We want to complete this in present article by cases of uncertain workloads which are represented by a vector of some characteristics (mean, dispersion, ranges, etc.) of profits.

2 Mathematical formulation

Throughout this paper we will denote by $\mathbf{A} = (a_{ijk})$ a nonnegative real matrix with m rows and n columns and p depths. Also we will denote by $I = \{1, 2, \ldots, m\}$, $J = \{1, 2, \ldots, n\}$ and $K = \{1, 2, \ldots, p\}$ the sets of row, column and depth indices. For each column $j \in J$ of matrix $\mathbf{A}$ we will use the notation π_j for the permutation of rows in that column.

Now, let us denote by $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ the vector of permutations of all columns of $\mathbf{A}$ and by $\mathbf{A}^\pi$ we will denote the permuted matrix itself. Thus the element in row i, column j and depth k of matrix $\mathbf{A}^\pi$ is $a_{\pi_j(i)jk}$ for $i \in I, j \in J, k \in K$. We will use the notation $\mathbf{S}^\pi = (s^\pi_{ik})$ for the matrix with m rows and p columns of row-sums of permuted matrix $\mathbf{A}^\pi$, so

$$\mathbf{S}^\pi = \begin{pmatrix} s^\pi_{11} & s^\pi_{12} & \cdots & s^\pi_{1p} \\ s^\pi_{21} & s^\pi_{22} & \cdots & s^\pi_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ s^\pi_{m1} & s^\pi_{m2} & \cdots & s^\pi_{mp} \end{pmatrix}, \quad \text{where } s^\pi_{ik} = \sum_{j \in J} a_{\pi_j(i)jk} \quad \text{for } i \in I,\ k \in K.$$

Let us have some irregularity measure f (see [13] for $p = 1$), i.e. the nonnegative real function defined on the set of nonnegative matrices in $\mathbb{R}^{m \times p}$. Following optimization problem will be called *uniform workload distribution problem* (UWDP)

$$f(\mathbf{S}^{\pi^*}) = \min \left\{ \sum_{k \in K} f(\mathbf{s}^\pi_k),\ \pi \in \Pi_{mn} \right\}, \tag{1}$$

where Π_{mn} is a set of all n-tuples of permutations of the row indices I and $\mathbf{s}^\pi_k$ is k-th column of matrix $\mathbf{S}^\pi$.

In this paper we will consider a basic irregularity measure defined on m dimensional vectors $\mathbf{x} = (x_1, x_2, \ldots, x_m)$

$$f_{var}(x_1, x_2, \ldots, x_m) = \sum_{i \in I} (x_i - \overline{x})^2; \quad \overline{x} = \frac{1}{m} \sum_{i \in I} x_i. \tag{2}$$

Note that for constant vector $\mathbf{x}^*$ with elements equal to mean value $\overline{x}$ of row sum we have $f_{dif}(\mathbf{x}^*) = f_{var}(\mathbf{x}^*) = 0$.

3 Example

The main idea of our heuristic approach will be explained in the following example where uncertain workloads are represented by interval matrix $\mathbf{A} = (a_{ij1}, a_{ij2})$, $a_{ij1} \leq a_{ij2}$, $i \in I$, $j \in J$.

Let us take the matrix $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2)$ with row-sums vector $\mathbf{S}$ and $f_{var}(\mathbf{S}) = 650$.

$$\mathbf{A}_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 1 \\ 2 & 5 & 5 \\ 2 & 5 & 6 \\ 4 & 7 & 10 \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} 2 & 1 & 1 \\ 6 & 4 & 5 \\ 4 & 13 & 8 \\ 6 & 5 & 11 \\ 5 & 8 & 16 \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} 1 & 4 \\ 3 & 15 \\ 12 & 25 \\ 13 & 22 \\ 21 & 29 \end{pmatrix}.$$

After replacing some adjacent elements marked in bold we get a solution with $f_{var}(\mathbf{S}^{'}) = 22$,

$$\mathbf{A}_1^{'} = \begin{pmatrix} \mathbf{4} & \mathbf{7} & \mathbf{1} \\ \mathbf{2} & \mathbf{5} & \mathbf{1} \\ \mathbf{0} & \mathbf{2} & \mathbf{6} \\ 2 & 5 & \mathbf{5} \\ \mathbf{0} & \mathbf{0} & 10 \end{pmatrix}, \quad \mathbf{A}_2^{'} = \begin{pmatrix} \mathbf{5} & \mathbf{8} & \mathbf{5} \\ \mathbf{4} & \mathbf{13} & \mathbf{1} \\ \mathbf{6} & 4 & \mathbf{11} \\ 6 & 5 & \mathbf{8} \\ \mathbf{2} & \mathbf{1} & 16 \end{pmatrix}, \quad \mathbf{S}^{'} = \begin{pmatrix} 12 & 18 \\ 8 & 18 \\ 8 & 21 \\ 12 & 19 \\ 10 & 19 \end{pmatrix}.$$

Next replacing gives us a better solution $\mathbf{A}''$ with $f_{var}(\mathbf{S}^{''}) = 8.0$,

$$\mathbf{A}_1^{''} = \begin{pmatrix} 4 & \mathbf{5} & \mathbf{1} \\ 2 & \mathbf{7} & \mathbf{1} \\ 0 & \mathbf{5} & \mathbf{5} \\ 2 & \mathbf{2} & \mathbf{6} \\ 0 & 0 & 10 \end{pmatrix}, \quad \mathbf{A}_2^{''} = \begin{pmatrix} 5 & \mathbf{13} & \mathbf{1} \\ 4 & \mathbf{8} & \mathbf{5} \\ 6 & \mathbf{5} & \mathbf{8} \\ 6 & \mathbf{4} & \mathbf{11} \\ 2 & 1 & 16 \end{pmatrix}, \quad \mathbf{S}^{''} = \begin{pmatrix} 10 & 19 \\ 10 & 17 \\ 10 & 19 \\ 10 & 21 \\ 10 & 19 \end{pmatrix}.$$

The solution $\mathbf{A}^{''}$ cannot be improved via changing independent adjacent elements. But there is an optimum solution $\mathbf{A}^* = (\mathbf{A}_1^*, \mathbf{A}_2^*)$ with $f_{var}(\mathbf{S})^* = 0$,

$$\mathbf{A}_1^* = \begin{pmatrix} 4 & 5 & 1 \\ 2 & 2 & 6 \\ 0 & 0 & 10 \\ 2 & 7 & 1 \\ 0 & 5 & 5 \end{pmatrix}, \quad \mathbf{A}_2^* = \begin{pmatrix} 5 & 13 & 1 \\ 4 & 4 & 11 \\ 2 & 1 & 16 \\ 6 & 8 & 5 \\ 6 & 5 & 8 \end{pmatrix}, \quad \mathbf{S}^* = \begin{pmatrix} 10 & 19 \\ 10 & 19 \\ 10 & 19 \\ 10 & 19 \\ 10 & 19 \end{pmatrix},$$

and so last solution $\mathbf{A}^{''}$ is only an approximation. For finding matrix minimizing the irregular objective function f_{var} via independent changed elements we use following mixed quadratic programming (MIQP) model.

4 MIQP model

Let us assume that a nonnegative real matrix $\mathbf{A} = (a_{ijk})$, $i \in I$, $j \in J$, $k \in K$ is given. In this paragraph we will use notation i^+ and i^- for $i \in I$ in following meaning

$$i^+ = \begin{cases} i+1, & \text{if } 1 \leq i < m, \\ 1, & \text{if } i = m. \end{cases} \qquad i^- = \begin{cases} i-1, & \text{if } 1 < i \leq m, \\ m, & \text{if } i = 1. \end{cases} \tag{3}$$

Let us define bivalent variables x_{ij} for $i \in I$, $j \in J$. Let us define also value $x_{ij} = 1$ if the elements of a_{ijk} and $a_{i^+\,j\,k}$ will be changed and $x_{ij} = 0$ otherwise. Non negative variables z_{ik} for $i \in I$, $k \in K$ give i-th row-sum in k-th depth of matrix matrix $\mathbf{A}^{\pi}$. Note that the column's permutatons $\pi_j = (\pi_{1j}, \pi_{2j}, \dots, \pi_{mj})$ are then defined by

$$\pi_{ij} = \begin{cases} i^{+}, & \text{if } x_{ij} = 1, \\ i^{-}, & \text{if } x_{i^- j} = 1, \\ i, & \text{otherwise.} \end{cases}$$

We denote mean row-sum of k-depth of matrix $\mathbf{A}$ by $\overline{s}_k = \frac{1}{m}\sum_{i\in I}\sum_{j\in J} a_{ijk}$.

Now we can formulate modifed model of mixed quadratic programming problem (MUWPD):

$$\sum_{i\in I}\sum_{k\in K}(z_{ik} - \overline{s}_k)^2 \to \min, \tag{4}$$

$$z_{ik} = \sum_{j\in J}\left[a_{i^+jk}\cdot x_{ij} + a_{i^-jk}\cdot x_{i^-j} + a_{ijk}\cdot(1 - x_{i^-j} - x_{ij})\right], \qquad \forall i\in I,\ \forall k\in K, \tag{5}$$

$$x_{ij} + x_{i^+j} \le 1, \qquad \forall i\in I,\ \forall j\in J,\ \forall k\in K, \tag{6}$$

$$x_{ij} \in \{0,1\}, \qquad \forall i\in I,\ \forall j\in J,\ \forall k\in K, \tag{7}$$

$$z_{ik} \ge 0, \qquad \forall i\in I,\ \forall k\in K. \tag{8}$$

Objective function (4) is irregularity measure f_{var} of vector of all row-sums of permuted matrix $\mathbf{A}^{\pi}$. Constraints (5) define new row-sums of this matrix. Constraints (6) define independence on changed elements. Constrains (7) and (8) are obligatory.

Consequently we can formulate a heuristic which is based on iterative improvements gained by solution the MUWPD model.

Step 1. Let $\mathbf{A} = (a_{ijk})$, $i \in I$, $j \in J$, $k \in K$ be an arbitrary nonnegative real matrix and a number of changed elements $no = n$ (fictive initialize value).

Step 2. Solve MUWPD model for matrix $\mathbf{A}$ with solution $X = (x_{ij})$ and set $no = \sum_{i\in I}\sum_{j\in J} x_{ij}$.

Step 3. If $no = 0$ then STOP, $\mathbf{A} = (a_{ijk})$ as approximate solution.

Step 4. Let matrix $\mathbf{B} = (b_{ijk})$, $i \in I$, $j \in J$, $k \in K$ where

$$b_{ijk} = \begin{cases} a_{i^+jk}, & \text{if } x_{ij} = 1, \\ a_{i^-jk}, & \text{if } x_{i^-j} = 1, \\ a_{ijk}, & \text{if } x_{ij} + x_{i^-j} = 0. \end{cases}$$

Set $\mathbf{A} = \mathbf{B}$ and GOTO **Step 2.**

5 Computation experiments

Our experiments were conducted on HP XW6600 Workstation (8-core Xeon 3GHz, RAM 16GB) with OS Linux (Debian/jessie). We used Python-based tools and the Python interface to commercial mathematical programming solver Gurobi [8].

It should be noted that Gurobi solver is able to use all processor cores, so the solution time will be much smaller comparing to computation on one-core processor. We used academic license of Gurobi, which is a full version of solver without any constraints on size of problem instances or the number of processors (on multicore computers).

Let us say that all matrices we tested were nonnegative integer-valued. In the real-world applications this can be achieved by quantifying the profits into several integer levels (seldom there are more than few hundreds levels).

We experimented with randomly generated matrices with value zero of f_{var} objective function for three types of uncertain elements given by

- intervals: $\mathbf{A}^I = (a_{ij1}, a_{ij2})$ where $0 \le a_{ij1} \le a_{ij2} \le 100$,

- mean and variance of the random variables from normal distribution: $\mathbf{A}^N = (\mu_{ij}, \sigma_{ij}^2)$ where $0 \leq \mu_{ij} \leq 10, 0 \leq \sigma_{ij}^2 \leq 100$,
- triangle fuzzy numbers: $\mathbf{A}^{TF} = (a_{ij1}, a_{ij2}, a_{ij3})$ where $0 \leq a_{ij1} \leq a_{ij2} \leq a_{ij3} \leq 1000$,

for the instances with number of rows $\times$ columns from 10×7, 10×14 and 20×7.

With runs of heuristic we achieved interesting results given below in table 1–3. Iterations in tables are numbers of solutions in **Step 2** of heuristic. Characteristic optimum gives percentage of optimal solutions from 100 generated instances.

Instances	10×7	10×14	20×7
Time (sec.)	4.3	68.6	311.7
Iterations	7.7	5.8	14.6
Optimum (%)	0	79	0
Objective f_{var}	30.8	0.7	97.8

Table 1 Mean computational characteristics for matrices $\mathbf{A}^I$

Instances	10×7	10×14	20×7
Time (sec.)	8.4	104.2	253.5
Iterations	8.2	6.1	8.1
Optimum (%)	0	87	0
Objective f_{var}	30.8	0.3	91.7

Table 2 Mean computational characteristics for matrices $\mathbf{A}^N$

Instances	10×7	10×14	20×7
Time (sec.)	12.0	564.2	239.8
Iterations	9.3	8.1	15.9
Optimum (%)	0	1	0
Objective f_{var}	137.3	13.1	588.0

Table 3 Mean computational characteristics for matrices $\mathbf{A}^{TF}$

6 Conclusion and future work

Computation experiments has shown that the presented model for the uniform workload distribution problems gives applied heuristic. Open question remains whether it is the exact method for same spacial real instances.

We believe that the presented heuristics can be improved via generalizations of exchange elements in columns of matrix not only in rows $1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow \cdots \leftrightarrow m \leftrightarrow 1$ with step 1 but also for steps $2, 3 \ldots, m-1$. That means only to redefine i^+ and i^- in (3) with i^{+p} and i^{-p} $(p = 2, 3, \ldots, m = 1)$ as use the MUWPD model. So we can construct more robust heuristic which begins with possible exchange of elements from step $m-1$, continues with step $m-2$, etc. and ends with 1.

Acknowledgements

The research was supported by the research grants APVV-14-0658 Optimization of urban and regional public personal transport.

References

[1] Boudt, K., Vandeuffel, S. and Verbeken, K.: Block Rearranging Elements within Matrix Columns to Minimize the Variability of Row Sums, Working paper (2015).

[2] Coffman. E. G. and Yannakakis, M.: Permuting Elements within Columns of Matrix in order to Minimize Maximum Row Sum, *Mathematics of Operations Research*, *9(3)*:384390 (1984).

[3] Černý, J.: How to minimize irregularity? (Slovak), *Pokroky matematiky, fyziky a astronomie*, *30*(3), 1985, 145–150.

[4] Czimmermann, P. and Peško,Š.: The Regular Permutation Scheduling on Graphs, *Journal of Information, Control and Management Systems*, *1*, (2003), 15–22.

[5] Czimmermann, P.: A Note on Using Graphs in Regular Scheduling Problems, *Communications*, *4*, (2003), 47–48.

[6] Czimmermann, P., Peško,Š. and Černý,J.: Uniform Workload Distribution Problems, *Communications*, *1A*, (2016), 56–59.

[7] DellOlmoa, P., Hansenb, P., Pallottinoc, S. and Storchid G.: On Uniform k-partition Problems, *Discrete Applied Mathematics*, *150*, Issues 1–3, (2005), 121–139.

[8] Gurobi Optimization Inc.: *Gurobi optimizer reference manual*, 2015, `http://www.gurobi.com`.

[9] Peško,Š. and Černý, J.: Uniform Splitting in Managerial Decision Making, *E+M, Economics and Management* IX, *4*, Technical university of Liberec, (2006), 67–71.

[10] Peško,Š.: The Matrix Permutation Problem in Interval Arithmetic, *Journal of Information, Control and Management Systems*, *4*, (2006), 35–40.

[11] Peško,Š. and Hajtmanek, R.: *Matrix Permutation Problem for Fair Workload Distribution*, Proceedings of mathematical methods in economics, 32nd international conference, Olomouc, 2014, 783–788.

[12] Peško,Š. and Kaukič, M.: *Stochastic Algorithm for Fair Workload Distribution Problem*, Proceedings of Informatics, 13nd international scientific conference of informatics, Poprad, IEEE, 2015, 206–2010.

[13] Tegze, M. and Vlach, M.: *On the matrix permutation problem*, Zeitschrift fr Operations Research, *30*, Issue 3, (1986), A155–A159.

Multidimensional stochastic dominance for discrete uniform distribution

Barbora Petrová [1]

Abstract. Stochastic dominance is a form of stochastic ordering, which stems from decision theory when one gamble can be ranked as superior to another one for a broad class of decision makers whose utility functions, representing preferences, have very general form. There exists extensive theory concerning one dimensional stochastic dominance of different orders. However it is not obvious how to extend the concept to multiple dimension which is especially crucial when utilizing multidimensional non separable utility functions. One possible approach is to transform multidimensional random vector to one dimensional random variable and put equivalent stochastic dominance in multiple dimension to stochastic dominance of transformed vectors in one dimension. We suggest more general framework which does not require reduction of dimensions of random vectors. We introduce two types of stochastic dominance and seek for their generators in terms of von Neumann - Morgenstern utility functions. Moreover, we develop necessary and sufficient conditions for stochastic dominance between two discrete random vectors with uniform distribution.

Keywords: Multidimensional stochastic dominance, stochastic ordering, multidimensional utility function, generator of stochastic dominance.

JEL classification: C44, D81
AMS classification: 90C15, 91B16, 91B06

1 Introduction

The concept of stochastic dominance is widely used in seeking for the optimal investment strategy. The reference strategy is represented by a random variable which corresponds, for instance, to revenue of a benchmark portfolio. Then the stochastic dominance constraints ensures that our strategy is at least as good as the reference strategy. Note that the comparison of two strategies depends on the order of considered stochastic dominance. There exists extensive theory concerning the stochastic dominance of different orders between random variables. For instance, one can find detailed description of the first order stochastic dominance in Dupačová and Kopa [3], of the second order stochastic dominance in Kopa and Post [4], of the third order stochastic dominance in Kopa and Post [5] and of the decreasing absolute risk aversion stochastic dominance in Post et al. [6]. For a higher order stochastic dominance we refer to Branda and Kopa [1]. When the investment strategy represented by a random variable is replaced by an investment strategy represented by a random vector, it is not clear how to extend the definition of the stochastic dominance to multiple dimension. One could put equivalent the multidimensional dominance to the one dimensional dominance between each coordinates of the random vectors. Dentcheva and Ruszczýnski [2] define linear stochastic dominance of random vectors which uses the mapping of the random vector to the space of the linear combination of its coordinates. Both concept are however inapplicable when considering a class of investors with multidimensional non-separable utility functions (for more detailed discussion on multidimensional utility functions see, for instance, Richard [7]). In this case there is a need of defining the stochastic dominance of random vectors without any reduction of their dimension. In the further text we restrict our attention only to the first order stochastic dominance.

The paper is organized as follows. Chapter 2 is devoted to approximation of multivariate utility function by a linear combination of very simple functions, which comes to use in further chapters when defining the stochastic dominance (to see more detailed procedure we refer to Sriboonchitta et al. [8]). Chapter 3 summarizes characteristics of the first order stochastic dominance in one dimension. Chapter 4

[1] Charles University in Prague, Faculty of Mathematics and Physics, petrova@karlin.mff.cuni.cz

extends the one dimensional first order stochastic dominance into multidimensional and describes the generator of stochastic dominance in terms of von Neumann - Morgenstern utility functions. Chapter 5 focuses on stochastic dominance between two discrete random vectors where we formulate necessary and sufficient conditions for one vector to be stochastically dominated by another one. Finally, Chapter 6 provides a brief concluding statement and considerations about the future work.

2 Approximation of multivariate utility functions

A function $u : \mathbb{R}^d \to \mathbb{R}$ is said to be nondecreasing if $\boldsymbol{x} \leq \boldsymbol{y}$ implies $u(\boldsymbol{x}) \leq u(\boldsymbol{y})$. In the whole text we employ the usual partial order relation on $\mathbb{R}^d$, that is for any $\boldsymbol{x} = (x_1, \ldots, x_d)$ and $\boldsymbol{y} = (y_1, \ldots, y_d)$ in $\mathbb{R}^d$, we say that $\boldsymbol{x} \leq \boldsymbol{y}$ if and only if $x_i \leq y_i$, $i = 1, \ldots, d$. The following definition specifies the space of all utility functions which represents non-satiated individuals.

Definition 1. We define by $\mathcal{U}_1$ the space of all nondecreasing functions $u : \mathbb{R}^d \to \mathbb{R}$.

The following definition introduces upper sets which are crucial for definition of the multidimensional stochastic dominance.

Definition 2. A subset $M \subset \mathbb{R}^d$ is called an upper set if for each $\boldsymbol{y} \in \mathbb{R}^d$ such that $\boldsymbol{y} \geq \boldsymbol{x}$ one also has $\boldsymbol{y} \in M$ whenever $\boldsymbol{x} \in M$. We denote as $\mathcal{M}$ the set of all upper sets in $\mathbb{R}^d$.

One can similarly define lower sets and the set of all lower sets in $\mathbb{R}^d$, however the former definition is sufficient for the further theory. We note in advance that all statements in the following chapters can be equivalently formulated using the concept of lower sets. The simplest upper sets in $\mathbb{R}^d$ are set with a single generator, i. e. there exists a unique point $\boldsymbol{m}^* \in \mathbb{R}^d$ such that the upper set can be given as $\{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{x} = \boldsymbol{m}^* + \boldsymbol{t}, \boldsymbol{t} \geq \boldsymbol{0}\}$. These sets are thus of the form $(\boldsymbol{m}^*, \infty) = (m_1^*, \infty) \times \cdots \times (m_d^*, \infty)$ where $\boldsymbol{m}^* = (m_1^*, \ldots, m_d^*)$.

Definition 3. We denote as $\mathcal{M}^* \subset \mathcal{M}$ the subset of all upper sets in $\mathbb{R}^d$ which have a single generator, i. e. there exists a unique point $\boldsymbol{m}^* \in \mathbb{R}^d$ such that the upper set can be given as $\{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{x} = \boldsymbol{m}^* + \boldsymbol{t}, \boldsymbol{t} \geq \boldsymbol{0}\}$.

Remark 1. In $\mathbb{R}$ all upper sets are intervals of the form (m, ∞), $m \in \mathbb{R}$ and thus the subset $\mathcal{M}^*$ coincides with the set $\mathcal{M}$.

Let $u : \mathbb{R}^d \to \mathbb{R}^+$ be a nonnegative nondecreasing function. As a measurable function, u can be approximated by the standard procedure in measure theory by the following procedure. Let

$$u_n(\boldsymbol{x}) = \begin{cases} \frac{i}{2^n} & \text{for } \frac{i}{2^n} \leq u(\boldsymbol{x}) < \frac{i+1}{2^n}, \quad i = 0, 1, \ldots, n2^n - 1, \\ n & \text{for } u(\boldsymbol{x}) \geq n. \end{cases}$$

Then $u_n(\boldsymbol{x})$ converges pointwise to $u(\boldsymbol{x})$ for each $\boldsymbol{x} \in \mathbb{R}^d$, as $n \to \infty$. Let $A_{i,n} = \{\boldsymbol{x} \in \mathbb{R}^d : u(\boldsymbol{x}) \geq i/2^n\}$ for each $n \geq 1$. These sets are upper sets, which are moreover nested, i. e., $A_{1,n} \supseteq A_{2,n} \supseteq \ldots \supseteq A_{n2^n,n}$ with possible equality for some inclusion. Thus one can rewrite u_n as a positive linear combination of indicator functions of these sets as follows:

$$u_n(\boldsymbol{x}) = \frac{1}{2^n} \sum_{i=1}^{n2^n} \mathbb{1}_{A_{i,n}}(\boldsymbol{x}). \tag{1}$$

If u is nonnegative nonincreasing function, then the above approximation still holds true. In this case, $\{A_{i,n}\}_{i=1}^{n2^n}$ is a sequence of nested lower sets.

Remark 2. If u is a one dimensional nonnegative nondecreasing function, then $A_{i,n}$ is of the form $(a_{i,n}, \infty)$, where $a_{i,n} \in \mathbb{R}$. Thus we can write each u_n as a positive linear combination of indicator functions of the form $\mathbb{1}_{(a_{i,n},\infty)}$. Specifically, we have

$$u_n(x) = \frac{1}{2^n} \sum_{i=1}^{n2^n} \mathbb{1}_{(a_{i,n},\infty)}(x).$$

3 Univariate stochastic dominance

In this part we shortly summarize characteristics of the first order stochastic dominance between random variables. More detailed description together with the proof of the beneath stated theorem can be find in Sriboonchitta et al. [8].

Definition 4. Let X and Y be two random variables with distribution functions F and G. Then X stochastically dominates Y in the first order, denotes as $X \succeq Y$, if $F(x) \leq G(x)$ for every $x \in \mathbb{R}$.

Theorem 1. *Let X and Y be two random variables with distribution functions F,G and quantile functions F^{-1}, G^{-1}. Then the following statements are equivalent:*

1. *X stochastically dominates Y in the first order,*
2. *$\mathbb{E}u(X) \geq \mathbb{E}u(Y)$ for any $u \in \mathcal{U}_1$,*
3. *$F^{-1}(\alpha) \geq G^{-1}(\alpha)$ for all $\alpha \in (0,1)$,*
4. *$VaR_X(\alpha) \leq VaR_Y(\alpha)$ for all $\alpha \in (0,1)$,*

where $\mathcal{U}_1$ is the space of all nondecreasing functions $u : \mathbb{R} \longrightarrow \mathbb{R}$.

The following theorem states several equivalent ways how to decide about stochastic dominance between two random variables.

4 Multivariate stochastic dominance

We introduce this chapter by definition of multivariate stochastic dominance formulated by Sriboonchitta et al. [8].

Definition 5. Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d-dimensional random vectors. Then $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$, denoted as $\boldsymbol{X} \succeq \boldsymbol{Y}$, if for every upper set $M \in \mathcal{M}$ one has $\mathbb{P}(\boldsymbol{X} \in M) \geq \mathbb{P}(\boldsymbol{Y} \in M)$.

In the next definition we establish a new type of stochastic dominance which requires survival distribution functions of considered random vectors. The survival distribution function $\bar{F}$ of a random vector $\boldsymbol{X}$ is defined as $\bar{F}(\boldsymbol{m}) = \mathbb{P}(\boldsymbol{X} \geq \boldsymbol{m}) = \mathbb{P}(\boldsymbol{X} \in (m_1, \infty) \times \cdots \times (m_d, \infty))$ for each $\boldsymbol{m} = (m_1, \ldots, m_d) \in \mathbb{R}^d$.

Definition 6. Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d-dimensional random vectors with survival distribution functions $\bar{F}$ and $\bar{G}$. Then $\boldsymbol{X}$ weakly stochastically dominates $\boldsymbol{Y}$, denoted as $\boldsymbol{X} \succeq_w \boldsymbol{Y}$, if for each $\boldsymbol{m} \in \mathbb{R}^d$ one has $\bar{F}(\boldsymbol{m}) \geq \bar{G}(\boldsymbol{m})$.

When we consider random variables instead of random vectors, both definition are equivalent and we only talk about stochastic dominance. Indeed, in $\mathbb{R}$ all upper sets have the form of intervals (m, ∞), $m \in \mathbb{R}$ and thus $\mathbb{P}(X \in (m, \infty)) = \mathbb{P}(X \geq m) = \bar{F}(m)$. In this case, we obtain a standard definition of one-dimensional stochastic dominance.

The next theorem describes the generator of stochastic dominance in terms of von Neumann Morgenstern utility functions.

Theorem 2. *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d-dimensional random vectors. Then $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$ if and only if $\mathbb{E}u(\boldsymbol{X}) \geq \mathbb{E}u(\boldsymbol{Y})$ for all $u \in \mathcal{U}_1$.*

Proof. We refer to Sriboonchitta et al. [8]. □

Our goal is to describe also the generator of the weak stochastic dominance in terms of von Neumann - Morgenstern utility functions. Consider a utility function $u : \mathbb{R}^d \to \mathbb{R}$ such that for each $c \in \mathbb{R}$ there exists some $\boldsymbol{x}_c \in \mathbb{R}^d$ such that $u(\boldsymbol{x}_c) = c$ and for the set of solution to equation $u(\boldsymbol{x}) = c$ the following inclusion holds true: $S = \{\boldsymbol{x} \in \mathbb{R}^d : u(\boldsymbol{x}) = c\} \in \{\boldsymbol{x}_c\} + K$, where K is the convex cone of the form $K = \{\boldsymbol{t} : \boldsymbol{t} \geq \boldsymbol{0}\}$. Moreover, sets $\{\boldsymbol{x} \in \mathbb{R}^d : u(\boldsymbol{x}) \geq c\}$ are upper sets which belong to set $\mathcal{M}^*$. Then obviously each such described utility function is included in the space $\mathcal{U}_1$ and we will show that the set of these utility functions induces the weak stochastic dominance.

Definition 7. We define by $\mathcal{U}_1^*$ the subspace of all nondecreasing functions $u : \mathbb{R}^d \to \mathbb{R}$ for which the set of solution to $u(\boldsymbol{x}) = c$ is included in the convex set $\{\boldsymbol{x}_c\} + K$, where $\boldsymbol{x}_c$ satisfies $u(\boldsymbol{x}_c) = c$ and $K = \{\boldsymbol{t} : \boldsymbol{t} \geq \boldsymbol{0}\}$.

Theorem 3. *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two d-dimensional random vectors. Then $\boldsymbol{X}$ weakly stochastically dominates $\boldsymbol{Y}$ if and only if $\mathbb{E}u(\boldsymbol{X}) \geq \mathbb{E}u(\boldsymbol{Y})$ for all $u \in \mathcal{U}_1^*$.*

Proof. Firstly, we assume that the condition $\mathbb{E}u(\boldsymbol{X}) \geq \mathbb{E}u(\boldsymbol{Y})$ is satisfied for all $u \in \mathcal{U}_1^*$. Let $\mathcal{C}_1^*$ be the subspace of $\mathcal{U}_1^*$ consisting of functions of the form $\mathbb{1}_{M^*}$, where M^* is an upper set in $\mathcal{M}^*$. Since $\mathcal{C}_1^* \subseteq \mathcal{U}_1^*$, according to the assumption $\mathbb{E}u(\boldsymbol{X}) - \mathbb{E}u(\boldsymbol{Y}) \geq 0$ for all $u \in \mathcal{C}_1^*$. Now we take an arbitrary $\boldsymbol{m} = (m_1, \ldots, m_d) \in \mathbb{R}^d$ and show that $\bar{F}(\boldsymbol{m}) \geq \bar{G}(\boldsymbol{m})$. Let $M^* = (m_1, \infty) \times \cdots \times (m_d, \infty)$ and define $u = \mathbb{1}_{M^*}$. Then we get

$$\begin{aligned}
\mathbb{E}\mathbb{1}_{M^*}(\boldsymbol{X}) \geq \mathbb{E}\mathbb{1}_{M^*}(\boldsymbol{Y}) \quad &\Longleftrightarrow \quad \mathbb{P}(\boldsymbol{X} \in M^*) \geq \mathbb{P}(\boldsymbol{Y} \in M^*) \\
&\Longleftrightarrow \quad \mathbb{P}(\boldsymbol{X} \in (m_1, \infty) \times \cdots \times (m_d, \infty)) \geq \mathbb{P}(\boldsymbol{Y} \in (m_1, \infty) \times \cdots \times (m_d, \infty)) \\
&\Longleftrightarrow \quad \bar{F}(\boldsymbol{m}) \geq \bar{G}(\boldsymbol{m}).
\end{aligned}$$

The last statement ensures the random vector $\boldsymbol{Y}$ to be weakly stochastically dominated the random vector $\boldsymbol{X}$.

To show the opposite implication, we employ the approximation of utility functions described in the previous section (Equation (1)). Let $u \in \mathcal{U}_1^*$ and let $u^+ = \max(u, 0)$ and $u^- = \min(u, 0)$ be the positive and negative parts of u. Both these functions are nonnegative, u^+ is nondecreasing and u^- is nonincreasing. The positive part u^+ can be approximated by u_n^+ which equals to a positive linear combination of indicator functions of upper sets $A_i = \{\boldsymbol{x} \in \mathbb{R}^d : u(\boldsymbol{x}) \geq i/2^n\}$. These upper sets however belong to the set $\mathcal{M}^*$ since $u \in \mathcal{U}_1^*$. According to our assumption, $\boldsymbol{X}$ weakly stochastically dominates $\boldsymbol{Y}$, which means that $\mathbb{P}(\boldsymbol{X} \in M^*) \geq \mathbb{P}(\boldsymbol{Y} \in M^*)$ for all $M^* \in \mathcal{M}^*$, and thus we must have $\mathbb{E}u_n^+(\boldsymbol{X}) \geq \mathbb{E}u_n^+(\boldsymbol{Y})$ for each n. We get, by the monotone convergence theorem,

$$\mathbb{E}u^+(\boldsymbol{X}) = \lim_{n\to\infty} \mathbb{E}u_n^+(\boldsymbol{X}) \geq \lim_{n\to\infty} \mathbb{E}u_n^+(\boldsymbol{Y}) = \mathbb{E}u^+(\boldsymbol{Y}).$$

For the negative part u^- we use approximation by u_n^- which equals to a positive linear combination of indicator functions of lower sets $A_i = \{\boldsymbol{x} \in \mathbb{R}^d : u(\boldsymbol{x}) \geq i/2^n\}$. These lower sets are elements of the set $\mathcal{L}^*$ since $u \in \mathcal{U}_1^*$. The alternative definition says that $\boldsymbol{X}$ weakly stochastically dominates $\boldsymbol{Y}$ if and only if $\mathbb{P}(\boldsymbol{X} \in L^*) \leq \mathbb{P}(\boldsymbol{Y} \in L^*)$ for all $L^* \in \mathcal{L}^*$, and thus $\mathbb{E}u_n^-(\boldsymbol{X}) \leq \mathbb{E}u_n^-(\boldsymbol{Y})$ for each n. Again, by the monotone convergence theorem,

$$\mathbb{E}u^-(\boldsymbol{X}) = \lim_{n\to\infty} \mathbb{E}u_n^-(\boldsymbol{X}) \leq \lim_{n\to\infty} \mathbb{E}u_n^-(\boldsymbol{Y}) = \mathbb{E}u^-(\boldsymbol{Y}).$$

Combining last two inequalities gives us

$$\mathbb{E}u(\boldsymbol{X}) = \mathbb{E}u^+(\boldsymbol{X}) - \mathbb{E}u^-(\boldsymbol{X}) \geq \mathbb{E}u^+(\boldsymbol{Y}) - \mathbb{E}u^-(\boldsymbol{Y}) = \mathbb{E}u(\boldsymbol{Y}),$$

which finishes the proof. □

Clearly, when $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$, then $\boldsymbol{X}$ also weakly stochastically dominates $\boldsymbol{Y}$. This stems from the fact that the simplest upper sets in $\mathbb{R}^d$ have the form $(m_1^*, \infty) \times \cdots \times (m_d^*, \infty)$, $\boldsymbol{m} = (m_1, \ldots, m_d) \in \mathbb{R}^d$ and thus if $\mathbb{P}(\boldsymbol{X} \in M) \geq \mathbb{P}(\boldsymbol{Y} \in M)$ for all upper sets $M \in \mathcal{M}$ the same inequality is valid also for special types of upper sets $M^* \in \mathcal{M}^*$. The opposite implication is however not true as shows the following example.

Example 1. Consider two random vectors with discrete joint distribution functions:

$$\begin{aligned}
\boldsymbol{X} &= (0,1) \quad \text{with probability } 1/2, \\
&= (1,0) \quad \text{with probability } 1/2, \\
\boldsymbol{Y} &= (0,0) \quad \text{with probability } 1/2, \\
&= (1,1) \quad \text{with probability } 1/2.
\end{aligned}$$

Then $F(\boldsymbol{m}) \leq G(\boldsymbol{m})$ for all $\boldsymbol{m} \in \mathbb{R}^2$. Indeed,

$$\begin{aligned}
&F(m_1,m_2) = G(m_1,m_2) = 0 && \text{if } m_1 < 0 \text{ or } m_2 < 0,\\
&0 = F(m_1,m_2) < G(m_1,m_2) = 1/2 && \text{if } 0 \leq m_1 < 1 \text{ and } 0 \leq m_2 < 1,\\
&F(m_1,m_2) = G(m_1,m_2) = 1/2 && \text{if } m_1 \geq 1 \text{ and } 0 \leq m_2 < 1,\\
&F(m_1,m_2) = G(m_1,m_2) = 1/2 && \text{if } 0 \leq m_1 < 1 \text{ and } m_2 \geq 1,\\
&F(m_1,m_2) = G(m_1,m_2) = 1 && \text{if } m_1 \geq 1 \text{ and } m_2 \geq 1,
\end{aligned}$$

and $\boldsymbol{X}$ weakly stochastically dominates $\boldsymbol{Y}$. Now consider the upper set $M = \{\boldsymbol{x} = (x_1, x_2) \in \mathbb{R} : x_1+x_2 \geq 3/2\}$, then $0 = \mathbb{P}(\boldsymbol{X} \in M) < \mathbb{P}(\boldsymbol{Y} \in M) = 1/2$ and $\boldsymbol{X}$ does not stochastically dominates $\boldsymbol{Y}$.

5 Stochastic dominance for discrete distribution

In this chapter we focus on the stochastic dominance of vectors that have a discrete distribution. The motivation is as follows. When one solves a optimization problem with stochastic dominance constraints in order to seek for an optimal strategy, the distribution of a potential and reference strategy are often of the discrete form. For instance, one simulate the possible outcomes by a scenario tree based on the observed history. We focus on the stochastic dominance of vectors that have a uniform discrete distribution. We consider a special case when both compared vectors takes on values from sets with equal cardinality.

Theorem 4. *Let $\boldsymbol{X}$ be a d-dimensional random vector with the uniform distribution on the set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ and let $\boldsymbol{Y}$ be a d-dimensional random vector with the uniform distribution on the set $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$, $\boldsymbol{x}_i \in \mathbb{R}^d$ for all $i = 1, \ldots, m$ and $\boldsymbol{y}_j \in \mathbb{R}^d$ for all $j = 1, \ldots, m$. Then $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$ if and only if there exists a permutation $\Pi : \{1, \ldots, m\} \to \{1, \ldots, m\}$ such that $\boldsymbol{x}_i \geq \boldsymbol{y}_{\Pi(i)}$ for all $i = 1, \ldots, m$.*

Proof. Firstly, assume that $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$, thus for each upper set $M \in \mathcal{M}$ one has $\mathbb{P}(\boldsymbol{X} \in M) \geq \mathbb{P}(\boldsymbol{Y} \in M)$. We show by contradiction that there necessarily exists a permutation Π such that $\boldsymbol{x}_i \geq \boldsymbol{y}_{\Pi(i)}$ for all $i = 1, \ldots, m$. Assume that such permutation does not exists, i. e. for each permutation $\Pi : \{1, \ldots, m\} \to \{1, \ldots, m\}$ there exists at least one index in $\{1, \ldots, m\}$, take an arbitrary one and denote it as i, for which $\boldsymbol{x}_i < \boldsymbol{y}_{\Pi(i)}$. We construct an upper set M as $M = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{x} = \boldsymbol{y}_{\Pi(i)} + \boldsymbol{t}, \boldsymbol{t} \geq \boldsymbol{0}\}$. Then clearly $\boldsymbol{x}_i$ does not belongs to M. Let us denote as $J \subset \{1, \ldots, m\} \setminus \{i\}$ the set of indices such that for each $j \in J$ we have $\boldsymbol{y}_{\Pi(j)} \in M$. Then we have

$$\mathbb{P}(\boldsymbol{Y} \in M) = \frac{|J|+1}{m} > \frac{|J|}{m} \geq \mathbb{P}(\boldsymbol{X} \in M),$$

where we use notation $|J|$ for cardinality of the set J. The above stated inequality however contradicts the assumption about the stochastic dominance.

Conversely, assume that there exists a permutation $\Pi : \{1, \ldots, m\} \to \{1, \ldots, m\}$ such that $\boldsymbol{x}_i \geq \boldsymbol{y}_{\Pi(i)}$ for all $i = 1, \ldots, m$ and show that then $\mathbb{P}(\boldsymbol{X} \in M) \geq \mathbb{P}(\boldsymbol{Y} \in M)$ for each upper set $M \in \mathcal{M}$. Take an arbitrary $M \in \mathcal{M}$. Let us denote as $J \subseteq \{1, \ldots, m\}$ the set of all indices such that for each $j \in J$ $\boldsymbol{y}_{\Pi(j)} \in M$. Then necessarily also $\boldsymbol{x}_j \in M$ since by assumption $\boldsymbol{x}_i \geq \boldsymbol{y}_{\Pi(i)}$ for all $i = 1, \ldots, m$ and M is an upper set. Let us, moreover, denote as $I \subseteq \{1, \ldots, m\}$ the set of indices such that for each $i \in I$ $\boldsymbol{x}_i \in M$. Then $J \subseteq I$, since $\boldsymbol{x}_i \geq \boldsymbol{y}_{\Pi(i)}$, and therefore we must have

$$\mathbb{P}(\boldsymbol{X} \in M) = \frac{|I|}{m} \geq \frac{|J|}{m} = \mathbb{P}(\boldsymbol{Y} \in M).$$

Since M was selected arbitrary, the random vector $\boldsymbol{X}$ stochastically dominates the random vector $\boldsymbol{Y}$. □

Remark 3. When $d = 1$ the theorem gives us a well known result. Consider two random vectors X and Y with m equiprobable scenarios $x_1, \ldots, x_m$ and $y_1, \ldots, x_m$. Reorder the realizations of both variables so as $x_{(1)} < x_{(2)} < \ldots < x_{(m)}$ and $y_{(1)} < y_{(2)} < \ldots < y_{(m)}$. Then the random variable X stochastically dominates the random variable Y in and only if $x_{(i)} \geq y_{(i)}$ for all $i = 1, \ldots, m$.

In practice, when we desire to decide about stochastic dominance of two random vectors which have discrete uniform distributions with the same number of possible realizations, we can proceed as follows. Assume that we want to show that the random vector $\boldsymbol{X}$ uniformly distributed on the set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ stochastically dominates the random vector $\boldsymbol{Y}$ uniformly distributed on the set $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$. For each $i = 1, \ldots, m$ we construct an upper set M_i as $M_i = \{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{x} = \boldsymbol{y}_i + \boldsymbol{t}, \boldsymbol{t} \geq \boldsymbol{0}\}$ and seek for some realization of the random vector $\boldsymbol{X}$ included in this upper set. As soon as some realization of $\boldsymbol{X}$ is assigned to some $\boldsymbol{y}_i$, the realization cannot be assigned to any other realization of $\boldsymbol{Y}$. If it is possible to assign to each $\boldsymbol{y}_i$ exactly one $\boldsymbol{x}_i$ (but not unique) then $\boldsymbol{X}$ stochastically dominates $\boldsymbol{Y}$. If there is no such solution then $\boldsymbol{X}$ cannot stochastically dominates $\boldsymbol{Y}$. In other words, we seek for the permutation described in the previous theorem.

6 Conclusion

In this paper we presented a multidimensional stochastic dominance as it was defined in Sriboonchitta et al. [8]. We establish another type of stochastic dominance, the weak stochastic dominance, and described its generator in terms of von Neumann - Morgenstern utility functions. We clarified the relationship between two types of stochastic dominance. We focused our attention to exploration of stochastic dominance between discrete random vectors since optimization problems seeking for an optimal strategy often consider the potential and reference strategies to be realizations of discrete random vectors. We stated necessary and sufficient conditions for one discrete random vector with uniform distribution to be stochastically dominated by another one and propose the algorithm verifying their relation.

In the future work, we plan to incorporate the multidimensional stochastic dominance constraints into the problem of finding the optimal investment strategy. We also intend to continue with examination of the stochastic dominance between random vectors with specific distribution. Namely, we would like to focus on random vectors with uniform distribution on convex sets and with multidimensional normal distribution. Another filed of interest is to extend the stochastic dominance to dominance of higher orders as it was done in one dimensional case.

Acknowledgements

I would like to express my gratitude to my advisor doc. RNDr. Ing. Miloš Kopa, PhD., who offered invaluable assistance and guidance. The paper is supported by the grant no. P402/12/G097 of the Czech Science Foundation and SVV 2016 no. 260334 of Charles University in Prague.

References

[1] M. Branda and M. Kopa. Dea models equivalent to general n-th order stochastic dominance efficiency tests. *Operations Research Letters*, 44:285–289, 2016.

[2] D. Dentcheva and A. Ruszczýnski. Optimization with multivariate stochastic dominance constraints. *Mathematical Programming*, 117(1):111–127, 2009.

[3] J. Dupačová and M. Kopa. Robustness of optimal portfolios under risk and stochastic dominance constraints. *European Journal of Operational Research*, 234(2):434–441, 2014.

[4] M. Kopa and T. Post. A general test for ssd portfolio efficiency. *OR Spectrum*, 37(3):703–734, 2015.

[5] M. Kopa and T. Post. Portfolio choice based on third-degree stochastic dominance. *Management Science*, 2016. doi: accepted.

[6] T. Post, Y. Fang, and M. Kopa. Linear tests for decreasing absolute risk aversion stochastic dominance. *Management Science*, 61(7):1615–1629, 2015.

[7] S. F. Richard. Multivariate risk aversion, utility independence and separable utility functions. *Management Science*, 22(1):12–21, 1975.

[8] S. Sriboonchitta, W. K. Wong, S. Dhompongsa, and H. T. Nguyen. *Stochastic Dominance and Applications to Finance, Risk and Economics.* Chapman and Hall/CRC, 2009.

The Intuitionistic Fuzzy Investment Recommendations

Krzysztof Piasecki[1]

Abstract. The return rate is considered here as an intuitionistic fuzzy probabilistic set. Then the expected return rate is obtained as an intuitionistic fuzzy subset in the real line. This result is a theoretical foundation for new investment strategies. All considered strategies are result of comparison intuitionistic fuzzy profit index and limit value. In this way we obtain an imprecise investment recommendation. Financial equilibrium criteria are a special case of comparison the profit index and the limit value. The following criteria are generalized: the Sharpe's Ratio, the Jensen's Alpha and the Treynor's Ratio. Moreover, the safety-first criteria are generalized here for the fuzzy case. The Roy Criterion, the Kataoka Criterion and the Telser Criterion are generalized. Obtained here results show that proposed theory is useful for the investment applications. Moreover, the results so obtained may be applied in behavioural finance theory as a normative model of investment's decisions.

Keywords: imprecise return rate, investment recommendation, intuitionistic fuzzy subset.

JEL Classification: C02, G11
AMS Classification: 03E72, 90B50, 91G99

1 Research Problem

Fuzzy probabilistic subset in the real line is the most common model of imprecisely estimated the return rate under quantified uncertainty. The knightian uncertainty is excluded here. From practical point-view, exclusion of knightian uncertainty is too strong assumption. Commonly recognized model of knightian uncertainty is intuitionistic fuzzy set (for short IFS) [1]. Therefore, the imprecise return rate under any uncertainty may be described by probabilistic IFS [5]. Then expected return rate will be given as IFS in the real line. In [2] and [7], we can find examples of such return rates described by probabilistic IFS which are well justified by economic reasons. Investment strategies determined by the fuzzy expected return rate are presented in [6]. The result of pursuing such strategy was assigning fuzzy financial investment recommendation to each security. The main goal of this article is consideration of investment recommendations given for securities with fuzzy probabilistic return. There will be considered the financial equilibrium criteria and the safety-first criteria. In this way, setting the investment goal will be able to take into account the imprecision risk.

2 Intuitionistic Fuzzy Expected Return Rate

Let us assume that the time horizon $t > 0$ of an investment is fixed. The basic characteristic of benefits from owning any security is its simple return rate $\tilde{r}\colon \Omega = \{\omega\} \longrightarrow \mathbb{R}$, where the set Ω is a set of elementary states of the financial market. The return rate is a random variable which is at uncertainty risk. In practice of financial markets analysis, the uncertainty risk is usually described by probability distribution of return rates $\tilde{r}_t$. This probability distribution is given by cumulative distribution function $F_r\colon \mathbb{R} \longrightarrow [0;1]$. On other side, the cumulative distribution function F_r determines probability distribution $P\colon 2^{\Omega} \supset \tilde{r}^{-1}(\mathcal{B}) \longrightarrow [0;1]$, where the symbol $\mathcal{B}$ denotes the smallest Borel σ-field containing all intervals in the real line $\mathbb{R}$.

Let us assume that considered return rate is estimated by means of probabilistic IFS $\mathcal{R}$ determined by its membership function $\rho_{\mathcal{R}} \in [0;1]^{\mathbb{R}\times\Omega}$ and nonmembership function $\varphi_{\mathcal{R}} \in [0;1]^{\mathbb{R}\times\Omega}$. For this case, expected return rate is IFS $R \in \mathcal{I}(\mathbb{R})$ given as follow

$$R = \{(x, \rho_R(x), \varphi_R(x))\colon x \in \mathbb{R}\}, \tag{1}$$

where the membership function $\rho_R \in [0;1]^{\mathbb{R}}$ and nonmembership function $\varphi_R \in [0;1]^{\mathbb{R}}$ are determined by the identities

$$\rho_R(x) = \int_{\Omega} \rho_{\mathcal{R}}(x,\omega)dP\ , \tag{2}$$

[1] Poznań University of Economics and Business, Department of Operations Research, al. Niepodleglosci 10, 61-875 Poznań, Poland, k.piasecki@ue.poznan.pl

$$\varphi_R(x) = \int_{\Omega} \varphi_{\mathcal{R}}(x,\omega)dP \tag{3}$$

and the symbol $\mathcal{I}(\mathbb{R})$ denotes the family of all IFS in the real line $\mathbb{R}$.

3 Investment Recommendations Dependent on Expected Return

The investment recommendation is the counsel given by the advisers to the investor. For convenience, these recommendations may be expressed by means of standardized advices. Many advisors use a different terminology and different number of words forming advice vocabulary [12]. In this article we will concentrate on the five-element adviser's vocabulary given as the set

$$\mathbb{A} = \{\mathcal{B}, \mathcal{A}, \mathcal{H}, \mathcal{R}, \mathcal{S}\} \quad , \tag{4}$$

where:

- $\mathcal{B}$ denotes the advice **Buy** suggesting that evaluated security is significantly undervalued,
- $\mathcal{A}$ denotes the advice **Accumulate** suggesting that evaluated security is undervalued,
- $\mathcal{H}$ denotes the advice **Hold** suggesting that evaluated security is fairly valued,
- $\mathcal{R}$ denotes the advice **Reduce** suggesting that evaluated security is overvalued,
- $\mathcal{S}$ denotes the advice **Sell** suggesting that evaluated security is significantly overvalued.

Let us take into account fixed security $\check{S} \in \mathbb{Y}$ with expected return rate $r_s \in \mathbb{R}$ where the symbol $\mathbb{Y}$ denotes the set of all considered securities. For such case advisor's counsel depends on expect return. Then the criterion for competent choice of advice can be presented as a comparison of the values $g(r_s)$ and $\hat{G}$ defined as follows:

- $g: \mathbb{R} \to \mathbb{R}$ is an increasing function of substantially justified form,
- $\hat{G}$ denotes the substantially justified limit value.

The function $g: \mathbb{R} \to \mathbb{R}$ serves as a profit index. Using this criterion we define the advice choice function $\Lambda: \mathbb{R} \longrightarrow 2^{\mathbb{A}}$ as follows

$$\begin{cases} \mathcal{B} \in \Lambda(\check{S}) \Leftrightarrow g(r_s) > \hat{G} \Leftrightarrow g(r_s) \geq \hat{G} \wedge \neg\, g(r_s) \leq \hat{G}, \\ \mathcal{A} \in \Lambda(\check{S}) \Leftrightarrow g(r_s) \geq \hat{G}, \\ \mathcal{H} \in \Lambda(\check{S}) \Leftrightarrow g(r_s) = \hat{G} \Leftrightarrow g(r_s) \geq \hat{G} \wedge\ g(r_s) \leq \hat{G}\ , \\ \mathcal{R} \in \Lambda(\check{S}) \Leftrightarrow g(r_s) \leq \hat{G} \\ \mathcal{S} \in \Lambda(\check{S}) \Leftrightarrow g(r_s) < \hat{G} \Leftrightarrow \neg\, g(r_s) \geq \hat{G} \wedge g(r_s) \leq \hat{G}. \end{cases} \tag{5}$$

In this way we assign the advice subset to the security $\check{S}$. The value $\Lambda(\check{S})$ is called the investment recommendation. This recommendation may be used as valuable starting points for equity portfolio strategies. On the other hand, the weak point of the proposed choice function is omitting the fundamental analysis result and the behavioural factors impact. Analyzing the above choice function is easy to see lack strong boundary between the advices **Buy** and **Accumulate** and between the advices **Reduce** and **Sell**. Justification for distinguishing between these advices, we can search on the basis of fundamental analysis and between the behavioral aspects of the investment process.

4 Recommendations Dependent on Intuitionistic Fuzzy Expected Return

Let us assume that expected return rate from the security $\check{S} \in \mathbb{Y}$ is represented by IFS $R \in \mathcal{I}(\mathbb{R})$ described by (1). In the first step, we extend the domain $\mathbb{R}$ of profit index to the set $\mathcal{I}(\mathbb{R})$. According to the Zadeh's Extension Principle [13], for any $R \in \mathcal{I}(\mathbb{R})$ the profit index value $g(R)$ of profit index is determined by its membership function $\gamma \in [0;1]^{\mathbb{R}}$ and nonmembership function $\delta \in [0;1]^{\mathbb{R}}$ determined as follows

$$\gamma(x) = \sup\{\rho_R(r): x = g(r)\} = \rho_R(g^{-1}(x)), \tag{6}$$

$$\delta(x) = \inf\{\varphi_R(r): x = g(r)\} = \varphi_R(g^{-1}(x)), \tag{7}$$

In this way we define intuitionistic fuzzy profit index $g: \mathcal{I}(\mathbb{R}) \longrightarrow \mathcal{I}(\mathbb{R})$.

In the second step, we extend the advice choice function domain $\mathbb{R}$ to the set $\mathcal{I}(\mathbb{R})$. In agree with the Zadeh's Extension Principle [13], the investment recommendation $\Lambda(\check{S})$ is IFS described by its membership function $\lambda(\cdot|\check{S}) \in [0;1]^{\mathbb{A}}$ and nonmembership function $\kappa(\cdot|\check{S}) \in [0;1]^{\mathbb{A}}$. Due (5), (6) the membership function $\lambda(\cdot|\check{S})$ is determined by identities

$$\lambda(\mathcal{B}|\check{S}) = \sup\{\gamma(x): x \geq \check{G}\} \wedge \inf\{\delta(x): x \leq \check{G}\}, \tag{8}$$

$$\lambda(\mathcal{A}|\check{S}) = \sup\{\gamma(x): x \geq \check{G}\} = \sup\{\rho_R(g^{-1}(x)): x \geq \check{G}\}, \tag{9}$$

$$\lambda(\mathcal{H}|\check{S}) = \sup\{\gamma(x): x \geq \check{G}\} \wedge \sup\{\gamma(x): x \leq \check{G}\}, \tag{10}$$

$$\lambda(\mathcal{R}|\check{S}) = \sup\{\gamma(x): x \leq \check{G}\} = \sup\{\rho_R(g^{-1}(x)): x \leq \check{G}\}, \tag{11}$$

$$\lambda(\mathcal{B}|\check{S}) = \sup\{\gamma(x): x \geq \check{G}\} \wedge \inf\{\delta(x): x \leq \check{G}\}, \tag{12}$$

In similar way by (5), (7) we determine the nonmembership function $\kappa(\cdot|\check{S})$. We have here

$$\kappa(\mathcal{B}|\check{S}) = \inf\{\delta(x): x \geq \check{G}\} \vee \sup\{\gamma(x): x \leq \check{G}\}, \tag{13}$$

$$\kappa(\mathcal{A}|\check{S}) = \inf\{\delta(x): x \geq \check{G}\} = \inf\{\varphi_R(g^{-1}(x)): x \geq \check{G}\}, \tag{14}$$

$$\kappa(\mathcal{H}|\check{S}) = \inf\{\delta(x): x \geq \check{G}\} \vee \inf\{\delta(x): x \leq \check{G}\}, \tag{15}$$

$$\kappa(\mathcal{R}|\check{S}) = inf\{\delta(x): x \leq \check{G}\} = \inf\{\varphi_R(g^{-1}(x)): x \leq \check{G}\}, \tag{16}$$

$$\kappa(\mathcal{B}|\check{S}) = \inf\{\delta(x): x \geq \check{G}\} \vee \sup\{\gamma(x): x \leq \check{G}\}. \tag{17}$$

In this way we define function $\tilde{\Lambda}: \mathbb{Y} \longrightarrow \mathcal{I}(\mathbb{A})$ of advice choice. The value $\Lambda(\check{S})$ is called imprecise investment recommendation. The value $\lambda(\mathcal{X}|\check{S})$ may be interpreted as a degree in which the advice $\mathcal{X} \in \mathbb{A}$ is recommended for the security $\check{S}$. Moreover, the value $\kappa(\mathcal{X}|\check{S})$ may be interpreted as a degree in which the advice $\mathcal{X} \in \mathbb{A}$ is rejected for the security $\check{S}$. Each of these values describes the investment recommendation issued to the security $\check{S} \in \mathbb{Y}$ by the advisor. Each advice is thus recommended to some extent. The investor shifts some of the responsibility to advisers. For this reason, the investors restrict their choice of investment decisions to advices which are recommended to the greatest degree and rejected to the smallest degree. In this way, the investors minimize their individual responsibility for financial decision making. This means that the final investment decision criteria may be criteria for maximizing the value $\lambda(\cdot|\check{S})$and minimizing the value $\kappa(\cdot|\check{S})$. However, final investment decision will be made by the investor. Guided by own knowledge and intuition, the investor can choose the advice which is recommended in the lower degree.

Let us note that we have here

$$\lambda(\mathcal{B}|\check{S}) = \lambda(\mathcal{A}|\check{S}) \wedge \kappa(\mathcal{R}|\check{S}), \tag{18}$$

$$\lambda(\mathcal{H}|\check{S}) = \lambda(\mathcal{A}|\check{S}) \wedge \lambda(\mathcal{R}|\check{S}), \tag{19}$$

$$\lambda(\mathcal{B}|\check{S}) = \lambda(\mathcal{R}|\check{S}) \wedge \kappa(\mathcal{A}|\check{S}), \tag{20}$$

$$\kappa(\mathcal{B}|\check{S}) = \kappa(\mathcal{A}|\check{S}) \vee \lambda(\mathcal{R}|\check{S}), \tag{21}$$

$$\kappa(\mathcal{H}|\check{S}) = \kappa(\mathcal{A}|\check{S}) \vee \kappa(\mathcal{R}|\check{S}), \tag{22}$$

$$\kappa(\mathcal{B}|\check{S}) = \kappa(\mathcal{R}|\check{S}) \vee \lambda(\mathcal{A}|\check{S}). \tag{23}$$

This shows that the functions values $(\mathcal{A}|\check{S})$, $\lambda(\mathcal{R}|\check{S})$, $\kappa(\mathcal{A}|\check{S})$, $\kappa(\mathcal{R}|\check{S})$are sufficient to determine the imprecise investment recommendation. For this reason, when considering the specific method of advice choice, we will only determine these values. Let us note that the use of imprecisely estimated return will allow to determine the differences between the advices **Buy** and **Accumulate** and between the advices **Reduce** and **Sell**.

5 Financial Equilibrium Criteria

Each financial equilibrium model is given as the comparison of expected return from considered security and expected return of the market portfolio. We will consider fixed security $\check{S}$.

5.1 The Sharpe's Ratio

In this section any security $\check{Y} \in \mathbb{Y}$ is represented by the pair $(r_y, \sigma_y^2) \in \mathbb{R}^2$, where r_y is expected rate of return from $\check{Y}$ and σ_y^2 is variance of return rate from $\check{Y}$. We assume that there exists the risk free bond represented by the pair $(r_0, 0)$ and the market portfolio represented by the pair (r_M, σ_M^2). If the security $\check{S}$ is represented by the pair (r_s, σ_s^2), then Sharpe [9] defines the profit index $g: \mathbb{R} \to \mathbb{R}$ and the limit value $\hat{G}$ as follows

$$g(r) = \frac{r - r_0}{\sigma_s}, \tag{24}$$

$$\hat{G} = \frac{r_M - r_0}{\sigma_M}. \tag{25}$$

The Sharpe's profit index estimates amount of the premium per overall risk unit. The Sharpe's limit value is equal to the unit premium of the market portfolio risk. Let us consider the case when expected return from the security $\check{S}$ is estimated as $R \in \mathcal{I}(\mathbb{R})$ described by (1). In agree with (9), (11), (14) and (16) we have here

$$\lambda(\mathcal{A}|\check{S}) = \sup\left\{\rho(\sigma_s \cdot x + r_0): x \geq \frac{r_M - r_0}{\sigma_M}\right\}, \tag{26}$$

$$\lambda(\mathcal{R}|\check{S}) = \sup\left\{\rho(\sigma_s \cdot x + r_0): x \leq \frac{r_M - r_0}{\sigma_M}\right\}, \tag{27}$$

$$\kappa(\mathcal{A}|\check{S}) = \inf\left\{\varphi(\sigma_s \cdot x + r_0): x \geq \frac{r_M - r_0}{\sigma_M}\right\}, \tag{28}$$

$$\kappa(\mathcal{R}|\check{S}) = \inf\left\{\varphi(\sigma_s \cdot x + r_0): x \leq \frac{r_M - r_0}{\sigma_M}\right\}. \tag{29}$$

5.2 The Jensen's Alpha

On the capital market we observe the risk-free return r_0 and the expected market return r_M. The security $\check{S}$ is represented by the pair (r_s, β_s), where β_s is the directional factor of the CAPM model assigned to this security. Jensen [3] defines the profit index $g: \mathbb{R} \to \mathbb{R}$ and the limit value $\hat{G}$ as follows

$$g(r|\sigma_s) = r - \beta_s \cdot (r_M - r_0), \tag{30}$$

$$\hat{G} = r_0. \tag{31}$$

The Jensen's profit index estimates amount of the premium of the market portfolio risk. The Jensen's limit value is equal to the risk-free return rate. Let us consider the case when expected return from the security $\check{S}$ is estimated as $R \in \mathcal{I}(\mathbb{R})$ described by (1). In agree with (9), (11), (14) and (16) we have here

$$\lambda(\mathcal{A}|\check{S}) = \sup\{\rho_s(r): r - \beta_s \cdot (r_M - r_0) \geq r_0\}, \tag{32}$$

$$\lambda(\mathcal{R}|\check{S}) = \sup\{\rho_s(r): r - \beta_s \cdot (r_M - r_0) \leq r_0\}, \tag{33}$$

$$\kappa(\mathcal{A}|\check{S}) = \inf\{\varphi_s(r): r - \beta_s \cdot (r_M - r_0) \geq r_0\}, \tag{34}$$

$$\kappa(\mathcal{R}|\check{S}) = \inf\{\varphi_s(r): r - \beta_s \cdot (r_M - r_0) \leq r_0\}. \tag{35}$$

5.3 The Treynor's Ratio

Let us assume the assumptions are the same as the assumptions used in the section 5.2. Additionally we assume that the security return is positively correlated with the market portfolio return. Treynor [11] defines the profit index $g: \mathbb{R} \to \mathbb{R}$ and the limit value $\hat{G}$ as follows

$$g(r) = \frac{r - r_0}{\beta_s}, \tag{36}$$

$$\hat{G} = r_M - r_0. \tag{37}$$

The Treynor's profit index estimates amount of the premium per market risk unit. The Treynor's limit value is equal to the premium of the market risk. Let us consider the case when expected return from the security $\check{S}$ is estimated as $R \in \mathcal{I}(\mathbb{R})$ described by (1). In agree with (9), (11), (14) and (16) we have here

$$\lambda(\mathcal{A}|\check{S}) = \sup\{\rho(\beta_s \cdot x + r_0): x \geq r_M - r_0\}, \tag{38}$$

$$\lambda(\mathcal{R}|\check{S}) = \sup\{\rho(\beta_s \cdot x + r_0): x \leq r_M - r_0\}, \tag{39}$$

$$\kappa(\mathcal{A}|\check{S}) = \inf\{\rho(\beta_s \cdot x + r_0): x \geq r_M - r_0\}, \tag{40}$$

$$\kappa(\mathcal{R}|\check{S}) = \inf\{\varphi(\beta_s \cdot x + r_0): x \leq r_M - r_0\}. \tag{41}$$

Investment recommendation made by means of the Treynor's Ratio is identical with investment recommendation made by the Jensen's Alpha.

6 The First Safety Criteria

We consider the simple return rate $\tilde{r}$ on the security $\check{S}$. For each assumed value $r \in \mathbb{R}$ of expected simple return rate the probability distribution of this return is given by the cumulative distribution function $F(\cdot\,|r): \mathbb{R} \longrightarrow [0;1]$ which is strictly increasing and continuous. Then the Safety Condition [8] is given in following way

$$F(L|r) = \varepsilon, \tag{42}$$

where:

– L denotes minimum acceptable return rate;

– ε is equal to probability realization of return below the minimum acceptable rate.

The realization of return below the minimum acceptable rate is identified with loss. Therefore, the variable ε denotes the loss probability. Let us consider the case when expected return from the security $\check{S}$ is estimated as $R \in \mathcal{I}(\mathbb{R})$ described by (1).

6.1 The Roy's Criterion

The Roy's Criterion [8] is that, for fixed minimum acceptable return rate L the investor minimizes the loss probability. Additionally in order to ensure financial security, the investor assumes the maximum level ε^* of loss probability. Then the profit index $g: \mathbb{R} \to [-1;0]$ and the limit value $\hat{G}$ are defined as follows

$$g(r) = -F(L|r)\,, \tag{43}$$

$$\hat{G} = -\varepsilon^*\,. \tag{44}$$

In agree with (9), (11), (14) and (16) we have here

$$\lambda(\mathcal{A}|\check{S}) = \sup\{\rho(r): F(L|r) \le \varepsilon^*\}, \tag{45}$$

$$\lambda(\mathcal{R}|\check{S}) = \sup\{\rho(r): F(L|r) \ge \varepsilon^*\}, \tag{46}$$

$$\kappa(\mathcal{A}|\check{S}) = \inf\{\varphi(r): F(L|r) \le \varepsilon^*\}, \tag{47}$$

$$\kappa(\mathcal{R}|\check{S}) = \inf\{\varphi(r): F(L|r) \ge \varepsilon^*\}. \tag{48}$$

6.2 The Kataoka's Criterion

`The Kataoki's Criterion [4] is that, for fixed loss probability ε the investor maximizes the minimum acceptable return rate. In addition, in order to ensure interest yield, the investor assumes the minimum level L^* of return. Then the profit index $g: [0;1] \to \mathbb{R}$ and the limit value $\hat{G}$ are defined in following way

$$g(r) = F^{-1}(\varepsilon|r)\,, \tag{49}$$

$$\hat{G} = L^*\,. \tag{50}$$

In agree with (9), (11), (14) and (16) we have here

$$\lambda(\mathcal{A}|\check{S}) = \sup\{\rho(r): F^{-1}(\varepsilon|r) \ge L^*\}, \tag{51}$$

$$\lambda(\mathcal{R}|\check{S}) = \sup\{\rho(r): F^{-1}(\varepsilon|r) \le L^*\}, \tag{52}$$

$$\kappa(\mathcal{A}|\check{S}) = \inf\{\varphi(r): F^{-1}(\varepsilon|r) \ge L^*\}, \tag{53}$$

$$\kappa(\mathcal{R}|\check{S}) = \inf\{\varphi(r): F^{-1}(\varepsilon|r) \le L^*\}. \tag{54}$$

6.3 The Telser's Criterion

Based on the safety and profitability of the investment, the investor assumes a minimum level L^* of acceptable return and the maximum level ε^* of loss probability. If for the security $\check{S}$ fulfils the condition

$$F(L^*|r_s) \le \varepsilon^*, \tag{55}$$

then it is called safe-haven security.

The Telsers's Criterion [10] is that the investor maximizes the return on safe-haven securities. In addition, to ensure the profitability of investments, the investor takes into account the rate $r^* > L^*$of financial equilibrium. Then the profit index $g: \mathbb{R} \to \mathbb{R}$ and the limit value $\hat{G}$ are defined in following way

$$g(r) = r\ , \tag{56}$$

$$\hat{G} = r^*\ . \tag{57}$$

In agree with (9), (11), (14) and (16) we have here

$$\lambda(\mathcal{A}|\breve{S}) = \sup\{\rho_s(r): r \geq r^*\}, \tag{58}$$

$$\lambda(\mathcal{R}|\breve{S}) = \sup\{\rho_s(r): r \leq r^*\}, \tag{59}$$

$$\kappa(\mathcal{A}|\breve{S}) = \inf\{\varphi_s(r): r \geq r^*\}, \tag{60}$$

$$\kappa(\mathcal{R}|\breve{S}) = \inf\{\varphi_s(r): r \leq r^*\}. \tag{61}$$

Summary

Imprecision is relevant to the investment process. Imprecise estimate of the expected return could be a consequence of taking into account behavioural aspects of the investment process. The results so obtained may be applied in behavioural finance theory as a normative model of investment's decisions. These results may provide theoretical foundations for constructing an investment decision support system. Thus, we have shown here that the behavioral premises can influence investment recommendations in a controlled manner.

Applications of the above normative models cause several difficulties. The main difficulty is the high formal and computational complexity of the tasks involved in determining the membership function for the imprecise recommendations. The computational complexity of these models is the price we pay for the lack of detailed assumptions about the return rate. On the other hand, the low logical complexity is an important attribute of each formal model.

In this paper, the main cognitive result is to propose general methodology for imprecise investments recommendations. Moreover, the paper also offers original generalization of the financial equilibrium criteria and of the first safety criteria to the intuitionistic fuzzy case.

Acknowledgements

The project was supported by funds of National Science Center - Poland granted on the basis the decision number DEC-2012/05/B/HS4/03543.

References

[1] Atanassov K., Stoeva S.: Intuitionistic fuzzy sets. In: *Proceedings of Polish Symposium on Interval and Fuzzy Mathematics* (Albrycht J., Wiśniewski H., eds.). Poznań, 1985, 23-26.

[2] Echaust K., Piasecki K.: Black-Litterman model with intuitionistic fuzzy posterior return, *SSRN Electronical Journal* (2016).

[3] Jensen M.C.: Risk and pricing of capital assets, and the evaluation of investments portfolios. *Journal of Business* **42**, 2 (1969), 167-247.

[4] Kataoka S.: A stochastic programming model. *Econometrica* **31**, 1/2 (1963), 181-196.

[5] Zhang, Q., Jia B., and Jiang S.: Interval-valued intuitionistic fuzzy probabilistic set and some of its important properties. In: *Proceedings of the 1st International Conference on Information Science and Engineering ICISE2009*. Guangzhou, 2009, 4038-4041.

[6] Piasecki K.: On imprecise investment recommendations. *Studies in Logic, Grammar and Rhetoric* **37**, 50 (2014), 179-194.

[7] Piasecki K.: On return rate estimated by intuitionistic fuzzy probabilistic set. In: *Mathematical Methods in Economics MME 2015* (Martincik D., Ircingova J., Janecek P., eds.). Publishing House of Faculty of Economics, University of West Bohemian, Plzen, 2015, 641-646.

[8] Roy A.D.: Safety-first and the holding of assets. *Econometrics* **20** (1952), 431-449.

[9] Sharpe W.F.: Mutual fund performance. *Journal of Business* **19** (1966), 119-138.

[10] Telser L.G. : Safety first and hedging. *The Review of Economic Studies* **23**, 2 (1965), 1-16.

[11] Treynor J.L.: How to rate management of investment fund. *Harvard Business Review* **43** (1965), 63-75.

[12] www.marketwatch.com/tools/guide.asp (access 01.09.2013).

[13] Zadeh L.: The concept of a linguistic variable and its application to approximate reasoning, Part I. *Inf Sci.* **8** (1975), 199-249.

A decent measure of risk for a household life-long financial plan – postulates and properties

Radoslaw Pietrzyk[1], Pawel Rokita[2]

Abstract. A measure of risk that is well suited to a life-long financial plan of a household shell differ from other popular risk measures, which have originally been constructed for financial institutions, investors or enterprises. It should address threats to accomplishment of the household's life objectives and must take into account its life cycle. The authors of this research have recently proposed several candidates for such measures. This article presents, in turn, a discussion about general properties that should be fulfilled by a risk measure in household financial planning. At the current stage of the discussion, the posited postulates are rather of a qualitative nature, but they may also serve in the future as a conceptual background of a set of more formalized ones. They may be a kind of analogue of the conditions set by Artzner, Delbaen, Eber and Heath (coherent measure of risk) or by Föllmer and Schied (convex measure of risk). They should, however, better comply with the life-long household financial planning perspective. This means, amongst others, including multiple financial goals and their timing. Also threats to the aim of life standard preservation should be reflected.

Keywords: household financial planning, coherent measures of risk, integrated risk measure, life cycle.

JEL Classification: D14
AMS Classification: 91B30

1 General types of risk in household financial planning

In [6] we presented a number of proposals of household financial plan risk systematizations, with regard to different sets of criteria. That part of the discussion was concluded with the following list of risk types:
1. Life-length risk;
2. Risk of investment and financing;
3. Income risk;
4. Risk of events (insurance-like events);
5. Risk of goal realization;
6. Operational risk of plan management (particularly risk of plan implementation);
7. Model risk.

The choice of this list is dictated, to some extent, by the construction of the household financial plan optimization model that is used in [6], but, at the same time, it well reflects real-life risk situation of households.

When formulating this systematization, the following elements were investigated for each type of risk: the most natural candidate for the risk variable, the way in which the given type of risk influences household finance, the way it is or may be taken into account in our household financial plan optimization model (described in [9], [8] and [6]), the way in which it is or may be measured, and the possibility of steering of this risk by the household.

The types of risk that are present in household financial planning are very different in their nature (different risk variables, different risk factors influencing the risk variables, different statistical properties of the risk factors). This heterogeneity is, however, not the only problem of household risk analysis. As with some other situations when the success of a project is threatened by many risk factors of different nature, it is very important to avoid falling into a trap of spurious local risk mitigation. That is, the fault of such local reducing of particular risk types, which does not contribute to a successful overall risk reduction. This problem is the main motivation of using integrated risk measures. The existing risk measures, that are used in integrated risk management of financial institutions and enterprises, cannot be directly transferred to household financial planning. The main reason of this

[1] Wroclaw university of Economics/Department of Financial Investments and Risk Management, ul. Komandorska 118/120, 53-345 Wrocław, radoslaw.pietrzyk@ue.wroc.pl.
[2] Wroclaw university of Economics/Department of Financial Investments and Risk Management, ul. Komandorska 118/120, 53-345 Wrocław, pawel.rokita@ue.wroc.pl.

fact is that household financial plans need different measure of success (or failure) than a financial institution or enterprise.

Important concepts underlying most risk models are *risk variable* and *benchmark*. A risk variable is a random variable, whose value may be interpreted as a measure of success (or failure) in a decision problem under risk. In the classical decision-making theory it is just the payoff (or sometimes regret – i.e., opportunity loss). In risk analysis for financial institutions it is the value (e.g., of an investment portfolio or of the whole institution). In risk analysis for enterprises it is usually net cash flow or earning. The benchmark, in turn, may be just the expected value of the risk variable, or some other distinguished value, like, for example, a desired or planned one. Generally speaking, vast majority of popular risk measures describe potential behaviour of the risk variable in relation to the benchmark. What is, then, the risk variable and the benchmark for measuring of risk in household financial planning?

We propose to use cumulated net cash flow as a risk variable. We do not impose any particular benchmark yet. Different measures of risk may use different benchmarks. Our concern is rather whether the risk measures fulfil our postulates on a decent integrated risk measure for household financial planning. Moreover, the nature of the problem is such that the behaviour of the cumulated net cash flow during the whole life of the household needs to be taken into consideration. And a measure of risk should, thus, look at the whole term structure of the cumulated net cash flow, and, of course, it must do it for various scenarios.

Certainly, from the point of view of risk analysis, the most important values of the net cash flow process are negative ones, called *shortfalls*. In the process of the cumulated net cash flow, there may also occur *cumulated shortfalls*, which are more serious than just one-period shortfalls. The most serious ones are *unrecoverable shortfalls*. An unrecoverable shortfall may be defined as such, after which, under a given scenario, the cumulated net cash flow is negative and it remains negative until the end of the household (for this scenario). The ability of the household to recover from shortfalls depends not only on their sizes, but also on their timing. This is because some moments in the life cycle of a human, as well as in the life cycle of a household, are better and some other are worse for facing a financial shortfall (age, stage of the career, responsibility and obligations towards other persons, etc.).

2 Specificity of household financial planning

Before presenting some ideas concerning the risk measures, let us discuss the general properties of household finance and household financial planning, that make risk measurement for them different from that of financial institutions and enterprises.

An important term in the discussion on risk in household financial planning is a financial goal. In [10] we proposed to distinguish a general, qualitative, category of life objectives and a quantitative, cashflow-based, category of financial gals. Life objectives are just the ideas on how the household members want their life to look like in the future. *Financial goals* are understood as the goals of being prepared for expenses that are necessary to accomplish the life objectives. They are defined much more precisely than the corresponding life objectives. Namely, they are given in terms of pre-planned magnitudes and pre-planned moments of negative cash flows. The actual making of the expenditures is, in turn, called *realization of the financial goals*.

Financial planning deals with financial goals, which are financial representations of life objectives.

To the main features of household finance, that distinguish households from other entities, belong the following:

1. Human life-cycle concerns, especially:
   a) long-term planning,
   b) human-capital-to-financial-capital transfer (see, e.g., [5]),
   c) the fact that labour-income-generating potential (i.e., human capital) is developed years before one may start to use it (childhood, youth, early career),
   d) parenthood considerations – in most cases, an important part of human life cycle is the period of parenthood, during which significant changes both to income structure and consumption patterns are imposed,
   e) the fact that labour-income-generating potential (human capital) is usually depleted long before human life ends (vast part of personal financial plan must assume non-labour sources of income),
   f) the specificity of retirement goal (relatively big magnitude and lack of post-financing possibility).
2. Multitude of financial goals and ways of their financing.
3. All-goal-oriented planning (the general aim of life-long financial planning is accomplishing of all financial goals, if possible, with a special emphasis on retirement goal, as pointed out in the point 1.e, rather than maximization of value or profit).

4. Multi-person concerns:
   a) the role of internal risk sharing and capital transfer between household embers (compare [7], [2]),
   b) the problems of multivariate survival process modelling (compare [4]).
5. Specificity of household preferences with regard to goals, especially:
   a) inapplicability of the assumption of separable preferences to household financial goals,
   b) lack of preference transitivity, and thus – impossibility of defining a hierarchy of goals (except some simple cases of few non-complementary goals).
6. Specificity of household preferences with regard to intertemporal choice
   a) inapplicability of the simple rule that the nearer future is more important than the distant future in preference modelling for cumulated net cash flow term structure (what makes sense with regard to consumption time preference, would be an absurd assumption when considering cumulated surplus or cumulated shortfall timing) – see [6],
   b) the role of age-related costs of financing (see [10]) in modelling of time preferences.
7. Specificity of household risk:
   a) many types of risk; risk factors of different nature (in the sense of their economic interpretation) and with different statistical properties (distributional models, dynamics, observability, etc.)
   b) no consensus on how to integrate the types of household risk in one model,
   c) no obvious success-failure measure(s) to serve as risk variable(s), nor any unambiguous candidate for a benchmark that could be interpreted as an indicator of a normal (typical) or desired state,
   d) special role of life-length risk (and of the survival models that are used in its modelling – compare Yaari [11] model, and also its numerous modifications and augmentations),
   e) very long risk management horizon.

Taking the above-listed features into consideration, and building on some more general requirements that are set for any risk measures, we made an attempt to formulate a list of postulates that a household financial plan risk measure should fulfil. Before, let us briefly describe the input and output of the cumulated net cash flow model, which underlies the model of risk.

3 The model of household's cumulated net cash flow

The first step in defining the measure of risk is describing the risk variable. We propose to use cumulated net cash flow in this role. The measure of risk will look at its whole term structure throughout the lifespan of the household. Of course, the length of the analysed period may be different in each survival scenario. A general scheme of the cumulated net cash flow model is presented in the figure 1.

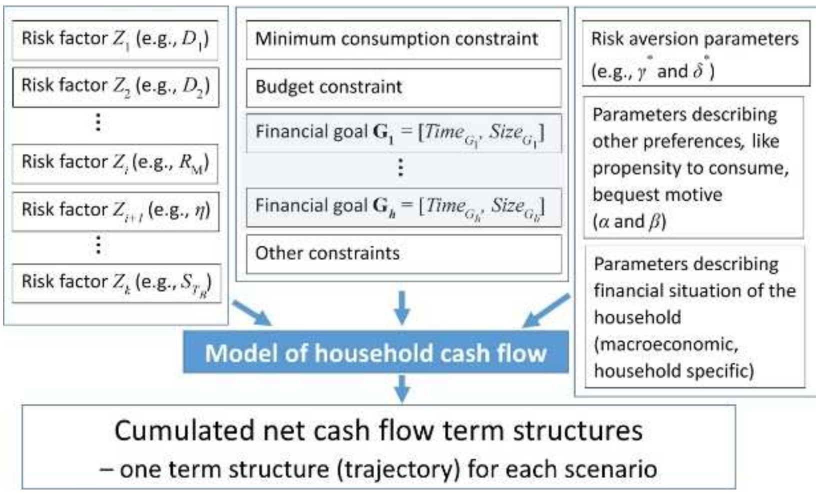


Figure 1 Household cumulated net cash flow model

Note (fig.1): Examples of risk factors: D_1 – the moment of death of the Person 1, D_2 – the moment of death of the Person 2, R_M – expected rate of return on equity portfolio, η – risk free rate, S_{TB} – real estate price at the moment of household end (moment of bequeathing). Risk aversion parameters: γ, δ – the numbers of years before and after the expected date of death, respectively, that determine bounds of the subset of all possible survival scenarios which the household wants to include into the optimization procedure. Other preferences: α, β – parameters describing propensity to consume and bequest motive.

One may distinguish three blocks of input in the scheme of input and output of the cumulated net cash flow model. These are:

- risk factors, which are random variables (left input segment of figure 1);
- constraints (central input segment in figure 1), amongst which the financial goals of the household belong;
- parameters (right input segment in figure 1).

Then, when the financial plan is optimized, some of the parameters from the right input segment are no longer fixed parameters, but they become decision variables of the optimization procedure. For instance, in the basic variant of the model, described in [9], the decision variables are:

- consumption rate on the first day of the plan,
- consumption-investment proportion,
- retirement investment contribution division (given as the proportion of the household's total retirement investment expenditure per period which is contributed to the private pension plan of the Person 1).

Each realization of the random vector from the left input segment (risk factors) is a *scenario*. As the scenarios contain realizations of several risk factors, corresponding to different types of risk, the resulting bunch of term structures (trajectories) of cumulated net cash flow also contains the information about these risk types. Moreover, as accomplishment of all financial goals is imposed as a constraint, any shortfalls occurring in the term structure of cumulated net cash flows indicate some problems with financial goal realization. But a negative cumulated net cash flow (cumulated shortfall) is not always a threat to the household. And even when it is, its severity is not always the same. The severity of the shortfalls depends not only on their sizes but also on their timing. Figures 2-5 show pairs of cumulated net cash flow term structures which show some similarities but are very different from the practical point of view.

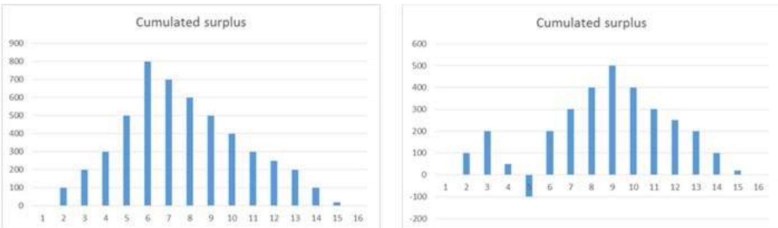


Figure 2 Cumulated net cash flow term structures with positive and equal final wealth

Note (fig. 2): The left plot shows a cumulated net cash flow term structure with no shortfall and the right plot – with a recoverable shortfall somewhere along the line; in both cases the household leaves the same bequest.

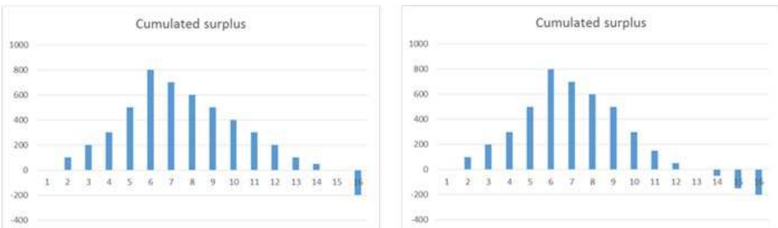


Figure 3 Cumulated net cash flow term structures with equal shortfalls at the end

Note (fig. 3): Both plots show cumulated shortfalls at the end, the shortfalls are of the same size, but the periods during which the household remains under shortfall are of different lengths (the situation of the household whose cumulated net cash flow is presented on the right panel starts to deteriorate two years earlier, but its deterioration speed is lower).

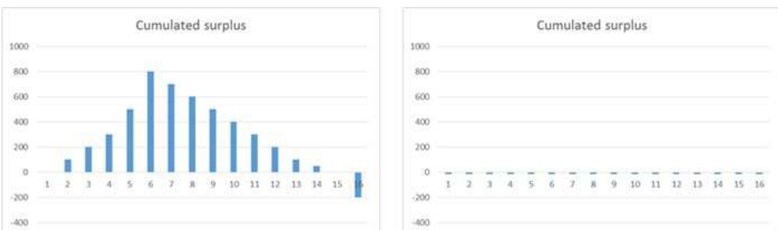


Figure 4 Cumulated net cash flow term structures with equal sums of discounted shortfalls

Note (fig. 4): Sums of shortfalls, taken time value of money into account, are equal, but severity of one big shortfall in an old age is much more serious than that of permanent small shortfalls, which might be, for instance, the result of a rolled-over debit of a modest size (not recommended but also not very dangerous).

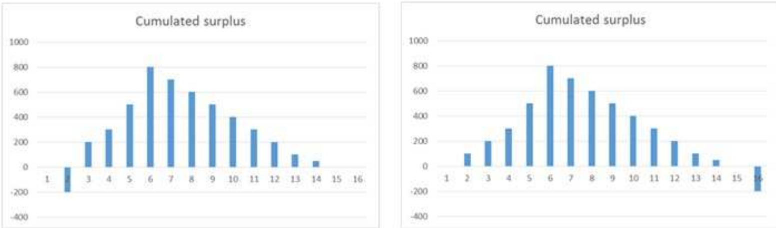


Figure 5 Cumulated net cash flow term structures with single instances of shortfalls of the same size (time value of money taken into account), but with different timing

Note (fig. 5): Here, the shortfall occurs only once and in both cases its value is the same (time value of money taken into account), but, in the case presented in the left plot, the household recovers from it (the plan is rather successful – at least under this particular scenario), whereas the shortfall from the left plot is unrecoverable (the plan is failed).

From the figures 2-5 one may grasp an idea of how important the shape of cumulated net cash flow term structure is. This means that the risk measure should be defined for term structures, not just for point values of cumulated surplus or shortfall.

4 Postulates on a household financial plan risk measure

In 1999 Artzner et al. [1] introduced the notion of *coherent measures of risk*, extended then by Föllmer and Schied [3] to the concept of *convex measures of risk*. The logical consistency and straightforward financial interpretation of the convex risk measure concept make it attractive both for theoreticians and practitioners of finance. Unfortunately, it has no direct implementation to the case of the life-long financial plan for a household. This is mainly because the risk realization in households is hardly ever understood as a loss of value of some assets. This may be a good interpretation for some subtypes of household risk, but does not apply to integrated risk of the whole financial plan. Here we make an attempt to identify the requirements that should be met by an integrated risk measure for a household whole-life financial plan. At the current stage of our research it is not a formalized definition of such measure yet, but rather a list of features that it should possess.

The features that a decent integrated risk measure for household financial planning should have are as stated by the Postulate1.

Postulate 1. *Postulates on integrated risk measure for household financial plans:*

Postulate 1.1. (An analogue to subadditivity) The value of the risk measure of a common plan for a two person household is lower or equal than the sum of values of this measure for two plans of the two persons (treated separately in the financial sense) – assuming that other conditions are unchanged.

Postulate 1.2. (Transitivity) If a plan *A* is not more risky than *B* and *B* is not more risky than *C*, then the plan *A* is also not more risky than *C*.

Postulate 1.3. The measure should reflect the size of a potential shortfall.

Postulate 1.4. The measure should be sensitive to the phase of the household life cycle in which the potential shortfall will be encountered.

Postulate 1.5. The measure should reflect the length of period (periods) during which the potential shortfall will be faced by the household.

Postulate 1.6. The measure should incorporate the information about probability of shortfalls (e.g., number of scenarios with shortfalls and probabilities of these scenarios).

Postulate 1.7. The measure should be in conformity with the interpretation that a scenario which is ending with a surplus (over a benchmark) is better than a scenario ending with a shortfall (bellow the benchmark).

To that, a postulate referring to plan optimization is needed. The risk optimization procedure must be consistent with the household preference model and with the risk measure.

Postulate 2. *The higher risk aversion declared by the household members, the less risky the optimal plan.*

The inverse does not hold, because a lower risk aversion should not automatically mean a more risky optimization outcome (the resulting plan may be more risky indeed, but only if the expected value of a measure of success for this more risky outcome is higher than for a less risky one).

The postulate 1.1. says that risk sharing within a household (obtained through internal capital transfers) causes reduction of risk in the financial plan, or at least does not increase it. The postulate 1.2. guarantees basic logical consistency. The postulates 1.3.-1.7. all together impose the requirement that the measure looks at the whole process of cumulated net cash flows, takes into account shapes of its trajectories under all considered scenarios, recognizes shortfalls as instances of risk realization, and depends on timing of shortfalls and on how likely a shortfall-generating scenario is. The last postulate (1.7.) is that the risk measure should reflect also the amount of residual wealth (that may be bequeathed). This, roughly speaking, means that such plans, for which the residual wealth is more likely to be negative (or below some benchmark) are more risky than those that rather end with a surplus.

5 Conclusions

A success or failure of a life-cycle-spanning financial plan depends on household's ability to accomplish multiple life objectives in the framework of the plan, under different scenarios. The life objectives may find their financial interpretation in the household's ability to cover the expenses that are necessary to realize them. The aim of being prepared to spend a defined sum of money at a pre-planned moment is called here a "financial goal", and the actual expense itself is called "realization of the financial goal". If, under a considered set of scenarios, realization of the financial goals does not harm household finance, the plan is not risky within this set of scenarios. Provided that the full realization of all financial goals is imposed, the plan is risky if there exists a subset of the considered set of scenarios in which the household would incur shortfalls. Severity of the shortfalls depends on how unlikely it is to recover from them, which, in turn, depends on their timing. The timing of shortfalls must be considered in relation to the phases of the household life cycle, taking into account the dates of some important events in the lives of its members. This refers, moreover, not only to the expected scenario, but to a whole bunch of considered scenarios. No classical measure of risk fulfils these conditions. We put forward some proposals of candidates for integrated risk measures of household financial plan in [10] and [6], but no postulates on such measures have been explicitly formulated so far. Therefore, there did not exist any criteria, according to which the measures might have been evaluated. Now, with the postulates that are proposed here, the direction of further research may be to check whether the proposed measures fulfil the requirements. If none of them conforms to all the conditions, then the next research objective will be to develop one that meets these requirements.

References

[1] Artzner, P., Delbaen, F., Eber, J. M., and Heath, D.: Coherent Measures of Risk. *Mathematical Finance* **9**(3), 1999, 203–228.

[2] Brown, J. R. and Poterba J. M.: Joint Life Annuities and Annuity Demand by Married Couples. *Journal of Risk and Insurance* **67**(4), 2000, 527–554.

[3] Föllmer, H. and Schied, A.:Convex measures of risk and trading constraints. *Finance and Stochastics* **6**(4), 2002, 429–447.

[4] Hurd, M. D.: Mortality Risk and Consumption by Couples. *NBER Working Paper Series*, 7048, 1999.

[5] Ibbotson, R. G., Milevsky, M. A., Chen, P., and Zhu, K. X.: *Lifetime financial advice: human capital, asset allocation, and insurance*. Research Foundation of CFA Institute, Charlottesville, Va., 2007.

[6] Jajuga, K., Feldman, Ł., Pietrzyk R., and Rokita, P.: *Integrated Risk Model in Household Life Cycle*. Publishing House of Wrocław University of Economics, Wrocław, 2015.

[7] Kotlikoff, L. J. and Summers, L. H.: The Role of Intergenerational Transfers in Aggregate Capital Accumulation. *Journal of Political Economy* **89**(4), 1981, 706–732.

[8] Pietrzyk, R. and Rokita, P.: Facilitating Household Financial Plan Optimization by Adjusting Time Range of Analysis to Life-Length Risk Aversion. In: *Analysis of Large and Complex Data* (Wilhelm, A. F. X. and Kestler, H. A., eds.), Springer International Publishing, Cham, 2016 (in press).

[9] Pietrzyk, R. and Rokita, P.: On a concept of household financial plan optimization model. *Research Papers of Wrocław University of Economics*, 381, 2015, 299–319.

[10] Pietrzyk, R. and Rokita, P.: Stochastic Goals in Household Financial Planning. *Statistics in Transition. New Series* **16**(1), 2015, 111–136.

[10] Pietrzyk, R. and Rokita, P.: Integrated measure of risk in a cumulated-surplus-based financial plan optimization model for a 2-person household. In *33rd International Conference Mathematical Methods in Economics. Conference Proceedings* (Martinčík, D., Ircingová, J. and Janeček, P., eds.), Cheb, 2015, 647–652.

[11] Yaari, M. E.: Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *The Review of Economic Studies* **32**(2), 1965, 137–150.

Optimum classification method of critical IT business functions for efficient business impact analysis

Athanasios Podaras[1]

Abstract. Business Impact Analysis (BIA) helps develop business recovery objectives by determining how disruptions affect various organizational activities. The criticality ranking of a given IT business function is an informally conducted task during the BIA procedure. Some experts underline that various market intermediaries relate the criticality of a given function to its estimated recovery time or to the impact that an outage would have from a financial, customer or other aspect. Other experts state that high criticality ranking is related to business functions that contain intricate and complex procedures. The current paper introduces a standard mathematical decision making approach for the classification of critical business functions based on their complexity. The method is based on the Use Case Points method for software complexity estimation. Its formal illustration is based on decision tree classification algorithms implemented with the R-Studio software. The approach aims to the simple and effective conducting of the BIA process for a given business function.

Keywords: business continuity management (BCM), business impact analysis (BIA), critical business functions, complexity, decision tree classification

JEL Classification: C88
AMS Classification: 68W01

1 Introduction

Business Continuity is the management of a sustainable process that identifies the critical functions of an organization and develops strategies to continue these functions without interruption or to minimize the effects of an outage or a loss of service provided by these functions [15]. Miller and Engemann [12] mention that the BCP process involves analyzing the possible threats, crisis events, the risks, and the resulting business impact, and from this analysis, developing and implementing strategies necessary to ensure that the critical components of an organization function at acceptable levels. The importance of a fully organized Business Continuity Management policy in modern organizations, is remarkably highlighted by multiple researchers and practitioners. Ashurst et al [2] state that Business Continuity (BC) ensuring is a major strategic objective for many organizations. Antlová et al [1] mention that BCM is considered to be one of the key elements of ICT competencies.

An important part of the BCM process is the implementation of a thorough Business Impact Analysis (BIA). Business Impact Analysis (BIA) helps develop business recovery objectives by determining how disruptions affect various organizational activities [11]. Moreover, one of the most important part of BIA is the efficient prioritization of the company's objectives. The specific task is performed so that the most critical business functions will be restored first and within an accepted timeframe. Despite the significance of the specific task, modern enterprises do not follow a standard classification method for their critical business functions. On the contrary, the criticality estimation of a given business function is a vaguely performed task during the Business Impact Analysis (BIA) process, since the BIA process itself requires interviews with senior managers who may have different perspectives [5]. The current paper introduces a modern approach for the standard and optimum classification of core enterprise IT business functions. The contribution involves mathematical equations as well as decision making algorithms for its standardization.

2 Background-Problem statement

One of the crucial steps of the business continuity management process is the determination of the most critical business functions of an organization, as well as the resumption timeframes of these functions in a proactive and predictive manner. Based on a definition of the European Banking Authority [4], a *function,* is a structured set of activities, services or operations that are delivered by the institution or group to third parties (including its clients, counterparties, etc.). The concept of "critical function" is intrinsically linked to the concept of the underlying services that are essential to the setting up of the deliverable services, relationships or operations to third parties.

[1] Technical University of Liberec, Economics Faculty, Department of Informatics, Voroněžská 13, Liberec 1,460 01, athanasios.podaras@tul.cz

When mathematical models and software tools are proposed towards the implementation of Business Continuity/Disaster Recovery policy, they do not include any accurate or standard procedure towards the criticality ranking of the business functions.

The International Organization of Securities Commissions (IOSCO) [9], underlines that various market intermediaries in different countries relate criticality of a given function with the estimated recovery time, while other support that criticality is related to the impact that an outage would have from a financial, customer, reputational, legal or regulatory. Another study performed by The Rowan University Business Continuity Management program (Rowan University BCM Policy, 2014) [14], underlines that high criticality ranking is related to business functions that contain intricate and complex procedures. Still, the diversity of opinions towards the criticality ranking of organizational functions reveals the lack as well as the necessity of the presence of standard classification procedures based on strong mathematical methods and decision making algorithms.

Considering the above statements, the author of the current work introduces a modern method for the optimal classification of IT business functions in an enterprise. The core idea behind the contribution is that the criticality ranking of a given business function is based on its complexity. Organizational literature has considered complexity as an important factor in influencing organizations [7]. The method follows the principles of the Use Case Points [10] method for estimating software complexity. The basis of the current contribution is to classify a business function according to the Unadjusted Points value as calculated by Karner. The algorithmic procedure for deriving precise criticality ranking for any business function follows the rules of the binary classification decision trees. The optimum criticality ranking of a given function will lead to its reasonable and objective business impact analysis including the estimation of its recovery time as well as the recommendation of appropriate types of business continuity exercises based on various types of recovery scenarios.

3 Methods and tools

3.1 Calculation of the unadjusted points of a business function according to Use Case Points

According to Karner [10], the first part of the Use Case Points method for estimating software complexity, is devoted to the calculation of the *Unadjusted Points.* The specific part of the approach classifies *Actors* and *Use Cases* and estimates the total value of unadjusted points. Of course the specific calculation is not enough for estimating the required time for the development of a software application, and Karner introduced various factors (Technical and Environmental) which are also taken into consideration throughout the specific process. However, the specific value is considered as an initial software complexity index.

The current contribution, applies similar principles to estimate an initial index for the complexity of an IT business function. The differentiation from the Use Case points, is that the Actors are further classified as *Human Level Actors* and *Application Level Actors.* Moreover, the Use Cases are now replaced by *Business Functions*. The classification of the actors and the business functions is depicted at *Table 1.*

Human Level Actors	**Application Level Actors**	**Actor's Weight**	**Business Function**	**Weight of a Business Function**
Simple (Employee)	Simple (External System with a defined API (Application Interface))	0.5	Simple (<=3 business processes)	0.5
Average (Supervisor)	Average (External system interacting through a protocol such as TCP/IP)	1	Average (>=4 and <=7 business processes)	1
Complex (Business Manager/Expert)	Complex (Person interacting via GUI (Graphical User Interface	1.5	Complex (>7 business processes)	1.5

Table 1 Classification of actors and business functions according to the contribution

The derived equations of the contribution are similar to the equations of the Use Case Points. Thus, the Unadjusted Points are formulated as follows:

$$UHW = \sum_{i=1}^{n} \left(A1_i \times W_i \right) \tag{1}$$

where UHW is the Unadjusted Human Weight value, $A1_i$ is Human Level Actor i, and W_i is the Actor's ith Weight, in order to compute Human Level Actors,

$$UAPW = \sum_{i=1}^{n} \left(A2_i \times W_i \right) \tag{2}$$

where UAPW is the Unadjusted Application Weight value, $A2_i$ is Application Level Actor i, and W_i is the Actor's Weight, and

$$UBFW = \sum_{i=1}^{n} \left(BP_i \times W_i \right) \tag{3}$$

where, UBFW= Unadjusted Business Function Weight, n = Number of included Business Processes, (BP_i) is the Type of the given Business Process i and W_i is the Weight of the corresponding Business Process. Based on the above equations it is concluded that the total number of unadjusted points should be provided by the sum of UHW, UAPW and UBFW values. The derived value is entitled Unadjusted Business Function Recovery Points (UBFRP). The word recovery indicates that the specific value will be considered for the further estimation of the demanded recovery time effort. The detailed description of the estimation of the recovery time is described by the author in [13].

The derivation of the UBFRP value helps in the initial, rapid, easy but standard classification of a given business function according to its complexity. The precise classification of the business function is described in the results section of the current paper. For the estimation of the approximate recovery time in case of an unexpected failure, more factors have to be considered and their description is beyond the scope of this article.

3.2 General classification of IT business functions

Gibson [6] classifies critical IT business functions into four (4) different levels, namely Impact Value Levels, based on two important values that is, the *Recovery Time Objective (RTO)* and the *Maximum Accepted Outage (MAO)*. These values indicate a logical and a maximum timeframe within which an IT business function should be brought back to its normal operation. The classification of Gibson is depicted at *Table 2*. The initial estimation of the complexity (UBFRP) value of a given function enables us to classify it rapidly according to these rules.

Levels	Importance/Criticality of BF	RTO	MAO
L1	BF should operate without interruption	<2 hours	=2 hours
L2	BF may be interrupted for a short period	<24 hours	=24 hours
L3	BF may be interrupted for 1 or more days	<72 hours (3 days)	=72 hours (3 days)
L4	BF may be interrupted for extended period	<168 hours (1 week)	=168 hours (1 week)

Table 2 Classification of IT business functions based on RTO and MAO timeframes

3.3 Algorithmic expression of the method via binary classification decision trees

The author utilizes a decision tree in order to depict the algorithmic decision making process towards the classification of an IT business function based on its criticality. Decision tree learning is one of the most widely used techniques. Its accuracy is competitive with other learning methods and it is very efficient. The learned classification model utilized in our case is the Classification and Regression Tree (CART) approach, which is introduced and described in detail by Breiman et al [3]. Grąbczewski [8], in his recent study, mentions that the method is designed to be applicable to both classification and regression problems. The algorithm is nonparametric and creates binary trees from data described by both continuous and discrete features. The same study [15] mentions that exhaustive search for the best splits estimates split qualities with the impurity reduction criterion with impurity defined as so called *Gini (diversity) index*. The decision tree built by the CART algorithm is always a binary decision tree (each node will have only two child nodes). The specific index is provided by the following formula:

$$Gini(T) = 1 - \sum_{j=1}^{n} p(j)^2 \tag{7}$$

where, p(j) is the relative frequency of class j in T, and T is the dataset which contains examples of n classes. For the induction of the decision tree, the node split is automatically derived with the help of the R-Studio software.

4 Results

4.1 Initial estimation of the UBFRP Values

The first part of the current work involves the estimation of specific values regarding the Unadjusted Business Function Recovery Points (UBFRP) as described in the 3.1 section of the paper. The calculations have been performed with the creation of formulas in Microsoft Excel 2013. The inferred results are depicted at *Table 3*. The table includes the most representative UBFRP values which indicate changes in the decision making process, regarding not only the classification of the process as critical or non-critical, but also the assignment of the precise Level of importance and approximate RTO/MAO values according to *Table 2*. The calculations include a Simple, an Average and a Complex representative recovery scenario regarding IT business functions in case of their operational failure. The Simple scenario involves 1 Simple Human Level Actor, 1 Average Human Level actor and 1 Complex Human Level Actor (1, 1, 1). In a similar way, the Average Scenario is a (1, 2, 2) Human Level Actor combination and the Complex Scenario involves a (1, 3, 3) Human Level Actor combination. The three scenarios have corresponding combinations for Application Level Actors and Business Function categories.

Recovery Scenario	UHW	UAPW	UBFW	UBFRP
Simple	(1,1,1)	(1,1,1)	(1,1,1)	9
Average	(1,2,2)	(1,2,2)	(2,2,2)	15
Complex	(1,3,3)	(1,3,3)	(3,3,3)	21

Table 3 Classification of IT business functions based on RTO and MAO timeframes

4.2 Decision Tree induction in R-Studio

The specific software tool has been selected by the author for decision tree induction, due to the fact that it supports tree induction with various algorithms, i.e. ID3, C4.5 and CART. For the implementation of the CART algorithm 2 main packages in R-Studio had to be utilized, that is (tree) and (rpart) package. The author of the current article has performed both approaches, and both inferred the same result. The purpose of the research was to use the specific software tool for implementing a binary final decision of whether a business function should be classified as critical for immediate recovery or not, according to the above inferred UBFRP values. The decision tree was formulated after importing a detailed dataset with all the UBFRP values in Ms Excel, from a corresponding .txt file. The code required for installing the tree package and generating the decision tree according to the CART algorithm is depicted below:

```
> install.packages("tree")
> library(tree)
> t3<-tree(IVL~UBFRPVAL, data=UBFRP_VALUES_2, method="recursive.partition")
> plot(t3)
> text(t3)
```

```
> print(t3)
node), split, n, deviance, yval, (yprob)
      * denotes terminal node

1) root 25 67.70 L4 ( 0.2000 0.2400 0.2000 0.3600 )
  2) UBFRPVAL < 14.5 14 18.25 L4(NO) ( 0.0000 0.0000 0.3571 0.6429 )
    4) UBFRPVAL < 9.5 9  0.00 L4(NO) ( 0.0000 0.0000 0.0000 1.0000 ) *
    5) UBFRPVAL > 9.5 5  0.00 L3(NO) ( 0.0000 0.0000 1.0000 0.0000 ) *
  3) UBFRPVAL > 14.5 11 15.16 L2(YES) ( 0.4545 0.5455 0.0000 0.0000 )
    6) UBFRPVAL < 20.5 6  0.00 L2(YES) ( 0.0000 1.0000 0.0000 0.0000 ) *
    7) UBFRPVAL > 20.5 5  0.00 L1(YES) ( 1.0000 0.0000 0.0000 0.0000 ) *
```

The corresponding data.frame, or dataset, as well as the inducted tree are depicted at Fig. 1 and Fig. 2 respectively. The final nodes of the decision tree provide information about the impact value level of the function, as it is proposed by Gibson, as well as the Boolean decision about whether the given business function is critical or not (YES/NO). Root node is the UBFRP Value, according to which (If UBFRPVAL<14,5 then BF_Non-Critical = YES, with Impact Value Levels L3, L4) the criticality of the given function is determined.

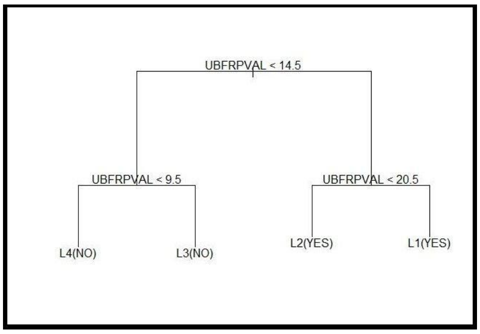
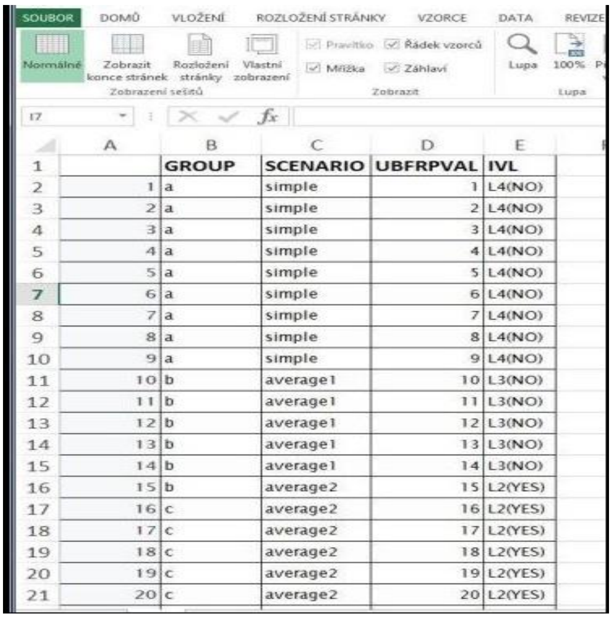


Figure 1 Decision tree inducted in R-Studio



	GROUP	SCENARIO	UBFRPVAL	IVL
1	a	simple	1	L4(NO)
2	a	simple	2	L4(NO)
3	a	simple	3	L4(NO)
4	a	simple	4	L4(NO)
5	a	simple	5	L4(NO)
6	a	simple	6	L4(NO)
7	a	simple	7	L4(NO)
8	a	simple	8	L4(NO)
9	a	simple	9	L4(NO)
10	b	average1	10	L3(NO)
11	b	average1	11	L3(NO)
12	b	average1	12	L3(NO)
13	b	average1	13	L3(NO)
14	b	average1	14	L3(NO)
15	b	average2	15	L2(YES)
16	c	average2	16	L2(YES)
17	c	average2	17	L2(YES)
18	c	average2	18	L2(YES)
19	c	average2	19	L2(YES)
20	c	average2	20	L2(YES)

Figure 2 Dataset in Excel file for generating decision tree in R-Studio

5 Discussion

The major challenge of the current paper is the determination of the recovery complexity of a business function based on software complexity estimation. Can the two complexities be compared and could we derive reasonable complexity measurements for an IT business function this way? From the author's standpoint the answer to this question is positive since IT business functions rely on information systems and applications whose complexity is estimated by the Use Case Points method. Another issue of major importance is the simplicity of an approach so as to achieve early and proactive, in terms of the formulation of the BIA document, classification of critical IT business functions. Miller and Engemann [11] state that, to be effective, methods to prioritize objectives should have some analytical rigor while at the same time, not being so complex that decision makers cannot use them.

The current method is standard, follows algorithmic principles and provides reasonable results of impact value levels through which efficient and timely binary classification of the enterprise business functions as critical/non-critical can be determined, at least in a primary level. These elements of the presented method are highly advantageous for IT managers in modern enterprises.

6 Conclusion – Future work

The current paper introduced a part of an ongoing research in the field of IT Business Continuity Management, which focuses on the classification of IT business functions based on their criticality. The criticality is determined based on the complexity of the business function and its corresponding recovery timeframe. The primary principle is “the more complex the business function, the less time should be afforded for its recovery”. The method is formulated according to the Use Case Points principles for software complexity estimation. Moreover, the classification of a specific IT business function is implemented according to the CART algorithm for decision tree classification. The contribution, which is addressed to the business impact analysis (BIA) formulation, is promising for enterprise business continuity management due to its simplicity and its preciseness, taking into consideration that rapid decisions for business impact analysis can be taken. Future work will include the software development for the specific functionality, because automation for deriving the criticality classification results is highly demanded by enterprises. The software tool will include not only a primary classification of a given IT business function, but also the precise time required for its recovery. The software tool will include both a desktop application as well as a web platform.

Acknowledgements

The current paper is supported by the SGS Project entitled “Advanced Information and Communication Services as a Key Tool during Emergency Situation Management”, No 21142, Technical University of Liberec.

References

[1] Antlová, K., Popelinský, L., & Tandler, J.: Long term growth of SME from the view of ICT competencies and web presentations, *E&M Ekonomie a Management*, 4 (2011), 125-139.
[2] Ashurst, C., Freer, A., Ekdahl, J., and Gibbons, C.: Exploring IT enabled innovation: a new paradigm? *International Journal of Information Management*, **32** (2012), 326–336.
[3] Breiman, L., Friedman, J.H., Olshen, R. and Stone, C.L.: *Classification and regression trees*. Chapman and Hall, 1984.
[4] European Banking Authority (EBA). *Technical advice on the delegated acts on critical functions and core business lines*. EBA/Op/2015/05, 2015
[5] Gallangher, M. *Business continuity management - How to protect your company from danger.* Prentice Hall, 2003.
[6] Gibson, D.: *Managing risks in information systems*. Jones & Bartlett Learning, 2010
[7] Gorzeń – Mitka, I., and Okręglicka, M.: Review of complexity drivers in enterprise. In: *Proceedings of the 12th International Conference Liberec Economic Forum’15* (Kocourek, A., ed.). Liberec, 2015, 253-260.
[8] Grąbczewski, K.: *Meta-Learning in Decision Tree Induction*. Springer International Publishing, Switzerland, 2014
[9] International Organization of Securities Commissions (IOSCO). *Market intermediary business continuity and recovery planning*. Consultation Report - CR04/2015, 2015
[10] Karner, G. (1993). Resource estimation for objectory projects. In: Systems SF AB.
[11] Miller, H.E., and Engemann, K.J.: Using analytical methods in business continuity planning. In: *Proceedings of MS 2012* (Engemann, K.J., Gil Lafuente, A.M. and Merigó, J.M., eds.). LNBIP, Springer, Heidelberg, 2012, 2-12.
[12] Miller, H.E., and Engemann, K.J.: Using reliability and simulation models in business continuity planning. *International Journal of Business Continuity and Risk Management* **5** (2014), 43-56
[13] Podaras, A., Antlová, K., and Motejlek, J.: Information management tools for implementing an effective enterprise business continuity strategy, *E&M Ekonomie a Management*, **1** (2016), 165-182.
[14] Rowan University. Business Continuity Management Policy, 2014
[15] Tucker, E.: *Business continuity from preparedness to recovery – a standards-based approach*. Elsevier Inc., 2015.

Regression Analysis of Social Networks and Platforms for Tablets in European Union

Alena Pozdílková[1], Pavla Koťátková Stránská[2], Martin Navrátil[3]

Abstract. In the last few years two operation systems for tablets: Android and iOS dominate in all states of European Union. Android is leading almost through whole Europe, it has a dominant position in all EU countries except the UK with a relatively small variations. Four social networks that are used on tablets, having the largest number of users in Europe include Facebook, Twitter, Stumble Upon and Pinterest. The number of portable electronic devices, whose major portion consists tables, is now experiencing worldwide growth. The aim of the paper is to create a regression model for individual model platforms and social networks for tablets in Europe, with a focus on the European Union and their analysis. The result of this paper will be also determination the degree of sensitivity of individual parameters in regression models.

Keywords: Social network, Regression, Tablet, European Union.

JEL Classification: C58, G21, C610
AMS Classification: 90C15

1 Introduction

The number of mobile devices more than doubled over the last two years in the Czech population. It results from a comparative analysis of data research of the Media project [11]. The fastest expansion is of distribution of tablets in last two years, whose numbers have tripled. 1.9 million people aged 12-79 years, representing 22 percent of the population currently own tablet according to their research [11].

In the European Union dominate two operating systems for tablets in last few years: Android and iOS. Other operating systems for tablets, namely Linux and Win8.1 RT, occupy only a small part of the market [1]. Android, which has a dominant position throughout Europe, also has a dominant position in all EU countries except the UK. According to standard deviations can be seen relatively small variations across the EU. Unlike iOS, this with a smaller variation achieves major deviations [1].

Facebook, Twitter, Stumble Upon and Pinterest are four social networks with the largest number of users using tablets in the European Union [11].

Social networks have been studied by many authors [6], [7].

The most important social network with many active users all around the world Facebook was founded in 2004. Facebook’s mission is to give people the power to share and make the world more open and connected. People use Facebook to stay connected with friends and family, to discover what’s going on in the world, and to share and express what matters to them. Facebook popularity and distribution of users by age shows Figure 1.

Statistics:

- 890 million daily active users on average for December 2015;
- 745 million mobile daily active users on average for December 2015;
- 1.39 billion monthly active users as of December 31, 2015;
- 1.19 billion mobile monthly active users as of December 31, 2015;
- approximately 82.4% of our daily active users are outside the US and Canada [2].

[1] University of Pardubice, Faculty of Electrotechnics and Informatics, Department of Mathematics and Physics, Studentská 95, 532 10 Pardubice 2, Czech Republic, e-mail: alena.pozdilkova@upce.cz.
[2] University of Pardubice, Faculty of Electrotechnics and Informatics, Department of Mathematics and Physics, Studentská 95, 532 10 Pardubice 2, Czech Republic, e-mail: pavla.kotatkovastranska@upce.cz.
[3] University of Pardubice, Faculty of Electrotechnics and Informatics, Department of Mathematics and Physics, Studentská 95, 532 10 Pardubice 2, Czech Republic, e-mail: martin.navratil@student.upce.cz.

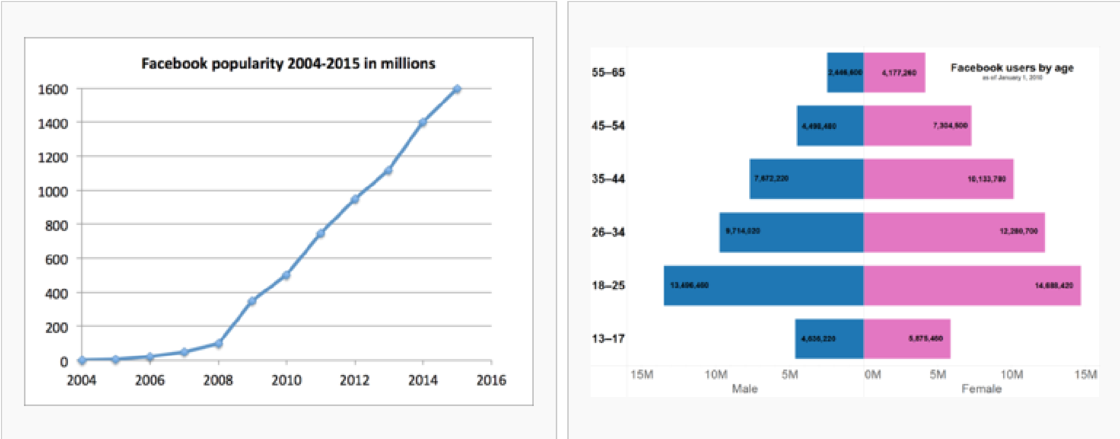


Figure 1 Facebook popularity and distribution of users by age [2]

Analysis in following chapters is focused on the European Union, states in Eurozone and states in European Union with their own currency. 19 countries of the European Union in Eurozone are: Belgium, Estonia, Finland, France, Ireland, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Portugal, Austria, Greece, Slovakia, Slovenia and Spain.

2 Data and methodology

Stats are based on aggregate data collected by StatCounter [4] on a sample exceeding 15 billion pageviews per month collected from across the StatCounter network of more than 3 million websites. Stats are updated and made available every 4 hours, however are subject to quality assurance testing and revision for 14 days from publication [4].

StatCounter is a web analytics service, tracking code is installed on more than 3 million sites globally. These sites cover various activities and geographic locations. It provides independent, unbiased stats on internet usage trends, which are not collated with any other information sources. No artificial weightings are used [4].

3 Analysis for Eurozone (19 states)

The correlation analysis will be performed in this part at first. Correlation matrix will be composed. This matrix will be established between different social networks and mobile platforms [3], [5]. Subsequently regression models will be developed where the dependent variable is the relevant social network, respectively mobile platform architecture and the independent variable is time (in years 2012-2015 for all countries of the Eurozone).

3.1 Correlation Analysis for Eurozone

The correlation analysis through all the years and states - marked correlations at a significance level of 0.05 are statistically significant. See Table 1 and Table 2.

variable	Facebook	Twitter	StumbleUpon	Pinterest	Other I
Facebook	1.000000	-0.813517	-0.756457	-0.858389	-0.668029
Twitter	-0.813517	1.000000	0.512300	0.505493	0.293331
Stumble Upon	-0.756457	0.512300	1.000000	0.662972	0.564610
Pinterest	-0.858389	0.505493	0.662972	1.000000	0.734851
Other I.	-0.668029	0.293331	0.564610	0.734851	1.000000

Table 1 Results of correlation analysis – correlation matrix for social networks

variable	iOs	Android	Other II.
iOs	1.000000	-0.994559	-0.029220
Android	-0.994559	1.000000	-0.045848
Other II.	-0.029220	-0.045848	1.000000

Table 2 Results of correlation analysis – correlation matrix for platforms

The significant correlation relations among variables can be seen from the correlation coefficients in Table 1 and Table 2. The significance was detected by Statistica software version 1.12. The insignificant relationship is between variables Other I. and Twitter or Other II and Android, iOS.

3.2 Regression Analysis for Eurozone

The regression models for variables Facebook, Twitter, StumbleUpon, Pinterest, Other I, iOS, Android and Other II will be analyzed in this part [8], [10], [12].

The regression models are properly assembled, the quality of the model is verified based on the p-values for the parameter beta b (the parameter is statistically significant at a level of 0.05, p-value is less than the significance level). The quality of the entire regression model is satisfied (p-value 0.0000), determination index is less than the statistic Durbin-Watson - it is not an apparent regression. This is applied to all models. Facebook and Android are expected a growing trend according to the regression model. StumbleUpon, Pinterest and iOS are expected conversely decreasing trend [9].

The regression coefficient indicates how much the variable changes when you increase the time unit by one. A positive sign means that the value will increase with the growth of time units, negative value of the variable will decrease with an increase in the time unit. See Table 3 and Table 4.

variable	Regression parameter	p-value	Confident interval -95%	Confident interval +95%
Facebook	0.573	0.00000	4.305	8.578
Twitter	-0.460	0.00000	-4.380	-1.656
Stumble Upon	-0.520	0.00000	-1.401	-0.636
Pinterest	-0.470	0.00000	-2.577	-1.006
Other I.	-0.330	0.00340	-1.02	-0.209

Table 3 Summary results of regression analysis for social networks

variable	Regression parameter	p-value	Confident interval -95%	Confident interval +95%
iOs	-0.610	0.00000	-7.917	-4.287
Android	0.575	0.00000	3.824	7.591
Other II.	0.450	0.00000	0.213	0.575

Table 4 Summary results of regression analysis for platforms

The tables above present the regression parameters "b" as well as the confidence intervals, which complement the conclusions regarding the statistical significance of the regression parameters. The regression parameters for all variables were identified as statistically significant (rejecting the null hypothesis H0: the regression parameter is insignificant), but the conclusion should be complemented by the confidence interval, respectively determining the interval in which the relevant parameter ranges.

The growth respectively average growth can be expected in the range values of 4.305 to 8.578 in the social network Facebook. The variable Twitter is expected to decline between the values of -4.380 to -1.656 in the next period. StumbleUpon, Pinterest and Other I. variables for mobile iOS platform also can be expected to decline in the next period, while Android is expected to grow.

4 Analysis for European Union without Eurozone (9 states)

The correlation analysis and regression models in years 2012-2015 for all countries out the Eurozone (9 states) will be performed in this part.

4.1 Correlation Analysis for states out of Eurozone

Table 5 and Table 6 show marked correlations at a significance level of 0.05, which are statistically significant.

variable	Facebook	Twitter	StumbleUpon	Pinterest	Other I
Facebook	1.000000	-0.863876	-0.758516	-0.888166	-0.836347
Twitter	-0.863876	1.000000	0.626506	0.602293	0.605628

Stumble Upon	-0.758516	0.626506	1.000000	0.614636	0.459498
Pinterest	-0.888166	0.602293	0.614636	1.000000	0.839374
Other I.	-0.836347	0.605628	0.459498	0.839374	1.000000

Table 5 Results of correlation analysis – correlation matrix for social networks

variable	iOs	Android	Other II.
iOs	1.000000	-0.996267	-0.251372
Android	-0.996267	1.000000	0.193457
Other II.	-0.251372	0.193457	1.000000

Table 6 Results of correlation analysis – correlation matrix for platforms

The significant correlation relations among variables can be seen from the correlation coefficients in Table 5 and Table 6. The significance was detected by Statistica software version 1.12. The insignificant relationship is between variables Other I. with StumbleUpon and Other II with Android and iOs.

4.2 Regression Analysis for states out of Eurozone

The regression models for variables Facebook, Twitter, StumbleUpon, Pinterest, Other I, iOS, Android and Other II will be analyzed in this part [8], [10], [12].

The regression models are properly assembled, the quality of the model is verified based on the p-values for the parameter beta b (the parameter is statistically significant at a level of 0.05; p-value is less than the significance level). The quality of the entire regression model is satisfied (p-value 0.0000), determination index is less than the statistic Durbin-Watson - it is not an apparent regression. This is applied to variables Facebook, StumbleUpon, Pinterest, iOS and Android. Twitter, Other I and Other II aren´t for regression models statistically significant. Facebook and Android are expected a growing trend according to the regression model. StumbleUpon, Pinterest and iOS are expected conversely decreasing trend [9].

The regression coefficient indicates how much the variable changes when you increase the time unit by one. A positive sign means that the value will increase with the growth of time units, negative value of the variable will decrease with an increase in the time unit. See Table 7 and Table 8.

variable	Regression parameter	p-value	Confident interval -95%	Confident interval +95%
Facebook	0.438	0.00810	1.262	7.615
Twitter	-0.310	0.063	-3.152	0.092
Stumble Upon	-0.600	0.00010	-1.507	-0.551
Pinterest	-0.390	0.02010	-2.468	-0.221
Other I.	-0.260	0.118	0.120	-1.217

Table 7 Summary results of regression analysis for social networks

variable	Regression parameter	p-value	Confident interval -95%	Confident interval +95%
iOs	-0.450	0.00560	-13.110	-2.410
Android	0.437	0.00780	2.162	13.104
Other II.	0.096	0.58080	-0.333	0.586

Table 8 Summary results of regression analysis for platforms

Similar conclusions are also found in countries that are not part of the Eurozone, despite the higher p-values and thus less difference compared to the 0.05 level of significance than was the case in the Eurozone countries. The regression parameters were not identified as statistically significant for all the variables (rejecting the null hypothesis H0: the regression parameter is insignificant). Parameters Twitter, Other I and Other II were labeled as statistically insignificant..

The growth respectively average growth can be expected in the range values of 1.262 to 7.615 in the social network Facebook. The variable Twitter is expected to decline between the values of -3.152 to 0.092 in the next period. StumbleUpon, Pinterest and Other I. variables for mobile iOS platform also can be expected to decline in the next period, while Android is expected to grow. Average increases, respectively decreases are shown in the confidence interval.

5 Graph results

The results from previous chapters are illustrated by the figures below. Figures 3 and 4 show the boxplot comparison for different parts of Europe in years 2012 and 2015 for all variables. Growing trend for most of the compared countries can be seen in tables in previous chapters. Values in Figure 4 are for many countries much higher than values in Figure 3.

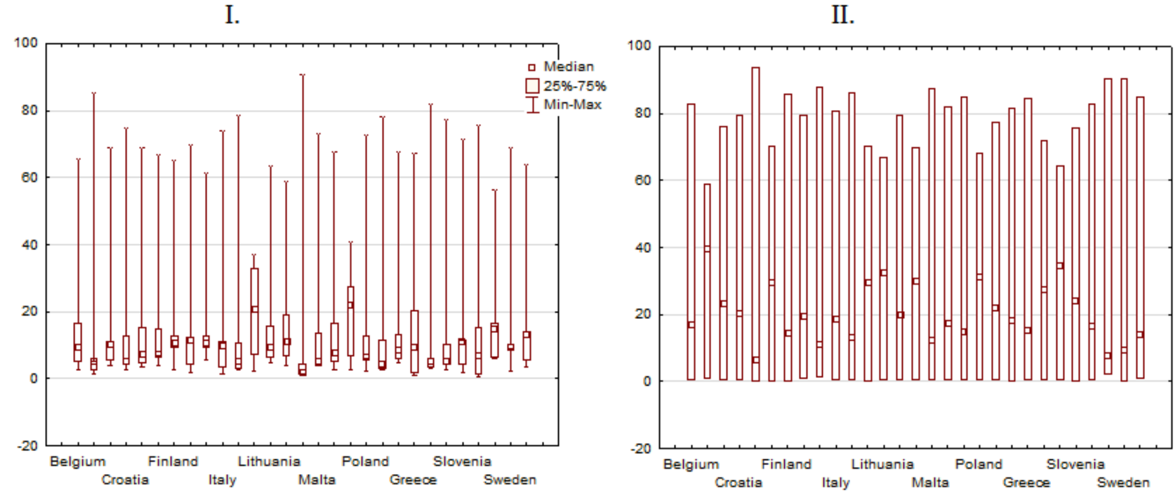


Figure 3 Boxplot comparisons for different parts of Europe in 2012

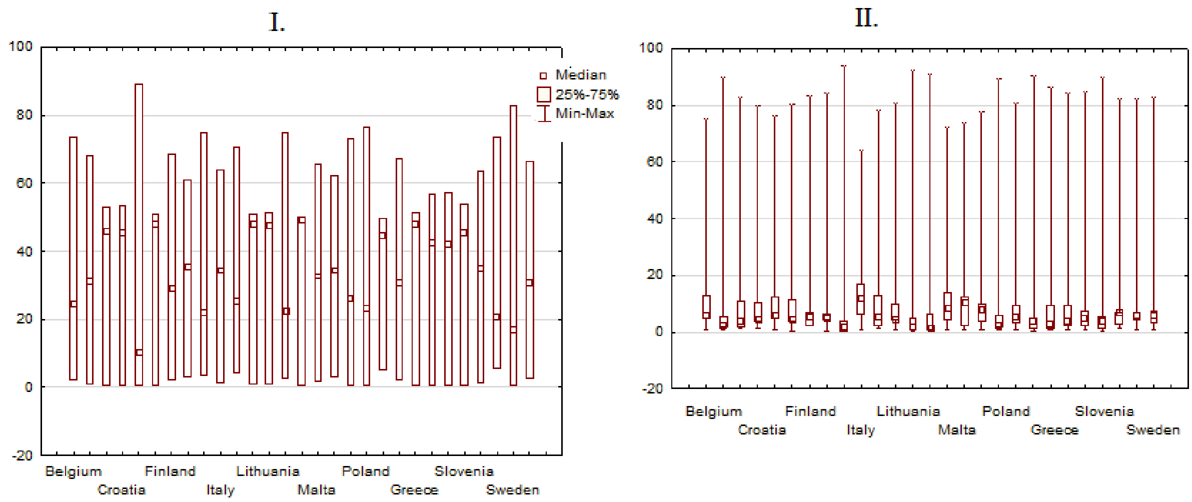


Figure 4 Boxplot comparisons for different parts of Europe in 2015

6 Discussions

The trend of social network Facebook is surprising in comparison with others social networks. Facebook is expected to grow, while Twitter, StumbleUpon and Pinterest are expected to decline. A similar phenomenon occurs in the case of mobile platforms. iOS is expected to decline in the upcoming period, while the growth of Android is anticipated. The argument proving the summary above could be a correlation matrix, respectively values of the individual correlation coefficients.

The compared confidence intervals have approximately the same values for both the Eurozone countries and countries outside the Eurozone except variable Facebook. Values for social network Facebook will grow faster in the Eurozone countries than in countries outside the Eurozone for the upcoming season.

7 Conclusions

The popularity of social networks is experiencing a worldwide growth (including people 60+). The number of users is increasing thanks to connectivity with mobile devices. It can be seen that Eurozone states and states with its own currency have the same trend in the upcoming period for both social networks and mobile platforms.

The price is the main reason, why cheap Android dominates the market of mobile devices against expensive iOS. Another reason is the brand of tables. Nowadays are popular products from Asia manufactures, which use in most cases Android operating system. Among other factors affecting the sale of various mobile devices are promotions, the purchasing power of the population, popularity, reliability, price, charging patent, easy to install software etc.

The analysis shows that for the social network Facebook is expected to grow for both compared areas – Eurozone and states outside Eurozone. The variable Twitter is expected to decline in next period. StumbleUpon, Pinterest and Other I. for the mobile iOS platform are also expected to decline in the coming period, while Android is expected to grow.

The results of regression models can be used as a recommendation for further development in the area. The results can be an incentive for additional marketing promotion and development within social networks. It can be also used for more complex regression models and further research.

Acknowledgements

This work has been mainly supported by the University of Pardubice via SGS Project (reg. no. SG660026) and institutional support.

References

[1] C news. URL: http://www.cnews.cz/clanky/android-stale-utika-konkurenci-ma-takovy-podil-ze-jej-mozna-zacne-vysetrovat-eu, online, [cit. 30. 4. 2016].
[2] Facebook informations and statistics. URL: https://www.facebook.com/about, online, [cit. 30.4.2016].
[3] Field, A.: *Discovering Statistics Using SPSS*. London, SAGE Publications, 2005.
[4] Global stats, Statcounter. URL: http://gs.statcounter.com/mobile-social-media-CZ-monthly-201303-201303-bar, online, [cit. 30. 4. 2016].
[5] Hebák, P. et al.: *Vícerozměrné statistické metody [3]*. Praha, Informatorium, 2007.
[6] Jackson, M., and Wolinsky, A.: A Strategic Model of Economic and Social Networks. *Journal of Economic Theory* **7** (1996), 44-74.
[7] Kvasnička, M.: Markets, Social Networks and Endogenous Preferences. *30th International Conference Mathematical Methods in Economics*, Palacký University Olomouc, 2014.
[8] Lavine, K. B.: Clustering and Classification of Analytical Data. *Encyclopedia of Analytical Chemistry*, Chichester, John Wiley & Sons Ltd, 2000, 9689- 9710.
[9] Meloun, M., Militký, J., and Hill, M.: *Počítačová analýza vícerozměrných dat v příkladech*. Praha, Academia, 2005.
[10] Romesburg, Ch.: *Cluster Analysis for Researcher*s. North Carolina, Lulu Press, 2014.
[11] Unie vydavatelů, Median, Stem/Mark. URL: http://www.mediaguru.cz/2015/07/tablety-a-chytre-telefony-se-v-populaci-rychle-siri/#.Vk3jUr8_aOB, online, [cit. 30.4.2016].
[12] Zou, H., Hastie, T., and Tibshirani, R.: Sparse principal component analysis. *Journal of computational and Graphical Statistics* **15**, 2 (2006), 265-280.

Dynamics Model of Firms' Income Tax

Pavel Pražák [1]

Abstract. The aim of this paper is to introduce a simple dynamic model that describes the tax revenue of a given government. The paper deals with a modification of Ramsey model that uses the neoclassical dynamical model and determines the optimal consumption as a result of optimal intertemporal choices of individual households. For simplicity the model considers only a tax on income of firms. This leads to a different selecting of optimal prices of labor and capital services. Euler equation of optimal consumption path of household is then introduced as well as its particular interpretation. We also study long-term effects of taxation of the given model. Hence the steady state of the dynamical model is introduced and then a formula for steady state Laffer curve is derived. Having this formula we further study the properties of optimization of tax revenue. First, it is shown that the optimal tax rate is dependent on the magnitude of elasticity of capital which is an exogenous parameter of the neoclassical production function. Then methods of comparative statics are used to find a sensitivity of tax revenue at steady state of the economic unit on some parameters.

Keywords: Laffer Curve, taxation, consumption, optimal control.

JEL classification: C61, H21, E62
AMS classification: 49J15, 39C15

1 Introduction

Governments use taxes to finance their public services. Although we understand this reality well we rather prefer low tax rates. In past decades some supply side economists suggested tax cuts and claimed that low tax rates can stimulate investment, spending and consumption that can support economic growth, [4], [3], [13]. Their argumentation was partially supported by popular Laffer curve, [8]. The fact that some taxes were reduced in the last two decades in European Union countries can be documented on Figure 1. Particularly in the Czech Republic the corporate income tax rate was gradually reduced from 40% in 1995 to 19% in 2010, [7]. Because lower corporate tax can attract more foreign investments it is preferred by government that plan to improve technology equipment or that want to improve employment in the given country. However, the cutting taxes without corresponding cutting government spending can result in fluent increase in budget deficit, [3], [9]. It seems that the recent debt crisis in some EU countries can stop tax cut trends and even it can reverse some of them. The debate on the level of tax rates is therefore still unfinished and current. It is known that taxes can change equilibrium prices and quantities of taxed goods, [9], which can influence tax revenue. We think that it is useful to construct models that allow us to study the influence of different endogenous or exogenous parameters of the model on tax revenue, [13]. The aim of this paper is to answer the partial question of how the behaviour of households and firms in an economic unit can adjust if fiscal policy of a government changes tax rate on firm's income.

2 Methods

A dynamic macroeconomic model that consists of household and production sectors will be considered. The production sector will be represented by a production function. The government sector will be represented only by an exogenous parameter given by tax rate on output of production sector. It is considered that households maximize their total consumption over the planning period and that they are constrained by their budgets at each period. It means that their purchases can be financed by

[1] University of Hradec Králové, Faculty of Informatics and management, Department of Informatics and Quantitative Methods, Rokitanského 62, Hradec Králové, Czech Republic, pavel.prazak@uhk.cz

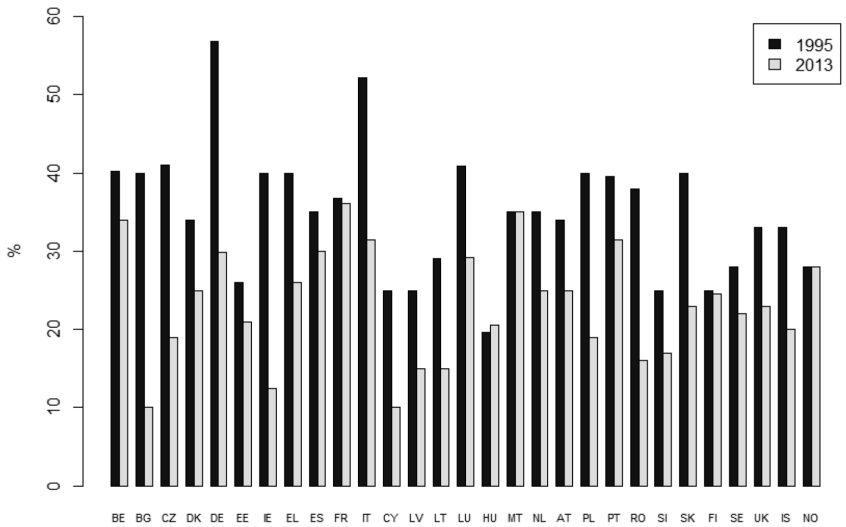


Figure 1 Reduction of tax rate on corporate income of EU countries in 1995 and 2013, in %., [7].

their incomes after tax. The problem of maximization the total utility of consumption will be given as a discrete time dynamic optimization problem, [12], [2]. Such a problem can be characterized by two principal features: a) the evolution of a given macroeconomic system is described by a discrete time dynamical system and b) a cost function that represents the demand of a planner is given. The dynamical system that describe the evolution of a state variable can be controlled by a control variable. In a model with discrete time a difference equation can be used, [11]. The aim of the optimization problems is to find maximum (or minimum) of a given cost function with respect to constraint that is represented by a difference equation. Since a long-last planning horizon will be considered, a dynamic optimization problem over an infinite horizon will be used. The advantage of this formulation is that it is not necessary consider what happens after a finite horizon is reached. Then necessary optimal conditions can be formulated, [6], [2], [10]. These conditions allow finding candidates for optimal solution.

3 Model

Two main sets of agents will be considered. Households own labor and assets and receive wages and rental payments. Firms produce homogeneous output and rent labor and capital. Instead of considering all particular households or all particular firms we introduce a representative household or a representative firm. The representative household is a model that is used for the decisions of many hidden small households. Similarly, the representative firm is a model that is used for the decisions of many hidden small firms. Government will be represented only indirectly by its tax rate on output of the representative firm. There are several markets. The rental market for labor L with price $w, w > 0$, and capital services K with price $R, R > 0$. Households can borrow and lend with the interest rate $r, 0 < r < 1$, at the assets market. Firms sell their products at the market for final output that is considered to be competitive. The commodity price in this market is normalized to unity. The objective of the representative firm is to maximize its after-tax profit at each time period. For further analysis we consider that the revenue is non-deductible and that the profit of the firm can be expressed in the following form

$$\Pi(K, L) = (1 - \tau)F(K, L) - RK - wL, \tag{1}$$

where $F(K, L)$ is a production function and $\tau, 0 < \tau < 1$, is a constant tax rate, which is an exogenous parameter given by government. This situation correspond with value-added tax. For alternative profit function and further discussion see section 5 of this paper. The first order conditions for the problem (1) are given by the following equations

$$(1 - \tau)\frac{\partial F}{\partial K} - R = 0, \ (1 - \tau)\frac{\partial F}{\partial L} - w = 0. \tag{2}$$

Notice that in this model the optimal price of capital services or wages are dependent on tax rate τ. If it is assumed that the production function is neoclassical with constant returns to scale the output can be written as

$$Y = F(K, L) = LF(K/L, 1) = Lf(k), \tag{3}$$

where $k = K/L$ is capital-labor ratio, $y = Y/L$ is per capita output and $f(k) = F(k, 1)$ is intensive neoclassical production function with corresponding properties, [1]. Using the given notations equations (2) can be rewritten as

$$(1-\tau)f'(k) = R,\ (1-\tau)(f(k) - kf'(k)) = w. \tag{4}$$

Assume further that $T(\tau)$ denotes the tax revenue of the government per one period when the tax rate is τ, then

$$T(\tau) = \tau \cdot Y = \tau \cdot Lf(k). \tag{5}$$

The household's preferences can be represented by an infinite series of additive intertemporal utility function with a constant rate of time preference

$$\sum_{t=0}^{\infty} \beta^t u(c_t), \tag{6}$$

where $u(.)$ is an intertemporal utility function with standard properties, [1], c_t is a consumption of household at period t, $\beta = 1/(1+\rho)$ is a rate of time preference and $\rho \geq 0$ is a discount rate. A low value of β indicates that the household is impatient and it prefers current consumption to future consumption. If the household cares equally about the present and the future consumption, then $\beta = 1$. The household chooses a consumption plan which maximizes its total life time utility subject to its budget constraint. This constraint can be expressed as a difference equation for household's assets a_t and has a form

$$a_{t+1} = w_t + (1 + r_t)a_t - c_t,\ t \geq 0, \tag{7}$$

with a given initial value a_0. Now the objective of the representative household can be shortly written as

$$\max_{c_t \geq 0} \left\{ \sum_{t=0}^{\infty} \beta^t u(c_t),\ a_{t+1} = w_t + (1 + r_t)a_t - c_t \right\}. \tag{8}$$

To solve this infinite time horizon problem (8) the discrete version of the maximum principle can be used, [10]. The household can freely choose the level of consumption ct in each period t in response to change of the tax rate τ. It means that consumption c_t is a control variable and its assets a_t is a state variable of the given problem. First we introduce the current value Hamiltonian

$$H(a_t, c_t, \lambda_{t+1}) = u(c_t) + \beta\lambda_{t+1}(w_t + (1 + r_t)a_t - c_t), \tag{9}$$

where λ_{t+1} is the adjoint variable. Then the first-order necessary conditions can be expressed as system of two difference equations

$$\frac{\partial H}{\partial c}(a_t, c_t, \lambda_{t+1}) = u'(c_t) - \lambda_{t+1} = 0,\ t \geq 0, \tag{10}$$

$$\frac{\partial H}{\partial a}(a_t, c_t, \lambda_{t+1}) = \beta\lambda_{t+1}(1 + r_t) = \lambda_t,\ t \geq 1. \tag{11}$$

If (10) is substituted into (11) we gain the following Euler equation

$$u'(c_t) = \beta(1 + r_t)u'(c_{t+1}), \tag{12}$$

that represents the necessary condition of the given optimal control problem. Equation (12) describes the optimal local trade-off between consumption in two succeeding periods and its intuitive interpretation is as follows. The household cannot gain anything from redistributing consumption between two succeeding periods. A one unit reduction of consumption in the first period lowers the utility in this period by $u'(c_t)$. The one consumption unit which is saved in the first period can be converted into $(1 + r_t)$ units of consumption in the second period and it can raise the second period utility by discounted value $\beta(1 + r_t)u'(c_{t+1})$. The Euler equation states that these two quantities are equal at the global optimum. We will return to this equation once again later.

4 Results and Interpretations

The equilibrium of the economy is a process that is described by the quartet (c_t, a_t, k_t, y_t) and corresponding prices (w_t, R_t, r_t) such that households and firms behave optimally and all markets clear. Optimization of the representative firm is given by (4). Optimization of the representative household is given by (7), (12). Because no arbitrage in asset market is considered, the return from capital stocks that depreciate at the constant rate $\delta, 0 < \delta < 1$, is equal to return from interest rate. It can be written as

$$R_t - \delta = r_t. \tag{13}$$

Consider a closed economy and clearing at the asset market. The only asset in positive net supply is capital, because all the borrowing and lending must cancel within the closed economy. It means that the equilibrium in this market is characterized by the following equality

$$k_t = a_t. \tag{14}$$

If we use this relation, relations (4), (13) and substitute them into (7), we get

$$k_{t+1} = (1-\tau)f(k_t) + (1-\delta)k_t - c_t. \tag{15}$$

If we use again (14), (4) and substitute them into Euler equation (12), we get

$$\frac{u'(c_t)}{\beta u'(c_{t+1})} = (1-\tau)f'(k_t) + (1-\delta). \tag{16}$$

The left side of this equation is the household's marginal rate of substitution (MRS) between consumption in the two succeeding periods. The right side of this equation is the household's marginal rate of transformation (MRT), which is how much extra output results the next period consumption from an additional unit of saving this period. The Euler equation states that in optimal consumption process MRS is equal to MRT. Assume now that a steady-state of the economy has been already reached and let us study long run effects of taxation. In the steady-state both the capital stock and the consumption remain constant in time. It means that $k_t = k^\circ$ and $c_t = c^\circ$ for all t, respectively. If we apply properties of steady state in equation (16) and use the fact that $\beta = 1/(1+\rho)$, we gain the following equation for stationary value of capital

$$f'(k^\circ) = \frac{\rho+\delta}{1-\tau}. \tag{17}$$

Knowing stationary value of capital relation (15) can be used and the stationary value of consumption can be found

$$c^\circ = (1-\tau)f(k^\circ) - \delta k^\circ. \tag{18}$$

To be able to solve equation (17) we introduce intensive Cobb-Douglas production function

$$y = f(k) = Ak^\alpha, \tag{19}$$

where $\alpha, 0 < \alpha < 1$, is an output elasticity of capital and $A, A > 0$, is the level of technology, [1]. If the derivative of (19) is substituted into (17) we get

$$k^\circ = (1-\tau)^{\frac{1}{1-\alpha}} \cdot \left(\frac{A\alpha}{\rho+\delta}\right)^{\frac{1}{1-\alpha}}. \tag{20}$$

If (20), (5) and the assumption that labor L has a constant magnitude the following closed form solution to the steady state of tax revenue can be found

$$T(\tau) = \tau L f(k^\circ) = \tau LA(k^\circ)^\alpha = \tau(1-\tau)^{\frac{\alpha}{1-\alpha}} \cdot LA^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{\rho+\delta}\right)^{\frac{\alpha}{1-\alpha}} \tag{21}$$

The graph of function (21) has reversed U shape and can be considered as a model of the Laffer curve for the steady state of the economic unit, see Figure 2. The settings for the left panel of this figure is $\delta = 0.3$, $\rho = 0.8$, $A = 1$, $L = 1$ and $\alpha = 0.6$. The settings for the right panel of this figure remains the same with one modification $\alpha = 0.4$. Using (21) it can be proved that the maximum tax revenue of government is reached for tax rate $\tau_{opt} = 1 - \alpha$. As a consequence of these observations we can summarize that the Laffer curve effect can be observed if $\tau > 1 - \alpha$.

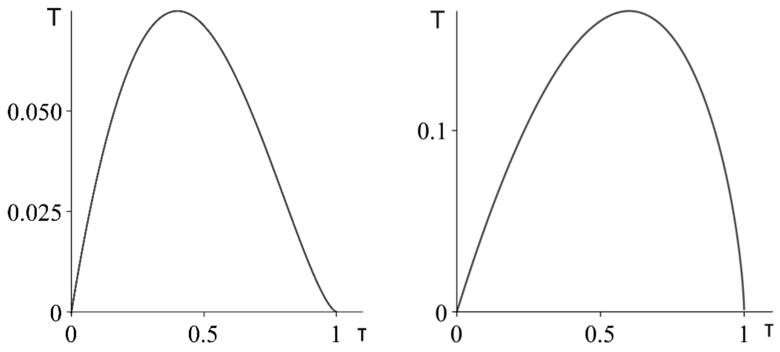


Figure 2 Laffer curve for different parameter specification.

5 Discussion

For the further analysis methods of comparative statics will be used. If the derivative of (21) with respect to A is considered and this result is used to find the total differential of $T(\tau)$, we successively get

$$\frac{\frac{\Delta T(\tau)}{T(\tau)}}{\frac{\Delta A}{A}} = \frac{1}{1-\alpha} = \eta_A, \tag{22}$$

which means that the point elasticity of steady state tax revenue for technology level A is $(1-\alpha)^{-1}$. Now we can estimate the effect of a change of the technology level: 1% increase in the level of technology A causes $(1-\alpha)^{-1}$% increase in steady state tax revenue. Similarly the point elasticity of steady state tax revenue for capital productivity α can be determined

$$\frac{\frac{\Delta T(\tau)}{T(\tau)}}{\frac{\Delta \alpha}{\alpha}} = \frac{\alpha}{(1-\alpha)^2} \cdot \left(1 - \alpha + \ln \frac{(1-\tau)\alpha A}{\rho + \delta}\right) = \eta_\alpha. \tag{23}$$

It means that 1% increase in the capital productivity α causes η_α% increase/decrease in steady state tax revenue. If the same setting as in Figure 2 (the left panel) is used, then for $\tau = 0,2$ we gain $\eta_\alpha = -1.6$. Hence 1% increase in the output elasticity of capital α causes 1.6% decrease in steady state tax revenue.

Another important note that has to be mention is the problem of profit function of the representative firm that is given by (1). Assuming that the revenue of the firm is deductible the after-tax profit of the firm can be written in a more appropriate way as

$$\Pi(K, L) = (1-\tau)(F(K, L) - RK - wL). \tag{24}$$

The study of the steady-state level of capital in the model with profit function (24) will be analyzed in another study.

6 Conclusion

A good tax system allows a government to withdraw enough money to finance its public services. In the past the question of maximization of tax revenue was the subject of series of static economic models. This question is contemporary again because there exists a fluent increasing of government's budget deficits in some European Union countries. In this paper, we discussed an optimal behaviour of households and firms under tax on firms' income. We consider that the problem of taxes setting is a contemporary because of the fluent increasing of government's budget deficits in some European Union countries. Results that are based on the analysis of the model can be summarized as follows: a) the optimal behaviour of household's consumption is given by Euler equation (12) or (16). b) the optimal prices of capital services and labor are given by equations (2); these prices are dependent on tax on income of firms, c) to study long run effects of the given tax rate a steady state of the economic unit was introduced; we find that the steady state tax revenue can exhibit Laffer effect, d) the optimal tax rate on firm's income at steady state given by Laffer curve (21) is indirectly proportional to elasticity of capital; the higher the magnitude of the elasticity of capital is the lower the optimal tax rate on firm's income, e) we also find that the higher level of the technology of firm is the higher tax revenue can gain. In the future we would like to discuss the following steps: a) to introduce a model with more complex fiscal tools focused on taxes, particularly we would like to consider not only tax on firms' earnings but also tax rates on wage income, private asset income and consumption, b) having introduced more complex model we would like to calibrate the parameters of the model for the case of the Czech Republic

Acknowledgements

Support of the Specific research project of the Faculty of Informatics and Management of University of Hradec Králové is kindly acknowledged.

References

[1] Barro, R. and Sala-i-Martin, X.: *Economic Growth.* MIT Press, Cambridge MA, 2004.

[2] Bertsekas, D. P.: *Dynamic Programming and Optimal Control.* Athena Scientific, Belmont MA, 2005.

[3] Blanchard, O.: *Macroeconomics.* Prentice Hall, New Jersey, 2000.

[4] Canto, V. A., Joines, D. H. and Laffer, A. B.: *Foundations of Supply-Side Economics: Theory and Evidence.* Academic Press, New York, 1983.

[5] Case, K. E. and Fair, R. C.: *Principles of Economics.* Prentice Hall, Upper Saddle River, NJ, 2007.

[6] Chow, G. C.: *Dynamic Economics, Optimization by Lagrange Method.* Oxford University Press, Oxford, 1997.

[7] Hemmelgarn, T.: *Taxation trends in the European Union.* Eurostat statistical books, European Union, 2013.

[8] Laffer, A. B.: Government exactions and Revenue deficiencies. *Cato Journal* **1**, 1 (1981), 1-21.

[9] Lowell, M. C.: *Economics with Calculus.* World Scientific, Singapore, 2004.

[10] Miao, J.: *Economic Dynamics in Discrete Time.* MIT Press, Cambridge MA, 2014.

[11] Prazak, P.: Discrete Dynamic Economic Models in MS Excel. In: *Proceedings of the 11th International Conference Efficiency and Responsibility in Education 2014* (Houska, M., Krejci, I. and Flegl, M. eds.), Czech University of Life Sciences, Prague, 2014, 595-601.

[12] Sydsaeter, K., Hammond, P., Seierstad, A. and Strøm, A.: *Further Mathematics for Economic Analysis.* Prentice Hall, Harlow, 2005.

[13] Trabandt, M. and Uhling, H.: The Laffer curve revisited. *Journal of Monetary Economics* **58** (2011), 305-327.

A Note on Economy of Scale Problem in Transport with Non-linear Cost Function of Transported Quantity

Vladimír Přibyl[1], Anna Černá[2], Jan Černý[3]

Abstract. The article deals with a problem, which is a modification of the known assignment problem: Given a set Q of mutually interchangeable objects to be transported. Given a road network, represented by a digraph G, where each arc is characterized by a transit time and by a nonlinear cost of joint transit of any number of elements from the set Q (the joint cost is supposed less than the sum of individual costs). Given the vertices, which represent the initial and the final positions of the elements from the set Q. The problem consists of the two following questions:
S1. How to assign the final positions to the initial ones?
S2. How to design a route connecting a pair resulting from S1, for each such pairs optimally?
Remark: It is supposed that if several routes resulting from S1 pass through a given arc, then the corresponding objects are transported together with the nonlinear costs depending on the number of the objects. Therefore, the set of routes is optimal if the sum of costs of joint transits of elements through all "exploited" arcs is minimal.
In the paper the optimization technique, which is a generalization of the one from [2], is presented and verified.

Keywords: optimization, assignment, routing, joint transport, nonlinear costs.

JEL Classification: C61, O18, R42
AMS Classification: 90B06, 90B10

1 Introduction

Mass production is more efficient than the individual one. This rule is valid for transport as well. E.g., an aircraft transports hundreds passengers together, a freight train carries many packages at once etc. Nevertheless, one can see persons or consignments transported individually, although they could be transported more efficiently together, at least in parts of their trips.

Extensive class of optimization problems focused on such situations is presented in articles [2] and [10], the latter of which also introduces a classification system in their structure. Furthermore, these articles contain the mathematical models and methods for solving a number of problems of belonging to this structure. However, the case of non-linear cost dependence on the number of jointly transported object has not been resolved yet and remains open. The main purpose of the present paper is to fill this gap for the case of mutually interchangeable objects.

The issue of the article is related to issues of non-linear costs network flows. A special case of convex nonlinearity is dealt in the paper [1] and the concave one in [7]. The difference is that the present paper does not work with flows but with individual objects to be transported and, moreover, no assumption concerning the type of nonlinearity is introduced. From systemic viewpoint, as well as because the nonlinearity may occur, it is also related to marshalling trains issue [4], [5]. However, it deals with at least thousands objects to be transported and therefore the methodologies are different from the current case dealing at most with tens. Another relationship, i.e. the one with truck routing (dispatching) [3], [6], [9], cannot serve as methodological source, since it does not deal with nonlinearity of costs. Hence an original approach is necessary.

1.1 Problem Formulation

Assume that a set of m mutually interchangeable (=equal-type) objects are located in the different *origin* vertices $o_1, \ldots, o_n$ of an (undirected) graph $G = (V, E, c_k)$ representing a network, where the value $c_k(e)$ is the transit cost of

[1] University of Economics/Faculty of Management, Jarošovská 1117, Jindřichův Hradec, Czech Republic pribyl@fm.vse.cz.
[2] University of Economics/Faculty of Management, Jarošovská 1117, Jindřichův Hradec, Czech Republic cerna@fm.vse.cz.
[3] University of Economics/Faculty of Management, Jarošovská 1117, Jindřichův Hradec, Czech Republic cerny@fm.vse.cz.

k objects through the edge e where $c_0(e) = 0$, $c_1(e) > 0$ and $c_{k+1}(e) \geq c_k(e)$ for each $e \in E$ and $k = 1, 2, ...$ Assume further that the objects are to be translocated to the different *destination* vertices $u_1, \ldots, u_n$ (and it does not matter which object to which vertex), determined by the one to one mapping d of the set where $d(o_i)$ is said the *destination* of the object located in o_i. The translocation from origins to destinations will take place along the routes determined by a mapping r of the set of all pairs $(o_i, d(o_i))$ into the set of all routes on G, connecting the pairs $(o, u) \in O \times U$ where $O = \{o_1, \ldots, o_n\}$, $U = \{u_1, \ldots, u_n\}$ and $r(o_i, d(o_i))$ is said the *transport route* of the object, located in o_i to its destination $d(o_i)$. Each such a pair (d, r) is said a *translocation plan* for the given origins $o_1, \ldots, o_n$ and destinations $u_1, \ldots, u_n$ in the graph G (briefly *plan*). **The problem** is to determine *optimal plan* (d, r), which meets the following condition:

$$c(d,r) = \sum_{e \in E} c_{k(e)}(e) \to \min \text{ where } k(e) = \text{card}\{q_i \in Q\text{: } e \in r(o_i, d(q_i))\} \tag{1}$$

In (1), the value $c(d, r)$ represents the *cost of the plan* (d, r).

1.2 Problem Discussion

Remark l. The requirement that all n translocated objects in the Problem 1.1 ought to have different origins and different destinations can be omitted in practice without any change of resolving method or algorithm: Assume e.g. that a vertex o is the origin of three objects. Then three different "dummy" vertices o_1, o_2, o_3 and edges (o_1, o), (o_2, o), (o_3, o) may be added with zero costs $c_k(e) = 0$ for all $e \in E$ and $k = 0, 1, \ldots$ Similar way can be adopted for destinations. I.e., this special formulation of the problem 1.1 does not limit its practical application.

It is suitable to define the digraph $D(G) = (V, A, c)$ *derived from* G in such a way that $[v, w] \in A$ and $[w, v] \in A$ if and only if $(v, w) \in E$ and, of course, $c_k[v, w] = c_k[w, v] = c_k(v, w)$ for $k = 0, 1, 2, \ldots$ It is better for expression of the fact that, e.g., the vertex v immediately follows the vertex w in a route r by the sentence: "r passes through $[v, w]$" or " $[v, w] \in r$ " which becomes ambiguous replacing $[v, w]$ by (v, w).

The following proposition is obvious, since for $k > 0$ it is $c_k > 0$ on E or A respectively:

Proposition 1. *Any solution (d, r) of a problem 1.1 contains only simple routes, i.e. the paths in the graph-theoretical sense.*

Consequently, all "candidate" routes for the solution of the problem 1.1 ought to be simple, i.e. are allowed to pass any vertex at most once.

It has no sense that within a plan (d, r) there exist a duplex passing of an edge, i.e. two routes r_1, r_2 and an edge $(v, w) \in E$ such that $[v, w] \in r_1$, and $[w, v] \in r_2$ concurrently. Assume $r_1 = r_{1h},v,w,r_{1t}$, $r_2 = r_{2h},w,v,r_{2t}$ and replace them by $r_1' = r_{1h},v,r_{2t}$, $r_2' = r_{2h},w,r_{1t}$ in the plan, denoted now (d', r'). All routes from (d, r) except two have remained in (d', r') as well and the two last have been shortened and therefore are cheaper which implies that $c(d', r') < c(d, r)$. That implies the following proposition:

Proposition 2. *Any solution (d, r) of a problem 1.1 is duplex edge passing free.*

Consequently, each "candidate" plan (d, r) solving the problem 1.1 ought to be duplex edge passing free.

The similar problems of network theory as 1.1 have been solved by many different methods. The most frequent is, probably, MILP, i.e. the mixed integer linear programming. However, the nonlinearity in objective function complicates this approach. Use of a piecewise linearization method [8] would lead to extremely large problems. Similarly, use of DfS (depth-first-search) technique seems too complicated as well since there are too many possible plans. The authors consider some modification of "negative cost circuit" approach more hopeful.

2 Solution Method for Problem 1.1

The method is a modification of the one of the "classic" problem of minimum cost flow problem.

2.1 Theoretical Basis

Use of definitions and denotations from 1.1 continues. The *increment* of cost is defined as $g_k(e) = c_k(e) - c_{k-1}(e)$ for each $e \in E$ and $k = 1, 2, \ldots$ Obviously, $c_k(e) = g_1(e) + \ldots + g_k(e)$ for each $e \in E$ and $k = 1, 2, \ldots$ The first aim of this subchapter is to investigate mutual relations of two given plans (d, r) and (d', r') in the digraph D(G).

Lemma 1. *Let (d, r) and (d', r') be two given plans of Problem 1.1 in the digraph $D(G)$. Then the mapping $d^* = d^{-1}d'(\bullet)$: $O = \{o_1, \ldots, o_n\} \to O$ is one-to-one and it defines a basic decomposition Δ: $O = O_1 \vee O_2 \vee \ldots \vee O_k$ .s.t. $d^*(O_i) = O_i$ for all $i = 1, \ldots, k$ and Δ cannot be refined more.*

Proof. Since the mapping d is one-to one, the inverse mapping d^{-1} is unambiguously defined and also one-to-on. Therefore, d^* is one-to-one as well. Thus one can define a retagging of the vertices $o_1, \ldots, o_n$ and $u_1, \ldots, u_n$ as follows: o_1 remains as before, $u_1 = d'(o_1)$, if $d^*(o_1) = d^{-1}(u_1) = o_1$ then it is trivial and $O_1 = \{o_1\}$, $\Delta = \{O_1, \ldots\}$. Hence, for this moment, it is supposed that $o_2 = d^{-1}(u_1) \neq o_1$ and $v_2 = d'(o_2)$, etc. Suppose that m is the lowest number for which $d^{-1}(u_m) = o_1$. Of course, then for all $j < m$ it holds $d^{-1}(u_j) = o_{j+1}$ and, consequently, the set $O_1 = \{o_1, \ldots, o_m\}$ may be called *indecomposable* with respect to the mappings d' and d. If $m = n$, then $O_1 = O = \{o_1, \ldots, o_n\}$ and O is called *indecomposable* as well. Otherwise, when $m < n$ the retagging procedure can be repeated starting with o_{m+1} etc. That concludes the proof.

Lemma 2. *If* (d, r) *and* (d', r') *are two solutions of the Problem 1.1, then*

either the set $O = \{o_1, \ldots, o_n\}$ *is indecomposable with respect to the mappings* d' *and* d,

or (d, r) *and* (d', r') *together with the Problem 1.1 itself can be considered a composition of k independent sub-plans* (d_i, r_i) *and* (d_i', r_i'), $i = 1, \ldots, k$ *of k sub-problems defined subsequently by origin sets* O_i *and destination sets* U_i, $i = 1, \ldots, k$ *in the same network G resp.* $D(G)$ where

$d(o) = d_i(o)$ *and* $d'(o) = d_i'(o) \Leftrightarrow o \in O_i$ *for* $i = 1, \ldots, k$ and

$r(o, d(o)) = r_i(o, d_i(o))$ and $r'(o, d'(o)) = r_i'(o, d_i'(o))$.

Proof. The assertion of Lemma 2 is an easy consequence of Lemma 1.

Corollary. *If* $c(d', r') < c(d, r)$ *when the conditions of Lemma 2 are met then there exist such* $i \in \{1, \ldots, \mathrm{k}\}$ *that* $c(d_i', r_i') < c(d_i, r_i)$.

The solution method of Problem 1.1 works with the following digraph $D(r, r') = (V(r, r'), A(r, r'), \gamma)$ which is called *derived from the digraph* $D(G)$ *and from the plans* (d, r), (d', r') s. t. the set $O = \{o_1, \ldots, o_n\}$ is indecomposable with respect to the mappings d' and d, using the denotations $k([v, w]) = card\{o \in O\colon [v, w] \in r(o, d(o))\}$ and $k'([v, w]) = card\{o \in O\colon [v, w] \in r'(o, d'(o))\}$:

$V(r, r') = \{v \in V$: there exists such $o \in O$ that $v \in r(o, d(o_i)$ or $v \in r'(o, d'(o))\}$

$[v, w] \in A(r, r') \Leftrightarrow$ either $[w, v] \in r(o, d(o))$ for some $o \in O$; then $\gamma_j([v, w]) = c_{k([w, v])}([w, v]) - c_{k([w, v])-j}([w,v])$

or $[v, w] \in r'(o, d'(o))$ *for* some $o \in O$; then $\gamma_j([v, w]) = c_j([v, w])$

As one can see, the arcs defined in the first rows have opposite orientation than in the routes r. Therefore, they are called *reversed* (oriented from U to O i.e. the arrow head is closer to O then the other end). The others are called *straight* (oriented from O to U). A circuit s (not necessary simple) in the digraph $D(r, r')$ is called *admissible*, if it transits through any reverse arc $[v, w]$ not more than $k([w, v])$ times and through any straight arc $[v, w]$ not more than $k'([v, w])$ times. The first question is, whether the digraph $D(r, r')$ is strongly connected.

Lemma 3. *If the set of origins O is indecomposable, then* $D(r, r')$ *is strongly connected*.

Proof. Let $v_o \in V(r, r')$, $v_d \in V(r, r')$ and $v_o \neq v_d$. It is necessary to prove that there exists a route from v_o to v_d in $D(r, r')$. If $v_o \in r(o_i, d(o_i))$ for $i < m$ then there exists a route $r(v_o, o_i)$ from v_o to o_i since the directions of arcs from $r(o_i, d(o_i))$ in $D(G)$ are reversed in $D(r, r')$. From the proof of the Lemma 1 it follows that there exist a route from o_i to u_i, then to $o_{i+1(\mathrm{mod}\,m)}$, to $u_{i+1(\mathrm{mod}\,m)}$, etc. which implies that there exist a route from v_o to each origin o_i and each destination u_i. The same remains obviously valid if $v_o \in r'(o_i, d'(o_i))$, since $d'(o_i) = u_i$. Similarly, if $v_d \in r(o_i, d(o_i))$ or $v_d \in r'(o_i, d'(o_i)$ then there exist a route from u_i or from o_i to v_d and the proof is complete.

Lemma 4. Let $A^+(v) = \{[v, w] \in A(r, r')\}$ and $A^-(v) = \{[w, v] \in A(r, r')\}$ for the given $v \in V(r, r')$. Let

$$\deg^+(v) = \sum_{[v,w]\in A(r,r')} \left(k([v,w]) + k'([v,w])\right) \text{ for all } v \in V(r, r') \tag{2}$$

$$\deg^-(v) = \sum_{[w,v]\in A(r,r')} \left(k([w,v]) + k'([w,v])\right) \text{ for all } v \in V(r, r') \tag{3}$$

Then $\deg^+(\mathrm{v}) = \deg^-(\mathrm{v})$ for each $v \in V(r, r')$.

Proof. The assertion is obviously valid for all vertices $o_1, \ldots, o_n, u_1, \ldots, u_n$ since both mappings d^{-1} and d' are one-to-one. The validity for other vertices follows from the construction of $D(r, r')$ and from the fact that any route of any object enters an intermediate vertex and then exits from it. The proof is complete.

The method of solution of Problem 1.1 is based on the following theorem.

Theorem 1. *Let* (d, r) *and* (d', r') *be plans of the problem 1.1 with indecomposable set O, let* $c(d', r') < c(d, r)$. *Let* $MD(r, r') = (V(r, r'), A^*(r, r'))$ *be the multi-digraph derived from the digraph* $D(r, r')$ *in such a way that each straight arc in* $D(r, r')$ *is multiplied* $k'(a)$ *times and each reverse one* $k(a)$ times. *Then*

A1. *An Euler circuit p exists in the multi-digraph* $MD(r, r')$ *and the corresponding circuit* p^* *passing the same vertices in the same order in* $D(r, r')$ *has total cost* $\gamma(p^*) = c(d', r') - c(d, r) < 0$.

A2. *Let the plan* (d, r) *is changed following the circuit* p^* *in such a way that*

(i) *if* $\gamma([w, v]) < 0$ *for some*$[w, v]$ *which is contained j times in the circuit* p^**,then j* transits of primal routes from the plan (d, r) *through the arc* $[v, w]$ *are omitted and this is done for all such arcs* $[v, w] \in A$; *afterwards*

(ii) *if* $\gamma([v, w]) > 0$ *for some*$[v, w]$ *which is contained j times in the circuit* p^**,then j* transits of primal routes from the plan (d', r') *through the arc* $[v, w]$ *are added and this is done for all such arcs* $[v, w] \in A$-

then it unambiguously determines the new plan equal to (d', r').

Proof of Theorem 1. The assertion A1 holds as a consequence of the construction of $MD(r, r')$ and Lemmas 3 and 4. A2 follows from the fact that the circuit p^* was derived from Euler circuit p and from the construction of the multigraph $MD(r, r')$ assuring that all routes from (d, r) are cancelled and all transits of all routes from (d', r') are introduced.

A *circuit* p^* from the proof of Theorem 1 is said *converting the solution* (d, r) *into* (d', r').

Corollary. If the plan (d, r) is not optimal, and thus there exist another plan (d', r') with $c(d', r') < c(d, r)$, then a) either the set O is indecomposable with respect to the mappings d' and d and then, due to Theorem 1, there exist a circuit p^* in $D(r, r')$ transforming (d, r) into (d', r'), i.e. into a better plan;

b) or the set O is decomposable, then due to Lemma 2, its corollary and Theorem 1 there exist a circuit p^* transforming (d, r) into a better plan (d'', r''), but, maybe, worse than (d', r').

Remark 2. Obviously, there exist always a transforming circuit p^* which does not contain a pair of arcs $[v, w]$ together with $[w, v]$, since its elimination has no influence on transforming operation.

Remark 3. A quite natural question is whether there exist a simple circuit (= passing any arc at most once) p^* transforming (d, r) into a better plan (d'', r'') in the case when (d, r) is not optimal. Unfortunately, it is not true, as one can see in Figure 1, using the costs determined by Table 1. All origin ($o_{1\ldots3}$) and destination ($u_{1\ldots3}$) vertices are considered the dummy ones (see Remark 1) and are drawn by a dotted line (of course with their dummy edges)

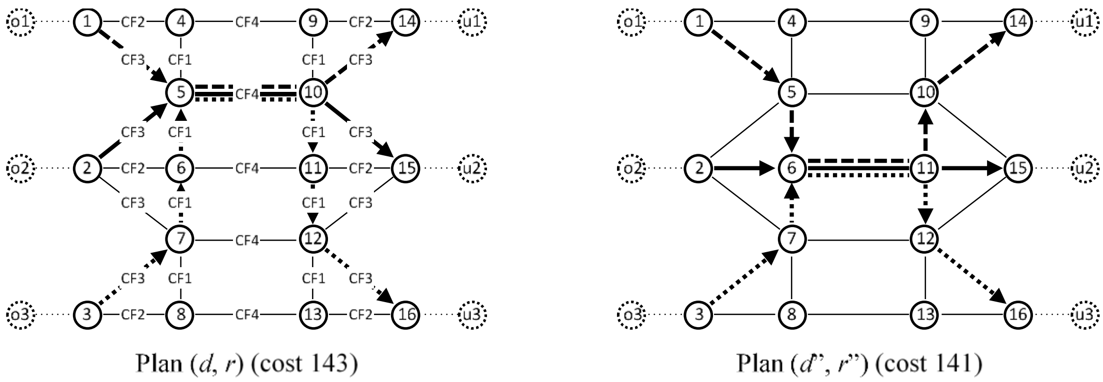


Figure 1 Demonstration of Remark 3

Let

$o_1 = o_1' = 1$, $d(1) = d'(1) = 14$; $o_2 = o_2' = 1$, $d(2) = d'(2) = 15$; $o_3 = o_3' = 3$, $d(3) = d'(3) = 16$;

$r(1, 14) = 1,5,10,14$; $r(2, 15) = 2,5,10,15$; $r(3, 16) = 3,7,6,5,10,11,12,16$; $c(d, r) = 6*13 + 4*5 + 45 = 143$

$r'(1, 14) = 1,5,6,11,10,14$; $r'(2, 15) = 2,6,11,15$; $r(3, 16) = 3,7,6,11,12,16$; $c(d', r') = 2*12 + 4*18 + 45 = 141$

Therefore, (d, r) is not optimal since (d', r') is better. However, a simple circuit p^* transforming (d, r) into a better plan (d'', r'') with $c(d'', r'') < 143$ does not exist. It can be proved indirectly: Let p^* be such a simple circuit. It cannot contain the arc [10, 5] with the cost –3 since then it ought to contain one of the arcs [4, 9], [6, 11], [7, 12], or [8, 13] with the cost 36 and the negative costs of circuit is not reachable. However, a circuit with negative costs obviously exists neither among the vertices 1-8 nor 9-16.

2.2 Description of the Method for Problem 1.1

The main idea of the method is based on the Corollary of Theorem 1: First to find any initial plan (d, r) for the sets O and U in the digraph $D(G)$ (e.g. by a greedy algorithm). Then cancel costs to reverse arcs from all $r(o, d(o))$ and give them costs similarly as in $D(r, r')$ (although any r' is unknown). The costs of all other arcs $[v, w]$ remain unchanged. Afterwards look for a *simple circuit p* with negative cost in $D(G)$ using any tagging method (e.g. Ford or something like this). After having found it cancel one transit through any arc with a negative cost in p and add a new transit through the ones with positive cost (of the first transit). Obviously, new plan with smaller costs is reached.

If no simple circuits exist more, then the construction of a non-simple circuit via tagging procedure will continue in a modified way: The multiplicity of passing of any arc in $D(G)$ by p^* is limited by $n = |O|$. After each passing of an arc a by the subsequently constructed p^* the cost $c(a)$ is changed similarly as defined in $D(r, r')$. The technical implementation of this idea reached by the authors until now is not suitable for its unacceptable computational complexity, but there are hopes for its acceleration.

3 Computational experience

For testing purposes, the authors have designed a simple graph with edges evaluated using 4 types of nonlinear cost functions. Cost functions (CF_1, CF_2, CF_3 and CF_4) are shown in Table 1. There are labels in Figure 1 showing the relevant cost functions of edges. Because of the cardinality $|O| = 4$ in all test cases, the functions are tabulated for a maximum of 4 objects.

Type of cost function	$k = 1$	$k = 2$	$k = 3$	$k = 4$
CF_1	5.0	5.8	6.3	6.5
CF_2	12.0	14.0	15.0	15.5
CF_3	13.0	15.0	16.0	16.5
CF_4	36.0	42.0	45.0	46.5

Table 1 Transit cost of k objects

The tagging procedure was tested on above mentioned graph (depicted in Figure 1) with the origin vertices $o_1 = 1$; $o_2 = o_3 = 2$; $o_4 = 3$ and destination vertices $u_1 = 14$; $u_2 = u_3 = 15$; $u_4 = 16$ and different initial plans.

Table 2 describes the circuits and their resulting plans with their costs in the case when the optimal plan is reachable by successive application of 6 simple circuits. The structure of a plan description is "…|k [arc]|…", where k is the number of objects transported via an arc [arc].

Step	Costs	Plan description	Cycle description
0	190	1 [1,4]|1 [4,9]|1 [9,14]|2 [2,6]|1 [3,8]|2 [6,11]| 1 [8,13]|2 [11,15]|1 [13,16]	-
1	188	1 [1,4]|1 [4,5]|1 [10,14]|2 [2,6]|1 [3,7]|1 [5,6]| 3 [6,11]|1 [7,12]|1 [12,13]|3 [11,15]|1 [13,16]|1 [15,10]	16,13,8,3,7,12,15,10,14,9,4,5,6,11,15,12,13,16
2	184	1 [1,5]|1 [10,14]|2 [2,6]|1 [3,7]|1 [5,6]|3 [6,11]|1 [7,12]| 1 [12,13]|3 [11,15]|1 [13,16]|1 [15,10]	16,13,8,3,7,12,15,10,14,9,4,1,5,4,9,14,10,15,12,7,3,8,13,16
3	175	1 [1,5]|1 [10,14]|2 [2,6]|1 [3,7]|1 [5,6]|3 [6,11]|1 [7,12]| 1 [12,13]|2 [11,15]|1 [13,16]|1 [11,10]	16,13,8,3,7,12,15,11,10,15,12,7,3,8,13,16
4	159,5	1 [1,5]|1 [10,14]|3 [2,6]|1 [3,7]|1 [5,6]|4 [6,11]|1 [11,12]|1 [12,13]|2 [11,15]|1 [13,16]|1 [7,2]|1 [11,10]	16,13,8,3,7,2,5,6,11,12,7,3,8,13,12,15,10,14,9, 4,5,2,6,5,1,4,9,14,10,15,11,10,5,4,1,5,10,11, 15,12,13,16
5	150,5	1 [1,5]|1 [10,14]|2 [2,6]|1 [3,7]|1 [5,6]|4 [6,11]| 1 [11,12]|1 [12,13]|2 [11,15]|1 [13,16]|1 [7,6]|1 [11,10]	16,13,8,3,7,6,2,7,3,8,13,16
6	146,5	1 [1,5]|1 [10,14]|2 [2,6]|1 [3,7]|1 [5,6]|4 [6,11]| 1 [11,12]|2 [11,15]|1 [12,16]|1 [7,6]|1 [11,10]	16,13,12,16

Table 2 Application of successive simple circuits leading to the optimal plan in 6 steps

Step 0 describes the initial plan obtained by the greedy algorithm, which for each destination vertex finds the cheapest route connecting it with an available origin vertex. The algorithm for searching simple circuits stops once the first circuit with negative costs is reached, transforms the plan and starts searching again. Experiments have shown that starting with another initial plan can lead to non-optimal plans, which cannot be improved by application of simple circuit. It corresponds with Remark 3.

Table 3 describes a non-simple circuit, which can transform the same initial plan as in previous case into the optimal plan in only one step. Unfortunately, this circuit was found intuitively. Although the implemented simple circuit tagging procedure is relatively fast (computational time in case described in Table 2 was about 1 second), extending it for non-simple circuits leads (until now) to the unacceptable computational times.

Step	Costs	Plan description	Cycle description
0	190	1 [1,4]|1 [4,9]|1 [9,14]|2 [2,6]|1 [3,8]|2 [6,11]| 1 [8,13]|2 [11,15]|1 [13,16]	-
1	146,5	1 [1,5]|1 [10,14]|2 [2,6]|1 [3,7]|1 [5,6]|4 [6,11]| 1 [11,12]|2 [11,15]|1 [12,16]|1 [7,6]|1 [11,10]	16,13,8,3,7,6,11,10,14,9,4,1,5,6,11,12,16

Table 3 Application of non-simple circuit leading to the optimal plan in one step

4 Conclusion and Outline of Future Research

The paper presents and solves the problem of optimal plan of, partially joint, transport of *n* mutually interchangeable objects through the given network. It is assumed that the objects are located in *n* origin vertices and ought to be transported to the given set of destination vertices when it does not matter which object will end in a specific destination. The main feature of the problem is that the joint transportation of *k* objects through an edge nonlinearly depends on *k* and it is much cheaper, that *k* times costs of individual transportation there.

The solution method is based on repeated looking for a circuit (simple if possible, but not necessarily) with negative cost in an auxiliary digraph derived from the network and an initial solution, which is easily found by the origin-destination matching resulting from the “classic” assignment problem.

The main direction of future research can be expected in the acceleration of the solution method, either based on quicker finding of a non-simple circuit with negative cost or looking for an absolutely different principle, e.g. mixed integer linear programming, but not based on piecewise linearization of cost.

Acknowledgements

The authors greatly appreciate the support of research, whose results are published in this paper, by the Faculty of Management, University of Economics in Prague within the resources for long-term institutional development, Project number IP600040.

References

[1] Bertsekas, D.P., Polymenakos L.C. and Tseng, P.: An ε-Relaxation Method for Separable Convex Cost Network Flow Problems. *SIAM Journal of Optimization* **7** (1997), 853-870.
[2] Černý, J. and Přibyl, V.: Partially Joint Transport on Networks. In *Mathematical Methods in Economics 2013. Jihlava*: College of Polytechnics Jihlava, 2013, 113-116.
[3] Cordeau, J.F., Desaulniers, G., Desrosiers, J. and Solomon, M.M.: The VRP with Time Windows. In: *The Vehicle Routing Problem*, Paolo Toth and Daniele Vigo (eds), SIAM Monographs on Discrete Mathematics and Applications, 2001.
[4] Flodr, F., Mojžíš, V. and Pokorný, J.: *Dopravní provoz železnic – vlakotvorba* (Railway Traffic - Marshalling Trains – in Czech), VŠDS Žilina, 1989.
[5] Gašparík, J. a kol.: *Vlakotvorba a miestne dopravné procesy* (Marshalling Trains and Local Transport Processes – in Slovak). 1st ed.. Univerzita Pardubice 2011.
[6] Golden, B. L., Raghavan, S. and Wasil, E. A. (Eds.). (2008). *The vehicle routing problem: latest advances and new challenges*. Vol. 43, Springer Science & Business Media.
[7] Guisewite, G.M. and Pardalos, P.M.: Minimum concave-cost network flow problems: Applications, complexity, and algorithms. *Annals of Operations Research* **25** (1990), 75-99.
[8] Moussourakis, J. and Haksever, C.: Project compression with nonlinear cost functions. *Journal of Construction Engineering and Management* **136** (2009), 251-259.
[9] Peško, Š, 2002. *Optimalizácia rozvozov s časovými oknam* (Time Constrained Routing Optimization – in Slovak). Konferencia KYBERNETIKA - história, perspektívy, teória a prax, zborník referátov, 76–81, Žilina.
[10] Přibyl, V., Černý, J. and Černá, A.: Economy of Scale in Transport Industry – New Decision Problems. Submitted to *Prague Economic Papers*, the preprint in electronic form can be requested from the authors.

Finite game with vector payoffs: how to determine weights and payoff bounds under linear scalarization

Miroslav Rada[1], Michaela Tichá[2]

Abstract. Assume a two-player finite noncooperative game with multiple known payoff functions (also called multi-criteria game) of each player. Furthermore, assume that utility of both the players can be obtained as a linear weighted scalarization of their payoffs. Given a strategy profile of the players, we are interested in the set of weights for which the strategy profile is the Nash equilibrium of the scalarized game. We use the standard quadratic program for finding Nash equilibria in bimatrix games to show that the set of the sought weights can be described as a system of linear inequalities. Furthermore, assuming that the given strategy profile is really a Nash equilibrium of the game, we show how to compute lower and upper bounds of players' utilities over all possible weights; this can be in fact realized using linear programming. We also discuss some further extensions and suggest some generalizations and robustifications of the approach.

Keywords: multi-criteria game, vector payoffs, weighted scalarization, payoff bounds.

JEL classification: C72
AMS classification: 91A10

1 Introduction

1.1 Multi-criteria games

Multi-criteria games are natural extension of classic game-theoretic models, this holds both for games in strategic form and for coalitional games.

Related work. The basics for multi-criteria games were laid independently by Blackwell [2] and Shapley [11, 12]. Work of both the authors addressed roughly the same problem: the question of existence of a Nash equilibrium in multi-criteria zero-sum game, however, Shapley's terminology is closer to the one that we are nowadays used to.

This work was followed by several other researchers, for example by Zeleny [16]. Zeleny's approach introduced a form of weighted linear scalarization, while still considering zero-sum games. This was generalized by several other authors, namely Borm et al. [3] or Kruś and Bronisz [7] to mention some. Recently, a survey on solution concepts of multi-criteria games appeared due to Anand et al. [1].

There are many more interesting works in the area of noncooperative multi-criteria games. We mention the work of Voorneveld et al. [14] and especially the work on stable equilibria by de Marco and Morgan [5], which is closely related to our paper. Recently, interesting topics in multi-criteria game theory were addressed also by Yu and Liu [15] (robustification), by Pusillo and Tijs [10] (E-equilibria), by de Marco [4] (generic approach to games) or by Fahem and Radjef [6] (proper efficiency of equilibria).

Our starting point. Most recently, several results on both the noncooperative and coalitional games were achieved in the doctoral thesis by Tichá [13]. One of the results – determining weights under which a given strategy profile is Nash equilibrium in a finite noncooperative game – serves as the starting point for this paper. We take the result, strengthen it slightly and formulate an approach for obtaining additional informations from the knowledge of a strategy profile.

The goal is formulated more precisely and formally in Section 2.

1.2 Notation and general results for noncooperative games

Convention. In the whole paper, superscripts serve to distinguish variables, not for exponentiation. For a natural n, the symbol $[n]$ denotes the set $\{1, \dots, n\}$. The symbol $\mathbf{1}$ is the vector $(1, \dots, 1)^{\mathrm{T}}$ of suitable dimension.

For the purposes of this paper, it is sufficient to consider games of two players. However, for sake of generality,

[1]University of Economics, Prague, Faculty of Finance and Accounting, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic, miroslav.rada@vse.cz
[2]University of Economics, Prague, Faculty of Informatics and Statistics, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic, michaela.ticha@vse.cz

we start with some known results from the area of noncooperative games, which sets our problem into broader perspective. First, we introduce the general definitions and propositions, then we switch to the two-player case.

We start with an exhaustive definition that shall fix the notation:

Notation 1 (Basic symbols)**.** We define and denote by

(a) n the *number of players*,
(b) $N := [n]$ the *set of players*,
(c) m^i the *number of pure strategies* of i-th player, for all $i \in N$,
(d) $S^i := [m^i]$ the *set of pure strategies* of i-th player, for all $i \in N$,
(e) $S := \prod_{i \in N} S^i$ the *set of pure strategy profiles*,
(f) $X^i := \{x^i \in \mathbb{R}^{m^i} : x^i \geq 0, 1^{\mathrm{T}} x^i = 1\}$ the *set of mixed strategies* of i-th player, for all $i \in N$,
(g) $X := \prod_{i \in N} X^i$ the *set of all strategy profiles*,
(h) $x^i \in X^i$ the *strategy* of i-th player, for all $i \in N$,
(i) $s^i_j \in X^i$ the *j-th pure strategy in the mixed space* of i-th player, for all $i \in N$,
(j) $x := (x^i)_{i \in N} \in X$ the *strategy profile* of the players,
(k) $x^{-i} := (x^1, \ldots, x^{i-1}, x^{i+1}, \ldots, x^n)$ the strategy profile x *without* the strategy of i-th player; this is to be used with some strategy $\xi^i \in X^i$ of i-th player; in particular, (ξ^i, x^{-i}) is the strategy profile x with strategy ξ^i instead of the strategy x^i,
(l) r^i the *number of payoffs* of i-th player;
(m) P^i_j the mapping $S \mapsto \mathbb{R}$, for each $i \in N$ and for each $j \in [r^i]$,
(n) $v^i_j(x) := \sum_{s^1 \in S^1} \cdots \sum_{s^n \in S^n} P^i_j(s^1, \ldots, s^n) \prod_{k \in N} x^i_{s^k}$ the mapping $X \mapsto \mathbb{R}$ called *j-th payoff of i-th player*, for $j \in [r^i]$ for $i \in N$,
(o) $v^i(x) := (v^i_j(x))_{j \in [r^i]}$ the *payoff vector of i-th player*, for all $i \in N$.

Note that the symbol P in the item (m) in the above definition has the similar role as the payoff matrices in the two-player games. Also, the letter P was chosen as a shortcut for "payoff".

Notation 1 provides us with notation that allows us to define the multi-criteria game in a very simple manner:

Definition 2 (Multi-criteria game in strategic form)**.** Given N, X and $((v^i_j)_{j \in [r^i]})_{i \in N}$, the triplet

$$G := (N, X, ((v^i_j)_{j \in [r^i]})_{i \in N}) \tag{1}$$

is called *(mixed extension of finite) multi-criteria game in strategic form.*

For definition of equilibria of the multi-criteria game, it will be helpful to introduce several ways of comparing vectors.

Convention. Let $a, b \in \mathbb{R}^m$ be given vectors. We define the following comparison operations:

- $a \geq b$ if $a_j \geq b_j$ for every $j \in [m]$,
- $a \gneqq b$ if $a \geq b$ and $a \neq b$,
- $a > b$ if $a_j > b_j$ for every $j \in [m]$.

Definition 3 (Weak equilibrium of multi-criteria game)**.** A strategy profile $x \in X$ is called *weak equilibrium* of multi-criteria game G, if, for all $i \in N$, there is no $\xi^i \in X^i$ such that

$$v^i(\xi^i, x^{-i}) > v^i(x). \tag{2}$$

Definition 4 (Strong equilibrium of multi-criteria game)**.** A strategy profile $x \in X$ is called *strong equilibrium* of multi-criteria game G, if, for all $i \in N$, there is no $\xi^i \in X^i$ such that

$$v^i(\xi^i, x^{-i}) \gneqq v^i(x). \tag{3}$$

Straightforward approach to solving a multi-criteria game is to *aggregate* the multi-ple criteria somehow into one criterion. This principle is called *scalarization* and is pretty standard in other areas of multi-criteria decision-making. Generally, there are many different ways how to scalarize; here, we are interested in linear scalarization, i.e. in aggregating the criteria via linear combination (or, more precisely, convex combination) of the criteria with some coefficients of the combination. To formalize this, Notation 5 follows:

Notation 5 (Scalarization stuff)**.** We define and denote by

(o) $W^i := \{w \in \mathbb{R}^{r^i} : w \geq 0, 1^{\mathrm{T}} w = 1\}$ the *set of possible weights* of the payoffs, for all $i \in N$,
(p) $w^i \in W^i$ the *weights* of i-th player, for all $i \in N$,

(q) $u^i(x, w^i) \coloneqq \sum_{j\in[r^i]} w^i_j v^i_j(x)$ the mapping $X \times W^i \mapsto \mathbb{R}$ called *utility function of i-th player*, for all $i \in N$.

Definition 6 (Linear scalarization of multi-criteria game in strategic form)**.** Given N, X and $(u^i)_{i\in N}$ (assume that knowledge of u^i implies the knowledge of w^i), the triplet

$$G^w \coloneqq (N, X, (u^i)_{i\in N}) \tag{4}$$

is called *linear scalarization of (mixed extension of finite) multi-criteria game in strategic form.*

Multi-criteria games and their scalarizations (and their Nash equilibria) are related by the classic result of Kruś and Bronisz [7]:

Theorem 1. *Assume $x \in X$ is a weak Nash equilibrium of the scalarized multi-criteria game*

$$G^w = \left(N, X, (u^i)_{i\in N}\right). \tag{5}$$

Then x is a weak Nash equilibrium of the multi-criteria game

$$G = \left(N, X, ((v^i_j)_{j\in[r^i]})_{i\in N}\right). \tag{6}$$

In other words, the theorem says that if we assign weights to each criterion and convert multi-criteria game to classical game, we will find a subset of the weak equilibria of the multi-criteria game corresponding to the preferences of players.

In this paper, we assume mixed extensions of finite games. However, Theorem 1 is even stronger: it holds for arbitrary payoff functions and strategy spaces.

For our specific case of payoffs that are linear in every strategy of each player, the converse of the implication in Theorem 1 also holds: for a Nash equilibrium x of the multi-criteria game G, there exist weights w such that x is Nash equilibrium of G^w. This justifies the importance of studying the scalarizations.

2 Problem formulation

From now on, we assume $n = 2$, i.e. we are restricted to games of two players. Standard problem in two-player games, both one-criterial and multi-criterial, is to find equilibria. For one-criterial games, there are several ways how to find equilibrium. Here, we come from the classic mathematical-programming formulation of Mangasarian and Stone [9]. However, we shall mention a Lemke-Howson algorithm [8] for finding *all* equilibria of this game.

Convention. The shortcut *NE* stands for weak Nash equilibrium. Since we have now only two players, it will be helpful to consider the mappings P^i_j (see Notation 1(m)) as matrices and to simplify v^i_j (see Notation 1(n)) to $x^{1\mathrm{T}} P^i_j x^2$, as it is usual for the case of bimatrix games (bimatrix multi-criteria games, in our case). This is formalized in Notation 7.

Notation 7 (Replacement for the case of two players)**.** We replace the following symbols from Notation 1 and 5 as follows and also add one more symbol:

- Not. 1(m) – every symbol P^i_j now represents a matrix of dimensions $m^1 \times m^2$ and denotes the payoffs of i-th player according to j-th criterion, for all $j \in [r^i]$, $i \in \{1, 2\}$,
- Not. 1(n) – the mapping v^i_j is now defined as $v^i_j(x) = x^{1\mathrm{T}} P^i_j x^2$,
- Not. 5(q) – the mapping u^i is now defined as $u^i(x, w^i) = \sum_{j\in[r^i]} x^{1\mathrm{T}} w^i_j P^i_j x^2 = x^{1\mathrm{T}} P^i x^2$, where
- $P^i = \sum_{j\in[r^i]} w^i_j P^i_j$, for all $i \in \{1, 2\}$,
- Not. 1(h) – the strategy x^1 of player 1 is from now on understood as row vector; this will be useful in matrix multiplications.

The result of Mangasarian and Stone is (for our game G^w with fixed w) stated in Theorem 2.

Theorem 2 (Quadratic program for bimatrix games [9])**.** *Given a one-criterial game $G^w = (\{1, 2\}, X^1 \times X^2, (u^1, u^2))$ and fixed weights $w^1 \in W^1, w^2 \in W^2$, the strategy profile $(x^1, x^2) \in X^1 \times X^2$ is a Nash equilibrium of G^w if and only if it is the global optimal solution of the quadratic program*

$$\begin{aligned} \max_{x^1\in X^1, x^2\in X^2, \alpha,\beta\in\mathbb{R}} & x^1 \left(P^1 + P^2\right) x^2 - \alpha - \beta, \\ & P^1 x^2 - \alpha \mathbf{1} \le 0, \\ & x^1 P^2 - \beta \mathbf{1} \le 0, \end{aligned} \tag{7}$$

with the optimal value of 0.

Corollary 3. *Given weights $w^1 \in W^1$ and $w^2 \in W^2$, one is able to find at least one Nash equilibrium of a multi-criteria game of two players using quadratic program* (7) *and Theorem 1.*

Our problem(s). The problem "*we are given multi-criteria game and weights, compute a Nash equilibrium*" is easily solvable due to Corollary 3. However, one can pose it in the opposite way:

$$\text{“given a strategy profile } x \text{, find weights } w^1 \in W^1, w^2 \in W^2 \text{ such that } x \text{ is a Nash equilibrium”.} \tag{P1}$$

Furthermore, if there are multiple such weights, one can ask, what is actually the payoff of each particular player. Since that cannot be answered without an additional information about weights, we will address the following weaker problem:

$$\text{“given a strategy profile } x \text{, find lower and upper bound of payoff of each player s.t. } x \text{ is a NE”.} \tag{P2}$$

Goal. The goal of this paper is simply to propose a way how to solve problems (P1) and (P2) efficiently.

Structure. The rest of the paper is structured as follows. In Section 3, we formulate an optimization problem for solving problem (P1), based on quadratic program (7); this is formalized in Lemma 5. The formulation (7) is further extended in Section 4 into form suitable for solving (P2), see Lemma 6. The Section 5 completes and overviews the paper.

3 Finding the set of weight vectors

We come out from the quadratic program (7). The Theorem 2 can be directly used for characterizing the set of weights for which a given $x \in X$ is NE – it is sufficient to swap variables – we optimize over w^1, w^2 instead of x^1, x^2. The formulation then reads

$$\max_{w^1 \in W^1, w^2 \in W^2, \alpha, \beta \in \mathbb{R}} x^1 \left(\sum_{j \in [r^1]} w_j^1 P_j^1 + \sum_{j \in [r^2]} w_j^2 P_j^2 \right) x^2 - \alpha - \beta, \tag{8a}$$

$$\sum_{j \in [r^1]} w_j^1 P_j^1 x^2 - \alpha \mathbf{1} \le 0, \tag{8b}$$

$$x^1 \sum_{j \in [r^2]} w_j^2 P_j^2 - \beta \mathbf{1} \le 0, \tag{8c}$$

Note that the formulation (8) is *linear* in w^1 and w^2. This significantly simplifies the situation. First, let us supplement the formulations (7) and (8) with several easy but useful propositions:

Proposition 4.
- *(a) Let $(x^1, x^2, \alpha, \beta) \in X \times \mathbb{R}^2$ be a feasible solution of* (7) *for some fixed $w^1 \in W^1$ and $w^2 \in W^2$. Then $x^1 P^1 x^2 - \alpha \le 0$ and $x^1 P^2 x^2 - \beta \le 0$ hold.*
- *(b) Let $(w^1, w^2, \alpha, \beta) \in W^1 \times W^2 \times \mathbb{R}^2$ be a feasible solution of* (8) *for some fixed $x \in X$. Then $x^1 \sum_{j \in [r^1]} w_j^1 P_j^1 x^2 - \alpha \le 0$ and $x^1 \sum_{j \in [r^2]} w_j^2 P_j^2 x^2 - \beta \le 0$.*
- *(c) $x \in X$ is NE of G^w if and only if the objective value of* (7) *or* (8) *is* 0.

Proof.
1) The items (a) and (b) follow from the way the programs (7) and (8) are derived. We focus on the program (8). In fact, the m^1 constraints (8b) tell that the payoff of player 1 for every pure strategy s_j^1 is less than α; making standard trick with the convex combination of these inequalities we can conclude that, for every x^2, the payoff of player 1 is less than α regardless of choice of x^1, hence, it is less than α for the solution of (8), too. The same holds for the player 2 and constraints (8c).
2) The item (c) is obvious according to the items (a) and (b) and Theorem 2. □

Remark. Note that α resp. β is the upper bound for payoff of player 1 resp. 2. If $x \in X$ is NE, then α resp. β is payoff of the player 1 resp. 2.

Lemma 5 (Linear program for (P1))**.** *Assume the game G^w and a strategy profile $x \in X, w^i \in W^i$ are given. The* *sets $\mathcal{W}^1(x), \mathcal{W}^2(x)$ of weights $w^i \in W^i$ such that x is NE of G^w are the convex polyhedra*

$$\mathcal{W}^1(x) := \left\{ w^1 \in W^1 : \sum_{j\in[r^1]} w^1_j P^1_j x^2 \leq \alpha \mathbf{1},\ x^1 \sum_{j\in[r^1]} w^1_j P^1_j x^2 = \alpha \right\},$$
$$\mathcal{W}^2(x) := \left\{ w^2 \in W^2 : x^1 \sum_{j\in[r^2]} w^2_j P^2_j \leq \beta \mathbf{1},\ x^1 \sum_{j\in[r^1]} w^1_j P^2_j x^2 = \beta \right\}. \tag{9}$$

Proof. We take the formulation (8). Due to Proposition 4(c), the objective function can be treated as constraint by setting the requested value to 0. Then, due to Proposition 4(b), this new constraint can be separated into two constraints.

Clearly, all constraints in definition of $\mathcal{W}^1(x)$ and $\mathcal{W}^2(x)$ are linear in w^1 resp. w^2. Since the sets are defined as intersection of linear constraints, they are given as convex H-polyhedra. □

4 Finding payoff bounds

Note that with unknown weights, the payoffs are also unknown. Here, we are interested in computation of lower and upper bound of payoff of each player. Since the sets $\mathcal{W}^1(x)$ and $\mathcal{W}^2(x)$ have linear description, this is for a given x possible simply with a linear program:

Lemma 6 (Linear program for (P2))**.** *Assume the game G^w is given and a strategy profile $x \in X$ that is known to be NE. Then, the lower bound of payoff of i-th player can be determined using the linear program*

$$\min_{w^i} x^1 \left(\sum_{j\in[r^i]} w^i_j P^i_j \right) x^2, \tag{10}$$
$$\text{s.t. } w^i \in \mathcal{W}^i(x). \tag{11}$$

Proof. Obvious. □

The upper bound can be found by replacing "min" by "max" in (10).

5 Conclusion

Given linear scalarization of a multi-criteria game with known strategy profile and unknown weights, we presented an approach (see Lemma 5) for determining weights of criteria under which the strategy profile is a Nash equilibrium. Also, we presented an approach (see Lemma 6) for computing upper and lower bound of payoffs of players; since under unknown weights they are also unknown.

Surprisingly, the above problems can be solved simply using linear programming and hence it is fully reasonable and ready to use in practical multi-criteria games – unfortunately, for space reasons, we cannot present the example of a multi-criteria game (for example the multi-criteria battle of sexes) to demonstrate the power of our results.

On the other hand, the simplicity of the presented results calls for some generalizations. Namely, it is apparent that the approach shall be generalized for the games of arbitrary number of players. Also, the results could be somewhat robustified. Namely, consider the following problem: from m observations of realizations of a game, we estimate the strategy profile x. What is the α-percent region of confidence for weights? What is the α-percent region of confidence for payoffs? These questions should be the subject of our interest for further research.

Acknowledgements

The work of the first author was supported by Institutional Support of University of Economics, Prague, VŠE IP 100040. The work of the second author is supported by internal grant agency VŠE IGA F4/54/2015.

References

[1] Anand, L., Shashishekhar, N., Ghose, D., and Prasad, U. R.: A survey of solution concepts in multicriteria games. *Journal of the Indian Institute of Science* **75** (2013), 141–174.

[2] Blackwell, D.: An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics* **6** (1956), 1–8.

[3] Borm, P., Vermeulen, D., and Voorneveld, M.: The structure of the set of equilibria for two person multicriteria games. *European Journal of Operational Research* **148** (2003), 480–493.

[4] De Marco, G.: On the genericity of the finite number of equilibria in multicriteria games: a counterexample. *International Journal of Mathematics in Operational Research* **5** (2013), 764–777.

[5] De Marco, G., and Morgan, J.: A Refinement Concept for Equilibria in Multicriteria Games Via Stable Scalarizations. *International Game Theory Review* **9** (2007), 169–181.

[6] Fahem, K., and Radjef, M. S.: Properly efficient Nash equilibrium in multicriteria noncooperative games. *Mathematical Methods of Operations Research* **82** (2015), 175–193.

[7] Kruś, L., and Bronisz, P.: On n-person noncooperative multicriteria games described in strategic form. *Annals of Operations Research* **51** (1994), 83–97.

[8] Lemke, C., and Howson, J.: Equilibrium Points of Bimatrix Games. *Journal of the Society for Industrial and Applied Mathematics* **12** (1964), 413–423.

[9] Mangasarian, O. L., and Stone, H.: Two-person nonzero-sum games and quadratic programming. *Journal of Mathematical Analysis and Applications* **9** (1964), 348–355.

[10] Pusillo, L., and Tijs, S.: E-equilibria for Multicriteria Games. In: *Advances in Dynamic Games*. Springer, 2013, 217–228.

[11] Shapley, L. S.: *Equilibrium Points in Games with Vector Payoffs*. Technical report, RAND, 1956.

[12] Shapley, L. S., and Rigby, F. D.: Equilibrium Points in Games with Vector Payoffs. *Naval Research Logistics Quarterly* **6** (1959), 57–61.

[13] Tichá, M.: *Vícekriteriální hry*. Ph.d. thesis, University of Economics, Prague, Prague, 2016.

[14] Voorneveld, M., Grahn, S., and Dufwenberg, M.: Ideal equilibria in noncooperative multicriteria games. *Mathematical methods of operations research* **52** (2000), 65–77.

[15] Yu, H., and Liu, H. M.: Robust multiple objective game theory. *Journal of Optimization Theory and Applications* **159** (2013), 272–280.

[16] Zeleny, M.: Games with multiple payoffs. *International Journal of Game Theory* **4** (1975), 179–191.

Strong Consistent Pairwise Comparisons Matrix with Fuzzy Elements and Its Use in Multi-Criteria Decision Making

Jaroslav Ramík [1]

Abstract. The decision making problem considered in this paper is to rank n distinct alternatives from the best to the worst, using the information given by the decision maker in the form of an n by n pairwise comparisons matrix with fuzzy elements from abelian linearly ordered group (alo-group) over a real interval. The concept of reciprocity and consistency of pairwise comparisons matrices with fuzzy elements have been already studied in the literature. We define stronger concepts, namely the strong reciprocity and strong consistency of pairwise comparisons matrices with fuzzy intervals, derive necessary and sufficient conditions for strong reciprocity and strong consistency and investigate some consequences to the problem of multi-criteria decision making.

Keywords: pairwise comparisons matrix, matrix with fuzzy elements, multi-criteria decision making.

JEL classification: C44
AMS classification: 90C15

1 Introduction

A *decision making problem (DM problem)* which forms an application background in this paper can be formulated as follows. Let $X = \{x_1, x_2, ..., x_n\}$ be a finite set of alternatives ($n > 1$). The decision maker's aim is to rank the alternatives from the best to the worst, using the information given by the decision maker in the form of an $n \times n$ pairwise comparisons matrix.

Fuzzy sets being the elements of the pairwise comparisons matrix can be applied whenever the decision maker is not sure about the preference degree of his/her evaluation of the pairs in question. Fuzzy elements may be taken also as the aggregations of crisp pairwise comparisons of a group of decision makers in the group DM problem. Decision makers acknowledge fuzzy pairwise preference data as imprecise knowledge about regular preference information. Usually, an ordinal *ranking* of alternatives is required to obtain the "best" alternative(s), however, it often occurs that the decision maker is not satisfied with the ordinal ranking among alternatives and a cardinal ranking i.e. *rating* is then required.

The former works that investigated the problem of finding a rank of the given alternatives based on some pairwise comparisons matrices are e.g. [3], [4], [7], and [8]. The recent paper is in some sense a continuation of [7].

2 Preliminaries

Here, it will be useful to understand fuzzy sets as special nested families of subsets of a set, see [6]. A *fuzzy subset* of a nonempty set X (or a *fuzzy set* on X) is a family $\{A_\alpha\}_{\alpha \in [0,1]}$ of subsets of X such that $A_0 = X, A_\beta \subset A_\alpha$ whenever $0 \leq \alpha \leq \beta \leq 1$, and $A_\beta = \cap_{0 \leq \alpha < \beta} A_\alpha$ whenever $0 < \beta \leq 1$. The *membership function of* A is the function μ_A from X into the unit interval $[0, 1]$ defined by $\mu_A(x) = \sup\{\alpha \mid x \in A_\alpha\}$. Given $\alpha \in]0, 1]$, the set $[A]_\alpha = \{x \in X \mid \mu_A(x) \geq \alpha\}$ is called the α-*cut of fuzzy set* A. If X is a nonempty subset of the n-dimensional Euclidean space $\mathbf{R}^n$, then a fuzzy set A of X is called *closed,*

[1]Silesian University in Opava/School of Business Administration in Karvina, University Sq. 1934/3, 73340 Karvina, Czech Republic, ramik@opf.slu.cz

bounded, compact or *convex* if the α-cut $[A]_\alpha$ is a closed, bounded, compact or convex subset of X for every $\alpha \in]0,1]$, respectively.

We say that a fuzzy subset A of $\mathbf{R}^* = \mathbf{R} \cup \{-\infty\} \cup \{+\infty\}$ is a *fuzzy interval* whenever A is normal and its membership function μ_A satisfies the following condition: there exist $a, b, c, d \in \mathbf{R}^*$, $-\infty \le a \le b \le c \le d \le +\infty$, such that

$$\begin{aligned} \mu_A(t) = 0 \quad & \text{if } t < a \text{ or } t > d, \\ \mu_A \quad & \text{is strict increasing and continuous on } [a,b], \\ \mu_A(t) = 1 \quad & \text{if } b \le t \le c, \\ \mu_A \quad & \text{is strict decreasing and continuous on } [c,d]. \end{aligned} \tag{1}$$

A fuzzy interval A is *bounded* if $[a,d]$ is a compact interval. Moreover, for $\alpha = 0$ we define the *zero-cut* of A as $[A]_0 = [a,d]$. A bounded fuzzy interval A is the *fuzzy number,* if $b = c$.

An *abelian group* is a set, G, together with an operation $\odot$ (read: operation *odot*) that combines any two elements $a, b \in G$ to form another element in G denoted by $a \odot b$, see [1]. The symbol $\odot$ is a general placeholder for a concretely given operation. $(G, \odot)$ satisfies the following requirements known as the *abelian group axioms*, particularly: *commutativity*, *associativity*, there exists an *identity element* $e \in G$ and for each element $a \in G$ there exists an element $a^{(-1)} \in G$ called the *inverse element to a.* The *inverse operation* $\div$ to $\odot$ is defined for all $a, b \in G$ as follows: $a \div b = a \odot b^{(-1)}$.

An ordered triple $(G, \odot, \le)$ is said to be *abelian linearly ordered group*, *alo-group* for short, if $(G, \odot)$ is a group, $\le$ is a linear order on G, and for all $a, b, c \in G$: $a \le b$ implies $a \odot c \le b \odot c$.

If $\mathcal{G} = (G, \odot, \le)$ is an alo-group, then G is naturally equipped with the order topology induced by $\le$ and $G \times G$ is equipped with the related product topology. We say that $\mathcal{G}$ is a *continuous alo-group* if $\odot$ is continuous on $G \times G$.

By definition, an alo-group $\mathcal{G}$ is a lattice ordered group. Hence, there exists $\max\{a, b\}$, for each pair $(a, b) \in G \times G$. Nevertheless, a nontrivial alo-group $\mathcal{G} = (G, \odot, \le)$ has neither the greatest element nor the least element. $\mathcal{G} = (G, \odot, \le)$ is *divisible* if for each positive integer n and each $a \in G$ there exists the (n)-th root of a denoted by $a^{(1/n)}$, i.e. $\left(a^{(1/n)}\right)^{(n)} = a$. The function $\|.\| : G \to G$ defined for each $a \in G$ by $\|a\| = \max\{a, a^{(-1)}\}$ is called a *$\mathcal{G}$-norm.* The operation $d : G \times G \to G$ defined by $d(a, b) = \|a \div b\|$ for all $a, b \in G$ is called a *$\mathcal{G}$-distance.*
Next, we present the well known examples of alo-groups, see also [2], or [7].

Example 1: Additive alo-group $\mathcal{R} = (]-\infty, +\infty[, +, \le)$ is a continuous alo-group with:
$e = 0,\ a^{(-1)} = -a,\ a^{(n)} = a + a + ... + a = n.a.$

Example 2: Multiplicative alo-group $\mathcal{R}^+ = (]0, +\infty[, \bullet, \le)$ is a continuous alo-group with:
$e = 1,\ a^{(-1)} = a^{-1} = 1/a,\ a^{(n)} = a^n$. Here, by $\bullet$ we denote the usual operation of multiplication.

Example 3: Fuzzy add. alo-group $\mathcal{R}_a = (]-\infty, +\infty[, +_f, \le)$, see [7], is a continuous alo-group with:
$a +_f b = a + b - 0.5,\ e = 0.5, a^{(-1)} = 1 - a, a^{(n)} = n.a - (n-1)/2.$

Example 4: Fuzzy multipl. alo-group $]\mathbf{0}, \mathbf{1}[_\mathbf{m} = (]0,1[, \bullet_f, \le)$, see [2], is a continuous alo-group with:

$$a \bullet_f b = \frac{ab}{ab + (1-a)(1-b)}, e = 0.5, a^{(-1)} = 1 - a.$$

3 PCF matrices

Notice that elements of PCF matrices may be crisp and/or fuzzy numbers, also crisp and/or fuzzy intervals, fuzzy intervals with bell-shaped membership functions, triangular fuzzy numbers, trapezoidal fuzzy numbers etc. Such fuzzy elements may be either evaluated by individual decision makers, or, they may be made up of crisp pairwise evaluations of decision makers in a group DM problem.

Now, we shall define reciprocity properties for PCF matrices. Reciprocity of a PC matrix is a natural property defining the evaluation of couples of alternatives in the reverse order. First, we define reciprocity for PCF matrices, then we define the concept of the strong reciprocity and we shall investigate some relationships between the two properties. We derive necessary and sufficient conditions for a PCF matrix to be strong reciprocal. Our approach will cover the classical definitions of reciprocity presented e.g. in [5], and [7].

Let $C = \{\tilde{c}_{ij}\}$ be an $n \times n$ PCF matrix, $\alpha \in [0,1]$. C is said to be *α-$\odot$-reciprocal*, if the following condition holds: For every $i,j \in \{1,2,...,n\}$ there exist $c_{ij} \in [\tilde{c}_{ij}]_\alpha$ and $c_{ji} \in [\tilde{c}_{ji}]_\alpha$ such that $c_{ij} \odot c_{ji} = e$, or, equivalently, $c_{ji} = c_{ij}^{(-1)}$. $C = \{\tilde{c}_{ij}\}$ is said to be *$\odot$-reciprocal*, if C is α-$\odot$-reciprocal for all $\alpha \in [0,1]$.

Remark 1. If $C = \{\tilde{c}_{ij}\}$ is a PCF matrix with crisp elements, then $\tilde{c}_{ij} = c_{ij}$ with $c_{ij} \in G$ for all i and j, and the above definition coincides with the classical definition of reciprocity for crisp PCF matrices: A crisp PCF matrix $C = \{c_{ij}\}$ is *$\odot$-reciprocal* if for all i and j: $c_{ji} = c_{ij}^{(-1)}$. Particularly, $C = \{c_{ij}\}$ is additive-reciprocal if $c_{ji} = -c_{ij}$ for all i and j; $C = \{c_{ij}\}$ is multiplicative-reciprocal if $c_{ji} = \frac{1}{c_{ij}}$ for all i and j. Clearly, if $C = \{\tilde{c}_{ij}\}$ is an α-$\odot$-reciprocal PCF matrix, $\alpha, \beta \in [0,1]$, $\beta \leq \alpha$, then $C = \{\tilde{c}_{ij}\}$ is β-$\odot$-reciprocal. We denote

$$[c_{ij}^L(\alpha), c_{ij}^R(\alpha)] = [\tilde{c}_{ij}]_\alpha.$$

For a PCF matrix with non-crisp elements the reciprocity condition is a rather weak condition. That is why we shall ask stronger conditions to define the concept of *strong* reciprocity.

Let $C = \{\tilde{c}_{ij}\}$ be an $n \times n$ PCF matrix, $\alpha \in [0,1]$. C is said to be *strong α-$\odot$-reciprocal*, if the following condition holds:
For every $i,j \in \{1,2,...,n\}$ and for every $c_{ij} \in [\tilde{c}_{ij}]_\alpha$ there exists $c_{ji} \in [\tilde{c}_{ji}]_\alpha$ such that

$$c_{ij} \odot c_{ji} = e. \tag{2}$$

$C = \{\tilde{c}_{ij}\}$ is said to be *strong $\odot$-reciprocal*, if C is strong α-$\odot$-reciprocal for all $\alpha \in [0,1]$.

Remark 2. Every strong α-$\odot$-reciprocal PCF matrix is α-$\odot$-reciprocal, but not vice-versa.

Proposition 1. *Let $C = \{\tilde{c}_{ij}\}$ be a PCF matrix, $\alpha \in [0,1]$. The following conditions are equivalent.*

(i) *C is strong α-$\odot$-reciprocal.*

(ii) $c_{ij}^L(\alpha) \odot c_{ji}^R(\alpha) = e$ *and* $c_{ij}^R(\alpha) \odot c_{ji}^L(\alpha) = e$, *for all* $i,j \in \{1,2,...,n\}$.

(iii) $[c_{ji}^L(\alpha), c_{ji}^R(\alpha)] = [(c_{ij}^R(\alpha))^{(-1)}, (c_{ij}^L(\alpha))^{(-1)}]$, *for all* $i,j \in \{1,2,...,n\}$.

Remark 3. When evaluating fuzzy elements of a PCF matrix $C = \{\tilde{c}_{ij}\}$, only one of the membership functions of elements $\tilde{c}_{ij}$ and $\tilde{c}_{ji}$, $i \neq j$, should be evaluated, the other should satisfy condition (ii), or (iii). Then the PCF matrix becomes strong $\odot$-reciprocal. If $C = \{\tilde{c}_{ij}\}$ is a strong α-$\odot$-reciprocal PCF matrix, $\alpha, \beta \in [0,1], \beta < \alpha$, then $C = \{\tilde{c}_{ij}\}$ need not be strong β-$\odot$-reciprocal. It is, however, β-$\odot$-reciprocal.

Rationality and compatibility of a decision making process can be achieved by the consistency property of PCF matrices.

Let $\mathcal{G} = (G, \odot, \leq)$ be a divisible and continuous alo-group, $C = \{\tilde{c}_{ij}\}$ be a crisp PCF matrix, where $c_{ij} \in G \subset \mathbf{R}$ for all $i,j \in \{1,2,...,n\}$. The following definition is well known, see e.g. [2]. A crisp PCF matrix $C = \{c_{ij}\}$ is *$\odot$-consistent* if for all $i,j,k \in \{1,2,...,n\} : c_{ik} = c_{ij} \odot c_{jk}$.

Now, we extend the definition to non-crisp PCF matrices as follows, see also [7].

Let $\alpha \in [0,1]$. A PCF matrix $C = \{\tilde{c}_{ij}\}$ is said to be *α-$\odot$-consistent*, if the following condition holds: For every $i,j,k \in \{1,2,...,n\}$, there exists $c_{ik} \in [\tilde{c}_{ik}]_\alpha$, $c_{ij} \in [\tilde{c}_{ij}]_\alpha$ and $c_{jk} \in [\tilde{c}_{jk}]_\alpha$ such that $c_{ik} = c_{ij} \odot c_{jk}$. The matrix C is said to be *$\odot$-consistent*, if C is α-$\odot$-consistent for all $\alpha \in [0,1]$. The following properties are evident.

- If $\alpha, \beta \in [0,1]$, $\beta \leq \alpha$, $C = \{\tilde{c}_{ij}\}$ is α-$\odot$-consistent, then it is β-$\odot$-consistent.
- If $C = \{\tilde{c}_{ij}\}$ is α-$\odot$-consistent, then it is α-$\odot$-reciprocal.
- If $C = \{\tilde{c}_{ij}\}$ is $\odot$-consistent, then it is $\odot$-reciprocal.

Now, the consistency property of PCF matrices will be strengthen. Similarly to α-$\odot$-strong reciprocity defined earlier, we define strong α-$\odot$-consistent PCF matrices and derive their properties.

Let $\alpha \in [0,1]$. A PCF matrix $C = \{\tilde{c}_{ij}\}$ is said to be *strong α-$\odot$-consistent*, if the following condition holds: For every $i, j, k \in \{1, 2, ..., n\}$, and for every $c_{ij} \in [\tilde{c}_{ij}]_\alpha$, there exist $c_{ik} \in [\tilde{c}_{ik}]_\alpha$ and $c_{jk} \in [\tilde{c}_{jk}]_\alpha$ such that $c_{ik} = c_{ij} \odot c_{jk}$. The matrix C is said to be *strong $\odot$-consistent*, if C is strong α-$\odot$-consistent for all $\alpha \in [0,1]$. The following properties are evident.

- Each strong α-$\odot$-consistent PCF matrix is α-$\odot$-consistent.
- If $C = \{\tilde{c}_{ij}\}$ is a strong α-$\odot$-consistent PCF matrix, $\alpha, \beta \in [0,1], \beta < \alpha$, then $C = \{\tilde{c}_{ij}\}$ need not be strong β-$\odot$-consistent. However, it must be β-$\odot$-consistent.
- If $C = \{\tilde{c}_{ij}\}$ is strong α-$\odot$-consistent, then $C = \{\tilde{c}_{ij}\}$ is strong α-$\odot$-reciprocal.
- If $C = \{\tilde{c}_{ij}\}$ is strong $\odot$-consistent, then $C = \{\tilde{c}_{ij}\}$ is strong $\odot$-reciprocal.

Now, we formulate two necessary and sufficient conditions for strong α-$\odot$-consistency of a PCF matrix. This property may be useful for checking strong consistency of PCF matrices.

Proposition 2. *Let $\alpha \in [0,1]$, $C = \{\tilde{c}_{ij}\}$ be a PCF matrix. The following three conditions are equivalent.*

(i) $C = \{\tilde{c}_{ij}\}$ *is strong α-$\odot$-consistent.*

(ii) $[c^L_{ik}(\alpha), c^R_{ik}(\alpha)] \cap [c^R_{ij}(\alpha) \odot c^L_{jk}(\alpha), c^L_{ij}(\alpha) \odot c^R_{jk}(\alpha)] \neq \emptyset,$ *for all* $i, j, k \in \{1, 2, ..., n\}$.

(iii) $c^L_{ik}(\alpha) \leq c^L_{ij}(\alpha) \odot c^R_{jk}(\alpha),$ *and* $c^R_{ik}(\alpha) \geq c^R_{ij}(\alpha) \odot c^L_{jk}(\alpha),$ *for all* $i, j, k \in \{1, 2, ..., n\}$.

Remark 4. Property (iii) in Proposition 2 is useful for checking strong α-$\odot$-consistency. For a given PCF matrix $C = \{\tilde{c}_{ij}\}$ it can be easily checked whether inequalities (iii) are satisfied or not.

Example 5: Consider the additive alo-group $\mathcal{R} = (\mathbf{R}, \odot, \leq)$ with $\odot = +$, see Example 1. Let PCF matrices $A = \{\tilde{a}_{ij}\}$ be given as follows:

$$A = \begin{bmatrix} (0;0;0) & (1;2;3) & (3.5;6;8) \\ (-3;-2;-1) & (0;0;0) & (2.5;4;5) \\ (-8;-6;-3.5) & (-5;-4;-2.5) & (0;0;0) \end{bmatrix}.$$

Here, A is a 3×3 PCF matrix, particularly a PCF matrix with triangular fuzzy number elements and the usual "linear" membership functions. Checking inequalities (iii), we obtain that A is strong α-$\odot$-consistent PCF matrix for all α, $0 \leq \alpha \leq 1$, hence, A is strong $\odot$-consistent.

4 Priority vectors, inconsistency of PCF matrices

Definition of the priority vector for ranking the alternatives will be based on the optimal solution of the following optimization problem:
(P1)

$$\alpha \longrightarrow \max; \tag{3}$$

$$\textit{subject to} \qquad c^L_{ij}(\alpha) \leq w_i \div w_j \leq c^R_{ij}(\alpha) \text{ for all } i, j \in \{1, 2, ..., n\}, \tag{4}$$

$$\bigodot_{k=1}^{n} w_k = e, \tag{5}$$

$$0 \leq \alpha \leq 1, w_k \in G, \text{ for all } k \in \{1, 2, ..., n\}. \tag{6}$$

If optimization problem (P1) has a feasible solution, i.e. system of constraints (4) - (6) has a solution, then (P1) has also an optimal solution. Let α^* and $w^* = (w^*_1, ..., w^*_n)$ be an optimal solution of problem (P1). Then α^* is called the *$\odot$-consistency grade of C*, denoted by $g_\odot(C)$, i.e. $g_\odot(C) = \alpha^*$. Here,

$w^* = (w_1^*, ..., w_n^*)$ is called the $\odot$-*priority vector of* C. This vector is associated with the ranking of alternatives as follows: If $w_i^* > w_j^*$ then $x_i \succ x_j$, where $\succ$ stands for "is better then".

If optimization problem (P1) has no feasible solution, then we define $g_\odot(C) = 0$. The priority vector will be defined in this section.

Generally, problem (P1) is a nonlinear optimization problem that can be efficiently solved e.g. by the dichotomy method, which is a sequence of optimization problems, see e.g. [5]. For instance, given $\alpha \in [0, 1]$, $\odot = +$, problem (P1) can be solved as an LP problem (with variables $w_1, ..., w_n$).

Let $C = \{\tilde{c}_{ij}\}$ be a PCF matrix, $\alpha \in [0, 1]$. If there exists a triple of elements $i, j, k \in \{1, 2, ..., n\}$ such that for any $c_{ij} \in [\tilde{c}_{ij}]_\alpha$, any $c_{ik} \in [\tilde{c}_{ik}]_\alpha$, and any $c_{kj} \in [\tilde{c}_{kj}]_\alpha : c_{ik} \neq c_{ij} \odot c_{jk}$, then the PCF matrix C is α-$\odot$-*inconsistent.* If for all $\alpha \in [0, 1]$ the PCF matrix C is α-$\odot$-inconsistent, then we say that C is $\odot$-*inconsistent.* By this definition, for a given PCF matrix C and given $\alpha \in [0, 1]$, C is either α-$\odot$-consistent, or, C is α-$\odot$-inconsistent.

Notice, that for a PCF matrix C problem (P1) has no feasible solution, if and only if C is $\odot$-inconsistent, i.e. C is α-$\odot$-inconsistent for all $\alpha \in [0, 1]$.

It is an important task to measure an intensity of $\odot$-inconsistency of the given PCF matrix. In some cases, a PCF matrix can be "close" to some $\odot$-consistent matrix, in the other cases $\odot$-inconsistency can be strong, meaning that the PCF matrix can be "far" from any $\odot$-consistent matrix.

The $\odot$-inconsistency of C will be measured by the minimum of the distance of the "ratio" matrix $W = \{w_i \div w_j\}$ to the "left" matrix $C^L = \{c_{ij}^L(0)\}$ and "right" matrix $C^R = \{c_{ij}^R(0)\}$, as follows.

Let $w = (w_1, ..., w_n), w_i \in G$, $i, j \in \{1, ..., n\}$. Denote

$$\begin{aligned} d_{ij}(C, w) \quad &= e \text{ if } c_{ij}^L(0) \le w_i \div w_j \le c_{ij}^R(0), \\ &= \min\{\|c_{ij}^L(0) \div (w_i \div w_j)\|, \|c_{ij}^R(0) \div (w_i \div w_j)\|\}, \text{ otherwise.} \end{aligned} \tag{7}$$

Here, by $\|.\|$ we denote the norm defined in Section 2. We define the maximum deviation to the matrix $W = \{w_i \div w_j\}$:

$$I_\odot(C, w) = \max\{d_{ij}(C, w) | i, j \in \{1, ..., n\}\}. \tag{8}$$

Now, consider the following optimization problem.
(P2)

$$I_\odot(C, w) \longrightarrow \min; \tag{9}$$

$$\textit{subject to} \qquad \bigodot_{k=1}^{n} w_k = e, w_k \in G, \text{ for all } k \in \{1, 2, ..., n\}. \tag{10}$$

The $\odot$-*inconsistency index of PCM matrix* C, $I_\odot(C)$, is defined as

$$I_\odot(C) = \inf\{I_\odot(C, w) | w_k \text{ satisfies (10) }\}. \tag{11}$$

If there exists a feasible solution of (P2), then $\odot$-inconsistency index of PCM matrix C, $I_\odot(C)$, is equal to e, i.e. $I_\odot(C) = e$. If $w^* = (w_1^*, ..., w_n^*)$ is an optimal solution of (P2), then $I_\odot(C) = I_\odot(C, w^*)$. It is clear that an optimal solution of (P2) exists, the uniqueness of the optimal solution of (P2) is, however, not saved. Depending on the particular operation $\odot$, problem (P2) may have multiple optimal solutions which is an unfavorable fact from the point of view of the decision maker. In this case, the decision maker should reconsider some (fuzzy) evaluations of pairwise comparison matrix.

Now, we define a priority vector also in case of $g_\odot(C) = 0$, i.e. if no feasible solution of (P1) exists. Let C be an $\odot$-inconsistent PCM matrix. The optimal solution $w^* = (w_1^*, ..., w_n^*)$ of (P2) will be called the $\odot$-*priority vector of* C.

For a PCF matrix $C = \{\tilde{c}_{ij}\}$ just one of the following two cases occurs:

- Problem (P1) has a feasible solution. Then consistency grade $g_\odot(C) = \alpha$, for some $\alpha, 0 \le \alpha \le 1$, $I_\odot(C) = e$. The $\odot$-priority vector of C is the optimal solution of problem (P1).
- Problem (P1) has no feasible solution. Then consistency grade $g_\odot(C) = 0$, C is $\odot$-inconsistent, and $I_\odot(C) > e$. The $\odot$-priority vector of C is the optimal solution of problem (P2).

Example 6: Let $X = \{x_1, x_2, x_3\}$ be a set of alternatives, let $F = \{\tilde{f}_{ij}\}$ be a PCF matrix on the fuzzy multiplicative alo-group $]\mathbf{0},\mathbf{1}[_\mathbf{m}$=$(]0,1[,\bullet_f,\leq)$, with:

$$a \bullet_f b = \frac{ab}{ab + (1-a)(1-b)}, e = 0.5, a^{(-1)} = 1 - a, \|a\| = max\{a, 1-a\}. \tag{12}$$

Fuzzy multiplicative alo-group $]\mathbf{0},\mathbf{1}[_\mathbf{m}$ is divisible and continuous. For more details and properties, see Example 4, eventually, [7]. Let

$$F = \begin{bmatrix} (0.5; 0.5; 0.5) & (0.6; 0.7; 0.8) & (0.75; 0.8; 0.9) \\ (0.2; 0.3; 0.4)) & (0.5; 0.5; 0.5) & (0.7; 0.75; 0.8) \\ (0.1; 0.2; 0.25) & (0.2; 0.25; 0.3) & (0.5; 0.5; 0.5) \end{bmatrix}.$$

Here, F is a 3×3 PCF matrix, particularly, PCF matrix with elements on $]0,1[$. F is a $\bullet_f$-reciprocal PCF matrix (noncrisp), the elements of F are triangular fuzzy numbers. There is an optimal solution of the corresponding problem (P1) with $C = F$, the consistency grade $g_{\bullet_f}(F) = 0.6$. The $\bullet_f$-priority vector w^* of F is $w^* = (0.586, 0.302, 0.112)$, hence $x_1 \succ x_2 \succ x_3$. The inconsistency index $I_{\bullet_f}(F) = e = 0.5$.

5 Conclusion

This paper deals with pairwise comparison matrices with fuzzy elements. Fuzzy elements of the pairwise comparison matrix are usually applied whenever the decision maker is not sure about the value of his/her evaluation of the relative importance of elements in question. In comparison with PC matrices investigated in the literature, here we investigate pairwise comparison matrices with elements from abelian linearly ordered group (alo-group) over a real interval (PCF matrices). We generalize the concept of reciprocity and consistency of pairwise comparison matrices with triangular fuzzy numbers to bounded fuzzy intervals of alo-groups (trapezoidal fuzzy numbers). We also define the concept of priority vector which is an extension of the well known concept in crisp case used here for ranking the alternatives. Such an approach allows for unifying the additive, multiplicative and also fuzzy approaches known from the literature. Then we solve the problem of measuring inconsistency of a PCF matrix C by defining corresponding indexes. Six numerical examples have been presented to illustrate the concepts and derived properties.

Acknowledgements

Supported by the grant No. 14-02424S of the Czech Grant Agency.

References

[1] Bourbaki, N.: *Algebra II.* Springer Verlag, Heidelberg-New York-Berlin, 1998.

[2] Cavallo, B., and D'Apuzzo, L.: A general unified framework for pairwise comparison matrices in multicriteria methods. *International Journal of Intelligent Systems* **24** 4 (2009), 377–398.

[3] Leung, L.C., and Cao, D.: On consistency and ranking of alternatives in fuzzy AHP. *European Journal of Operational Research* **124** (2000), 102–113.

[4] Mikhailov, L.: Deriving priorities from fuzzy pairwise comparison judgments. *Fuzzy Sets and Systems* **134** (2003), 365–385.

[5] Ohnishi, S., Dubois, D. et al.: A fuzzy constraint based approach to the AHP. In: Uncertainty and Intelligent Inf. Syst., World Sci., Singapore, 2008, 217–228.

[6] Ramik, J., and Vlach, M.: *Generalized concavity in optimization and decision making.* Kluwer Academic Publishers, Boston-Dordrecht-London 2001.

[7] Ramik, J.: Isomorphisms between fuzzy pairwise comparison matrices. *Fuzzy Optim. Decis. Making* **14** (2014), 199–209.

[8] Xu, Z.S., and Chen, J.: Some models for deriving the priority weights from interval fuzzy preference relations. *European Journal of Operational Research* **184** (2008), 266–280.

The housing market and credit channel of monetary policy

Vlastimil Reichel[1], Miroslav Hloušek[2]

Abstract. This paper is focused on a credit channel of monetary policy in the housing market. We test the relevance of the credit channel in the Czech housing sector in period 2006:Q1 - 2015:Q4. We use two types of structural VAR models to identify monetary policy shocks and to investigate if they have impacts on behaviour of the real variables. These VAR models are estimated using seven variables: real GDP, consumer price inflation, real house prices, volumes of mortgages, and total loans to households, mortgage interest rate and money market interest rate. Point estimates of impulse responses show that monetary policy shock has a (i) negative effect on the real GDP and mortgages of households; (ii) temporary positive effect on the real house price and (iii) positive effect on the spread between the mortgage rate and the money market interest rate. These results indicate that the monetary policy is able to significantly affect the housing market in the Czech economy.

Keywords: Housing sector, VAR model, credit channel.

JEL classification: E51, E52, C22
AMS classification: 91B84

1 Introduction

Housing sector plays an important role in the business cycle, because changes in house prices can have important wealth effects on consumption and residential investment, which is the most volatile component of aggregate demand (see Bernanke and Gertler [1]). The housing market in the Czech economy was relatively small, but it has been expanding rapidly. The size of mortgage market in 2002 was 4.18 % of GDP, but it has increased to 21.72 % of GDP in 2015 (See Czech National Bank [3]). One of the strong impulses for development of the housing market in the Czech Republic was a rent deregulation law implemented in 2006.

A larger share of the housing sector allows to observe typical relationships among macroeconomic variables. These relationships are called credit channel of monetary policy. Iacoviello and Minetti [7] argue that the relevance and form of the credit channel depends on structural features of the housing finance system of each country. Typically, countries with more developed market have stronger connections between monetary policy and the housing sector. Concretely, monetary policy can affect the housing market through behaviour of households and mortgage banks. To identify the credit channel in the Czech housing market, we use two types of structural VAR models. These models are estimated and studied by impulse response functions. The results show that there is evidence of credit channel in the Czech economy and thus the monetary policy is able to significantly affect the housing market variables.

2 Methodology and Data

2.1 Data

We use seven macroeconomic variables for the identification of the VAR models. They are: real GDP, consumer price inflation, real house prices, volume of mortgages and total loans to households, mortgage interest rate and money market interest rate. The data are obtained from the Czech Statistical Office

[1]Masaryk University, Lipová 41a, Brno, Czech Republic, reichel.v@mail.muni.cz
[2]Masaryk University, Lipová 41a, Brno, Czech Republic, hlousek@econ.muni.cz

(CZSO) and the Czech National Bank (CNB) databases and cover period 2006:Q1 – 2015:Q4. More detailed description, data sources and transformation of the variables is quoted in Table 1.

Variable	Description	Data Source	Transformation
Y	real GDP	CZSO	log, Δ
CPI	Consumer Price Index	CZSO	log, Δ
HP	Offering price of flats deflated by CPI	CZSO	log, Δ
HL	Housing loans (mortgages) to households	CNB	log, Δ
TL	Total loans to households	CNB	log, Δ
R	3M PRIBOR	CNB	–
M	Mortgage interest rate	CNB	–
SP	Spread between M and R	own constr., CNB	–

Table 1 Data for VAR models

We transform following variables: Y, CPI, HP, HL, TL into stationary form using first differences (Δ) of logarithm. Other variables R, M and SP are assumed stationary and are used without transformation.

2.2 VAR model

Two structural VAR models are used for identification of monetary policy shocks. With variables of interest collected in the k-dimensional vector Y_t the reduced-form VAR model can be written as:

$$Y_t = A_1 \cdot Y_{t-1} + A_2 \cdot Y_{t-2} + \cdots + A_p \cdot Y_{t-p} + u_t, \tag{1}$$

where Y_{t-p} is vector of variables lagged by p periods, A_i is a time-invariant $k \times k$ matrix and u_t is a $k \times 1$ vector of error terms.

As the error terms are allowed to be correlated, the reduced-form model is transformed into structural model. Pre-multiplying both sides of equation 1 by the $(k \times k)$ matrix A_0, yields the structural form:

$$A_0 Y_t = A_0 A_1 \cdot Y_{t-1} + A_0 A_2 \cdot Y_{t-2} + \cdots + A_0 A_p \cdot Y_{t-p} + B\epsilon_t. \tag{2}$$

The relationship between structural disturbances ϵ_t and reduced-form disturbances u_t is described by equation of AB model (see Lutkepohl [8] or Lutkepohl and Kratzig [9]):

$$A_0 u_t = B\epsilon_t, \tag{3}$$

where A_0 also describes the contemporaneous relation among the endogenous variables and B is a $(k \times k)$ matrix. In the structural model, disturbances are assumed to be uncorrelated with each other. In other words, the covariance matrix of structural disturbances Σ_e is diagonal.

The model described by equation (2) can not be identified. Therefore, first the matrix B is restricted to $(k \times k)$ diagonal matrix. As a result, diagonal elements of matrix B represent estimated standard deviations of the structural shocks.

The next step is to impose restrictions to matrix A_0. The matrix A_0 is lower triangular matrix with additional restrictions that are obtained from statistical properties of the data. The starting point of the process is computation of partial correlations among the variables in the model and subsequent elimination of statistically insignificant relations. This methodology was used by Fragetta and Melina [5] who analyzed influence of fiscal policy shock in the US economy. This methodology is based on well defined statistical rules and allows to derive many possible structural VAR models that are later evaluated by statistical information criteria.

2.3 Credit channel

The credit channel can take many forms. It theoretically describes how monetary policy can affect the amount of credit that banks provide to firms and consumers which in turn affects the real part of economy. Our idea (inspirated by Lyziak et Al. [10]) of the monetary transmission mechanism is captured in Figure 1.

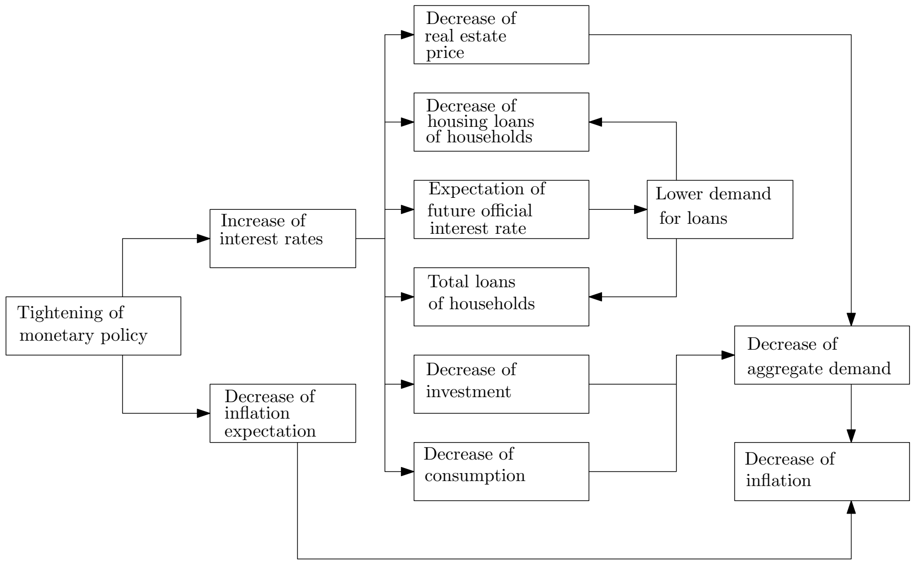


Figure 1 Monetary transmission mechanism

Monetary policy restriction represented by rise in short-term interest rates is followed by increase of lending and deposit rates in commercial and mortgage banks. If we assume price rigidities it causes rise in the real interest rates which affects consumption, savings and investment decisions. Increase in interest rates has an impact on a behaviour of households whose willingness to borrow declines. This leads to a drop in a growth of total household debt. This is a type of credit channel that affects the economy by altering the amount of credit of households and firms. The change in households behaviour may affect the real estate market because lower demand for loans also means lower demand for mortgage loans. It negatively influences demand for real estates and causes a decline of the real estate prices growth.

To find an evidence of monetary transmission mechanism and credit channel in the Czech housing finance system we are inspired by work of Iacoviello and Minetti [7] and use two structural VAR models. First SVAR model includes six variables: GDP growth rate, CPI inflation, money market interest rate, growth rate of real house price index, growth rate of households housing loans and growth rate of households total loans. This VAR model (with vector $Y_t = [\Delta Y, \pi, R, \Delta HP, \Delta HL, \Delta TL]$) is denoted by Iacoviello and Minetti [7] as uninformative model for detecting a credit channel. The model is based on the idea that increase in money market interest rate leads to a fall in loan demand implying a reduction in households loans. The idea is consistent with the traditional monetary transmission mechanism.

The second VAR includes: GDP growth rate, CPI inflation, money market interest rate, growth rate of real house prices and the spread between a mortgage interest rate and a money market interest rate. Iacoviello and Minetti [7] assume that this VAR model (with vector $Y_t = [\Delta Y, \pi, R, \Delta HP, SP]$) should capture the increase in the external finance premium associated with a credit channel.

3 Results

3.1 Identifications of SVAR models

We identify the monetary policy shocks in both types of VAR models following Iacoviello and Minetti [7] with algorithm proposed by Fragetta and Melina [5]. First, the VAR models described in equation 1 are estimated by OLS method. For both VAR models the lag length of four quarters was determined according to Akaike and Hannan-Quinn information criteria. Then vector of residual u_t for each variable is acquired and partial correlations among the series of innovations are computed. These partial correlations are quoted in Table 2.

VAR1

	u_t^y	u_t^π	u_t^R	u_t^{HP}	u_t^{HL}	u_t^{TL}
u_t^y	1					
u_t^π	-0,11	1				
u_t^R	-0,04	0,63	1			
u_t^{HP}	0,04	-0,32	-0,27	1		
u_t^{HL}	-0,20	0,39	0,45	0,11	1	
u_t^{TL}	-0,40	0,35	0,43	0,14	0,96	1

VAR2

	u_t^y	u_t^π	u_t^R	u_t^{HP}	u_t^{SP}
u_t^y	1				
u_t^π	-0,13	1			
u_t^R	0,14	0,18	1		
u_t^{HP}	0,51	0,09	0.34	1	
u_t^{SP}	-0,07	-0,24	-0,95	-0,34	1

Table 2 Estimated partial correlations of the series innovations

A_0 **of SVAR1**

$$\begin{pmatrix} 1 & & & & & \\ -a_{21} & 1 & & & & \\ 0 & -a_{32} & 1 & & & \\ 0 & -a_{42} & -a_{43} & 1 & & \\ -a_{51} & -a_{52} & -a_{53} & -a_{54} & 1 & \\ -a_{61} & -a_{62} & -a_{63} & -a_{64} & -a_{65} & 1 \end{pmatrix}$$

A_0 **of SVAR2**

$$\begin{pmatrix} 1 & & & & \\ -a_{21} & 1 & & & \\ -a_{31} & -a_{32} & 1 & & \\ -a_{41} & 0 & -a_{43} & 1 & \\ 0 & -a_{52} & -a_{53} & -a_{54} & 1 \end{pmatrix}$$

Table 3 Identification of matrix A_0 is SVAR models

The next step is to determine the SVAR model with AB specifications. Many authors use lower triangular structure of matrix A_0 based on a Cholesky decomposition. In this paper we use a data oriented method (see Fragetta and Melina [5]) based on the analysis of partial correlations among the variable's residuals. In other words, if partial correlation (in Table 2) is statistically insignificant[1], we assume that a matrix element corresponding to this correlation is equal to zero. The resulting matrix is shown in Table 3.

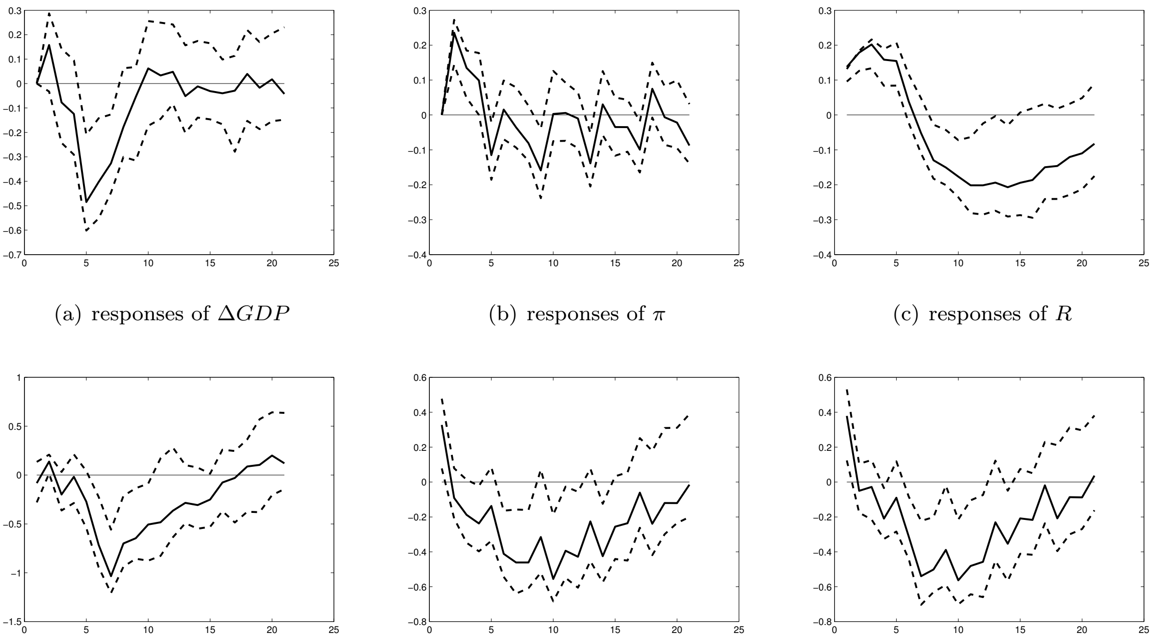
3.2 Impulse responses function SVAR1

For evidence of monetary policy transmission mechanism we simulate reaction to identified monetary policy shock (increase in nominal interest rate by one standard deviation). Figure 2 shows the responses of variables to a monetary contraction, along with percentile confidence interval calculated according to Hall [6]. The impulse responses show that after the shock the GDP growth at first increases but in the second quarter decreases and reaches the minimum in the fifth quarter. Inflation also increases at impact but then slows down and after four periods becomes negative. Slightly similar reaction was observed by Iacoviello and Minetti [7] in economy of Germany, Norway and Finland or by Lyziak et Al. [10] for Poland or by Ciccarelli, Peydro and Maddaloni [2] for euro area. The initial increase of inflation is usually interpreted as prize puzzle problem. A monetary contraction leads to a significant decline in housing and total bank loans. From the figure it is clear, that tight monetary policy has similar effects on housing and total loans of households. The change of household's behaviour (through a change in demand for real estate) affects the real estate market. Real house prices significantly react with the expected negative sign although with some lag. The though of decline is in seventh quarter, almost two years after the monetary restriction. Relatively delayed responses of this variable to monetary policy shock may be caused by: (i) still evolving real estate market and (ii) boom and bust in the Czech economy during estimated period.

3.3 Impulse responses function SVAR2

Figure 3 shows the impulse responses of the GDP growth, inflation, growth rate of real house prices and spread between mortgage rate on new housing loans and interbank interest rate to monetary policy shock.

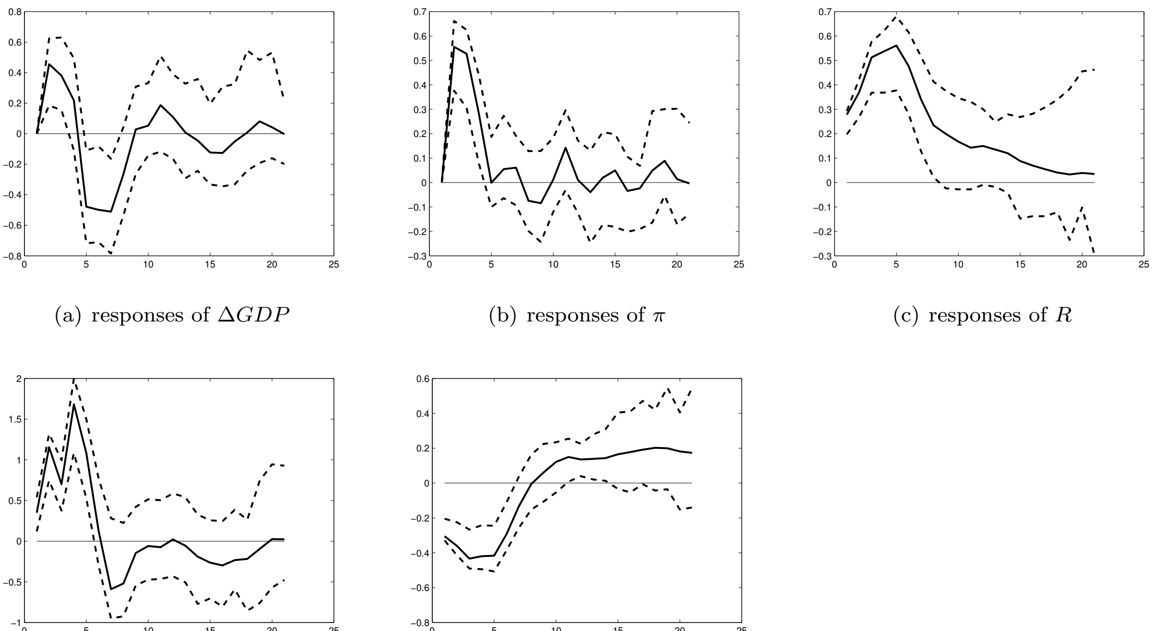
[1] If it is lower than ± 0.11.



(a) responses of ΔGDP (b) responses of π (c) responses of R

(d) responses of ΔHP (e) responses of ΔHL (f) responses of ΔTL

Figure 2 SVAR type 1 - Impulse responses functions to monetary restriction shock



(a) responses of ΔGDP (b) responses of π (c) responses of R

(d) responses of ΔHP (e) responses of SP

Figure 3 SVAR type 2 - Impulse responses functions to monetary restriction shock

Similarly as in the first model we can see firstly positive and after four quarters negative and statistically significant reaction of GDP growth. Along the same line react also inflation and real house price growth. The initial increase in both variables can be caused by time that economy needs to adapt to increase in the interest rate, or by some type of prize puzzle effect.

The spread SP decreases significantly and after five quarters it starts to increase. The initial decline may be caused by long-term fixation of mortgage rate (see Iacoviello and Minetti [7] p. 8). The increase after five quarters can be interpreted as an adaptation of mortgage interest rate. In other words, the mortgage banks adjust the mortgage rate to higher interbank interest rate. The significant increase of spread after the monetary contraction points at presence of a broad credit channel.

4 Conclusion

In this paper we tested the importance of credit channel in the housing sector of the Czech economy in period 2006:Q1 – 2015:Q4. We used two structural VAR models for identification of monetary policy shocks and studied reaction of macroeconomic and housing sector variables to increase in interest rate. The impulse responses showed that monetary policy can significantly affect the GDP growth, inflation, households loans, real house prices and the spread between mortgage interest rate and benchmark interest rate. Presence of the credit channel in the Czech housing market thus demonstrates that monetary policy can substantially influence important part of the economy. Existence of such mechanism should be taken into consideration for formation of monetary policy and building economic models.

Acknowledgements

This paper was supported by the specific research project no. MUNI/A/1049/2015 at Masaryk University.

References

[1] Bernanke, B., and Gertler, M.: Inside the black box: The credit channel of monetary transmission. *Journal of Economic Perspectives*, 1995.

[2] Ciccarelli, M., Peydro, J.L. and Maddaloni, A.: Trusting the bankers: a new look at the credit channel of monetary policy. *Working Paper Series 1228*, European Central Bank, 2010.

[3] Czech National Bank.: *CNB ARAD*, 2016. http://www.cnb.cz/

[4] Czech Statistical Office.: *CZSO*, 2016. http://www.czso.cz/

[5] Fragetta, M. and Melina.G.: The Effects of Fiscal Shocks in SVAR Models: A Graphical Modelling Approach, *Birkbeck Working Papers in Economics and Finance 1006*. Department of Economics, Mathematics and Statistics, Birkbeck, 2010.

[6] Hall, P.: *The Bootstrap and Edgeworth Expansion*. Springer, New York, 1992.

[7] Iacoviello, M. and Minetti, R.: The Credit Channel of Monetary Policy: Evidence from the Housing Market, *Journal of Macroeconomics* 2008.

[8] Lutkepohl, H.: *New introduction to multiple time series analysis*. Springer, Berlin, 2006.

[9] Lutkepohl, H. and Kratzig M.: *Applied time series econometrics*. Cambridge University Press, Cambridge, 2004.

[10] Lyziak T., Demchuk O., Przystupa J., Sznajderska A. and Wrobel E.: Monetary policy transmission mechanism in Poland. What do we know in 2011?. *National Bank of Poland Working Papers 116*, National Bank of Poland, Economic Institute, 2012.

Computing Fuzzy Variances of Evaluations of Alternatives in Fuzzy Decision Matrices

Pavla Rotterová[1], Ondřej Pavlačka[2]

Abstract. Decision matrices are a common tool for solving decision-making problems under risk. Elements of the matrix express evaluations if a decision-maker chooses the particular alternative and the particular state of the world occurs. In practice, the states of the world as well as the decision-maker's evaluations are often described only vaguely.

In the literature, there are considered fuzzy decision matrices, where alternatives are compared on the basis of the fuzzy expected values and fuzzy variances of their evaluations. However, the fuzzy variances of alternatives are calculated without considering dependencies between fuzzy evaluations under particular fuzzy states of the world and the fuzzy expected evaluation. So the final ranking of alternatives can be wrong. In the paper, we propose a proper way of how should the fuzzy variances of evaluations of alternatives be obtained. The problem is illustrated by the example of stocks comparisons.

Keywords: decision matrices, decision-making under risk, fuzzy states of the world, fuzzy expected values, fuzzy variances.

JEL classification: C44
AMS classification: 90B50

1 Introduction

Decision matrices are one of decision-making models to ordering alternatives. Elements of the matrices express evaluations of alternatives if the decision-maker chooses the particular alternative and the particular state of the world occurs.

In decision-making models, there are usually considered crisp (precisely defined) states of the world like e.g. "economy will grow more than 5 % next year", and crisp decision-maker's evaluations like e.g. 0. Crisp decision matrices are considered e.g. in [1], [2], [6] and [8]. However, in practice, states of the world as well as the decision maker's evaluations are sometimes described vaguely, like e.g. "economy will grow next year" or "approximately zero", respectively. Fuzzy decision matrices are considered e.g. in [7].

Talašová and Pavlačka [7] proposed the ranking of alternatives in a fuzzy decision matrix based on fuzzy expected evaluations of the alternatives and fuzzy variances of the alternatives evaluations. However, they suggested an approach to the computation of fuzzy variances without considering the dependencies between fuzzy evaluations under particular fuzzy states of the world and the fuzzy expected evaluation. Finally, Talašová and Pavlačka [7] selected the best alternative according to a decision-making rule of the expected value and the variance. Based on this rule, a decision-maker selects the alternative that maximises his/her expected evaluation of the alternative and minimises the variance of his/her evaluation of the alternative.

The aim of this paper is to propose the proper way of the fuzzy variance computation. The paper is organised as follows. Section 2 is devoted to the extension of a crisp probability space to the case of fuzzy events. In section 3, a decision matrix with fuzzy states of the world and with fuzzy evaluations under particular states of the world is described. The approach of the fuzzy variance computation proposed

[1]Palacký University Olomouc, Faculty of Science, Department of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic, pavla.rotterova01@upol.cz
[2]Palacký University Olomouc, Faculty of Science, Department of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/12, 771 46 Olomouc, Czech Republic, ondrej.pavlacka@upol.cz

in [7] is also described here, and the proper way of the fuzzy variance computation is introduced . The differences of these two approaches are illustrated in the example in section 4. Some concluding remarks are given in Conclusion.

2 Fuzzy events and their probabilities

Let *a probability space* $(\Omega, \mathcal{A}, P)$ be given, where Ω is a non-empty universal set of all elementary events, $\mathcal{A}$ denotes a σ-algebra of subsets of Ω, i.e. $\mathcal{A}$ is a set of all considered random events, and P denotes a probability measure. In this section, the σ-algebra $\mathcal{A}$ is extended to the case of fuzzy events and crisp probabilities of fuzzy events are defined.

A *fuzzy set* A on a non-empty set Ω is determined by its membership function $\mu_A : \Omega \to [0, 1]$. $\mathcal{F}(\Omega)$ denotes the family of all fuzzy sets on Ω. Let us note that any crisp set $A \subseteq \Omega$ can be viewed as a fuzzy set of a special kind; the membership function μ_A coincides in such a case with the characteristic function χ_A of A.

In accordance with Zadeh [9], a *fuzzy event* $A \in \mathcal{F}(\Omega)$ is a fuzzy set whose α-cuts $A_\alpha := \{\omega \in \Omega \mid \mu_A(\omega) \geq \alpha\}$ are random events, i.e. $A_\alpha \in \mathcal{A}$ for all $\alpha \in [0, 1]$. The family of all such fuzzy events, denoted by $\mathcal{A}_F$, forms a σ-algebra of fuzzy events, which was shown in [4]. Let us note that in a decision matrix described further, fuzzy events represent fuzzy states of the world.

Zadeh [9] extended the probability measure P to the case of fuzzy events. This extended measure is denoted by P_Z. He defined a probability $P_Z(A)$ of a fuzzy event A as follows:

$$P_Z(A) := E(\mu_A) = \int_\Omega \mu_A(\omega)\, dP. \tag{1}$$

The mapping P_Z possesses the important properties of a probability measure, see e.g. [5] and [9]. However, there is a problem with interpretation of the probability measure P_Z as was discussed in [5].

3 Fuzzy variance calculations in a fuzzy decision matrix

In this section, we introduce fuzzification of the decision matrix. Then, we recall the formulas for calculations of the fuzzy expected value and the fuzzy variance of the decision-maker's evaluations proposed by Talašová and Pavlačka [7]. Finally, we show the proper way of the computation of the fuzzy variance.

A vaguely defined quantity can be described by a fuzzy number. A *fuzzy number* A is a fuzzy set on the set of all real numbers $\mathbb{R}$ such that its core $Core\, A := \{\omega \in \Omega \mid \mu_A(\omega) = 1\}$ is non-empty, whose α-cuts are closed intervals for any $\alpha \in (0, 1]$, and whose support $Supp\, A := \{\omega \in \Omega \mid \mu_A(\omega) > 0\}$ is bounded. The family of all fuzzy numbers on $\mathbb{R}$ is denoted by $\mathcal{F}_N(\mathbb{R})$. In the paper, we consider an α-cut of any fuzzy number A expressed by $A_\alpha = [A^L_\alpha, A^U_\alpha]$ for any $\alpha \in (0, 1]$.

	S_1 $P_Z(S_1)$	S_2 $P_Z(S_2)$	$\cdots$ $\cdots$	S_m $P_Z(S_m)$	Characteristics	
x_1	$H_{1,1}$	$H_{1,2}$	$\cdots$	$H_{1,m}$	$(\mathcal{F})\, EH_1$	$(\mathcal{F})\, var\, H_1$
x_2	$H_{2,1}$	$H_{2,2}$	$\cdots$	$H_{2,m}$	$(\mathcal{F})\, EH_2$	$(\mathcal{F})\, var\, H_2$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_n	$H_{n,1}$	$H_{n,2}$	$\cdots$	$H_{n,m}$	$(\mathcal{F})\, EH_n$	$(\mathcal{F})\, var\, H_n$

Table 1 Fuzzy decision matrix

Now, let us introduce a fuzzy extension of the decision matrix, where states of the world and also evaluations under particular states of the world are modelled by fuzzy sets, given in Table 1. In the matrix, $x_1, x_2, \ldots, x_n$ represent alternatives of a decision-maker, $S_1, S_2, \ldots, S_m$ fuzzy states of the world that form a fuzzy partition of a non-empty universal set Ω, i.e. $\sum_{j=1}^m \mu_{S_j}(\omega) = 1$ for any $\omega \in \Omega$, $P_Z(S_1), P_Z(S_2), \ldots, P_Z(S_m)$ stand for the probabilities of the fuzzy states of the world $S_1, S_2, \ldots, S_m$ defined by (1), and for any $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$, $H_{i,j}$ means the fuzzy evaluation if the decision-maker chooses the alternative x_i and the fuzzy state of the world S_j occurs. Moreover, we assume that probability distribution on Ω is known. Further, we consider that fuzzy elements in the matrix are expressed by fuzzy numbers.

Let $(H_{i,j})_\alpha = [(H_{i,j})^L_\alpha, (H_{i,j})^U_\alpha]$ for any $\alpha \in (0,1]$. The fuzzy expected evaluations, denoted by $(\mathcal{F})\,EH_1, (\mathcal{F})\,EH_2, \ldots, (\mathcal{F})\,EH_n$, are calculated as

$$(\mathcal{F})EH_i = \sum_{j=1}^{m} P_Z(S_j) \cdot H_{i,j}, \tag{2}$$

i.e. for any $i \in \{1, 2, \ldots n\}$ and any $\alpha \in (0,1]$, the α-cut of the fuzzy expected evaluation $(\mathcal{F})\,EH_{i\alpha} = [EH^L_{i\alpha}, EH^U_{i\alpha}]$, is calculated as follows:

$$EH^L_{i\alpha} = \sum_{j=1}^{m} P_Z(S_j) \cdot (H_{i,j})^L_\alpha,$$
$$EH^U_{i\alpha} = \sum_{j=1}^{m} P_Z(S_j) \cdot (H_{i,j})^U_\alpha.$$

Talašová and Pavlačka [7] considered fuzzy variances denoted by $(\mathcal{F})\ var_t\ H_1$, $(\mathcal{F})\ var_t\ H_2$, $\ldots$, $(\mathcal{F}\ )var_t\ H_n$ of alternatives evaluations defined for any $i \in \{1, 2, \ldots n\}$ as

$$(\mathcal{F})\ var_t\ H_i = \sum_{j=1}^{m} P_Z(S_j) \cdot \left(H_{i,j} - (\mathcal{F})EH_i\right)^2. \tag{3}$$

For any $\alpha \in (0,1]$, the α-cut $(\mathcal{F})\ var_t\ H_{i\alpha} = [var_t\ H^L_{i\alpha}, var_t\ H^U_{i\alpha}]$ is computed as follows:

$$var_t\ H^L_{i\alpha} = \sum_{j=1}^{m} P_Z(S_j) \cdot e_j,$$

where

$$e_j = \begin{cases} 0 & \text{if} \quad 0 \in [(H_{i,j})^L_\alpha - EH^U_{i\alpha}, (H_{i,j})^U_\alpha - EH^L_{i\alpha}], \\ \min\left\{\left((H_{i,j})^{y_1}_\alpha - EH^{y_2}_{i\alpha}\right)^2 | y_1 = L, U, y_2 = L, U\right\} & \text{otherwise.} \end{cases}$$

$$var_t\ H^U_{i\alpha} = \sum_{j=1}^{m} P_Z(S_j) \cdot \max\left\{\left((H_{i,j})^{y_1}_\alpha - EH^{y_2}_{i\alpha}\right)^2 | y_1 = L, U, y_2 = L, U\right\}$$

However, this approach, based on the concept of constrained fuzzy arithmetic (see e.g. [3]), does not takes into account dependencies between fuzzy evaluations under particular fuzzy states of the world $H_{i,j}$ and the fuzzy expected value $(\mathcal{F})EH_i$. Let us propose a proper way of the calculation of the α-cut of the fuzzy variance denoted by $(\mathcal{F})\ var_s\ H_i$ for any $i \in \{1, 2, \ldots n\}$. Let us denote for any $\alpha \in (0,1]$

$$(\mathcal{F})\ var_s\ H_{i\alpha} = [var_s\ H^L_{i\alpha}, var_s\ H^U_{i\alpha}].$$

The lower and upper bounds of the $(\mathcal{F})\ var_s\ H_{i\alpha}$ are defined as follows:

$$var_s\ H^L_{i\alpha} = \min\left\{\sum_{j=1}^{m} P_Z(S_j) \cdot \left(h_{i,j} - \sum_{k=1}^{m} P_Z(S_k) \cdot h_{i,k}\right)^2 \mid h_{i,j} \in (H_{i,j})_\alpha, j = 1, 2, \ldots, m\right\}, \tag{4}$$

$$var_s\ H^U_{i\alpha} = \max\left\{\sum_{j=1}^{m} P_Z(S_j) \cdot \left(h_{i,j} - \sum_{k=1}^{m} P_Z(S_k) \cdot h_{i,k}\right)^2 \mid h_{i,j} \in (H_{i,j})_\alpha, j = 1, 2, \ldots, m\right\}. \tag{5}$$

In the formulas (4) and (5), the resultant fuzzy expected value is not employed in contrast to (3). Instead of this, the way how the fuzzy expected value is obtained, is taken into account. Thus, this approach takes into consideration the relationships among variables.

The resultant fuzzy variances obtained by the formulas (4) and (5), are subsets of fuzzy variances obtained by formula (3). So for each $\alpha \in (0,1]$, $(\mathcal{F})\ var_s\ H_{i\alpha} \subseteq (\mathcal{F})\ var_t\ H_{i\alpha}$. Using the formulas (4) and (5), the uncertainty of the fuzzy variance is not falsely increased as in the case of the variance computation by the formula (3). Moreover, alternatives can be differently ordered if fuzzy variances are calculated according to the formulas (4) and (5) instead of (3). This is illustrated by the example in the following section.

4 Example of the fuzzy decision matrix of stock comparisons

Let us consider a problem of comparing two stocks, A and B, with respect to their future yields. The considered states of the world are great economic drop, economic drop, economic stagnation, economic growth and great economic growth denoted by GD, D, S, G, and GG, respectively. They are expressed by a trapezoidal fuzzy numbers that form a fuzzy scale shown in Figure 1. A *trapezoidal fuzzy number* is determined by its *significant values* a_1, a_2, a_3, and a_4, where $a_1 \leq a_2 \leq a_3 \leq a_4$. The membership function of any trapezoidal fuzzy number $A \in \mathcal{F}_N(\mathbb{R})$ is defined for any $x \in \mathbb{R}$ as follows:

$$\mu_A(x) = \begin{cases} \frac{x-a_1}{a_2-a_1} & \text{if } x \in [a_1, a_2), \\ 1 & \text{if } x \in [a_2, a_3], \\ \frac{a_4-x}{a_4-a_3} & \text{if } x \in (a_3, a_4], \\ 0 & \text{otherwise.} \end{cases}$$

A trapezoidal fuzzy number A determined by its significant values is denoted by (a_1, a_2, a_3, a_4).

Let us assume that the economy states are given by the development of the gross domestic product, abbreviated as GDP. We also assume that next year prediction of GDP development shows a normally distributed growth of GDP with parameters $\mu = 1.5$ and $\sigma = 2$.

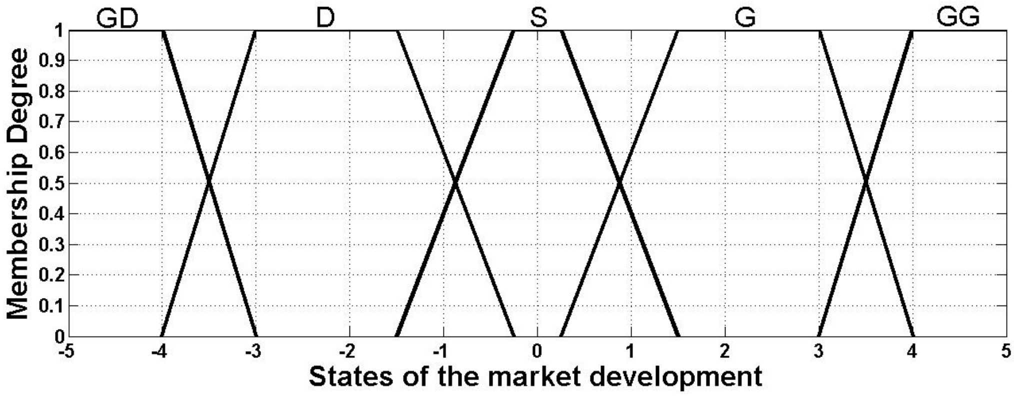


Figure 1 Fuzzy scale of the economy states

The resultant fuzzy expected values and fuzzy variances can be compared based on their centres of gravity. The *centre of gravity* of a fuzzy number $A \in \mathcal{F}_N(\mathbb{R})$ is a real number cog_A given as follows:

$$cog_A = \frac{\int_{-\infty}^{\infty} \mu_A(x) \cdot x \, dx}{\int_{-\infty}^{\infty} \mu_A(x) \, dx}.$$

Economy states	GD = (-∞, -∞, -4, -3)				D = (-4, -3, -1.5, -0.25)				S = (-1.5, -0.25, 0.25, 1.5)			
Probabilities	0.0067				0.1146				0.2579			
A yield (%)	-48	-37	-31	-20	-24	-17	-10	-3	-5	-3	3	8
B yield (%)	-46	-38	-32	-24	-22	-17	-11	-6	-4	-3	3	4
Economy states	G = (0.25, 1.5, 3, 4)				GG = (3, 4, ∞, ∞)							
Probabilities	0.4596				0.1612							
A yield (%)	6	12	17	24	20	27	34	41				
B yield (%)	6	12	16	22	24	29	35	40				

Table 2 Fuzzy decision matrix of stocks yields

Stock Characteristic	Fuzzy Stock Yield (%)				Centre of Gravity
$(\mathcal{F})\ EA$	1.62	6.90	12.71	19.22	10.17
$(\mathcal{F})\ EB$	2.72	7.49	12.45	17.22	9.75
$(\mathcal{F})\ var_t\ A$	5.66	79.11	346.67	857.79	287.76
$(\mathcal{F})\ var_t\ B$	22.07	97.84	335.49	702.75	267.34
$(\mathcal{F})\ var_s\ A$	41.87	117.61	257.13	447.2	210.57
$(\mathcal{F})\ var_s\ B$	74.55	134.71	259.79	392.80	213.26

Table 3 Fuzzy expected values and fuzzy standard deviations

The fuzzy expected values $(\mathcal{F})\ EA$ and $(\mathcal{F})\ EB$, computed by formula (2), are trapezoidal fuzzy numbers. Their significant values are given in Table 3. Fuzzy variances $(\mathcal{F})\ var_t\ A$ and $(\mathcal{F})\ var_t\ B$, obtained by formula (3), and $(\mathcal{F})\ var_s\ A$ and $(\mathcal{F})\ var_s\ B$, calculated by (4) and (5) , are not trapezoidal fuzzy numbers. Their membership functions are shown in Figure 2. We can see that properly computed fuzzy variances are significantly narrower than variances calculated by the original formula. In this example, we can also see that the proper calculation of the fuzzy variance can cause a change in the decision-maker's preferences, i.e. based on the results calculated by the formula (3), the decision-maker is not able to make a decision on the basis of the rule of the expected value and the variance, while based on $(\mathcal{F})\ var_s\ A$ and $(\mathcal{F})\ var_s\ B$, the decision-maker should select the stock A (the higher expected value and the lower variance than the stock B). The expected values and the variances are compared based on their centres of gravity.

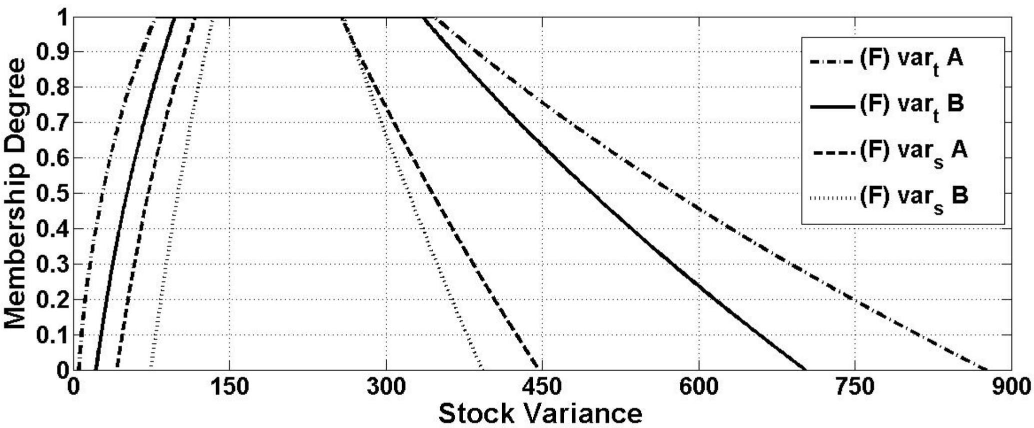


Figure 2 Membership functions of fuzzy variances $(\mathcal{F})\ var_t\ A$, $(\mathcal{F})\ var_t\ B$, $(\mathcal{F})\ var_s\ A$, $(\mathcal{F})\ var_s\ B$

5 Conclusion

We have proposed the proper way of the fuzzy variance computation in the fuzzy decision matrix. The original way of the fuzzy variance computation does not consider dependencies between fuzzy evaluations under particular fuzzy states of the world and the fuzzy expected evaluation. Contrary, the proper one takes the dependencies into account, which means that obtained fuzzy variances do not falsely increase the uncertainty of this variance. In the numerical example, we have shown, that considering dependencies between fuzzy evaluations under particular fuzzy states of the world and the fuzzy expected evaluation can affect final ranking of the alternatives.

Acknowledgements

Supported by the grant IGA_PrF_2016_025 *Mathematical Models* of the *Internal Grant Agency of Palacký University Olomouc* and by the project No. GA 14-02424S of the Grant Agency of the Czech Republic.

References

[1] Ganoulis, J.: *Engineering Risk Analysis of Water Pollution: Probabilities and Fuzzy Sets.* John Wiley & Sons, 2008.

[2] Huynh, V. et. al.: A Fuzzy Target Based Model for Decision Making Under Uncertainty. *Fuzzy Optimization and Decision Making* **6**, 3 (2007), 255–278.

[3] Klir,G. J., Pan, Y.: Constrained fuzzy arithmetic: Basic questions and some answers. *Soft Computing* **2** (1998), 100–108.

[4] Negoita, C.V., Ralescu, D.: *Applications of Fuzzy Sets to System Analysis.* Birkhuser Verlag–Editura Technica, Stuttgart, 1975.

[5] Pavlačka, O., Rotterová, P.: Probability of fuzzy events. In: *Proceedings of the 32nd International Conference Mathematical Methods in Economics* (J. Talašová, J. Stoklasa, T. Talášek, eds.), Palacký University Olomouc, Olomouc, 2014, 760–765.

[6] Özkan,I., Türken,IB.: Uncertainty and fuzzy decisions. In: *Chaos Theory in Politics*, Springer Netherlands, 2014, 17–27.

[7] Talašová, J., Pavlačka, O.: Fuzzy Probability Spaces and Their Applications in Decision Making. *Austrian Journal of Statistics* **35**, 2&3 (2006), 347–356.

[8] Yoon, K.P., Hwang, Ch.: *Multiple Attribute Decision Making: An Introduction.* SAGE Publications, California, 1995.

[9] Zadeh, L.A.: Probability Measures of Fuzzy Events. *Journal of Mathematical Analysis and Applications* **23**, 2 (1968), 421–427.

The main inflationary factors in the Visegrád Four

Lenka Roubalová[1], David Hampel[2]

Abstract. This paper deals with the analysis of the main inflationary factors and their changes caused by the onset of the economic crisis in 2008. We attempt to uncover the factors causing inflation in the Visegrád Four from both demand and supply sides in the pre-crisis period as well as in the following one. To determine main inflation drivers in each country, multiple regression models are estimated. To describe the common factors causing inflation among all member countries we use panel data models. We found out that the most significant inflationary factors have been changed after 2008 but these factors do not differ significantly in observed countries. The results prove the influence of the domestic factors in the pre-crisis period. The impact of the external factors, especially commodity prices, is significant in the period during crisis. In conclusion, supply side factors have major influence on inflation and the inflationary factors have been changed over time.

Keywords: demand and supply side inflation factors, domestic and external inflation factors, inflation, multiple regression models, panel data models, Visegrád Four.

JEL classification: E31
AMS classification: 91B84

1 Introduction

Inflation has been a common theme and subject of expert analysis for years. The economic structure has been changing at both, global and national level and so it is interesting to examine factors influencing inflation in different countries and compare the results obtained among different periods. The knowledge about main inflation drivers can be useful for individuals and businesses to anticipate further economic development. Especially, it is important for monetary authorities in countries where inflation targeting regime is implemented.

As Stock and Watson reported in [16], there have been significant changes in the dynamic of inflation causing changes in the economic structure for the past five decades. Nowadays, for instance, energy is not so important part of the total expenditure in compare with the period of oil shocks. The labour union become smaller and the production of services increases. This development is associated with the monetary policy changes as well. The rules are more flexible and take into account several factors like inflation expectations, for example.

There are not many publications dealing with inflationary factors in countries of Central Europe. We can mention the results reported by Golinelli and Orsi in [6] who proved the inflationary impact of the production gap and exchange rate in V4 countries during the transformation process. Further study deals with the inflation drivers development from the transformation process to the financial crisis in 2008, see [1], and points out the importance of cost-pushed inflation. In [2] Time-Varying Parameter VAR model with stochastic volatility and VAR Neural Network model are used for predicting inflation based on selected inflationary factors.

Financial crisis can break the economic structure significantly and affect the efficiency of monetary policy, see for example [4], [9] and [14] and monetary authorities can be forced to accommodate these changes. The crisis can affect the behavior of individuals and businesses which react more sensitively to

[1]Mendel University in Brno, Department of Statistics and Operation Analysis, Zemědělská 1, 613 00 Brno, Czech Republic, xkravack@mendelu.cz

[2]Mendel University in Brno, Department of Statistics and Operation Analysis, Zemědělská 1, 613 00 Brno, Czech Republic, qqhampel@mendelu.cz

some economic changes during crisis. BVAR model with the specifically defined prior is used for modeling monetary transmission mechanism in the Czech Republic and Slovakia from January 2002 to February 2013 and the results point to the inadequacy of the common VAR model, see [17]. In [3] asymmetries in effects of monetary policy and their link to labour market are investigated. Time-Varying Parameter VAR model with stochastic volatility is employed there to capture possible changes in transmission mechanism and in underlying structure of economy over time.

The main aim of this paper is to find out the main inflationary factors in V4 countries in the period before the onset of the crisis in 2008 and in the period during crisis. We attempt to verify the thesis that inflationary factors have been changed over time. This analysis should also provide some information about the impact of external and domestic factors or supply and demand side factors among V4 countries.

2 Material

The selection of macroeconomic variables used in the econometric analysis is based on macroeconomic theory and on the information published in inflation reports of monetary authorities. A dependent variable is an inflation rate measured by CPI changes compared to the same period of the previous year. For Slovak Republic estimates we use HICP changes, which do not differ substantially but they are suitable for comparing and discussion to NBS publications. As the independent variables we use 23 macroeconomic indicators which represent two groups of variables, the external and the domestic factors.

The group of domestic factors include GDP of each country which represents aggregate demand and household consumption which represent consumer demand. There are many notices about total aggregate demand and consumer demand in the inflation report, so both variables are included. Then we use price factors like food prices, energy prices, prices of fuels, unprocessed food prices and industrial production prices. The labour market is represented by the level of unemployment, nominal wages, labour productivity per worker and other labour costs. Then we use the government debt to GDP rate, monetary aggregates, three month interbank offered rates, 3M PRIBOR, 3M BUBOR, 3M WIBOR a 3M (BRIBOR) EURIBOR and inflation expectations.

The external factors represent a group of variables containing exchange rate of domestic currency to EUR, nominal effective exchange rate, current account balance and import prices. We also use variables representing global development like oil prices, European industrial production prices and average global agricultural prodcution prices for EU and USA. Data of macroeconomic variables – quarterly time series – are expressed as an index, change against the same quarter from the last year. This form of expression is advantageous because it eliminates the effect of seasonality. The number of time series observations $T = 51$ corresponds to the period from the first quarter of 2003 to the second quarter of 2015, while Q1 2003 – Q3 2008 represents the period before the crisis and Q4 2008 – Q2 2015 represents the period during the crisis.

Analyzed data were obtained from ARAD – database managed by the Czech National Bank, macroeconomic database of the National Bank of Slovakia, database of the Polish Statistical Office, database of the National Bank of Poland and from Eurostat database. Time series of Brent oil prices and global agricultural production prices were obtained from the OECD database and from the database of the European Central Bank.

3 Methods

The first part of this paper deals with creating multivariate regression models for each country of the Visegrád Group separately. We start with testing stationarity of the aforementioned time-series: approximately using the graph of time series and exactly by unit root tests, specifically we use Augmented Dickey-Fuller (ADF) test [5] and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test [12]. The consequence of modeled time series nonstationarity can be spurious regression. Therefore, it is necessary to test whether examined time series are cointegrated. This task we perform by verification of residues of cointegration regression stationarity.

Before carrying out the parameters estimate we verify the presence of multicollinearity using pairwise correlation coefficients and Variance Inflation Factors; variables causing multicollinearity are eliminated. Further we apply backward elimination of variables according to their significance given by t-test. When selecting the best model from several candidate models we compare adjusted coefficient of determination and information criteria (AIC, BIC, HQC). In particular, we look to whether the parameter estimates are consistent with economic theory. We also provide specification tests (LM test and Ramsey RESET test), serial correlation test (significance of residual autocorrelations, Durbin-Watson and Ljung-Box test), heteroskedasticity tests (Breusch-Pagan and White test). Normality of the error term we verify graphically using a histogram and Q-Q plot; for exact testing we use the Shapiro-Wilk test.

Inflationary factors common to all V4 countries can be detected by the analysis of panel data representation. We deal with non-stationary time series, therefore we test cointegration of panel representation of data by the Kao test [11]. Generally, it is possible to choose model with fixed or random effects of objects. This decision is based on Hausman test [7] usually. In our case, when number of objects is lower than number of variables, we cannot use model with random effects, so only models with fixed effects are estimated. Tests of different intercepts for particular objects are conducted. Statistical tests are provided at the significance level $\alpha = 5$ %. Estimates of multivariate regression models and panel data models are obtained using software Gretl 1.9.91, EViews 8 and Matlab R2015b.

4 Results

To estimate the models we used quarterly time series. Although some of them are non-stationary, we found this series cointegrated. Almost all estimated parameters are statistically significant, see Tab. 1. The models are well specified and all classical assumptions were fulfilled. With the respect to the fact that predicting inflation is very difficult, the coefficients of determination between 73 % and 96 % indicate that following variables describe the variability of inflation rate very well. Estimated inflation is illustrated graphically in Fig. 1, dashed line.

	Before crisis period					**During crisis period**				
Variable	**CZ**	**HU**	**PL**	**SK**	**V4**	**CZ**	**HU**	**PL**	**SK**	**V4**
3M interbank rate	×	×	×	<0.001	×	×	<0.001	×	×	×
Agricultural prices	0.001	<0.001	×	×	×	×	<0.001	<0.001	<0.001	<0.001
Consumption	×	×	0.004	×	×	×	×	×	0.050	×
Energy prices	×	<0.001	×	<0.001	<0.001	<0.001	×	×	×	<0.001
Food & energy prices	×	×	0.002	×	×	×	×	×	×	×
Food prices	×	×	×	<0.001	<0.001	0.003	×	×	×	×
GDP	0.004	×	×	0.008	0.003	×	×	×	×	×
GDP EA	×	0.099	×	×	×	×	×	0.068	×	×
Inflation expectations	×	<0.001	0.019	×	×	<0.001	×	×	×	<0.001
Monetary base	×	×	×	×	×	<0.001	×	<0.001	<0.001	×
Nominal wages	0.065	×	×	×	×	×	×	×	×	×
Oil prices	<0.001	×	×	×	×	×	<0.001	<0.001	<0.001	0.043
Unproc. food prices	×	×	×	×	×	×	0.028	<0.001	×	×

Table 1 Variables included in models for particular countries (CZ, HU, PL, SK) and in panel data model (V4) in both pre-crisis and crisis periods. P-values of the t-test are given, the symbol '×' means that insignificant variable is omitted in concrete model.

For the Czech Republic, the positive signs of all numerical estimates determine positive influence of these variables on inflation. In the case of Hungary we found all variables affecting inflation positively. The negative impact of the interest rate is related to the loans denominated in euro during this period. For Poland, the negative impact of the Eurozone aggregate demand results from the declining aggregate demand in Eurozone that caused increasing demand for food production form Poland. It caused, that Polish export increased and the price level increased as well. All numerical estimates in model for Slovak Republic have positive signs which indicate that these variables affect inflation rate positively.

The panel data analysis results indicate factors that are the most convenient to describe inflation development during both periods in the Visegrád Four countries collectively. The most appropriate

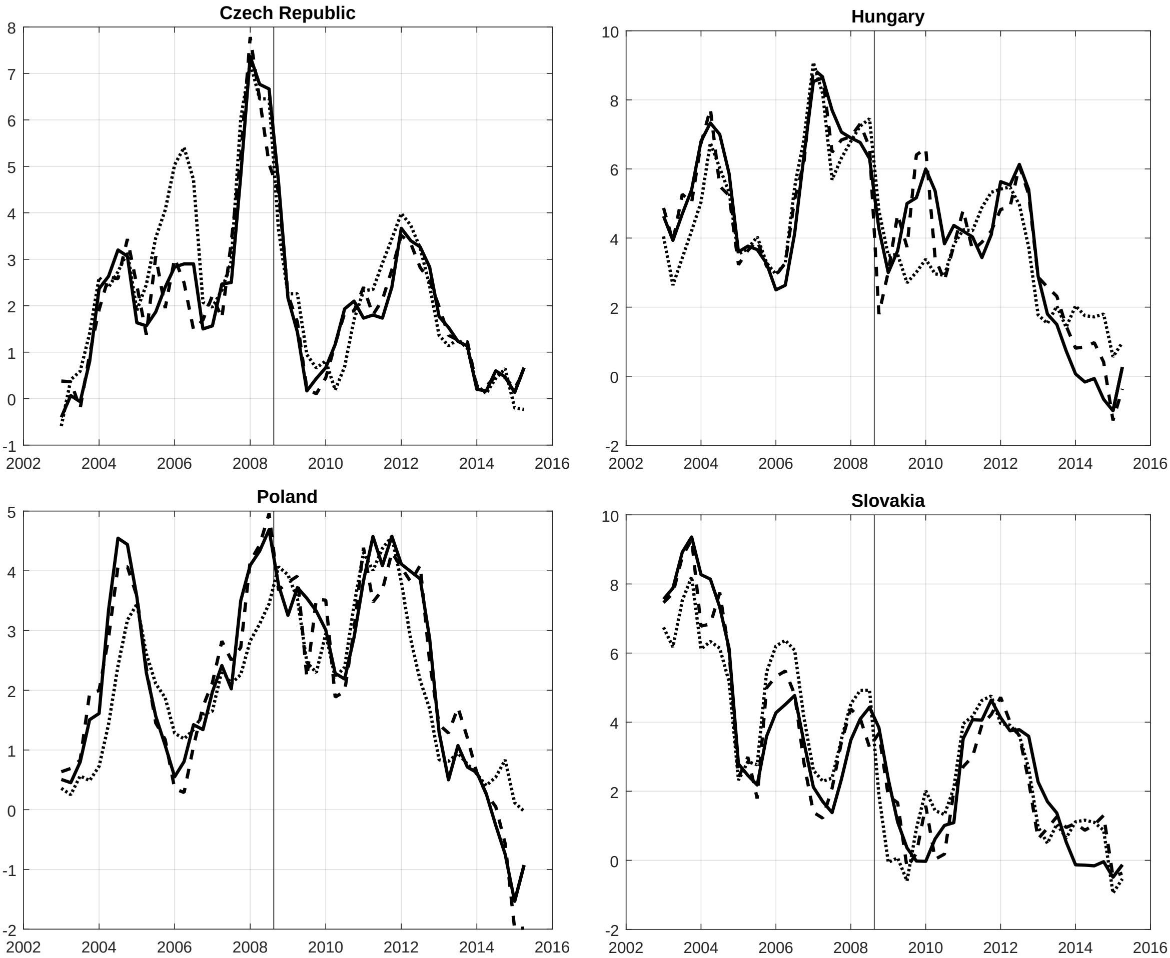


Figure 1 Real inflation (solid line), multivariate regression estimate (dashed) and panel Fixed Effects model estimate (dotted) for V4 countries. Y-axis inflation rate, x-axis time. Vertical line separates periods before and during crisis.

resulting model with $R^2_{adj} = 0.832$ in the pre-crisis period is the fixed effect model containing regressors like energy prices, food prices and domestic aggregate demand. With the use of Kao test of panel cointegration with $p = 0.0029$ we can reject the null hypothesis of non-cointegration. We can reject the existence of common intercepts using the test with $p = 0.0021$, and we cannot reject homoskedasticity and normality of random errors. Estimated inflation for both pre-crisis and during-crisis periods is illustrated graphically in Fig. 1, dotted line.

The period during crisis can be collectively described with the fixed effects model with $R^2_{adj} = 0.691$ where the inflation rate is described with the use of regressors like global agricultural prices, crude oil prices, inflation expectations and energy prices. Kao test of panel cointegration with $p < 0.001$ proved we can reject the null hypothesis of non-cointegration. With the use of the test of common intercepts with $p < 0.001$ we can reject the null hypothesis about common intercept among groups of countries. We cannot reject the null hypothesis about normality of random errors, but we reject their homoskedasticity (we can see greater variance for Hungary and Slovakia than for Czech republic and Poland).

5 Discussion

Generally our results prove that inflation in V4 countries can be successfully described with food prices, energy prices and with the national aggregate demand development in the pre-crisis period. We have found the importance of domestic factors in this period. However, there are some differences among individual countries. It is worth pointing out the case of the Czech Republic where the impact of crude oil prices and global agricultural prices are detected. The changes in oil prices in 2002 affected the industrial prices development significantly through the changes in the import prices in 2005. This changes affected primary oil processing prices thereby the energy production and distribution prices increased. The changes in crude oil prices and in global agricultural prices finally rise domestic food prices. We should also mention the impact of nominal wages that contributed to this development of price level in the Czech Republic. The impact of labour costs in the Czech Republic results form the research of M. Alexová. She found the significant impact of this factor on inflation in the Czech Republic and Hungary in 1996–2011, see [1].

Eurozone aggregate demand is one of the main inflation drivers in Hungary. It can be connected with the rising output of the main trading partners in this period. According to MNB evidence the net export was the most important component contributing to the domestic aggregate demand increase at the end of this period[1]. The results indicate the importance of inflation expectations in this period individually in Poland and Hungary. It can be connected wit the accession to EU in this period. In the case of Hungary there is significant impact of the Eurozone demand. The economic growth in Eurozone probably affected inflation expectation in Hungary as well.

The results prove that the external factors prevailed during crisis. Generally in all countries the oil prices, global agricultural prices, inflation expectations and energy prices are the main inflation drivers during this period. Neely and Rapach found that international factors affect inflation of more than 50 % and the impact of this group of factors still has been increasing, see [15]. Multiple regression models point some detail information about inflation factors not typical for all V4 members. In the case of Poland, the results show negative impact of the Eurozone demand on inflation in Poland. Due to the loss of output and weak aggregate demand in Eurozone these countries increased the amount of imported food production from Poland[2]. In Hungary a total of more than two thirds of all household loans were denominated in Swiss Francs and Euros in 2010 [10]. However, the Forint[3]-denominated loan repayment increased with the depreciation of Forint. Secondarily, the depreciation contributed to an increase in inflation.

In all countries together we can see considerable impact of inflation expectations. These results are in accordance with Levy (1981), who found that the inflation expectations and the monetary base are significant inflation factors during crisis. Although there should not be significant effect of monetary base in the inflation targeting regime, we found the importance of money indicators in Poland and in the Czech Republic. Another research points that monetary base is not very important in predicting inflation in V4 countries, but in the case of individual countries, especially Poland, the results indicate that some money indicators improve the inflation forecast [8]. This conclusion confirms our results concerning the fact that monetary base can be important within individual countries but it does not play the role in general development of price level in the V4 countries together. In spite of the general summary which indicates that inflationary factors have been changed over time, it is necessary to mention energy prices.

Panel data models indicate the impact of domestic energy prices on inflation in both periods. There are various sources affecting this variable. Except for oil prices development there are many domestic influences like VAT tax, regulations and weather, which affect the consumption of energy. According to Czech National Bank[4] evidence we assume that in the pre-crisis period the oil prices development affected inflationary pressure of energy prices. Secondly, it could be strengthened by the extremely low temperature measured in Central Europe in winter 2005. During the crisis the impact of energy prices on inflation is lower and multiple regression models do not prove the importance of this variable in all V4 countries individually. Our results imply that the inflation development is mostly affected by factors on the supply side that cause cost-push inflation. The research of M. Alexová partly confirms our results but her results put more weight on labor costs, see [1].

6 Conclusions

We used multiple regression models and panel data models to identify the main inflationary factors in the Visegrád Four. Estimated models proved there are some specific characteristics in the inflation development in each country, but generally there is a group of factors common to all V4 countries in each period. Despite these individual differences we can conclude that the main inflationary factors do not

[1]Magyar Nemzeti Bank Quarterly Report on Inflation (November 2007).
[2]Economic newsletter from Poland 2014/2 of the Czech-Polish Chamber of Commerce.
[3]Magyar national currency.
[4]Czech National Bank Inflation Report July 2005.

differ significantly in observed countries. We found out that the impact of domestic factors prevail in the pre-crisis period. The impact of external factors and inflation expectations is significant during crisis. From this estimates we derive that impact of external factors increase and the cost-pushed inflation has prevailed. The inflation factors have been changed by the onset of economic crisis in 2008. The results imply that national banks are not able to drive all inflationary factors directly, because a significant part of total inflation development is based on external factors.

References

[1] Alexová, M.: Inflation Drivers in New EU Members, *NBS Working Paper* **6**, (2012), 1–42.

[2] Dobešová, A. and Hampel, D.: Predicting Inflation by the Main Inflationary Factors: Performance of TVP-VAR and VAR-NN Models. In: *Proceedings of the 33rd International Conference Mathematical Methods in Economics*, (2015), 133–138.

[3] Dobešová, A. and Hampel, D.: Asymmetries in Effects of Monetary Policy and Their Link to Labour Market – Application of TVP-VARs to Three European Countries. In: *Proceedings of the 33rd International Conference Mathematical Methods in Economics*, (2014), 156–161.

[4] Dobešová, A. and Hampel, D.: Inflation Targeting During Financial Crisis in Visegrád Group Countries, *Economics and Management* **2**, (2014), 12–19.

[5] Elliott, G., Rothenberg, T. J. and Stock, J. H.: Efficient Tests for an Autoregressive Unit Root, *Econometrica* **64**, 4 (1996), 813–836.

[6] Golinelli, R. and Orsi, R.: Modelling Inflation in EU Accession Countries: The Case of the Czech Republic, Hungary and Poland, *Eastward Enlargement of the Euro-zone Working Papers* **9**, (2002), 1–43.

[7] Hausman, J. A.: Specification Tests in Econometrics, *Econometrica* **46**, 6 (1978), 1251–1271.

[8] Horvath, R. Komarek, L. and Rozsypal, F.: Does Money Help Predict Inflation? An Empirical Assessment for Central Europe, *CNB Working paper series 2010* **5**, (2010), 1–27.

[9] Jannsen, N., Potjagailo, G. and Wolters M. H.: The Monetary Policy During Financial Crises: Is the Transmission Mechanism Impaired?, *Kiel Working Paper*, (2015), 1–37.

[10] Jílek, J.: *Finance v globální ekonomice I. Peníze a platební styk*. Grada, Praha, 2013.

[11] Kao, C.: Spurious Regression and Residual Based Tests for Cointegration in Panel Data, *Journal of Econometrics* **90**, 1 (1999), 1–44.

[12] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y.: Testing the Null Hypothesis of Stationarity Against the Alternative of an Unit Root, *Journal of Econometrics* **54**, 1–3 (1992), 159–178.

[13] Levy, M.: Factors Affecting Monetary Policy in an Era of Inflation, *Journal of Monetary Economics* **8**, 3 (1981), 351–373.

[14] Mishkin, F. S.: Monetary Policy Strategy: Lessons from the Crisis, *NBER Working Paper* **16755**, (2011), 1–62.

[15] Neely, C. and Rapach, D.: International Comovements in Inflation Rates and Country Characteristics, *Journal of International Money and Finance* **30**, 7 (2011), 1471–1490.

[16] Stock, J. H. and Watson, M. W.: Modeling Inflation After the Crisis, *NBER Working Paper* **16488**, (2010), 1–59.

[17] Vaněk, T., Dobešová, A. and Hampel, D.: Selected Aspects of Modeling Monetary Transmission Mechanism by BVAR Model. In *AIP Conference Proceedings* **1558**, (2013), 1855–1858.

Posteriori setting of the weights in multi-criteria decision making in educational practice

Jiří Rozkovec[1], Jan Öhm[2], Kateřina Gurinová[3]

Abstract. The aim of this paper is to propose an alternative way of classification of students writing tests, when here we mean subjects like mathematics, physics etc. The classification is a multi-criteria decision making problem. The tasks to solve there are criteria and the results of each student are values of individual variant and the weights are numbers of points assigned to examples. Usually, the teacher sets these points before the test according to his meaning about a difficulty of particular example and also according to a mutual comparison of all the examples. A disadvantage of this approach could be the individual experience of the teacher who does not have to estimate the difficulties for a given group of students fittingly. Then easy examples are appraised by high number of points and vice versa. The presented method calculates the weights (i.e. the points) according to the results achieved by the students and then the total score. The score of a particular example is calculated only in percent and then its average. The reciprocal value of the average is the weight. Afterwards the total score of every test is evaluated and also identified the difficulties of assigned tasks.

Keywords: multi-criteria decision making, weights, goodness-of-fit test.

JEL classification: C44
AMS classification: 90B50

1 Introduction

Pedagogy, in its normative character, systematically dictates teachers to evaluate their students. Evaluation is a complex and difficult process in which some action (act, product) is evaluated using verbal or non-verbal methods, not only by a grade (i.e. grading is less complex and figures as a result of evaluation - see [3]). Evaluation of student's performance in achievement tests is one of the key operations in an achievement test creation process. It has many functions and is arguably divided into many typologies.

The only indisputable division of evaluation is by two elementary aspects − absolute and relative. Absolute aspect is formulated by an observed student's result compared to a desirable result, potentially to the ideal result. Relative aspect expresses student's level (e.g. knowledge) related to his colleagues. It tells which of them are better, equal or worse (see [6]).

Evaluation process is subject of many critical essays. It is considered a moral aspect of assessment [7], economic criteria of testing or the cumulation of evaluation acts in some methods [6]. Teachers are often under pressure and unfavourable judgements for their presumptive stressful grading. Čapek in [2] warns againts normative evaluation (in his classification similar to relative) and discusses an abrogation of evaluation by grades. Subjectivism and cumulation of evaluation acts can be decreased by the usage of achievement tests. Generally are achievement test in opposition to psychological test (IQ test, test of personality, etc.) because they measure knowledge and skills [5].

Following the functional classification [5], achievement tests can be divided into these categories:

[1]Technical University of Liberec/Faculty of Economics, department of Economic Statistics, Voroněžská 13, 46001 LIBEREC, Czech Republic, jiri.rozkovec@tul.cz
[2]Technical University of Liberec/Faculty of Economics, department of Economic Statistics, Voroněžská 13, 46001 LIBEREC, Czech Republic, jan.ohm@tul.cz
[3]Technical University of Liberec/Faculty of Economics, department of Economic Statistics, Voroněžská 13, 46001 LIBEREC, Czech Republic, katerina.gurinova@tul.cz

- INITIAL. Used at the beginning of lesson in order to evaluate initial level of students' knowledge.
- DIAGNOSTIC. Used for analysis of knowledge and skills, detection of errors and their elimination.
- OUTPUT. Its purpose is for evaluation of lesson outputs.
- TEMPORARY. Has time limit and contain equally difficult tasks. Success is determined by achievement realised in particular time interval.
- SCALAR. Opposite to temporary. Contain variously difficult tasks. Success is determined by achieved level of difficulty.

These tests are considered as objective although it is the misleading statement. Achievement tests are focused on objective examination of students' content of curriculum mastery.

2 Grading as a multiple-criteria decision making

Grading could be considered as a multiple-criteria decision making problem, where the aim is to set ranking of alternatives, not to find any optimal solution. The alternatives are in fact the results of particular tests. For the evaluation will be used *weighted sum model* (WSM) [4]. Let's assume that there is n alternatives and k criteria, where i^{th} alternative is a vector $\vec{X_i} = (x_{i1}, x_{i2}, \ldots, x_{ik})$. Vectors $\vec{X_1}, \vec{X_2}, \ldots, \vec{X_n}$ then create rows of decision matrix $X_{n\times k}$. The decision criteria create the columns of X. The corresponding weights are in vector $\vec{w} = (w_1, \ldots, w_k)$.

However, this model slightly differs from the practical life where teacher sets the points for all examples in advance - denote these values as a vector $\vec{b} = (b_1, \ldots, b_k)$. Then he/she evaluates the proportion which has been solved for each task by absolute number of points. The total score, in other words *importance* of the particular alternative, is given by the sum of achieved points for all k examples. To use WSM it is necessary to rearrange this data this way.

In the decision matrix X there are again the proportions mentioned in the previous paragraph but as a relative ones, i. e. $x_{ij} \in [0, 1]$ for $i = 1, \ldots, n$, $j = 1, \ldots, k$. The weights w_j, $j = 1, \ldots, k$ then imply from:

$$w_j = \frac{b_j}{\sum_{i=1}^{k} b_j} \tag{1}$$

In this case the importance of alternative $U_A(X_i) = \vec{w}\vec{X_i}$ where $U_A(X_i) \in [0, k]$ (index A means that a priori weights were used for the calculation of the importance). But here a small problem appears - and it is the setting of weights.

2.1 A priori weights

It was mentioned above that the weights from (1) were set in advance - it means before writting of the test. How were they set? Accidentally? Intentionally based on teacher's experience? Either way, they express an assumption about the difficulty of the examples which is at the same time a transformed expectation about the students' abilities and knowledge. And it could cause errors, because the teacher can overestimate or underestimate them.

In the first case some tasks would have greater weights than they should have and vice versa. To sum it, these *a priori* weights do not fit the difficulty of the tasks. In further text the proposal of construction of *posteriori* weights will be presented.

2.2 Posteriori weights

The weights which are being constructed should ensure the ranking of the students, but not only this. Also they should identify the tasks which are able to distinct good and poor students from the knowledge point of view. In other words, tasks solved either by everybody or nobody do not have such ability and from the teacher's point of view they do not have any additional profit.

The method is based just on the values of matrix $X_{n\times k}$. The measure of the difficulty d_j of j^{th} task is expressed as a reciprocal value of average $\overline{x}_{\cdot j}$ of j^{th} column of X:

$$\overline{x}_{\cdot j} = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \tag{2}$$

$$d_j = \frac{1}{\overline{x}_{\cdot j}} \tag{3}$$

for $j = 1,\ldots,k$. Naturally, $\overline{x}_{\cdot j} \in [0,1]$, so $d_j \in [1,\infty)$. It is necessary to add that the criteria where $\overline{x}_{\cdot j}$ equals to 0 or 1 will be excluded from the processing, because they do not have, besides other things, just the ability of distinction as was mentioned above.

Further d_j, $j = 1,\ldots,k$ have to be standardized to the weights v_j:

$$v_j = \frac{d_j}{\sum_{i=1}^{k} d_j} \tag{4}$$

Now the importance U_P of alternative X_i can be evaluated in a common way, i. e. $U_P(X_i) = \vec{v}\vec{X_i}$ (index P means that posteriori weights were used for the calculation of the importance U).

Now the a priori and posteriori weights $\vec{w}$ and $\vec{v}$ can be compared by χ^2 *goodness-of-fit* test [1]. Proportions in $\vec{w}$ are considered as theoretical distribution, while $\vec{v}$ is the empirical one.

3 Practical examples

The following two examples are the practical demonstration of the method described above. The both concern with tests of mathematical analysis from our educational practice in the last year. Let's note that the significance level for all test is $\alpha = 0,01$.

Example 1. This sample has following parameters: $n = 69, k = 6$, i. e. there was six examples and 69 students. The success rates of examples were: $\vec{\overline{x}}. = (0,2638,\ 0,3237,\ 0,4928,\ 0,5246,\ 0,6928,\ 0,1246)$. The prior weights are:

$$\vec{w} = (0,125,\ 0,375,\ 0,125,\ 0,125,\ 0,125,\ 0,125)$$

If they are compared with the success rates, then, roughly said, the higher weight, the lower success. From $\vec{\overline{x}}.$ the *posteriori* weights $\vec{v}$ were evaluated using formula (4):

$$\vec{v} = (0,1869,\ 0,1523,\ 0,1001,\ 0,094,\ 0,0712,\ 0,3955)$$

The χ^2 goodness-of-fit test has the test statistic $\chi^2 = 54,13$. The critical value is $\chi^2_5(0,99) = 15,09$. So the null hypothesis that the proportions are indentical *is rejected*. Now the question is how the importances U_A and U_P differ. At first the test of their correlation will be performed. Sample correlation $r(U_A, U_P) = 0,9549$, P-Value = 0,00. So there is a strong positive correlation between them. Let's define now the difference $Z(X_i)$ as:

$$Z(X_i) = U_A(X_i) - U_P(X_i) \quad i = 1,\ldots,n \tag{5}$$

The average and variance of the difference Z is $\overline{Z} = 0,088$, $s^2_Z = 0,0064$. Z is normally distributed (P-Value of χ^2 goodness-of-test is 0,049). The hypothesis $H_0 : \mu = 0$ was rejected and $H_1 : \mu > 0$ was accepted (P-Value=0). So there is a statistically significant positive difference $U_A - U_P$.

Example 2. This sample has following parameters: $n = 72, k = 5$. The success rates were: $\vec{\bar{x}}. = (0,4792,\ 0,3281,\ 0,4219,\ 0,4427,\ 0,599)$.

The weights $\vec{w}$ and $\vec{v}$ are:

$$\vec{w} = (0,2,\ 0,2,\ 0,2,\ 0,2,\ 0,2)$$

$$\vec{v} = (0,1825,\ 0,2666,\ 0,2073,\ 0,1976,\ 0,146)$$

The χ^2 goodness-of-fit test has the test statistic $\chi^2 = 2,78$. The critical value is $\chi^2_4(0,99) = 13,28$. So the null hypothesis that the proportions are indentical *is not rejected*. Sample correlation $r(U_A, U_P) = 0,9947$, P-Value = 0,00. So there is also strong positive correlation between U_A and U_P.

The average and variance of the difference Z for this example is $\overline{Z} = 0,017$, $s^2_Z = 0,0007$. Z is normally distributed (P-Value of χ^2 goodness-of-test is 0,042). The hypothesis $H_0 : \mu = 0$ was again rejected and $H_1 : \mu > 0$ was accepted (P-Value = 10^{-7}). So there is a statistically significant positive difference $U_A - U_P$.

4 Conclusion

The aim of this paper was to propose an alternative approach of posteriori setting of the weights for multiple-criteria decision making problems in educational practice. Concretely it concerns with the grading of written tests. Usually the weights are given in advance, but they would not have to correspond to reality. So the motivation is to assess the difficulty of the solved tasks more precisely in comparison with the a priori weights. The estimates are based on the achivements of tested students and this is the main difference - the weights are, in fact, generated by the values of the decision matrix. The idea how to express the difficulty is to use a reciprocal value of the average achievement for the particular task. In presented examples the weights $\vec{w}$ and $\vec{v}$ were significantly different (Example 1), but in Example 2 their fit was not rejected. It means that in the first case teacher's prediction of the difficulty was not so accurate like in the second one. It was the main goal - the construction and comparison of weights.

Consequently, the relation between $U_A(X_i)$ and $U_P(X_i)$ was briefly analyzed. The importance $U_A(X_i)$ calculated with a priori weights $\vec{w}$ was greater than $U_P(X_i)$ using posteriori weights $\vec{v}$. Their difference was normally distributed. The mean of difference Z is positive in the both cases and it is statistically significant. Further the correlation of U_A and U_P was tested. It is nearly perfect and positive in the both cases as well.

Reference

[1] Anděl, J.: *Statistické metody.* Matfyzpress, Praha, 2007.

[2] Čapek, R.: *Odměny a tresty ve školní praxi.* Grada Publishing, Havlíčkův Brod, 2008.

[3] Dittrich, P.: *Pedagogicko-psychologická diagnostika.* H & H, Jinočany, 1992.

[4] Jablonský, J.: *Operační výzkum.* Professional Publishing, Praha, 2002.

[5] Kratochvíl, M.: *Kontrola, zkoušení a hodnocení ve školní praxi.* Technická univerzita v Liberci, Liberec, 1998.

[6] Solfronk, J.: *Zkoušení a hodnocení žáků.* Pedagogická fakulta Univerzity Karlovy, Praha, 1996.

[7] Wiggins, G. P.: *Assessing Student Performance: Exploring the Limits and Purpose of Testing.* Jossey-Bass Publishers, San Francisco, 1993.

A client's health from the point of view of the nutrition adviser using operational research

Lucie Schaynová [1]

Abstract. In this paper, we analyse daily nutrient requirements of an individual person from the point of view of the nutrition adviser. The goal is to simplify the adviser's menu planning for a client as much as possible. After that, we design an individual eating plan for a week using a new linear optimization model. The model respects eating habits and it follows the client's or adviser's recommended recipes taking the compatibility of foods into account. The model involves linear constraints to ensure that two incompatible foods are not used in the same course. The model comprises further constraints to guarantee the diversity of the courses. The purpose of other constraints is to use an exact amount of some food, e.g. one whole egg or 100 grams of cheese, during the week. The model is made up so that the final dietary plan for the client is as natural as possible. The model gives recommended amounts of foods for weekly recipe planning.

Keywords: nutrient requirement, linear programming, dietary plan

JEL classification: C44
AMS classification: 90C05

1 Introduction

Proper nutrition is important for the proper development and functioning of our organism and to maintain a good state of health. It also represents a very important role in the prevention of civilization diseases.

Nutrition from the perspective of linear programming has always the same foundation: the fulfillment of nutritional requirements for health of a larger group of people or the population. The objective functions of the linear programming models often vary. For example minimize greenhouse gas emmisions in the United Kingdom [8]. The cheapest eating is mainly connected with helping the people in developing countries. The goal is to design a nutritionally adequate diet suitable for the complementary feeding period using locally available food with the minimal price [3], [2]. For countries without restriction of the availability of food, models that include popularity of foods (the most frequently bought foods), with minimal price are used again [11]. On the basis of these favorite foods studies for inhabitants, in which they observe current and new optimal diets, are designed. Both kinds of diets are compared to make the people realize the difference between them and to give them recommendations to change their eating habits. That was done for the people in Italy [4] and in France [9]. More individual is a treatise of the eating model for a group of people designed for one week, which uses information about popularity of recipes in Hawaii. There is minimized either the price or total time of preparation of meals [7].

In this article we are going to present the diet problem for an individual person from the point of view of the nutrition adviser. We develop a diet plan tailored to a clinent's needs using linear programming. The model follows the adviser's steps when creating the diet plan. For the adviser the model is very helpful from the financial and the time point of view.

2 The problem

The nutrition adviser offers an individual consultation to a client who suffers from some malfunction nutrition or is in danger of some diseases that can be affected by proper nutrition, such as obesity,

[1] University of Ostrava, Faculty of Science, Department of Mathematics, 30. dubna 22, 701 03 Ostrava, lucie.schaynova@osu.cz

diabetes, malnutrition, celiac disease, high blood pressure, etc. Nutrition consultancy is also used in prevention of diseases and can help healthy people too. A healthy client can have a new diet plan in the context of normal (rational) or sports nutrition. Consultancy also includes clients who adhere to an alternative way of eating and want to make sure in their eating habits. This is intended for vegetarians, vegans, etc. In this article we will focus on the rational way of eating.

The adviser's methodology includes the anamnestic and analytical part.

Anamnestic part

In this part the adviser is informed about the client's collection of basic personal data, family and personal medical history, mental health status in the present and in the past, surgical procedures, diet or physical intolerance. It is also necessary to know whether the client is taking any medication. This information is necessary to be known if the adviser has to consult the new diet with the client's doctor. Further information is about the lifestyle (smoking, alcohol), eating habits, time stress or physical activity. When some information is missing it can have a negative impact on the client's health during the new diet. In this article we assume that the client is healthy and has no indispositions so that we can create a diet plan focused on rational eating.

Analytical part

This part includes anthropometric measurements. The adviser needs to know the client's body height, weight, circumference of hip and waist, measurement of subcutaneous fat, visceral fat, BMI, etc. The adviser evaluates the client's physical condition and sets up a goal that the client should achieve.

The adviser determines the *ideal body weight* [6] as follows:

$$\begin{aligned} w_m &= 0.655\, h_m - 44.1\,, \\ w_f &= 0.593\, h_f - 38.8\,, \end{aligned}$$

where w_m or w_f is the ideal weight of a man or a woman, respectively, in kilograms and h_m or h_f is the height of the man or the woman, respectively, in centimeters.

Then the *basal energy* is needed to be determined. This is energy expenditure to ensure basic living functions and temperature of the body. The basal energy can be calculated by using indirect calorimetry or by using the following Harris-Benedict equations [6]:

$$\begin{aligned} v_m &= 4.184\,(66.5 + 13.8\, w_m + 5.0\, h_m - 6.76\, a_m)\,, \\ v_f &= 4.184\,(655.1 + 9.56\, w_f + 1.85\, h_f - 4.68\, a_f)\,, \end{aligned}$$

where v_m or v_f is the basal energy of the man or the woman, respectively, in kilojoules (kJ) per day and a_m or a_f is the age of the man or the woman, respectively, in years.

Then the *factor activity* p is needed to be determined. This is a coefficient of average physical activity. In the table which can be found in [12, p. 27], there are some values assigned to some particular activities. For example the coefficient of sleeping is 0.95, the coefficient of hard working is 2.5, etc. The client has to record all his or her activites during the week and their duration. Then the adviser calculates the value of p.

Then the *total daily energy* requirement can also be calculated by using indirect calorimetry, monitoring heart rate or by using the equation

$$t = vp + d\,,$$

where t is the total daily energy requirement in kJ per day, v is the basal energy and d is the postprandial thermogenesis, see [1].

Then the adviser sets recommended amounts of some nutrients for the client's ideal weight from the total energy intake. The nutrients should be composed so that the 15 %, 30 %, and 55 % of the total daily energy intake comes from proteins, fats, and saccharides, respectively.

Those proportions mentioned are recommended for a healthy adult with normal physical activity. In our calculations we use the fact that one gram of fats yields 37 kJ of energy, one gram of proteins or saccharides yield 17 kJ of energy [13].

The client has to tell the adviser all the eating habits, such as which foods or dishes are the client's favourite or unfavourite, and the client should evaluate his or her normal eating habits. The client has to record all consumed food items during one week along with their amounts. The adviser can calculate from the data which amounts of nutrients are necessary to be increased or reduced. Then the adviser assesses the client's nutritional situation and gives recommendation for positive change.

Then the adviser can create a new diet with the help of some software. The principle of the design is always the same. The adviser selects foods from a database according to the adviser's recipes or the client's preferred food items, respectively. It is necessarry to assignt some quantity to each food. The adviser usually works with the most important nutrients only. In addition, the database does usually not contain all values of nutrients and many of them can be missing. When the adviser compiles a list of all the food items for one day, then some nutrients may exceed or do not reach the required amounts. Therefore, the adviser has to go back in the list of the food items and to change the quantity of some items. Sometimes the adviser has to replace some food items by other ones and changes the food quantities until the nutrients are at the optimal levels. This procedure can be time consuming. The adviser has to do the same procedure for at least the next 6 days. A diet plan is usually prepared for one week. Finally the adviser gives the client a final recommendation and principles which the client should follow.

3 Mathematical model

The basic components of the diet are nutrients. The optimal intake of nutrients is necessary for our body and life. Every nutrient is a carrier of some type of biochemical function in our organism and any lack or surplus of nutrients is dangerous.

Let us consider food items $i = 1, \ldots, m$ and nutrients $j = 1, \ldots, n$. We will work with 5 courses (breakfast, first snack, lunch, second snack and dinner) and we design a diet plan for 7 days, $d = 1, \ldots, 7$, considering 5 courses, $l = 1, \ldots, 5$, a day. So in total we have $k = 1, \ldots, 35$ courses during a week. Denote the set of all 35 courses during the 7 days as G and let the non-empty disjoint sets G_d for all days be such that $G = G_1 \cup G_2 \cup \cdots \cup G_7$. In addition let the set of all breakfasts E_1, first snacks E_2, lunches E_3, second snacks E_4 and dinners E_5 be such that $G = E_1 \cup E_2 \cup \cdots \cup E_5$.

Consider a binary matrix $\boldsymbol{A} = (a_{if})$ for all $i, f = 1, \ldots, m$ which means compatibility between food items (1 if two food items are compatible, 0 otherwise) and binary matrix $\boldsymbol{B} = (b_{ik})$ for all $i = 1, \ldots, m$, $k = 1, \ldots, 35$ which means compatibility between food items and courses (1 if a food item and a course are compatible, 0 otherwise). The columns of the matrix $\boldsymbol{B}$ for every course during the week can be the same. Let us have a real matrix of nutritive values $\boldsymbol{C} = (c_{ij})$ for all $i = 1, \ldots, m$, $j = 1, \ldots, n$. Vectors of minimal and maximal daily recommended values of nutrients are $\boldsymbol{b}^- = (b_j^-)$ and $\boldsymbol{b}^+ = (b_j^+)$ for all $j = 1, \ldots, n$. Vectors of minimal and maximal quantity of food are $\boldsymbol{v}^- = (v_i^-)$ and $\boldsymbol{v}^+ = (v_i^+)$ for all $i = 1, \ldots, m$.

Let y_{ik} be a binary variable. The variable means if the food item i is used in the course k. We forbide two incompatible food items in the same course as follows

$$y_{ik} + y_{fk} \leq 1\,, \tag{1}$$

for all $i, f = 1, \ldots, m$, $k = 1, \ldots, 35$, where $i \neq f$ and $a_{if} = 0$.

It is undesirable to have a food item in an incompatible course so

$$y_{ik} = 0 \tag{2}$$

for all $i = 1, \ldots, m$ and $k = 1, \ldots, 35$ such that $b_{ik} = 0$.

We want to have at least three food items in every course

$$\sum_i y_{ik} \geq 3 \tag{3}$$

for all $k = 1, \ldots, 35$.

Let $x_{ik} \geq 0$ be a real variable that means the amount of food item i used in course k. We need to satisfy daily nutrient recommendation as follows

$$\sum_i \sum_{k \in G_d} c_{ij} x_{ik} \geq b_j^- \qquad \text{and} \qquad \sum_i \sum_{k \in G_d} c_{ij} x_{ik} \leq b_j^+ \tag{4}$$

for all $j = 1, \ldots, n$ and $d = 1, \ldots, 7$.

Inspired by [10], we can add distribution of energy intake within the daily nutritional amounts, it means 20 % of total energy is needed for breakfast, 30 % for lunch, 25 % for dinner and 12,5 % for first and second snack. Inequalities for breakfast look as follows

$$\sum_i \sum_{k \in E_1} c_{i1} x_{ik} \geq 0.2\, b_1^- \qquad \text{and} \qquad \sum_i \sum_{k \in E_1} c_{i1} x_{ik} \leq 0.2\, b_1^+ \,, \tag{5}$$

where b_1^- or b_1^+ means minimum or maximum of daily recommended value of energy, respectively. The inequalities for other daily courses are analogous.

We do not want to waste the food so we can add condition to ensure that we consume a whole package (100 grams, 150 grams, etc.) of a food item or a whole hen egg (the average weight of an egg is 50 grams). So we add equalities like

$$\sum_k x_{ik} = 100\, z_i^{100} + 50\, z_i^{150} \,, \tag{6}$$

for some food items i that are supplied in packages of certain sizes, where variables like z_i^{100}, z_i^{150} are integer variables such that $0 \leq z_i^{100} \leq z_i^{150}$. The coefficient 50 by z_i^{150} is the difference between the size of the packages of 100 and 150 grams.

We would like to use some minimum or maximum value of food item every day

$$x_{ik} \geq v_i^- y_{ik} \qquad \text{and} \qquad x_{ik} \leq v_i^+ y_{ik} \,, \tag{7}$$

for all $i = 1, \ldots, m$ and $k = 1, \ldots, 35$.

There can happen that some food item is repeated in some courses. Let $M_{l,g}$ be a set of similar foods used in the same course l. For example the set of milk products for breakfast such as milk, curd or yogurt. Another example is the set of bakery products for breakfast. We do not want to have the same milk product nor the same bakery product for breakfast in two consecutive days. We can achieve that by inequalities

$$\sum_{i \in M_{l,g}} |y_{iq} - y_{ir}| \geq 1 \tag{8}$$

for all $q, r \in E_l$, $|q - r| = 5$, and $l = 1, \ldots, 5$ and the respective groups g of foods (such as milk products, bakery products, etc.). The inequality (8) is equivalent to

$$-\epsilon_{i,q,r} \leq y_{iq} - y_{ir} \leq \epsilon_{i,q,r} \qquad \text{and} \qquad \sum_{i \in M_{l,g}} \epsilon_{i,q,r} \geq 1 \,. \tag{9}$$

We have the new real variable $\epsilon_{i,q,r}$ for all $q, r \in E_l$, $|q - r| = 5$, and $l = 1, \ldots, 5$ and the respective groups of foods g.

In this model composed of conditions (1)–(9) we do not use any objective function. We minimize value 0. We are interested in finding a feasible solution only.

4 Results

We need input data described in Section 2 for our model described in Section 3. Let us consider a woman. She is 26 years old and 173 cm tall. The ideal body weight is $w_f = 64$ kg, the basal energy is $v_f = 6166$ kJ, the factor activity is $p = 1.5$ and the postprandial thermogenesis is $d = 919$ kJ. The total daily energy requirement is $t = 10109$ kJ. The composition of required macroelements is as follows: 89 or 82 or 327 grams of proteins or fats or saccharides, respectively. It is not necessary that the diet plan has the exact values mentioned above, tolerance of 10 % can be used in the model so we can work with intervals.

Nutrient	Food items [100 g]			Recommended amounts		
	Chicken	Potatoes	⋯	Min	Solution	Max
Energy [kJ]	694	322	⋯	9098	10364	11122
Proteins [g]	20	2	⋯	80	94	98
Fats [g]	10	0	⋯	73	90	90
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 1 Input data and calculated amounts of nutrients for one day

In the model we work with 28 nutrients and 65 food items (see [5]). The amounts of microelements are taken from [5], [12], [13].

Our model works with the values from Table 1. There are food items and the respective amounts of nutrients in each of the items. The minimal and maximal amounts of nutrients for one day are also given. Then, our model consists of 4432 variables, out of which 10 are integer and 2205 are binary, and 53411 constraints. The calculated diet for the client for one day of the week is shown in Table 2. The amounts of the nutrients in the one-day diet plan are shown in Table 1 in the Solution column. This model was computed in the specialized optimization software FICO$^{(R)}$ Xpress Optimization Suite. On a Windows XP SP3 computer with 0.99 GB RAM and Intel Atom 1.60 GHz CPU, the computation took about 5 seconds.

Course	Food items	Energy [kJ]
Breakfast	120 g yogurt, 30 g apple, 35 g muesli, 20 g orange marmelade, 34 g wholemeal biscuit, 6 g syrup, 130 g mandarin	2224.4
Snack	60 kaiser rolls (1 piece), 9 g margarine, 70 g cheese	1137.5
Lunch	Soup: 74 g whole wheat pasta, 33 g green beans, broth The main course: 150 g chicken, 130 g potatoes, 20 g cucumber, 33 g eggplant, 25 g lentils, 8 g sunflower oil	3336.6
Snack	50 g yogurt, 61 g cornflakes, 18 g honey, 7 g syrup	1390.3
Dinner	120 g kaiser roll (2 pieces), 13 g margarine, chives Salad: 24 g tomatoe, 94 g bell pepper, 100 g cucumber, 32 g ricotta, 13 g olive oil, basil	2275.0

Table 2 Optimal diet plan for one day

The client with this diet plan can be sure that all the nutritional values for her day are fulfilled.

5 Discussion

The composition of the food items in Table 2 corresponds to the client's eating habits.

It may be very time consuming for the adviser to create a one-day diet plan that respects the optimal amounts of all 28 nutrients; it may even be impossible for the adviser to do that for the whole week.

Our model gave us the solution in a couple of seconds. Time consuming in our case can be creating the matrices of compatibility $\boldsymbol{A}$ and $\boldsymbol{B}$. But the matrices can be the foundations of the adviser's data. The adviser can do only some small changes for every client in the matrices. In addition the model can give us a diet plan for period longer than one week.

The advisers usually do not care about wasting the food. If the clients are eating according to the adviser's diet plan, the clients can have a lot of wasted food items in the end of the week. Inequalities (8) of our model prevent that.

This article gives an idea how the working steps of the adviser and the care of the client's health can be improved.

Acknowledgements

The author thanks to doc. David Bartl for comments that helped to improve the article. The use of the FICO$^{(R)}$ Xpress Optimization Suite under the FICO Academic Partner Program is gratefully acknowledged. This research was supported by the internal grant no. SGS12/PřF/2016–2017, Geometric mechanics, optimization and number theory. The author is grateful to the two anonymous referees for their useful comments that helped to improve the manuscript.

References

[1] Bender, D., A.: *Introduction to nutrition and metabolism.* CRC Press, London, 2002.

[2] Briend, A., Ferguson, E., Darmon, N.: Local food price analysis by linear programming: a new approach to assess the economic value of fortified food supplements. *Food and Nutrition Bulletin* **22.2** (2001), 184–189.

[3] Briend, A., et al.: Linear programming: a mathematical tool for analyzing and optimizing children's diets during the complementary feeding period. *Journal of Pediatric Gastroenterology and Nutrition* **36.1** (2003), 12–22.

[4] Conforti, P., D'amicis, A.: What is the cost of a healthy diet in terms of achieving RDAs? *Public Health Nutrition* **3.3** (2000), 367–373.

[5] Heseker, H., Heseker, B.: *Die Nährwerttabelle: [über 40.000 Nährstoffangaben; einfache Handhabung; Tabellen zu Laktose, Fruktose, Gluten, Purin, Jod und trans-Fettsäuren].* Umschau, Neustadt an der Weinstraße, 2016.

[6] Kastnerová, M.: *Poradce pro výživu.* Nová Forma, České Budějovice, 2011.

[7] Leung, P., Wanitprapha, K., Quinn, L., A.: A recipe-based, diet-planning modelling system. *British Journal of Nutrition* **74.2** (1995), 151–162.

[8] Macdiarmid, J., et al.: Sustainable diets for the future: can we contribute to reducing greenhouse gas emissions by eating a healthy diet? *The American Journal of Clinical Nutrition* **96.3** (2012), 632–639.

[9] Maillot, M., Vieux, F., Amiot, M., J., Darmon, N.: Individual diet modeling translates nutrient recommendations into realistic and individual-specific food choices. *The American Journal of Clinical Nutrition* **91.2** (2010), 421–430.

[10] Reihserová, R.: Jak poskládat stravu v průběhu dne? *Svět potravin.* Available from: http://www.svet-potravin.cz/clanek.aspx?id=4177 (cited 25 April 2016).

[11] Sklan, D., Dariel, I.: Diet planning for humans using mixed-integer linear programming. *British Journal of Nutrition* **70.1** (1993), 27–35.

[12] Společnost pro výživu: *Referenční hodnoty pro příjem živin.* Výživaaservis s.r.o., Praha, 2011.

[13] Svačina, Š.: *Klinická dietologie.* Grada Publishing, Praha, 2008.

[14] Zavadilová, V.: *Výživa a zdraví.* Ostravská univerzita v Ostravě, Ostrava, 2014.

Wavelets Comparison at Hurst Exponent Estimation

Jaroslav Schürrer[1]

Abstract. In this paper we present Discrete Wavelet Transformation based on Hurst exponent estimation and compare different wavelets used in the process. Self-similar behavior mostly associated with fractals can be found in broad range of areas. For self-affine processes the local properties are reflected in the global ones and the Hurst exponent is related to fractal dimension, where fractal dimension is a measure of the roughness of a surface. For usually non-stationary time series the Hurst exponent is a measure of long term memory of time series.

From former works mentioned in references we know that Discrete Wavelet Transformation provides better accuracy compared to Continuous Wavelet Transformation and that it outperforms methods based on the Fourier spectral analysis and R/S analysis.

Keywords: Hurst exponent, Wavelet Transformation, signal power spectrum.

JEL classification: C44
AMS classification: 90C15

1 Hurst exponent

Many processes of interest in finance modeling are considered as self-similar processes. For these processes we estimate dimensionless estimator, called Hurst exponent and denoted by H usually, for self-similarity of a time series. Hurst's exponent estimation for real world data plays an important role in the study of processes that exhibit properties of self-similarity. Hurst exponent was firstly defined by Harold Edwin to develop law for regularities of the Nile water level. Hurst exponent is one of the fractal measures, which varies from $0 < H < 1$. There are three important regions of Hurst exponent values within meaningful range $[0, 1]$. For persistent time series $H > 0.5$ and $H < 0.5$ for anti-persistent time series and finally $H = 0.5$ for uncorrelated series. There are many methods for Hurst exponent estimation using time series. Most commonly used method are classical Rescaled Range analysis (R/S analysis), variance-time analysis, detrended fluctuation analysis (DFA) and wavelet based estimator. In this article we focus on the wavelet based estimator method.

It is widely accepted, that many stochastic processes exhibit a long-range dependence and fractal structure. The most suitable mathematical method for research of the dynamics and structure of such processes is fractal analysis. Let define self-similar stochastic process $X(t)$ as the process $a^{-H}X(at), a > 0$ which shows the same second-order statistical properties as $X(t)$. Hurst exponent is a measure of self-similarity or a measure of duration of long-range dependence of a stochastic process. Example of fractal stochastic structures is the modern financial market which uses information about past events to affect decisions in the present, and contains long-term correlations and trends. The market remains stable, as long as it retains its fractal structure.

2 Discrete Wavelet Transformation and signal power spectrum

Wavelets represent way of analyzing time series in particular non-stationary time series where they ensure time and frequency localization. It means that they provide wavelet coefficients that are local both in time and frequency. Main features of wavelets, multi-resolution and localization are well suited for extraction fluctuations at various scales from local trends over appropriate window sizes. Extracted fluctuations are influenced by the choice of wavelet.

[1] Czech Technical University in Prague/Masaryk Institute of Advanced Studies, Kolejni 2637/2a, Prague, Czech Republic, schurjar@cvut.cz

Multiresolution analysis decomposes signal into subsignals with different size resolution levels. From practical point of view, it is not possible to calculate wavelet coefficients at every scale. So we choose only subset of scales and positions to make calculations. Dyadic scales and positions (based on powers of two) are used for Discrete Wavelet Transformation (DWT). DWT is linear transformation where in our case input data is represented by vector of length equal to power of two and output is an another vector of the same length which separates two different frequency components.

An efficient way to implement this scheme using filters was developed in 1988 by Mallat. This algorithm is known as Pyramidal algorithm or two-channel sub-band coder. The input vector (signal) is passed through a series of high pass filters to analyze the high frequencies, and it is passed through a series of low pass filters to analyze the low frequencies. The resolution of the signal is changed by the filtering operations, and the scale is changed by upsampling and downsampling (subsampling) operations.

The DWT analyzes the signal at different frequency bands with different resolutions by decomposing the signal into an approximation and detail coefficients. DWT employs two sets of functions, called scaling functions and wavelet functions, which are associated with low pass and high pass filters. The original signal $x[n]$ is first passed through a halfband high pass filter $g[n]$ and a low pass filter $h[n]$ with subsampling by 2, simply by discarding every other sample. This constitutes one level of decomposition and can be mathematically expressed as follows:

$$y_{high}[k] = \sum_n x[n].g[2k-n] \tag{1}$$

$$y_{low}[k] = \sum_n x[n].h[2k-n] \tag{2}$$

where $y_{high}[k]$ and $y_{low}[k]$ are the outputs of the high pass and low pass filters respectively, after subsampling by 2. This procedure can be repeated for further decomposition where at every level, the filtering and subsampling will result in half the number of samples thus half the time resolution and half the frequency resolution. Following figure 1 schematically depicts this procedure. Because of filtering the lenght of input signal playes an important role. Wavelet algorithms expect the input length to be a power of two. If length is not suitable we have to use some extrapolation of the input data in order to extend the signal before computing the Discrete Wavelet Transform using the cascading filter banks algorithm. There are several methods of signal extrapolation that are mainly used: zero-padding, constant-padding, symmetric-padding and periodic-padding. Length of signal also influences maximum number of decomposition levels which is equal to $floor(log_2(data\_len/(filter\_len-1))/(log_2(2)))$.

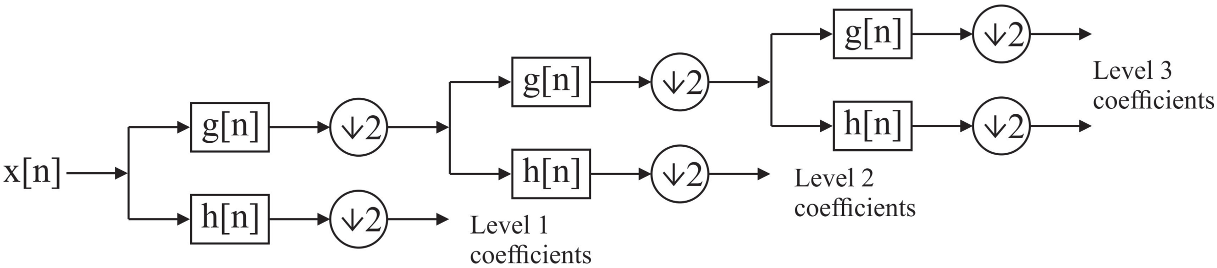


Figure 1 Wavelet decomposition scheme.

For a given mother wavelet ψ and scaling function φ with approximate coefficient $a(j,k)$ and detail coefficients $d(j,k)$ these coefficients are defined as:

$$a(j,k) = \int_{-\infty}^{\infty} X(t)\varphi_{j,k}(t)dt, d(j,k) = \int_{-\infty}^{\infty} X(t)\psi_{j,k}(t)dt \tag{3}$$

where $\varphi_{j,k} = 2^{-j/2}\varphi(2^{-j}t-k)$, $\psi = 2^{-j/2}\psi(2^{-j}t-k)$. Original signal $X(t)$ is then represented in terms of DWT with following equation:

$$X(t) = \sum_k a_{j,k}\varphi_{j,k}(t) + \sum_{j=1}^{J}\sum_k d_{j,k}\psi_{j,k}(t) \tag{4}$$

The wavelet power is calculated by summing the squares of the coefficient values (approximation or detailed) for each level:

$$E(j) = \sum_{k=0}^{\frac{N}{2^j}-1} W_{j,k}^2 \tag{5}$$

3 Hurst exponent estimation

During last decade there were several methods published which use wavelet transformation to estimate Hurst exponent. The averaged wavelet coefficient method described by Simonsen and Hansen in [5] is based on the statistical equality of continuous wavelet transformation of self affine function $h(x)$ expressed by equation $W[h(x)](a,b) \simeq W[\lambda^{-H}h(\lambda x)](a,b)$ where continuous wavelet transformation (CWT) is defined by equation

$$W[h](a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi_{a,b}^*(x)h(x)dx \tag{6}$$

Here $\psi^*(x)$ denotes the complex conjugate. We also mention that DWT can be understood as a critically sampled CWT.

Another method published by Faggini and Parziale in [1] is the wavelet-based Hurst parameter estimator based on a spectral estimator obtained by performing a time average of the wavelet detail coefficients $|d_{j,k}|^2$ at a given scale

$$S_r = \frac{1}{N_i} \sum_j |d_{j,k}|^2 \tag{7}$$

where N_i is the number of wavelet coefficients at a scale i and N is number of data points. First DWT is computed $log_2 E[d(j,k)]^2$ followed by variance of these estimates and then linear regression is performed to find slope γ. H is then calculated as: $H = 0.5(1+\gamma), 0 < \gamma < 1$.

Next, we mention the fluctuation functions based on discrete wavelet coefficients proposed by Manimaran,Paniggrahi and Parikh in [4] that use wavelet power also referred as energy to calculate fluctuation function at a level s:

$$F(s) = \left[\sum_{j=1}^{s} E(j) \right]^{\frac{1}{2}} \tag{8}$$

The scaling behavior is then obtained through $F(s) \sim s^H$. Hurst exponent is then obtained from slope of the log-log plot of $F(s)$ versus scales s. Here H is the Hurst scaling exponent, which can be obtained from the slope of the log-log plot of $F(s)$ vs scales s.

At the end we describe the Gloter, Hoffmann method published in [2] which is based on the energy level estimator. This estimator is defined by the equation:

$$E\left[\frac{Q_{j+1}}{Q_j}\right] = C2^{-2H} \tag{9}$$

where Q is the detailed coefficient energy of certain decomposition level and C is an arbitrary unknown constant. Computing this for several j and using a linear least squares fit, the Hurst exponent can be estimated as $\gamma * (-0.5)$ where γ is slope of E_j.

4 Comparison of Hurst exponent estimation

In this part we compare different wavelets in Hurst exponent estimation. The fractal Brownian motion (fBM) has been chosen as a model of random process which exhibits fractal properties. We generate two realisations with known Hurst exponent equal to 0.5 and 0.8. The length of data sample was chosen so that the border effect is eliminated. In Figure 2 we can see generated realizations of fBM.

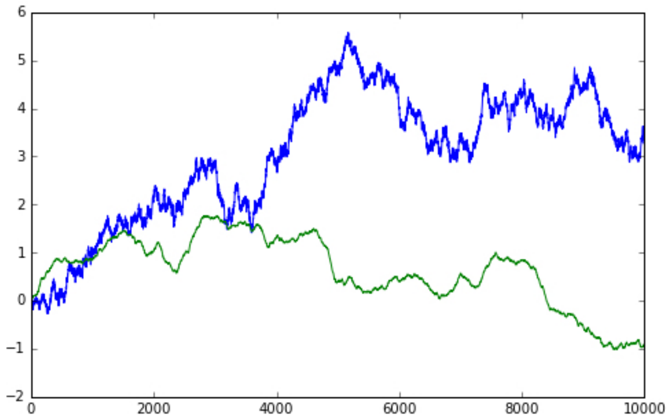


Figure 2 fBM with $H_1 = 0.5$ $(blue), H_2 = 0.8$ $(green)$.

All numerical computation were done by own code written in Python with the use of pyWavelets and NumPy packages.

The following algorithms have been compared:

- R/S analysis
- The wavelet energy level estimator Haar
- The wavelet energy level estimator DB2, DB4, DB5
- The wavelet energy level estimator Coiflet 1
- The wavelet energy level estimator Biorthogonal 1.3

First we analyze fBm with $H = 0.5$ and two different lengths of input vector $N = 10000$ and $N = 5000$. Following tables sumarizes results from R/S analysis and WEL estimators using different wavelet families. We can observe that Coiflet wavelets are pretty good in Hurst estimation. Another important observation is that compact support length of wavelet influences estimation accurancy (compare Haar, DB2 - DB5). Time series number of datapoints also plays important role.

Hurst estimate method	**H (N=10 000)**	**Difference**	**H (N=5000)**	**Difference**
RS analysis	0.51	-0.01	0.482	0.018
WEL estimator Haar	0.47	0.03	0.445	0.055
WEL estimator DB2	0.46	0.04	0.344	0.156
WEL estimator DB4	0.46	0.04	0.420	0.08
WEL estimator DB5	0.43	0.07	0.335	0.165
WEL estimator Coif1	0.499	0.001	0.463	0.037
WEL estimator Bior1.3	0.479	0.021	0.476	0.024

Table 1 Hurst estimator methods comparison for fBM with $H = 0.5$

Following table summarize result of the same procedure for input vector with Hurst exponent $H = 0.8$.

Hurst estimate method	**H (N=10 000)**	**Difference**	**H (N=5000)**	**Difference**
RS analysis	0.8	0.00	0.817	-0.017
WEL estimator Haar	0.78	0.02	0.756	0.044
WEL estimator DB2	0.78	0.02	0.671	0.029
WEL estimator DB4	0.77	0.03	0.763	0.037
WEL estimator DB5	0.69	0.11	0.699	0.101
WEL estimator Coif1	0.78	0.02	0.747	0.053
WEL estimator Bior1.3	0.79	0.01	0.75	0.05

Table 2 Hurst estimator methods comparison for fBM with $H = 0.8$

If we further decrease input vector length we notice that efficiency of Hurst estimator based on the DWT dramatically goes down. So we can conlude that estimator is efficient above $N = 4096$.

5 Conclusion

Discrete Wavelet Transformation approach yields correct values for the Hurst exponent. But we have to take into account several elements. First of all, border effect due filtering can lead to wrong values. This can be eliminated with smallest possible filter selection and signal extension. The length of input data also influences maximum level of decomposition during DWT and in our case the bigger is better. We can also observe some underestimation of H due to sampling especially in early stages of decomposition. Big advantage of DWT approach to Hurst exponent estimation is that Wavelet transformation can operate on non-stationary data (some methods can be applied only to stationary time series).

Acknowledgements

Research described in the paper was supervised by Prof. M. Vosvrda, Institute of Information Theory and Automation of the ASCR.

References

[1] Faggini, M., Parziale, A.: *Complexity in Economics: Cutting Edge Research*. Springer, 2014.

[2] Gloter, A., Hoffmann, M.: Estimation of the Hurst parameter from discrete noisy data. *Annals of Statistics* **35** (2007), 1947-1974.

[3] Gneiting, T., Schlather, M.: Stochastic models which separate fractal dimension and Hurst effect. *SIAM Review* **46** (2001), 269–282.

[4] Manimaran, P., Paniggrahi, P. K., Parikh, J. C.: On Estimation of Hurst Scaling Exponent through Discrete Wavelets. *ArXiv Physics e-prints* **physics/0604004** (2006),

[5] Simonsen, I., Hansen, A., Nes, O. M.: Determination of the Hurst Exponent by use of Wavelet Transforms. *Physical Review E* **53** (1998), 2779–2787.

[6] Walnut, D.: *An Introduction to Wavelet Analysis*. Birkhäuser, Boston, 2004.

Possibilities of Regional Input-Output Analysis of Czech Economy

Jaroslav Sixta[1], Jakub Fischer[2]

Abstract. Regional Input-Output Analysis belongs to the group of data demanding models. Such models can be used for detailed description of regional economy and for advanced regional studies. Theoretically, this issue is well founded and there is lots of available literature for these issues. But in reality, pure regional Input-Output Tables are difficult to find. Only some official statistical authorities compile and publish so detailed regional data or even completed Regional Input-Output Tables. Construction of these tables is demanding for data sources, i.e. statistical surveys and capacities. Therefore, we used different approach based on regionalization models and we constructed pure regional Input-Output Tables for the Czech Republic for 2011. Such work lasted three years (2013-2015) and we finally published product by product tables for NUTS 3 level, regions “Kraje” of the Czech Republic. The paper follows the work and illustrate regional specifics with simple tools of input-output analysis. Czech regions are compared on the level of input coefficients and its contribution to the national aggregates. Mutual dependency of regions is discussed, as well.

Keywords: Regional, Input-Output, Multipliers.

JEL Classification: C67, R10, O11
AMS Classification: 65C20

1 Introduction

Regional Input-Output Analysis (RIOA) is a powerful tool for analyzing regional economy. The structure of regional output and intermediates allows identification of regional dependency on imports and regional sensitivity to final demand shocks. Advanced economic models used for the preparation of economic policy need detailed and reliable data. In specific cases, regularly published national Input-Output Tables (IOTs) based on country averages are not adequate for the analysis. The fear of deviation from reality grows with analytical description of the regional impact. Even though the demand for regional data is quite high and frequent, the availability of such tables is somewhat lacking or not very timely trail behind this demand. A discussion of the necessity of regional policies and their efficient application require both detailed data and skilled users. A regional economy consists of both regional businesses and regional branches of national or multinational companies. When discussing regional economic growth, employment and living conditions, effective tools for the measurement of policy impacts are usually missing. Fortunately, compilation of Input-Output Tables has a long tradition in the Czech Republic, see Sixta [13] and construction of regional tables is therefore possible.

Regional input-output models face the principal difficulty with the lack of available data, which makes them more demanding for compilation. Since 1950s, input-output analysis was applied mostly in the form of single-region models to solve regional challenges. Comparing inter-regional (IRIO) and multiregional (MRIO) models proves IRIO models much more data demanding. Papers aimed at the level of real regions (of a state) occur rather rarely. Some attempts can be found, e.g. in Italy (Benvenuti, Martellato and Raffaelli, [1]), Finland (see [6]) or the Netherlands [2]. Some other examples may be found in Wiedmann [16]. The case of derivation of regional input-output tables based on GRIT method is deeply discussed in Miller and Blair [7]. GRIT is easy to apply and therefore it is widely popular among economists. It was used for numbers of countries, for example also for the Czech Republic [10], China (Wen and Tisdel [15]), etc.

The aim of this paper is to compare Input-Output Analysis based on national IOTs with the analysis based on regional IOTs. For this case we chose Moravian-Silesian Region, the region aimed at mining industry and facing serious economic problems. The issue is illustrated by simple statistic IOA. Since Regional Input-Output Tables (RIOTs) are not officially compiled in the Czech Republic, we used RIOTs constructed by the Department of Economic Statistics of the University of Economics in Prague for the year 2011. These tables are available for the

[1] University of Economics in Prague, Department of Economic Statistics, Nam W. Churchilla 4, Prague, Czech Republic, sixta@vse.cz
[2] University of Economics in Prague, Department of Economic Statistics, Nam W. Churchilla 4, Prague, Czech Republic, fischerj@vse.cz

use of domestic output (RIOTd) and imported products (RIOTi), product-by-product type (82) and valued at basic prices. There are lots of available information sources connected to RIOTs, e.g. Miller and Blair [7].

2 Methodology

Regional Input-Output Tables were compiled within the research project, see Sixta and Fischer [12] and detailed methodology can be found also in Sixta and Vltavská [13]. The main principle lies in decomposing of national symmetric IOTs into regions. The key role is played by estimates of output vectors, regional trade and published regional accounts. The crucial point significantly influencing the quality of regional input-output analysis is the quality of RIOTs. Especially, it is connected with the movement of goods and services; both by international and interregional trade. The most import part, regional output vectors, is not officially published. The Czech Statistical Office publishes only gross value added. The estimates of all 14 regional output vectors are based on the decomposition of national output vectors in specific institutional sector. National input coefficients were used for the estimates of regional intermediate consumption. It means that the national technology is applied on regional level. Expenditure approach is also not officially published and therefore we reconstructed the use side of IOTs in line with Kramulová and Musil [9].

Regional input-output tables are available for all 14 regions of the Czech Republic in ESA 1995 [4] methodology. Despite ESA 2010 [5] was implemented in September 2014, RIOTs based on ESA 1995 are fully usable for input-output analysis. Methodical changes between ESA 1995 and ESA 2010 do not affect results significantly. Such analysis is based on the technological relationships that were not changed. The key methodical differences between input-output tables compiled according to ESA 1995 and ESA 2010 lie in the recording of processing. With respect to improved approach to ownership concept, processing in international trade is excluded from imports, intermediate consumption, export and output. For the purposes of static input-output analysis, intermediates that originate only from domestic output are use and therefore the changes in import matrices do not significantly influence the results.

The benefits of regional input-output tables are demonstrated by simple static input-output analysis. Even if this tool is rather simple, it is very convenient for illustration purposes. It covers mainly complete impact of the change of final use on output, gross value added and employment. Suppose that all necessary IOA models requirements are fulfilled (see EU [3]) and then the change of vector of output (x) is estimated on the basis of the change of final demand (y), see formula 1:

$$\Delta x = (I - A_d)^{-1} \Delta y \tag{1}$$

where

x	output vector,
y	final demand vector,
A_d	matrix of input coefficients computed from domestically produced intermediates.

Intermediate consumption (z_j) is estimated by given input coefficients ($a_{i,j}$) from output vector:

$$z_j = \sum_i x_j a_{ij} \tag{2}$$

gross value added (g_j) is derived as difference

$$g_j = x_j - (z_j + mi_j) \tag{3}$$

where

g_j	gross value added of product j,
x_j	output of product j,
z_j	intermediates used in product j,
mi_j	imported intermediates used in product j.

Distinguishing between domestically produced (z) and imported (mi) intermediates allows identification of relationship between regional output and regional imports. The emphasis is put on the identification of the impact on the particular region. It means that import matrices that are available for all fourteen Czech regions to cover both international and interregional trade. The trade matrix connecting the regions has not been finished yet and we are still working on this issue. For every region, it is possible to separate international trade and interregional trade (both imports and exports). Imports from foreign countries were allocated according to the use of particular products and export according to output. Interregional flows were obtained by balancing of supplies and uses on the level of product. The sum of interregional exports must equal the sum of interregional imports, for more details see Sixta and Vltavská [13] or Kahoun and Sixta [8].

3 Regional Input-Output Model

The effect of using of regional input-output tables is presented on the case of Moravian-Silesian Region (MSR) and the decline of final demand for mining industry products. Moravian-Silesian Region amounts about 10% of Czech gross value added in 2011. This region is significantly dependent on the mining industry that forms 7% of regional gross value added. In the industries of mining and energy (B, D and E) about 26 thousand workers work, consisting 5% of the region. Mining industry is very much linked with other industries and possible negative effects of final demand on output of mining industry exceeds the region itself. The dependency of the region is given by the structure of intermediates.

For the illustration of negative final demand shock, let's suppose the decrease in export of mining industry by 20 CZK bn. From the perspective of the region, it is not important whether the shock is given by the other regions or situation abroad. It means that:

$$\Delta y = - 20\ 000. \tag{4}$$

When using static input-output analysis described by formulas 1 to 3, the overall change of the vector of output is dependent on the matrix of input coefficients. The model is computed for two cases; the first is based on national symmetric input-output tables and the second one is based on regional input-output tables for Moravian-Silesian Region. In the first case, matrix A_d is obtained from national IOTs and in the second case from regional input-output tables. The changes in output vector (x) are presented on Figure 1.

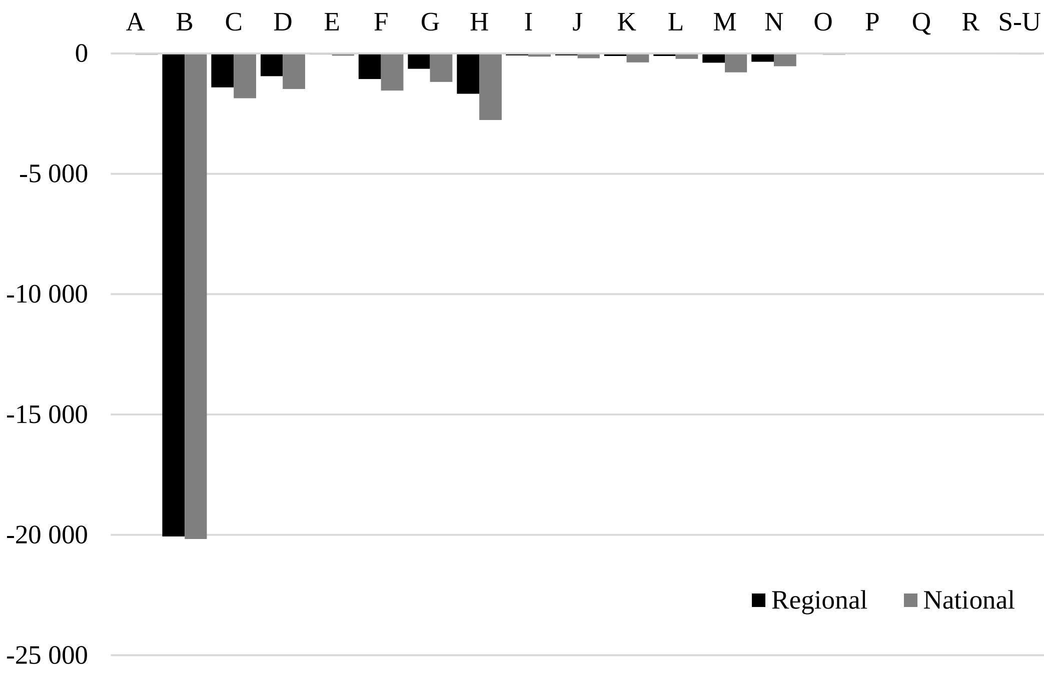


Figure 1 Change of output, CZK mil.

Overall difference between changes in the structure of output given by two mentioned approached is not very significant. It is caused by the very specific input structure of mining industry. In other words, mining industry is significantly influencing national input-output table. The most significant difference is obtained for transport and storage industry (H) since these products are highly imported for Moravian-Silesian Region. Similarly, the imports of manufacturing products can explain the differences in manufacturing industry (C).

The key benefit when using regional input-output tables for input-output analysis lies in the estimates of regional production and regional imports. When using national IOTs (NIOTs), it may not be clear which part of the output is imported from other regions and which part comes from abroad. Figure 2 describes total impact of imported products into the regions.

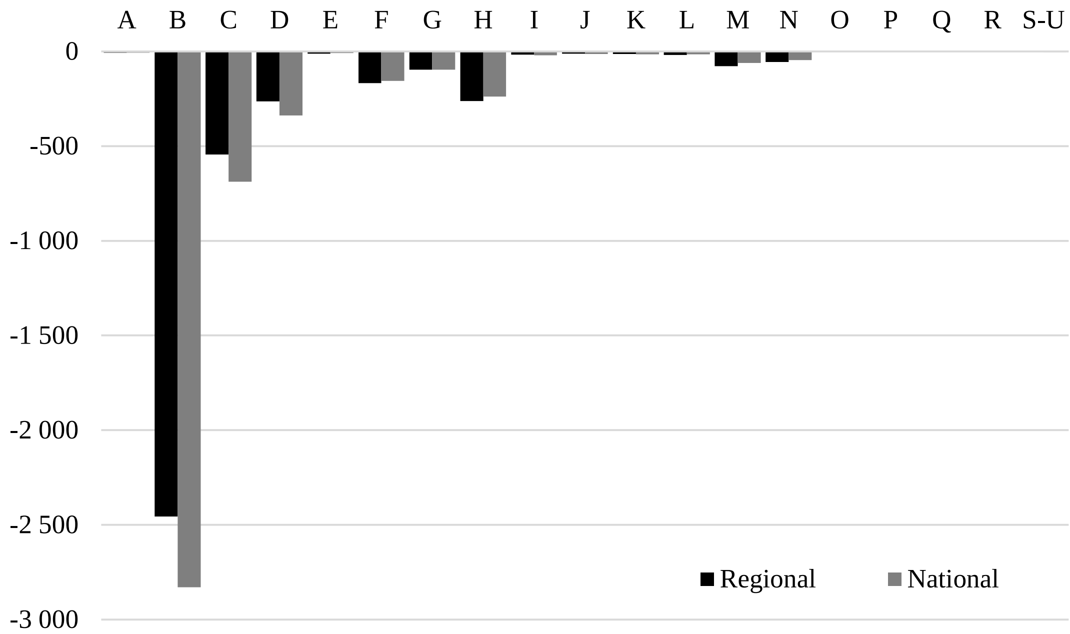


Figure 2 Change of imports, CZK mil.

Similarly, to the case of output, differences between both coefficients are relatively moderate. The most important impact is obtained for products used in mining industry by 2.5 CZK bn. that represents 36% of the decrease of intermediate consumption in this industry, see Table 1.

Indicator	Total industries		Mining industry	
	RIOT	NIOT	RIOT	NIOT
Output	-27 053	-31 497	-20 062	-20 170
Domestic intermediates	-7 053	-11 497	-4 328	-6 383
Imported intermediates	-4 005	-4 538	-2 455	-2 829
Gross value added	-15 995	-15 462	-13 279	-10 957
Imports	-6 710	-7 868	x	x

Table 1 Comparison of Input-Output Analysis based on RIOTs and NIOTs, CZK mil.

The impact on the decrease in external demand on gross value added is highly influenced by imported intermediates. If we use RIOT for Moravian-Silesian Region, the decrease in value added counts 16 CZK bn. In the case of national IOT, it is about 15.5 CZK bn. The most significant difference is observed in the change of total output, intermediate consumption and imports.

When using regional input-output table, the distribution of imports between regional flows and flows from abroad is possible. The impact is described on figure 3. The overall decrease of imports is about 4 CZK bn. of which 13% (500 CZK mil.) belongs to regional flows and 87% (3.5 CZK bn.) to flows from abroad. Mostly, the affected products cover mining, manufacturing and energy, trade and transport. Finally, it means that the decrease of imports in Moravian-Silesian Region will affect other Czech regions by 500 CZK mil.

The above described impacts deals with static input-output analysis only and therefore these effects can be assessed to immediate reaction of the economy. In reality, the impacts are more spread in the economy and multiplication effects takes place. In more advanced models (e.g. Zbranek and Sixta [17]) the initial shock will be distributed via economic processes in many rounds. The decrease of exports will cause the drop of output, value added, wages, operating profits. Decreased wages and profits in the regions mean weaker initiatives for investments (gross fixed capital formation) and household consumption.

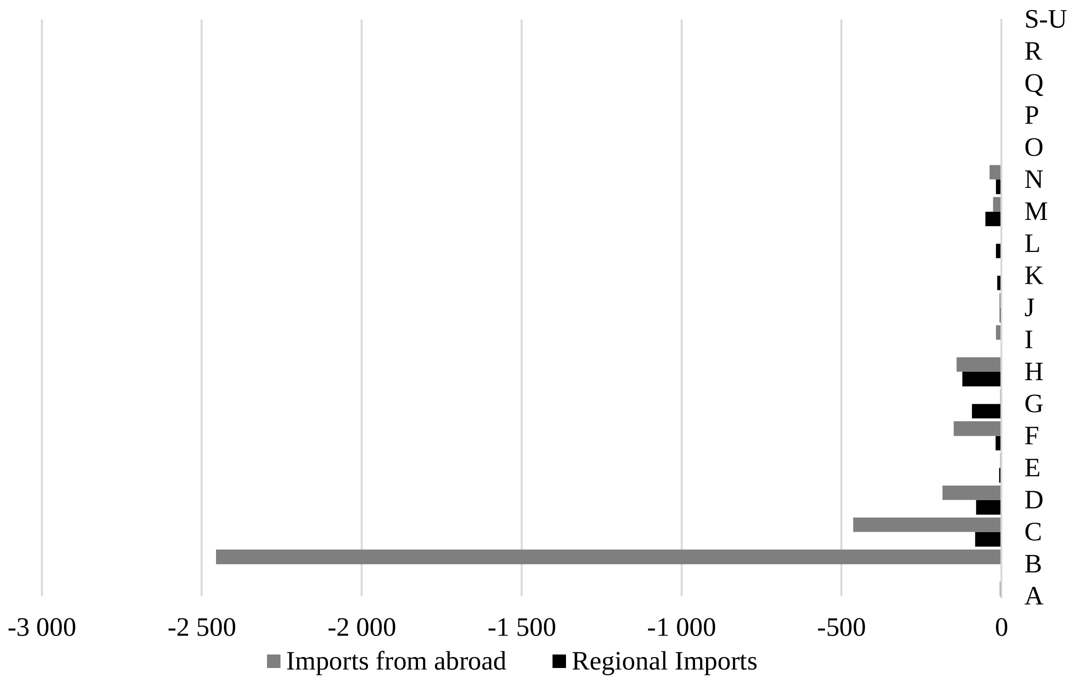


Figure 3 Distribution of imports

4 Conclusion

The paper briefly describes the benefits of Regional Input-Output Tables for Input-Output Analysis. It was demonstrated that using national averages (national Symmetric Input-Output Tables) leads to different estimates of output and subsequently imports and exports. The most important benefit of using Regional Input-Output Table lies in the possibility of estimates of the impacts on individual regions of the Czech Republic. For demonstration purposes, we selected Moravian-Silesian Region and we modeled the impact of the decrease in final demand for mining products production (output of mining industry). The analysis was conducted in two variants. The first variant is based on Regional Input-Output Tables and the second variant is based on National Input-Output Table. Even though these results are not very different in the case of value added, there are significantly different for output, intermediate consumption and imports.

Regional Input-Output tables are not usually published by official statistical authorities. These tables are available only for some countries for selected years, e.g. the U.S., Spain or Finland. International databases are usually aimed at groups of countries and pure regional tables focusing on the regions of a particular country are very scarce. We prepared Regional Input-Output Tables for 14 regions of the Czech Republic at basic prices for 2011. Even though we used ESA 1995 methodology, the tables can be still easily used for various approaches to Input-Output Analysis. The aim of the demonstration described within this paper was to illustrate possibilities and promote their usage. These tables are available for free download from our webpage, see http://kest.vse.cz/veda-a-vyzkum/vysledky-vedecke-cinnosti/regionalizace-odhadu-hrubeho-domaciho-produktu-vydajovou-metodou/

Acknowledgements

Supported by the grant No. 13-15771S "Regionalization of Estimated Gross Domestic Product by Expenditure Method" of the Czech Science Foundation and also with the long term institutional support of research activities by the Faculty of Informatics and Statistics of the University of Economics

References

[1] Benvenuti S. C., Martellato D., Raffaelli C.: A Twenty-Region Input-Output Model for Italy. *Economic Systems Research* (1995), 101-116.
[2] Eding G., Oosterhaven J., Vet De B., Nijmeijer H.: Constructing Regional Supply and Use Tables: Dutch Experiences. In: Hewings G. J. D., Sonis M., Madden M., Kimura. Y.: Understanding and Interpreting Economic Structure, Berlin, 1999, 237-263.
[3] European Union: *Input-Output Manual*. Luxembourg, 2008.
[4] Eurostat: *European System of Accounts – ESA 1995*. Luxembourg, 1996.
[5] Eurostat: *European System of Accounts – ESA 2010*. Luxembourg, 2013.

[6] Flegg. A. T. and Tohmo T.: Regional Input-Output Tables and the FLQ Formula: A Case Study of Finland. *Regional Studies* **47** (2013), 703-721.
[7] Kahoun, J. and Sixta, J.: Regional GDP Compilation: Production, Income and Expenditure Approach. *Statistika* (2013), 24–36.
[8] Kramulova, J. and Musil, P.: Experimentální odhad složek výdajové metody regionálního HDP v ČR. *Politická ekonomie* (2013), 814–833.
[9] Miller, R. E. and Blair, P. D.: *Input-Output Analysis: Foundations and Extensions*, 2009.
[10] Semerák V., Zigic K., Loizou E., Golemanova-Kuharova A.: *Regional Input-Output Analysis: Application on Rural Regions in Germany, the Czech Republic and Greece*. In: *European Association of Agricultural Economists*, Rural development: governance, policy design and delivery. Ljubljana, 2010.
[11] Sixta, J.: Development of Input-Output Tables in the Czech Republic. *Statistika* (2013), 4–14.
[12] Sixta, J. and Fischer, J.: Regional Input-Output Models: Assessment of the Impact of Investment in Infrastructure on the Regional Economy. In: *Proceedings of Mathematical Methods in Economics*, Cheb, 2015.
[13] Sixta, J. and Vltavská, K.: Regional Input-output Tables: Practical Aspects of its Compilation for the Regions of the Czech Republic. *Ekonomický časopis* **64** (2016), 56–69.
[14] Wen. J. J. and Tisdell C. A.: *Tourism and China's Development: Policies, Regional Economic Growth and Ecotourism*, Singapore, 2001.
[15] Wiedmann T. A: Review of Recent Multi-region Input-Output Models Used for Consumption-Based Emission and Resource Accounting. *Ecological Economics* (2009), 211-222.
[16] Zbranek, J. and Sixta J.: Possibilities of Time Input-Output Tables. In: *Proceedings of Mathematical Methods in Economics,* Jihlava, 2013, 1046–1051.

A Comparison of Integer Goal Programming Models for Timetabling

Veronika Skocdopolova[1]

Abstract. Timetabling is a wide-spread problem that every school or university deals with. It is an NP-hard problem. Optimisation of a timetable in general leads to cost reducing or avoiding time wasting. Goal programming is a widely used technique for solving not only multi-criteria decision making problems. It became very popular since it was formulated more than 60 years ago. Using goal programming for timetabling enables including soft constraints in the mathematical model. There are two main approaches for solving the timetabling problem. The first one is to split the problem into several interrelated stages. The other one is solving the problem with one complex model. In this paper there are compared three integer goal programming models for timetabling that were presented at previous MME conferences. The compared models are a three-stage, a four-stage, and a complex integer goal programming model. All three models were formulated for timetable construction at the University of Economics, Prague.

Keywords: Goal programming, integer programming, timetabling, soft constraints.

JEL Classification: C61
AMS Classification: 90C29

1 Goal Programming and Timetabling

Goal programming is a widely used technique not only for multi-criteria decision making. Due to quite easy formulation of the goal programme and good understanding its methodology by decision makers, many papers and books dealing with goal programming applications appeared since its first formulation in 1955. The applications to education are summarized in [15]. One of them is timetabling, a yearly problem of each school. The main aim of this paper is a comparison of three integer goal programming models prepared for the department of econometrics at the University of Economics, Prague.

In this chapter the basics of goal programming are summarized. There are also introduced the main approaches to the problem of timetabling. In the next part the three models are briefly described and then they are compared in the last chapter.

1.1 Goal Programming

Goal programming is based on assumption that the main decision-maker objective is to satisfy his or her goals. The aim of goal programming models is not to find the best solution, but to find a solution that meets the decision maker's goals. Therefore the chosen solution can be dominated.

The generic goal programming model can be formulated as follows [11]:

Minimise

$$z = f(\mathbf{d}^-, \mathbf{d}^+) \tag{1}$$

subject to

$$\mathbf{x} \in X\ , \tag{2}$$

$$f_k(\mathbf{x}) + d_k^- - d_k^+ = g_k\ , \quad k = 1, 2, \ldots, K, \tag{3}$$

$$d_k^-, d_k^+ \geq 0\ , \qquad k = 1, 2, \ldots, K,$$

[1] University of Economics, Prague, Faculty of Informatics and Statistics, Department of Econometrics, W. Churchill sq. 4, Prague, Czech Republic, skov01@vse.cz.

where (1) is a general penalisation function of vectors of positive and negative deviations $\mathbf{d}^-$ and $\mathbf{d}^+$, X is the set of feasible solution satisfying all of the constraints including non-negativity constraints (2), $f_k(\mathbf{x})$ is an objective function that represents the k-th criterion, g_k is the k-th goal, the decision maker wants to meet, d_k^- is the negative deviation from the k-th goal, which represents the underachievement of the k-th goal, d_k^+ is the positive deviation from the k-th goal, which shows the overachievement of the k-th goal; and (3) are the goal constraints. The goal constraints are often titled soft constraints, while the constraints that form the set of feasible solutions can be called hard constraints.

Time to time in practice certain small deviation from the goal is acceptable with just a little penalisation and greater deviation is allowed but with higher penalisation. To preserve the linear form of the model we can use the multistage penalisation function. In this case we have to replace the goal constraints (3) with the following one

$$f_k(\mathbf{x}) + d_k^- - d_{k1}^+ - d_{k2}^+ = g_k \text{, } k = 1, 2, \ldots, K, \tag{4}$$

$$0 \le d_{k1}^+ \le h_{k1}^+, 0 \le d_{k2}^+ \le h_{k2}^+, d_k^- \ge 0 \text{, } k = 1, 2, \ldots, K, \tag{5}$$

where d_{k1}^+ is the 1st stage positive deviation that is non-negative and has an upper bound h_{k1}^+ , and d_{k2}^+ is the 2nd stage positive deviation that is also non-negative with an upper bound h_{k2}^+. The two-stage penalisation function can be formulated as follows:

$$z = \sum_{k=1}^{K} \left(u_k d_k^- + v_{k1} d_{k1}^+ + v_{k2} d_{k2}^+ \right), \tag{6}$$

where v_{k1} and v_{k2} are the penalty constants of the 1st and 2nd stage positive deviation and u_k is the penalty constant of the negative deviation.

1.2 Timetabling

Timetabling is a wide-spread problem that every school deals with. There are many ways how to construct a university timetable. From the time-tested scheduling board, over various heuristic and metaheuristic methods to sophisticated optimisation models.

There are two main approaches to construction of a timetable using mathematical modelling. The first of them is creating a complex model usually using binary variables (e.g. [3] or [9]). However, solving integer programming models is, in general, NP-hard problem [7], so it leads to utilizing various heuristic or metaheuristic methods. Heuristic and metaheuristic methods give us solutions that are relatively close to optimal solution in relatively reasonable time. The most common metaheuristic methods used for timetabling are evolutionary algorithms (e.g. genetic [10] or memetic [13]), algorithms based on graph colouring (e.g. [1] or [5]), local and tabu search (e.g. [6] or [14]) or simulated annealing (e.g. [8]). The other approach is based on a decomposition of the problem into several interrelated stages (e.g. [2], [4], or [12]). This means outputs of one stage are inputs for the next stage.

2 Integer Goal Programming Models for Timetabling

During the past four years three integer goal programming models were presented at the conference Mathematical Methods in Economics. These models were formulated to help with timetable construction at the department of econometrics, University of Economics, Prague. In the next sections all three models are briefly described.

2.1 Three-stage model

This was the first model formulated for the department. It was inspired by [2]. In the model there are three interrelated stages. In the first stage each course (lectures and seminars) is assigned to a teacher. There are three goals in this stage – maximisation of total preference of courses, and achieving maximum seminar and lecture loads. The hard constraints of the first stage ensure, that

- any teacher is not assigned to a course he or she evaluated with 0 preference;
- all courses are assigned to a teacher; and
- any teacher does not teach more than two courses of selected subjects.

The result of this stage is the assignment of all courses to teachers according to their preferences.

In the second stage each course (assigned to a teacher) is assigned to a time window. The only goal of this stage is, that the total number of courses assigned to a time window cannot exceed the number of classrooms available for that time slot. The hard constraints of the second stage assure, that

- the teacher assigned to the course will be available in certain time window
- every teacher can be assigned most to one time window; and
- every course have to be assigned exactly to one time window.

The other constraints deal with cohesion of input from the previous stage and variables of this stage.

The last stage of the model assigns courses (assigned to teachers and time windows) to classrooms. There is only one goal in this stage, which deals with location of the course to a classroom with adequate capacity. The hard constraints of the third stage ensure, that

- each course has to be assigned exactly to one classroom;
- each classroom can be assigned most to one course; and
- any course can be assigned to a classroom only in case the classroom is available in the time window assigned to the course.

The result of the third stage is a complete timetable for the department.

This model was presented in [16]. The main problem of this model was that small seminars were often assigned to classrooms with large capacity. Therefore the four-stage model was formulated. The three-stage model also does not deal with the requirement to assign certain seminars to computer classrooms. This requirement is included in the four-stage model.

2.2 Four-stage model

The four-stage model is an extension of the three-stage model. In the first stage a complete timetable of lectures is prepared. To this timetable the seminars are added in the next three stages that are the same as in three-stage model. The only difference is that hard constraints assuring assignment of computer classrooms were added to the third and fourth stage. This model was presented in [18].

With these multi-stage models we have to assure the continuity of all the stages. It is necessary to secure in each stage that there will be a feasible solution in the next stage. This means eg. the model has to respect number of classrooms available in each time window while assigning courses to time windows. Due to the necessity of the continuity of all stages it was complicated to include some special requirements to the model. Hence the complex model was formulated.

2.3 Complex model

The complex model presented in [17] assures all the necessary features of timetable and includes all the special requirements of the department of econometrics. These requirements are

- considering teachers' time preferences;
- assignment of certain courses to computer classrooms every week or every other week;
- teaching certain courses in block;
- assignment of certain courses to particular classrooms; and
- teaching in two distant campuses.

The complex model has six goals:

- maximisation of total preference of courses;
- maximisation of total time preference;
- achieving adequate course loads of teachers;
- utilisation of classrooms' capacities;
- achieving a given number of days per week each teacher has at least one course; and
- achieving a given number of courses per day each teacher has.

The principle of two-stage penalization function is utilized in the third and fourth goal.

In addition to the hard constraints of the multi-stage models the constraints in the complex model assure, that

- each teacher has courses only in one campus each day (no transfers during one day);
- the courses that need computers are assigned to the computer classrooms every or every other week;
- the courses that are thought in block are assigned to immediately following time windows, to the same classroom, and to the same teacher; and
- certain courses are assigned to particular classrooms.

This complex model creates the timetable in just one stage. Although it is an NP-hard problem, the optimisation via software does not take a long time. The problem with ca 7.5 million variables and 6 thousands constraints was solved using *Gurobi 6.0* solver (www.gurobi.com) in less than 10 minutes (on a notebook with quad-core processor Intel® CoreTM i7-4702MQ 2.2GHz, operation memory 16 GB DDR3 and 64bit operation system Windows 8.1).

The complex model have been implemented into an application written in *Visual Basic for Applications* (*VBA*). The application can use anybody even without any knowledge of either mathematical modelling or optimisation systems. The user just fill in all the input data to an *MS Excel* worksheet and click a button. The results of the optimisation are then exported back to the *MS Excel* worksheet and the final timetable is created via *VBA* procedures.

3 Comparison of the Models

The input data for all three models are collected and adjusted in *MS Excel* worksheet using *VBA*. The multi-stage models were solved via optimisation system *LINGO 14.0* (www.lindo.com). Solving each stage took just a few seconds. The data were transformed between the stages using procedures written in *VBA*. For solving the complex model was chosen the solver *Gurobi* and the model was written in *MPL* (*Mathematical Programming Language*), which enables using many different solvers, such as *CPLEX*, *XPRESS, CONOPT* and many others.

All of the models were verified on the same data set for better comparison. It was the data from the summer term 2011/2012. In that term the models had to deal with 32 teachers, 80 courses, 35 time windows and 84 classrooms. In the Table 1 there are numbers of variables and constraints for each model.

		3-stage model	**4-stage model**	**Complex model**
1st stage	all variables	2,690	23,092	7,437,642
	integer variables	2,560	22,860	7,436,956
	constraints	2,962	4,395	5,893
2nd stage	all variables	95,100	2,624	x
	integer variables	92,400	1,984	x
	constraints	7,751	2,144	x
3rd stage	all variables	6,969	73,664	x
	integer variables	6,640	71,610	x
	constraints	330	6,527	x
4th stage	all variables	x	185,422	x
	integer variables	x	5,146	x
	constraints	x	7,711	x

Table 1 Number of variables and constraints

The difference in the number of variables between multi-stage models and the complex model is obvious. What is remarkable is the number of constraints is much higher for the multi-stage models than for the complex model, although there are not included all the necessary requirements in the multi-stage models. The reason for that is, that the multi-stage models need more constraints to ensure the integrity of the model. Moreover some of the constraints are used repeatedly in each stage.

The multi-stage models create applicable timetables. This means that the timetable avoids time conflicts, each course is assigned to a teacher, time window and classroom, and there is at most one course in each classroom. The problem of the multi-stage models was too complicated formulation of the special requirements such as assigning certain courses to computer classrooms every week or every other week, teachers' time preferences, teaching courses in block or teaching in two distant campuses. The main reason for choosing the multi-stage model at first was an assumption of the computational complexity of the complex model. However, solving the complex model with ca 7.5 million variables took less than 10 minutes. Therefore the complex model is more useful for creating the department timetable. The user does not have to transform the data between the stages. Although the transformation is done automatically by *VBA* procedures, there is an eventuality of making a mistake. The question is how it would be, if we would like to use the model for creating a timetable for the whole university. Rough estimate of the number of integer variables in such a model is 18.5 billion. It can result in lack of operation memory while preparing the model for solving or it can take several days to compute any result. In this case it is worth considering an adjustment of the multi-stage models. That is an issue for the future research.

4 Conclusion

In this paper three integer goal programming models for timetabling were briefly described and compared. The multi-stage models do not ensure all the specific requirements of the timetable at the department. The complex model assures all the necessary requirements. On the other hand if we would like to use the complex model for creation of a timetable for the whole university, it can lead to unsolvable model. In this case would be better to adjust one of the multi-stage model.

Acknowledgements

The research is supported by the Internal Grant Agency of the University of Economics, Prague, grant No. F4/62/2015.

References

[1] Abdul-Rahman, S., Burke, E. K., Bargiela, A., McCollum, B., and Özcan, E.: A constructive approach to examination timetabling based on adaptive decomposition and ordering. *Annals of Operations Research* **218** (2014), 3–21.

[2] Al-Husain, R., Hasan, M. K., and Al-Qaheri, H.: A Sequential Three-Stage Integer Goal Programming (IGP) Model for Faculty-Course-Time-Classroom Assignments. *Informatica* **35** (2011), 157–164.

[3] Asratian, A. S., and Werra, D.: A generalized class teacher model for some timetabling problems. *European Journal of Operational Research* **143** (2002), 531–542.

[4] Badri, M. A.: A two-stage multiobjective scheduling model for [faculty-course-time] assignments. *European Journal of Operational Research* **94** (1996), 16–28.

[5] Burke, E. K., Pham, N., Qu, R., and Yellen, J.: Linear combinations of heuristics for examination timetabling. *Annals of Operations Research* **194** (2012), 89–109.

[6] Caramia, M., Dell'Olmo, P., and Italiano, G. F.: Novel Local-Search-Based Approaches to University Examination Timetabling. *INFORMS Journal on Computing* **20** (2008), 86–99.

[7] Cerny, M.: *Computations, Volume III* (In Czech). Professional Publishing, Prague, 2012.

[8] Gunawan, A., Ng, K. M., and Poh, K. L.: A hybridized Lagrangian relaxation and simulated annealing method for the course timetabling problem. *Computers and Operations Research* **39** (2012), 3074–3088.

[9] Ismayilova, N. A., Sagir, M., and Rafail, N.: A multiobjective faculty–course–time slot assignment problem with preferences. *Mathematical and Computer Modelling* **46** (2005), 1017–1029.

[10] Jat, S. N., and Yang, S.: A hybrid genetic algorithm and tabu search approach for post enrolment course timetabling. *Journal of Scheduling* **14** (2011), 617–637.

[11] Jones, D., and Tamiz, M.: *Practical Goal Programming*. International Series in Operations Research & Management Science 141. Springer, New York, 2010.

[12] Ozdemir, M. S., and Gasimov, R. N.: The analytic hierarchy process and multiobjective 0–1 faculty course assignment. *European Journal of Operational Research* **157** (2004), 398–408.

[13] Qaurooni, D.: A memetic algorithm for course timetabling. In: *GECCO '11: Proceedings of the 13th annual conference on Genetic and evolutionary computation.* Association for Computing Machinery, Dublin, 2011, 435–442.

[14] Santos, H. G., Ochi, L. S., and Souza, M. J. F.: A Tabu search heuristic with efficient diversification strategies for the class/teacher timetabling problem. *Journal of Experimental Algorithmics* **10** (2005), 1–16.

[15] Skocdopolova, V.: Application of Goal Programming Models to Education. In: *Efficiency and Responsibility in Education 2012*. Czech University of Life Sciences Prague, Prague, 2012, 535–544

[16] Skocdopolova, V.: Construction of time schedules using integer goal programming. In: *Mathematical Methods in Economics 2012*, Silesian University in Opava, Karvina, 2012, 793–798.

[17] Skocdopolova, V.: Optimisation of University Timetable via Complex Goal Programming Model. In: *Mathematical Methods in Economics 2015*, University of West Bohemia, Plzeň, Cheb, 2015, 725–730.

[18] Skocdopolova, V.: University timetable construction – computational approach. In: *Mathematical Methods in Economics 2013*, College of Polytechnics Jihlava, Jihlava, 2013, 808–813.

Transient and Average Markov Reward Chains with Applications to Finance

Karel Sladký [1]

Abstract. The article is devoted to Markov reward chains with finite state space. Since the usual optimization criteria examined in the literature on Markov reward chains, such as a total discounted, total reward up to reaching some specific state (called transient models) or mean (average) reward optimality, may be quite insufficient to characterize the problem from the point of a decision maker. It seems that it may be preferable if not necessary to select more sophisticated criteria that also reflect variability-risk features of the problem. Perhaps the best known approaches stem from the classical work of Markowitz on mean variance selection rules, i.e. we optimize the weighted sum of average or total reward and its variance. In the article explicit formulae for calculating the variances for transient and average models are reported along with sketches of algorithmic procedures for finding policies guaranteeing minimal variance in the class of policies with a given transient or average reward. Application of the obtained results to financial models is indicated.

Keywords: dynamic programming, transient and average Markov reward chains, reward-variance optimality, optimality in financial models.

JEL classification: C44, C61, C63
AMS classification: 90C40, 60J10, 93E20

1 Introduction

The usual optimization criteria examined in the literature on stochastic dynamic programming, such as a total discounted or mean (average) reward structures, may be quite insufficient to characterize the problem from the point of a decision maker. To this end it may be preferable if not necessary to select more sophisticated criteria that also reflect variability-risk features of the problem. Perhaps the best known approaches stem from the classical work of Markowitz (cf. [2]) on mean variance selection rules, i.e. we optimize the weighted sum of average or total reward and its variance. In the present paper we restrict attention on transient and average models with finite state space and in the class of optimal policies we find the policy with minimal variance.

2 Notation and Preliminaries

In this note, we consider at discrete time points Markov decision process $X = \{X_n, n = 0, 1, \ldots\}$ with finite state space $\mathcal{I} = \{1, 2, \ldots, N\}$, and compact set $\mathcal{A}_i = [0, K_i]$ of possible decisions (actions) in state $i \in \mathcal{I}$. Supposing that in state $i \in \mathcal{I}$ action $a \in \mathcal{A}_i$ is chosen, then state j is reached in the next transition with a given probability $p_{ij}(a)$ and one-stage transition reward r_{ij} will be accrued to such transition. (We assume that $p_{ij}(a)$ is a continuous function of $a \in \mathcal{A}_i$.)

A (Markovian) policy controlling the decision process, $\pi = (f^0, f^1, \ldots)$, is identified by a sequence of decision vectors $\{f^n, n = 0, 1, \ldots\}$ where $f^n \in \mathcal{F} \equiv \mathcal{A}_1 \times \ldots \times \mathcal{A}_N$ for every $n = 0, 1, 2, \ldots$, and $f^n_i \in \mathcal{A}_i$ is the decision (or action) taken at the nth transition if the chain X is in state i. Let $\pi^m = (f^m, f^{m+1}, \ldots)$, hence $\pi = (f^0, f^1, \ldots, f^{m-1}, \pi^m)$, in particular $\pi = (f^0, \pi^1)$. The symbol E^π_i denotes the expectation if $X_0 = i$ and policy $\pi = (f^n)$ is followed, in particular, $\mathsf{E}^\pi_i(X_m = j) = \sum_{i_j \in \mathcal{I}} p_{i,i_1}(f^0_i) \ldots p_{i_{m-1},j}(f^{m-1}_{m-1})$; $\mathrm{P}(X_m = j)$ is the probability that X is in state j at time m.

[1] Institute of Information Theory and Automation of the Czech Academy of Sciences, Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic, sladky@utia.cas.cz

Policy π which selects at all times the same decision rule, i.e. $\pi \sim (f)$, is called stationary, hence following policy $\pi \sim (f)$ X is a homogeneous Markov chain with transition probability matrix $P(f)$ whose ij-th element is $p_{ij}(f_i)$. Then $r_i^{(1)}(f_i) := \sum_{j\in\mathcal{I}} p_{ij}(f_i) r_{ij}$ is the expected one-stage reward obtained in state i. Similarly, $r^{(1)}(f)$ is an N-column vector of one-stage rewards whose i-the elements equals $r_i^{(1)}(f_i)$. The symbol I denotes an identity matrix and e is reserved for a unit column vector.

Considering standard probability matrix $P(f)$ the spectral radius of $P(f)$ is equal to one. Recall that (the Cesaro limit of $P(f)$) $P^*(f) := \lim\limits_{n\to\infty} \frac{1}{n}\sum_{k=}^{n-1} P^k(f)$ (with elements $p^*_{ij}(f)$) exists, and if $P(f)$ is aperiodic then even $P^*(f) = \lim\limits_{k\to\infty} P^k(f)$ and the convergence is geometrical. Then $g^{(1)}(f) = P^*(f)\, r^{(1)}(f)$ is the (column) vector of average rewards, its $i$the entry $g_i^{(1)}(f)$ denotes the average reward if the process starts in state i. Moreover, if $P(f)$ is unichain, i.e. $P(f)$ contains a single class of recurrent states, then $p^*_{ij}(f) = p^*_j(f)$, i.e. limiting distribution is independent of the starting state and $g^{(1)}(f)$ is a constant vector with elements $\bar{g}^{(1)}(f)$. It is well-known (cf. e.g. [3, 7]) that also $Z(f)$ (fundamental matrix of $P(f)$), and $H(f)$ (the deviation matrix) exist, where $Z(f) := [I - P(f) + P^*(f)]^{-1}, H(f) := Z(f)\,(I - P^*(f))$.

Transition probability matrix $\tilde{P}(f)$ is called *transient* if the spectral radius of $\tilde{P}(f)$ is less than unity, i.e. it at least some row sums of $\tilde{P}(f)$ are less than one. Then $\lim_{n\to\infty}[\tilde{P}(f)]^n = 0$, $\tilde{P}^*(f) = 0$ $g^{(1)}(f) = \tilde{P}^*(f)\, r^{(1)}(f) = 0$ and $\tilde{Z}(f) = \tilde{H}(f) = [I - \tilde{P}(f)]^{-1}$. Observe that if $P(f)$ is stochastic and $\alpha \in (0,1)$ then $\tilde{P}(f) := \alpha P(f)$ is transient, however, if $\tilde{P}(f)$ is transient it may happen that some row sums may be even greater than unity. Moreover, for the so-called first passage problem, i.e. if we consider total reward up to the first reaching of a specific state (resp. the set of specific states), the resulting transition matrix is transient if the specific state (resp. the set of specific states) can be reached from any other state.

3 Reward Variance for Finite and Infinite Time Horizon

Let $\xi_n(\pi) = \sum_{k=0}^{n-1} r_{X_k,X_{k+1}}$ be the stream of rewards received in the n next transitions of the considered Markov chain X if policy $\pi = (f^n)$ is followed. Supposing that $X_0 = i$, on taking expectation we get for the first and second moments of $\xi_n(\pi)$

$$v_i^{(1)}(\pi,n) := \mathsf{E}_i^\pi(\xi_n(\pi)) = \mathsf{E}_i^\pi \sum_{k=0}^{n-1} r_{X_k,X_{k+1}}, \quad v_i^{(2)}(\pi,n) := \mathsf{E}_i^\pi(\xi_n(\pi))^2 = \mathsf{E}_i^\pi(\sum_{k=0}^{n-1} r_{X_k,X_{k+1}})^2. \tag{1}$$

It is well known from the literature (cf. e.g. [1],[3],[7],[8]) that for the time horizon tending to infinity policies maximizing or minimizing the values $v_i^{(1)}(\pi,n)$ for transient models, as well as policies maximizing or minimizing the value $g_i^{(1)}(\pi) = \lim_{n\to\infty} n^{-1} v_i^{(1)}(\pi,n)$ can be found in the class of stationary policies, i.e. there exist $f^*, \hat{f}, \bar{f}^*, \bar{f} \in \mathcal{F}$ such that for all $i \in \mathcal{I}$ and any policy $\pi = (f^n)$

$$v_i^{(1)}(f^*) := \lim_{n\to\infty} v_i^{(1)}(f^*,n) \ge \limsup_{n\to\infty} v_i^{(1)}(\pi,n), \quad v_i^{(1)}(\hat{f}) := \lim_{n\to\infty} v_i^{(1)}(\hat{f},n) \le \liminf_{n\to\infty} v_i^{(1)}(\pi,n), \tag{2}$$

$$g(\bar{f}^*) := \lim_{n\to\infty} \frac{1}{n} v_i^{(1)}(\bar{f}^*,n) \ge \limsup_{n\to\infty} \frac{1}{n} v_i^{(1)}(\pi,n), \;\; g(\bar{f}) := \lim_{n\to\infty} \frac{1}{n} v_i^{(1)}(\bar{f},n) \le \liminf_{n\to\infty} \frac{1}{n} v_i^{(1)}(\pi,n). \tag{3}$$

3.1 Finite Time Horizon

If policy $\pi \sim (f)$ is stationary, the process X is time homogeneous and for $m < n$ we write for the generated random reward $\xi_n = \xi_m + \xi_{n-m}$ (here we delete the symbol π and tacitly assume that $\mathrm{P}(X_m = j)$ and ξ_{n-m} starts in state j). Hence $[\xi_n]^2 = [\xi_m]^2 + [\xi_{n-m}]^2 + 2\cdot\xi_m\cdot\xi_{n-m}$. Then for $n > m$ we can conclude that

$$\mathsf{E}_i^\pi[\xi_n] = \mathsf{E}_i^\pi[\xi_m] + \mathsf{E}_i^\pi\Big\{\sum_{j\in\mathcal{I}} \mathrm{P}(X_m = j)\cdot \mathsf{E}_j^\pi[\xi_{n-m}]\Big\}. \tag{4}$$

$$\mathsf{E}_i^\pi[\xi_n]^2 = \mathsf{E}_i^\pi[\xi_m]^2 + \mathsf{E}_i^\pi\Big\{\sum_{j\in\mathcal{I}} \mathrm{P}(X_m = j)\cdot \mathsf{E}_j^\pi[\xi_{n-m}]^2\Big\} + 2\cdot \mathsf{E}_i^\pi[\xi_m]\sum_{j\in\mathcal{I}} \mathrm{P}(X_m = j)\cdot \mathsf{E}_j^\pi[\xi_{n-m}]. \tag{5}$$

In particular, from (2), (4) and (5) we conclude for $m = 1$

$$v_i^{(1)}(f, n+1) \;=\; r_i^{(1)}(f_i) + \sum_{j\in\mathcal{I}} p_{ij}(f_i)\cdot v_j^{(1)}(f,n) \tag{6}$$

$$v_i^{(2)}(f, n+1) \;=\; r_i^{(2)}(f_i) + 2\cdot \sum_{j\in\mathcal{I}} p_{ij}(f_i)\cdot\; r_{ij}\cdot v_j^{(1)}(f,n) + \sum_{j\in\mathcal{I}} p_{ij}(f_i)\, v_j^{(2)}(f,n) \tag{7}$$

where $r_i^{(1)}(f_i) := \sum_{j\in\mathcal{I}} p_{ij}(f_i)\; r_{ij}$, $\quad r_i^{(2)}(f_i) := \sum_{j\in\mathcal{I}} p_{ij}(f_i)[r_{ij}]^2$.

Since the variance $\sigma_i(f,n) = v_i^{(2)}(f,n) - [v_i^{(1)}(f,n)]^2$ from (6),(7) we get

$$\begin{aligned}\sigma_i(f,n+1) \;&=\; r_i^{(2)}(f_i) + \sum_{j\in\mathcal{I}} p_{ij}(f_i)\cdot\sigma_j(f,n) + 2\sum_{j\in\mathcal{I}} p_{ij}(f_i)\cdot\; r_{ij}\cdot v_j^{(1)}(f,n)\\ &\quad -[v_i^{(1)}(f,n+1)]^2 + \sum_{j\in\mathcal{I}} p_{ij}(f_i)[v_j^{(1)}(f,n)]^2 \end{aligned}\tag{8}$$

$$= \sum_{j\in\mathcal{I}} p_{ij}(f_i)[r_{ij} + v_j^{(1)}(f,n)]^2 - [v_i^{(1)}(f,n+1)]^2 + \sum_{j\in\mathcal{I}} p_{ij}(f_i)\cdot\sigma_j(f,n). \tag{9}$$

Using matrix notations (cf. [5]) equations (6),(7),(8) can be written as:

$$v^{(1)}(f,n+1) \;=\; r^{(1)}(f) + P(f)\cdot v^{(1)}(f,n) \tag{10}$$

$$v^{(2)}(f,n+1) \;=\; r^{(2)}(f) + 2\cdot P(f)\circ R\cdot v^{(1)}(f,n) + P(f)\;\cdot v^{(2)}(f,n) \tag{11}$$

$$\begin{aligned}\sigma(f,n+1) \;&=\; r^{(2)}(f) + P(f)\sigma(f,n) + 2\cdot P(f)\circ R\cdot v^{(1)}(f,n)\\ &\quad -[v^{(1)}(f,n+1)]^2 + P(f)\;\cdot [v^{(1)}(f,n)]^2\end{aligned} \tag{12}$$

where $R = [r_{ij}]_{i,j}$ is an $N\times N$-matrix, and $r^{(2)}(f) = [\; r_i^{(2)}(f_i)]$, $\quad v^{(2)}(f,n) = [v_i^{(2)}(f,n)]$, $v^{(1)}(f,n) = [(v_i^{(1)}(f,n)]$, $\sigma(f,n) = [\sigma_i(f,n)]$ are column vectors.
The symbol $\circ$ is used for Hadamard (entrywise) product of matrices. Observe that $r^{(1)}(f) = (P(f)\circ R)\cdot e$, $\quad r^{(2)}(f) = [P(f)\circ(R\circ R)]\cdot e$.

3.2 Infinite-Time Horizon: Transient Case

In this subsection we focus attention on transient models, i.e. we assume that the transition probability matrix $\tilde P(f)$ with elements $p_{ij}(f_i)$ is substochastic and $\rho(f)$, the spectral radius of $\tilde P(f)$, is less than unity.

Then on iterating (10) we easily conclude that there exists $v^{(1)}(f) := \lim\limits_{n\to\infty} v^{(1)}(f,n)$ such that

$$v^{(1)}(f) = r^{(1)}(f) + \tilde P(f)\cdot v^{(1)}(f) \Longleftrightarrow v^{(1)}(f) = [I - \tilde P(f)]^{-1} r^{(1)}(f). \tag{13}$$

Similarly, from (11) (since the term $2\cdot P(f)\circ R\cdot v^{(1)}(f,n)$ must be bounded) on letting $n\to\infty$ we can also verify existence $v^{(2)}(f) = \lim\limits_{n\to\infty} v^{(2)}(f,n)$ such that

$$v^{(2)}(f) = \; r^{(2)}(f) + 2\cdot\tilde P(f)\circ R\cdot v^{(1)}(f) + \tilde P(f)\, v^{(2)}(f) \tag{14}$$

hence

$$v^{(2)}(f) = [I - \tilde P(f)]^{-1}\;\Big\{\; r^{(2)}(f) + 2\cdot\tilde P(f)\circ R\cdot v^{(1)}(f)\Big\}. \tag{15}$$

On letting $n\to\infty$ from (8), (9) we get for $\sigma_i(f) := \lim\limits_{n\to\infty}\sigma_i(f,n)$

$$\begin{aligned}\sigma_i(f) \;&=\; r_i^{(2)}(f_i) + \sum_{j\in\mathcal{I}} p_{ij}(f_i)\cdot\sigma_j(f) + 2\sum_{j\in\mathcal{I}} p_{ij}(f_i)\cdot\; r_{ij}\cdot v_j^{(1)}(f)\\ &\quad -[v_i^{(1)}(f)]^2 + \sum_{j\in\mathcal{I}} p_{ij}(f_i)[v_j^{(1)}(f)]^2\end{aligned} \tag{16}$$

$$= \sum_{j\in\mathcal{I}} p_{ij}(f_i)[r_{ij} + v_j^{(1)}(f)]^2 - [v_i^{(1)}(f)]^2 + \sum_{j\in\mathcal{I}} p_{ij}(f_i)\cdot\sigma_j(f). \tag{17}$$

Hence in matrix notation

$$\sigma(f) = r^{(2)}(f) + \tilde{P}(f) \cdot \sigma(f) + 2 \cdot \tilde{P}(f) \circ R \cdot v^{(1)}(f) - [v^{(1)}(f)]^2 + \tilde{P}(f) \cdot [v^{(1)}(f)]^2. \tag{18}$$

After some algebra (18) can be also written as

$$\sigma(f) \quad = \quad [I - \tilde{P}(f)]^{-1} \cdot \{\, r^{(2)}(f) + 2 \cdot \tilde{P}(f) \circ R \cdot v^{(1)}(f) - [v^{(1)}(f)]^2\}. \tag{19}$$

In particular, for the discounted case, i.e. if for some discount factor $\alpha \in (0, 1)$ the transient matrix $\tilde{P}(f) := \alpha P(f)$ then (19) reads

$$\sigma(f) \quad = \quad [I - \alpha P(f)]^{-1} \cdot \{\, r^{(2)}(f) + 2 \cdot \alpha P(f) \circ R \cdot v^{(1)}(f) - [v^{(1)}(f)]^2\}. \tag{20}$$

(20) is similar to the formula for the variance of discounted rewards obtained by Sobel [6] using different methods.

3.3 Infinite-Time Horizon: Average Case

We make the following

Assumption 1. There exists state $i_0 \in \mathcal{I}$ that is accessible from any state $i \in \mathcal{I}$ for every $f \in \mathcal{F}$.

Obviously, if Assumption 1 holds then for every $f \in \mathcal{F}$ the transition probability matrix $P(f)$ is *unichain* (i.e. $P(f)$ have no two disjoint closed sets).

As well known from the literature (see e.g. [3]), if Assumption 1 holds, then the growth rate of $v^{(1)}(f, n)$ is linear and independent of the starting state. In particular, there exists constant vector $g^{(1)}(f) = P^*(f)r(f)$ (with elements $\bar{g}^{(1)}(f)$) along with vector $w^{(1)}(f)$ (unique up to an additive constant) such that

$$w^{(1)}(f) + g^{(1)}(f) = r(f) + P(f)w^{(1)}(f). \tag{21}$$

In particular, it is possible to select $w^{(1)}(f)$ such that $P^*(f)w^{(1)}(f) = 0$. Then $w^{(1)}(f) = H(f)r(f) = Z(f)r(f) - P^*(f)r(f)$. On iterating (21) we can conclude that

$$v^{(1)}(f, n) = g^{(1)}(f) \cdot n + w^{(1)}(f) + [P(f)]^n w^{(1)}(f). \tag{22}$$

To simplify the limiting behavior we make also

Assumption 2. The matrix $P(f)$ is aperiodic, i.e. $\lim_{n\to\infty}[P(f)]^n = P^*(f)$ exists for any $P(f)$.

Then for n tending to infinity $v^{(1)}(f,n) - ng^{(1)}(f) - w^{(1)}(f)$ tends to the null vector and the convergence is geometrical. In particular, by (22) we can conclude that for $\varepsilon(n) = P(f)^n w^{(1)}(f)$

$$v^{(1)}(f, n) = g^{(1)}(f) \cdot n + w^{(1)}(f) + \varepsilon(n). \tag{23}$$

The symbol $\varepsilon(n)$ is reserved for any column vector of appropriate dimension whose elements converge geometrically to the null vector.

Employing the above facts we can conclude that by (6),(21),(22)

$$\begin{aligned}
v_i^{(1)}(f, n+1) + v_j^{(1)}(f, n) \quad &= \quad r_i(f) + \sum_{k\in\mathcal{I}} p_{ik}(f) \cdot v_k^{(1)}(f, n) + v_j^{(1)}(f, n) \\
&= \quad r_i(f) + 2n\bar{g}^{(1)}(f) + \sum_{k\in\mathcal{I}} p_{ik}(f) w_k^{(1)}(f) + w_j^{(1)}(f) + \varepsilon(n) \\
&= \quad (2n+1)\bar{g}^{(1)}(f) + w_i^{(1)}(f) + w_j^{(1)}(f) + \varepsilon(n) \qquad (24) \\
v_i^{(1)}(f, n+1) - v_j^{(1)}(f, n) \quad &= \quad r_i(f) + \sum_{k\in\mathcal{I}} p_{ik}(f) \cdot v_k^{(1)}(f, n) - v_j^{(1)}(f, n) \\
&= \quad r_i(f) + \sum_{k\in\mathcal{I}} p_{ik}(f) w_k^{(1)}(f) - w_j^{(1)}(f) + \varepsilon(n) \\
&= \quad \bar{g}^{(1)}(f) + w_i^{(1)}(f) - w_j^{(1)}(f) + \varepsilon(n) \qquad (25)
\end{aligned}$$

From (23),(24),(25) we get

$$\begin{aligned}
\sum_{j\in\mathcal{I}} p_{ij}(f) \quad & [v_i^{(1)}(f,n+1) + v_j^{(1)}(f,n)][v_i^{(1)}(f,n+1) - v_j^{(1)}(f,n)] \\
= & \sum_{j\in\mathcal{I}} p_{ij}(f)[2n\bar{g}^{(1)}(f) + \bar{g}^{(1)}(f) + w_i^{(1)}(f) + w_j^{(1)}(f)][\bar{g}^{(1)}(f) + w_i^{(1)}(f) - w_j^{(1)}(f)] + \varepsilon(n) \\
= & \; 2n\bar{g}^{(1)}(f) \sum_{j\in\mathcal{I}} p_{ij}(f)[\bar{g}^{(1)}(f) + w_i^{(1)}(f) - w_j^{(1)}(f)] \\
+ & \sum_{j\in\mathcal{I}} p_{ij}(f) \left\{ [\bar{g}^{(1)}(f) + w_i^{(1)}(f)]^2 - [w_j^{(1)}(f)]^2 \right\} + \varepsilon(n) \\
= & \; 2n\bar{g}^{(1)}(f)\cdot r_i(f) + \sum_{j\in\mathcal{I}} p_{ij}(f) \left\{ [\bar{g}^{(1)}(f) + w_i^{(1)}(f)]^2 - [w_j^{(1)}(f)]^2 \right\} + \varepsilon(n). \qquad (26)
\end{aligned}$$

Similarly by (23) for the third term on the RHS of (8) (and also for the third term on the RHS of (12)), we have

$$\begin{aligned}
\sum_{j\in\mathcal{I}} p_{ij}(f)\cdot r_{ij}\cdot v_j^{(1)}(f,n) &= \sum_{j\in\mathcal{I}} p_{ij}(f)\cdot r_{ij}\cdot [n\cdot \bar{g}^{(1)}(f) + w_j^{(1)}(f) + \varepsilon(n)] \\
= \quad n\cdot \bar{g}^{(1)}(f)\cdot r_i(f) &+ \sum_{j\in\mathcal{I}} p_{ij}(f)\cdot r_{ij}\cdot w_j^{(1)}(f) + \varepsilon(n). \qquad (27)
\end{aligned}$$

Substitution from (26), (27) into (8) yields after some algebra

$$\begin{aligned}
\sigma_i(f,n+1) \quad = \quad & \sum_{j\in\mathcal{I}} p_{ij}(f)\cdot \sigma_j(f,n) + r_i^{(2)}(f) + 2\cdot \sum_{j\in\mathcal{I}} p_{ij}(f)\cdot r_{ij}\cdot w_j^{(1)}(f) \\
& + \sum_{j\in\mathcal{I}} p_{ij}(f)[w_j^{(1)}(f]^2 - [\bar{g}^{(1)}(f) + w_i^{(1)}(f)]^2 + \varepsilon(n) \\
= \quad & \sum_{j\in\mathcal{I}} p_{ij}(f)\cdot \{\sigma_j(f,n) + [r_{ij} + w_j^{(1)}(f)]^2\} - [\bar{g}^{(1)}(f) + w_i^{(1)}(f)]^2 + \varepsilon(n) \qquad (28)
\end{aligned}$$

Hence, in matrix form we have:

$$\sigma(f,n+1) = \sigma(f) + s(f) + \varepsilon(n), \qquad (29)$$

where elements $s_i(f)$ of the (column) vector $s(f)$ are equal to

$$\begin{aligned}
s_i(f) \quad &= \quad \sum_{j\in\mathcal{I}} p_{ij}(f)[r_{ij} + w_j^{(1)}(f)]^2 - [g^{(1)}(f) + w_i^{(1)}(f)]^2 \qquad (30) \\
&= \quad \sum_{j\in\mathcal{I}} p_{ij}(f)[r_{ij} + w_j^{(1)}(f) - g^{(1)}(f)]^2 - [w_i^{(1)}(f)]^2. \qquad (31)
\end{aligned}$$

Observe that by (31) follows immediately from (30) since by (21)
$-2\sum_{j\in\mathcal{I}} p_{ij}(f)(r_{ij} + w_j^{(1)}(f))g^{(1)}(f) - [g^{(1)}(f)]^2 \; = \; -2w_i^{(1)}(f)g^{(1)}(f) - [g^{(1)}(f)]^2$.
Employing (22) and the analogy between (9) and (29) we can conclude that

$$G(f) = \lim_{n\to\infty} \frac{1}{n}\sigma(f) = P^*(f)s(f) \qquad (32)$$

is the average variance corresponding to policy $\pi \sim (f)$.

4 Finding Optimal Policies

For finding second order optimal policies, at first it is necessary to construct the set of optimal transient or optimal average policies (cf. e.g. [1, 3, 7]). Since optimal policies can be found in the class of stationary policies, i.e. there exist $f^*, \bar{f}^* \in \mathcal{F}$ such that

$$v^{(1)}(f^*) \ge v^{(1)}(\pi) \quad \text{resp.} \quad g^{(1)}(\bar{f}^*) \ge g^{(1)}(\pi) \quad \text{for every policy} \;\; \pi = (f^n). \qquad (33)$$

Let $\mathcal{F}^* \subset \mathcal{F}$ be the set of all transient optimal stationary policies, $\bar{\mathcal{F}}^* \subset \mathcal{F}$ be the set of all average optimal stationary policies. Stationary optimal policies minimizing total or average variance can be constructed on applying standard policy or value iteration procedures in the class of policies from $\mathcal{F}^*$ or $\hat{\mathcal{F}}^*$.

5 Specific Example: Credit Management

The state of the bank is determined by the bank liabilities, i.e. deposits and the capital and is also influenced by the current state of the economy (cf. [7, 9]). It is a task for expert to evaluate each possible state of the bank by some value, say $i \in \mathcal{I}$, we assume that the set $\mathcal{I}$ is finite. A subset of the state space $\mathcal{I}$, say $\mathcal{I}^*$, is called optimal; the decision maker tries to reach this set. To this end at each time point the decision maker receives some money amount depending on the current state of the bank, say c_i, to improve the state of the bank. The decision maker has the following options:
1. advertise the activity of the bank,
2. assign small reward as a courtesy to the non-problematic credit holders,
3. warn and penalized the problematic credit holders.
Based on the experience of the bank, suitable advertising improves the state of the bank by reaching from state i some more suitable state $j \in \mathcal{I}$ with probability $p_{ij}(1)$. Similarly a courtesy reward in the total amount c_i can help to reach a more suitable state $j \in \mathcal{I}$ with probability $p_{ij}(2)$. Finally, warning and penalizing the problematic credit holders changes the state by reaching state j with probability $p_{ij}(3)$.

Using the above mentioned approach the problem of optimal credit-granting policy is formulated as a problem of finding optimal policy of a controlled Markov chain. Observe that the transient model can also grasp models with discount factor depending on the current state. Moreover, if the discount factor is very close to unity we try to optimize the long run average reward.

6 Conclusions

The article is devoted to second order optimality in transient and average Markov reward chains. To this end, formulas for the variance of total, discounted and average rewards are derived. In the class of optimal policies, procedures for finding policies with minimal total or average variance are suggested. Application of the obtained results for finding an optimal credit-granting policy of a bank is discussed.

Acknowledgements

This research was supported by the Czech Science Foundation under Grant 15-10331S and by CONACyT (Mexico) and AS CR (Czech Republic) under Project 171396.

References

[1] Mandl, P.: On the variance in controlled Markov chains. *Kybernetika* **7** (1971), 1–12.

[2] Markowitz, H.: *Portfolio Selection - Efficient Diversification of Investments.* Wiley, New York, 1959.

[3] Puterman, M.L.: *Markov Decision Processes – Discrete Stochastic Dynamic Programming.* Wiley, New York, 1994.

[4] Sladký, K.: On mean reward variance in semi-Markov processes. *Mathematical Methods of Operations Research 62* (2005), 387–397.

[5] Sladký, K.: Second order optimality in transient and discounted Markov Decision Chains. In: *Proceedings of the 33th International Conference Mathematical Methods in Economics 2015* (D. Martinčák et al., eds.), University of West Bohemia, Plzeň, Cheb 2015, pp. 731–736.

[6] Sobel, M.: The variance of discounted Markov decision processes. *J. Applied Probability* **19** (1982), 794–802.

[7] Bäuerle, N. and Rieder, U.: *Markov Decision Processes with Application to Finance.* Springer, Berlin, 2011.

[8] Veinott, A.F.Jr.: Discrete dynamic programming with sensitive discount optimality criteria. *Annals Math. Statistics* **40** (1969), 1635–1660.

[9] Waldmann, K.-H.: On granting credit in a random environment. *Mathematical Methods of Operations Research* **47** (2005), 99–115.

Interval data and sample variance: How to prove polynomiality of computation of upper bound considering random intervals?

Ondřej Sokol[1], Miroslav Rada[2]

Abstract. We deal with the computation of sample variance over interval data. Computing the upper bound of the sample variance is an NP-hard problem, however, there exist some efficient algorithms for various special cases. Using Ferson's algorithm, the computation of the maximal possible variance over interval-valued dataset can be realized in polynomial time in the maximal number of narrowed intervals intersecting at one point; narrowed means that the intervals are shrinked proportionally to the size of the dataset.
Considering random datasets, experiments conducted by Sokol allowed for conjecturing that the maximal number of narrowed intervals intersecting at one point is at most of logarithmic size for a reasonable choice of the data-generating process.
Here, we assume uniform distribution of centers and constant radii. Under this setting, we show that the computation of expected value of the maximal number of narrowed intervals intersecting at one point (which is random variable here) is reducible to the evaluation of the volume of some special simplicial cuts.

Keywords: Interval data, sample variance, simplicial cuts.

JEL classification: C44
AMS classification: 90C15

1 Introduction

Interval data. Assume an interval-valued dataset. We do not know the exact values of data. Instead of that, we know the intervals which these values belong to. This kind of uncertainty occurs naturally when we process rounded, censored or categorized data. Similar data-related issues arise also when we work with measurements which have known tolerances or when we work directly with intervals, e.g. interval predictions or intervals of minimal and maximal daily values in the stock market.

In our paper, we deal with the computation of statistics over interval data. When only interval data are at our disposal, we are interested in determining lower and upper bounds of selected statistics. From these lower and upper bounds, we can draw conclusions over the entire dataset.

Related work. Under interval uncertainty, even some of the basic statistics are not easy to compute. This paper is a contribution to the case of sample variance when the true mean is unknown. This statistic will be investigated in the next sections. However, other statistics have been studied too. For example t-ratio was studied by Černý and Hladík [2], entropy by Kreinovich [9] and Xiang et al. [13], higher moments by Kreinovich et al. [10] and others. Summaries of approaches to computing various statistics under interval uncertainty can be found in [7] and [11].

Goal. In this paper, we deal with the expected computational complexity of sample variance over random data. In our previous paper [3], a conjecture based on simulations experiment about computational complexity has been stated. According to this conjecture, the upper bound of computational complexity in average over *reasonable* random data is polynomial (for details, see Conjecture 3). The conjecture has not been proven yet. The goal of this paper is to outline an approach to do so by interconnecting nice computational geometry tools with the statistical background of the conjecture.

2 Problem statement

A general framework. Consider an unobservable one-dimensional dataset $x_1, \dots, x_n$ (for example, it can be a random sample from a certain distribution). We are given an (observable) collection of intervals $\boldsymbol{x}_1 = [\underline{x}_1, \overline{x}_1], \dots, \boldsymbol{x}_n = [\underline{x}_n, \overline{x}_n]$ such that it is guaranteed that

$$x_i \in \boldsymbol{x}_i, \quad i = 1, \dots, n. \tag{1}$$

[1]University of Economics in Prague, Department of Econometrics, Winston Churchill Square 4, CZ13067 Prague, Czech Republic, ondrej.sokol@vse.cz
[2]University of Economics in Prague, Department of Econometrics, Winston Churchill Square 4, CZ13067 Prague, Czech Republic, miroslav.rada@vse.cz

We emphasize that, given the observed values $\underline{x} = (\underline{x}_1, \ldots, \underline{x}_n)^{\mathrm{T}}$ and $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n)^{\mathrm{T}}$, the only information about the distribution of $x = (x_1, \ldots, x_n)^{\mathrm{T}}$ is that the axiom (1) holds a.s.

Let a statistic $S(x_1, \ldots, x_n)$ be given. Formally, $S : \mathbb{R}^n \to \mathbb{R}$ is a continuous function. Due to our weak assumptions, the main information (and in some sense the only information) we can infer about $S(x_1, \ldots, x_n)$ from the observable dataset $\underline{x}, \overline{x}$ is the lower and upper bound, respectively, of the form

$$\underline{S} = \min\{S(\xi) : \underline{x} \le \xi \le \overline{x}\}, \quad \overline{S} = \max\{S(\xi) : \underline{x} \le \xi \le \overline{x}\},$$

where the inequality "$\le$" between two vectors in $\mathbb{R}^n$ is understood componentwise.

Bounds $\underline{S}, \overline{S}$ for many important statistics S have been extensively studied in literature. Sometimes, the situation is easy: for the example of sample mean (i.e. $S \equiv \widehat{\mu} = \frac{1}{n}\sum_{i=1}^n x_i$) we immediately get

$$\underline{\widehat{\mu}} = \frac{1}{n}\sum_{i=1}^n \underline{x}_i, \quad \overline{\widehat{\mu}} = \frac{1}{n}\sum_{i=1}^n \overline{x}_i. \tag{2}$$

Sample variance under interval data. On the other hand, some statistics have been shown to be very hard to compute. In this paper, we will work with one of them: the sample variance. Sample variance is computed as

$$\widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \widehat{\mu})^2, \tag{3}$$

where the true mean μ is unknown and is replaced by the sample mean $\widehat{\mu} = \frac{1}{n}\sum_{i=1}^n x_i$. The computation of the upper and lower bounds, respectively, reduces to the optimization problems

$$\underline{\widehat{\sigma}^2} = \min_{x \in \mathbb{R}^n} \left\{ \frac{1}{n}\sum_{i=1}^n \left(x_i - \frac{1}{n}\sum_{j=1}^n x_j \right)^2 : \underline{x} \le x \le \overline{x} \right\}, \text{ and} \tag{4}$$

$$\overline{\widehat{\sigma}^2} = \max_{x \in \mathbb{R}^n} \left\{ \frac{1}{n}\sum_{i=1}^n \left(x_i - \frac{1}{n}\sum_{j=1}^n x_j \right)^2 : \underline{x} \le x \le \overline{x} \right\}. \tag{5}$$

It is obvious that computation of $\underline{\widehat{\sigma}^2}$ is a convex quadratic problem and thus it can be solved in weakly polynomial time. Furthermore, Ferson et al. [6] proposed another interesting method, yielding a strongly polynomial algorithm. Yet another strongly polynomial algorithm has been formulated in [13].

The computation of $\overline{\widehat{\sigma}^2}$ is an NP-hard problem. A proof of the NP-hardness can be found in [5]. Moreover, it is known to be inapproximable with an arbitrary absolute error, see Černý and Hladík [2]. The latter paper also gives a useful positive statement: computation of $\overline{\widehat{\sigma}^2}$ can be done in pseudopolynomial time.

The computational properties of $\overline{\widehat{\sigma}^2}$ have been studied extensively, see e.g. Ferson [6], Xiang [13], Dantsin [4]. The most interesting fact is that for many special cases the problem can be solved efficiently.

Our problem. In this paper, we address the question whether the efficiently solvable cases are "rare" or "frequent" in practice. In particular, the statistical approach usually assumes that the observed data $\underline{x}, \overline{x}$ are generated by an underlying random process. Then, the question is whether the hard instances—such as the instances resulting from NP-hardness proofs—occur with a high or low probability.

The paper by Ferson et al. [6] presents an algorithm for computation of $\overline{\widehat{\sigma}^2}$ with the property stated by Theorem 1. To fix notation, for an interval $\boldsymbol{x} = [\underline{x}, \overline{x}]$ we denote $x^C = \frac{1}{2}(\overline{x} + \underline{x})$ its center and $x^\Delta = \frac{1}{2}(\overline{x} - \underline{x})$ its radius. For an $\alpha \in (0, 1)$, let $\alpha\boldsymbol{x}$ denote the α-narrowed interval $[x^C - \alpha x^\Delta, x^C + \alpha x^\Delta]$.

Theorem 1. *There exists a polynomial $p(n)$ for which the following holds true: when $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are given such that*

every k-tuple of distinct indices $i_1, \ldots, i_k \in \{1, \ldots, n\}$ satisfies $[\underline{\widehat{\mu}}, \overline{\widehat{\mu}}] \cap \frac{1}{n}\boldsymbol{x}_{i_1} \cap \frac{1}{n}\boldsymbol{x}_{i_2} \cap \cdots \cap \frac{1}{n}\boldsymbol{x}_{i_k} = \emptyset, \quad (6)$

then the algorithm by Ferson et al. makes at most $2^k p(n)$ steps. [Here, $\underline{\widehat{\mu}}, \overline{\widehat{\mu}}$ are given by (2).]

Remark 2. The original paper presents a weaker statement, in particular the intersection property from (6) is written as $\frac{1}{n}\boldsymbol{x}_{i_1} \cap \frac{1}{n}\boldsymbol{x}_{i_2} \cap \cdots \cap \frac{1}{n}\boldsymbol{x}_{i_k} = \emptyset$.

So, we are interested in the maximum k for which the condition (6) holds. Let K_n denotes such k.

We assume that the centers $x_1^C, \ldots, x_n^C$ are sampled from a distribution Φ and that the radii $x_1^\Delta, \ldots, x_n^\Delta$ are sampled from a nonnegative distribution Ψ and the samples are independent. Then we are looking for estimated value of K_n.

Based on simulation experiments, Conjecture 3 has been stated (see also [3]):

Conjecture 3. *If Φ is a continuous distribution with finite first and second moments and its density function is limited from above and Ψ has finite first and second moments, then $\mathsf{E}(K_n) = O(\sqrt{\log n})$ and $\mathrm{Var}(K_n) = O(1)$.*

Remark 4. The paper [3] presented a much weaker formulation of Conjecture 3, namely it conjectured that $\mathsf{E}(K_n) = O(\log n)$; also no conjecture on variance of K_n was stated.

This Conjecture 3 has not been proven yet. To rephrase the goal, we outline the approach that could be helpful in proving the first part of hypothesis that $\mathsf{E}(K_n) = O(\sqrt{\log n})$ for suitable choice of the data generating processes Φ and Ψ.

3 Simplex as the space of gaps between centers

Convention. From now on, when we speak about i-th interval or about interval with index i, we mean the *narrowed* one, i.e. $\frac{1}{n}\boldsymbol{x}_i$.

We set Φ as uniform distribution on interval $(0, 1)$ and $\Psi = 1$. Therefore, the computational complexity is affected by the distance of centers of each other. We remind that the radius of narrowed interval decreases with n.

Firstly, since the distributions of centers x_i^C for all $i \in [n]$ are independent, we can w.l.o.g. assume that (7) holds:

$$x_1^C \leq x_2^C \leq \cdots \leq x_n^C. \tag{7}$$

Furthermore, we set $x_0^C = 0$ and $x_{n+1}^C = 1$

We denote the distance d_i of adjacent points x_i^C and x_{i-1}^C:

$$d_i = x_i^C - x_{i-1}^C, \quad i = 1, 2, \ldots, n+1. \tag{8}$$

From the definitions it follows that $\sum_{i=1}^{n+1} d_i = 1$ and $d_i \geq 0$ for every i. The Theorem 5 holds:

Theorem 5. *Assume $d = (d_1, \cdots, d_{n+1})$ is given by* (8).
1. *The support of the random variable d is the set $\{d \in \mathbb{R}^{n+1} : d \geq 0, \sum_{i=1}^{n+1} d_i = 1\}$.*
2. *The random variable d has Dirichlet distribution with parameters $\alpha = (1, \ldots, 1)$.*

Hence, random variable $d = (d_1, d_2, \ldots, d_{n+1})$ is uniformly distributed over an $(n+1)$-dimensional simplex $\mathcal{S}_{n+1}$ with vertices $a_1 = (1, 0, \ldots, 0), \ldots, a_{n+1} = (0, \ldots, 0, 1)$.

The simplex $\mathcal{S}_{n+1}$ has affine dimension n, since it lives in hyperplane $\sum_{i=1}^{n+1} d_i = 1$. It can be mapped to n-dimensional space using projection $d_{n+1} = 1 - \sum_{i=1}^{n} d_i$. The resulting simplex will be denoted by Δ_n. It reads

$$\Delta_n = \{d \in \mathbb{R}^n : d \geq 0, \sum_{i=1}^{n} d_i \leq 1\}. \tag{9}$$

This projection means no loss of information. The resulting simplex is fulldimensional.

Remark 6. Note, however, that we are not interested in d_1 and d_{n+1} as they do not affect the number of intersecting narrowed intervals.

4 Probability

Denote by $\mathrm{P}(K_n = k)$ the probability that, given a fixed n, the maximum number of intersecting intervals is k. Similarly we define $\mathrm{P}(K_n \geq k)$.

We are interested in the expected value of maximal number of intersecting intervals

$$\mathsf{E}(K_n) = \sum_{k=1}^{n} k\mathrm{P}(K_n = k) \tag{10}$$

for different n, in particular, we want to examine its behaviour when $n \to \infty$.

Observation 7 forms a basis for the rest of the text.

Observation 7. *Assume an index set $I \subseteq [n]$ with $|I| = k$ for some $k \in \mathbb{N}$ is given. The following holds:*

$$\bigcap_{i\in I} \frac{1}{n}\boldsymbol{x}_i \neq \emptyset \quad \Longleftrightarrow \quad \max_{i\in I} x_i^C - \min_{i\in I} x_i^C \leq \frac{2}{n} \quad \Longleftrightarrow \quad \sum_{i=1+\min\{j\in I\}}^{\max\{j\in I\}} d_i \leq \frac{2}{n}.$$

Roughly speaking, k intervals have a nonempty intersection if and only if the farthest centers (of these intervals) are close enough. Nevertheless, for us, it will be sufficient to consider sets of adjacent indices, i.e. $I \subseteq [n]$ such that $|I| = k$ and $\max\{i \in I\} - \min\{i \in I\} = k - 1$. An index set with this property will be called *connected*.

For computation of probability that all intervals in a connected index set I have a nonempty intersection (have a common point), we can use the geometric view introduced in the last section. We use the fact that differences between adjacent centers are uniformly distributed on Δ_n. We define $\mathcal{H}_I^n := \{d \in R^n : \sum_{j=1+\min\{i\in I\}}^{\max\{i\in I\}} d_j \leq \frac{2}{n}\}$. Since I is connected, the probability of nonempty intersection depends only on n and $k = |I|$, and neither on the minimal nor the maximal index in I. Therefore, definition of $\mathcal{H}_I^n$ can be simplified to $\mathcal{H}_k^n := \{d \in R^n : \sum_{j=2}^{k} d_j \leq \frac{2}{n}\}$. The halfspace $\mathcal{H}_k^n$ cuts off a part of Δ_n. The bigger the cut off is, the more probable is intersection of intervals in I.

We denote by $C(n, k)$ the set $\Delta_n \cap \mathcal{H}_k^n$. Analogously, we define $C(n, I)$. The following Lemma holds:

Lemma 8. *Assume we have n intervals in total. The probability that k adjacent intervals share a common point is denoted by* $\mathrm{P}(n, k)$ *and can be computed as*

$$\mathrm{P}(n,k) = \frac{\mathrm{Vol}(C(n,k))}{\mathrm{Vol}(\Delta_n)} = n!\,\mathrm{Vol}(C(n,k)). \tag{11}$$

The last equality follows from the fact that volume of Δ_n is $\frac{1}{n!}$.

The volume of $C(n, k)$. The volume of $C(n, k)$ can be computed or bounded in several ways. Gerber [8] and Varsi [12] studied the exact volume of simplex cut off. Gerber's work is especially interesting, since he proposed not only the formulas for volume of simplex cut off, but also the recursive formula for computing the fraction of volume of the simplex to be cut – the probability $\mathrm{P}(n, k)$ directly.

Since the recursive formula is not directly usable for our purpose, we formulated the following upper bound for the volume of $C(n, k)$.

Lemma 9. *The following holds:*

$$\mathrm{Vol}(C(n,k)) \leq \left(\frac{2}{n}\right)^k \frac{1}{(n-k+1)!(k-1)!}. \tag{12}$$

As the last instance, we mention excellent survey on algorithms for computation of volume of a polytope by Büeler et al. [1].

However, for our purposes, an analytic formula for volume would be the most appropriate. From this point of view, the only suitable formula is the upper bound stated by Lemma 9.

Two approaches to expressing $\mathrm{P}(K_n = k)$. Utilizing the formula for $P(n, k)$, there are two approaches to expressing $\mathrm{P}(K_n = k)$. We discuss them in Sections 4.1 and 4.2.

4.1 Inclusion-exclusion principle

We use the standard formula for computing probability of exact value of a variable: $\mathrm{P}(K_n = k) = \mathrm{P}(K_n \geq k) - \mathrm{P}(K_n \geq k + 1)$.

Then, note that $K_n \geq k$ occurs if and only if there are k adjacent intervals that share common point. Hence, we can consider all adjacent k-tuples of intervals, for each such k-tuple we cut off Δ_n and compute volume of the cut area:

$$\mathrm{P}(K_n \geq k) = n!\,\mathrm{Vol}\left(\bigcup_{I\subseteq[n],|I|=k} C(n,I)\right). \tag{13}$$

Clearly, one can't simply sum up the volumes of all the $C(n, I)$, since the cut off areas are overlapping. We shall utilize the inclusion-exclusion principle. However, we face the problem that we need to compute volume of much more complex polytopes, since the simplex is cut by more halfspaces (the number of halfspaces varies from 1 to $n-k+1$). Currently, we do not know how to even compute a tight upper bound for volume of the union of $C(n, I)$ in (13).

4.2 Complementary probability

Here, we take another approach. Instead of computing volume of union of cuts (formed by halfspaces in form $\mathcal{H}_I^n$), we compute volume of intersection of the opposite cuts and Δ_n. That allows us to get rid of the inclusion-exclusion framework, however, the main problem with "complex" polytope remains unchanged.

To be more formal, we define $\mathcal{H}_I^{-n}$ as the set $\{d \in RR^n : \sum_{i\in I} d_i \geq 2/n$ for connected index set I with $|I| = k$. Then,

$$\mathrm{P}(K_n < k) = n! \operatorname{Vol}\left(\Delta_n \bigcap \bigcap_{I\subseteq[n],|I|=k} \mathcal{H}_I^{-n}\right). \tag{14}$$

Then, $\mathrm{P}(K_n = k) = \mathrm{P}(K_n < k+1) - \mathrm{P}(K_n < k)$.

Note, that in (14) we need to compute the volume of simplex cut off by $n-k+1$ hyperplanes, i.e. we need an analytic expression of volume of a polytope given as an intersection $2n-k+2$ hyperplanes. As we also need send n to ∞, this task is untractable in full generality. We can only hope that the structure of the polytope is "special" and will allow for an efficient computation.

5 Conclusion

In this paper we presented an approach to compute the expected value of the maximal number of narrowed intervals that have a common point under specific stochastic setup. We connected computational geometry tools with the statistical background of the problem; our approach is based on geometric interpretation of the problem as we compute the probabilities using the volume of simplex cut-offs.

Although we sketched some ideas for the computation, the goal is still quite far away. Although the volume of simplex cut by one halfspace can be either computed directly or approximated (e.g. by Lemma 9), the evaluation of the volume of simplex cut by multiple halfspaces is computationally much more expensive task. Nevertheless, using the fact that distances between centers of intervals are uniformly distributed over n-dimensional simplex, the inclusion-exclusion principle seems to be a viable option.

Further research should focus on exact formulation of the probabilities using inclusion-exclusion principle, as sketched in Section 4.1.

Acknowledgements

The work of Ondřej Sokol was supported by the grant No. F4/63/2016 of the Internal Grant Agency of the University of Economics in Prague. The work of Miroslav Rada was supported by project GA ČR P403/16-00408S and by Institutional Support of University of Economics, Prague, IP 100040.

References

[1] Büeler, B., Enge, A., and Fukuda, K.: Exact Volume Computation for Polytopes: A Practical Study. In: *Polytopes — Combinatorics and Computation* (Kalai, G., and Ziegler, G. M., eds.). Birkhäuser Basel, Basel, 2000, 131–154.

[2] Černý, M., and Hladík, M.: The Complexity of Computation and Approximation of the t-ratio over One-dimensional Interval Data. *Computational Statistics & Data Analysis* **80** (2014), 26–43.

[3] Černý, M., and Sokol, O.: Interval Data and Sample Variance: A Study of an Efficiently Computable Case. In: *Mathematical Methods in Economics 2015* (Martinčík, D., Ircingová, J., and Janeček, P., eds.). University of West Bohemia, Plzeň, 2015, 99–105.

[4] Dantsin, E., Kreinovich, V., Wolpert, A., and Xiang, G.: Population Variance under Interval Uncertainty: A New Algorithm. *Reliable Computing* **12** (2006), 273–280.

[5] Ferson, S., Ginzburg, L., Kreinovich, V., Longpré, L., and Aviles, M.: Computing Variance for Interval Data is NP-hard. *ACM SIGACT News* **33** (2002), 108–118.

[6] Ferson, S., Ginzburg, L., Kreinovich, V., Longpré, L., and Aviles, M.: Exact Bounds on Finite Populations of Interval Data. *Reliable Computing* **11** (2005), 207–233.

[7] Ferson, S., Kreinovich, V., Hajagos, J., Oberkampf, W., and Ginzburg, L.: *Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty*. Sandia National Laboratories, 2007.

[8] Gerber, L.: The Volume Cut Off a Simplex by a Half-space. *Pacific Journal of Mathematics* **94** (1981), 311–314.

[9] Kreinovich, V.: Maximum Entropy and Interval Computations. *Reliable Computing* **2** (1996), 63–79.

[10] Kreinovich, V., Longpré, L., Ferson, S., and Ginzburg, L.: *Computing Higher Central Moments for Interval Data*. Technical report, University of Texas at El Paso, 2004.

[11] Nguyen, H. T., Kreinovich, V., Wu, B., and Xiang, G.: *Computing Statistics under Interval and Fuzzy Uncertainty*, volume 393 of *Studies in Computational Intelligence*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[12] Varsi, G.: The Multidimensional Content of the Frustum of the Simplex. *Pacific Journal of Mathematics* **46** (1973), 303–314.

[13] Xiang, G., Ceberio, M., and Kreinovich, V.: Computing Population Variance and Entropy under Interval Uncertainty: Linear-time Algorithms. *Reliable computing* **13** (2007), 467–488.

On the use of linguistic labels in AHP: calibration, consistency and related issues

Jan Stoklasa[1], Tomáš Talášek[2]

Abstract. Saaty's AHP method is a frequently used method in decision support. It provides its users with a simple and effective way to input necessary information concerning his/her preferences through pair-wise comparisons, as well as various consistency measures of the expressed preferences. The linguistic level attached to Saaty's fundamental scale provides a seemingly easy-to-use interface between the model (pair-wise comparison matrices) and its users. In this paper we focus on the linguistic level of Saaty's fundamental scale and analyze the consequences of its use as it was proposed by Saaty. We show that even in the context of the consistency condition as defined by Saaty, the linguistic level seems to be improperly defined and calibration of the linguistic labels might be in place. Also proper transformations into different languages rather than simple translations of the terms suggested in English are necessary. We argue, that in the context of the existence of a transformation between the multiplicative and additive form of Saaty's matrices it is reasonable to investigate whether the additive or the multiplicative form should be used to confirm (or appropriately set) the meanings of the linguistic terms. We illustrate on a practical example how the calibration of the meanings of linguistic terms in the additive form can affect the performance of the linguistic level in terms of consistency in the multiplicative form.

Keywords: linguistic label, linguistic scale, AHP, pairwise comparison, calibration, consistency.

JEL classification: C44
AMS classification: 91B74

1 Introduction

AHP is a widely used multiple criteria decision making tool proposed by T. L. Saaty [7, 8] and it has received much attention of practitioners and researchers since its introduction. Let us consider a single decision maker who needs to evaluate each alternative from a given set of n alternatives $\{A_1, \dots, A_n\}$ and let us assume that only one evaluation criterion is considered. This criterion can be either quantitative (measurable or numerical) or qualitative. It is convenient to express the preference structure on the set of alternatives by comparing pairs of alternatives and assessing which one is more preferred to the other and also assessing the strength of this preference.

The aim of the AHP is to obtain evaluations $h_1, \dots, h_n$ of these alternatives. Based on these evaluations a reciprocal square matrix H of the dimension $n \times n$ can be constructed, $H = \{h_{ij}\}_{i,j=1}^n$, such that $h_{ij} = h_i/h_j$ and obviously $h_{ij} = 1/h_{ji}$. The value h_{ij} then represents the relative preference of alternative i over alternative j and can be linguistically interpreted as A_i *being* h_{ij} *times more important than* A_j. The usual multiple-criteria decision-making problem is one of finding such evaluations $h_1, \dots, h_n$, as they are not known in advance. To do so an estimation of the matrix H, a reciprocal square matrix of preference intensities $S = \{s_{ij}\}_{i,j=1}^n$, $s_{ij} = 1/s_{ji}$, can be constructed by a decision maker. In case of more decision makers the matrices of preference intensities can be aggregated into a single overall

[1]Palacký University Olomouc, Department of Applied Economics, Křížkovského 12, 77180 Olomouc, jan.stoklasa@upol.cz.
[2]Palacký University Olomouc, Deptartment of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/12, 77146 Olomouc and Lappeenranta University of Technology, Skinnarilankatu 34, 53850 Lappeenranta, Finland, tomas.talasek@upol.cz.

matrix of intensities of preferences. This can be done for each element of this overall matrix by computing a geometrical mean of all the values provided by various decision makers for the respective pairwise comparison. Based on the matrix S the evaluations $h_1, \ldots, h_n$ can be computed as the arguments of minimum of the following expression $\sum_{i=1}^{n}\sum_{j=1}^{n}(s_{ij} - \frac{h_i}{h_j})^2$. The solution to this problem (the evaluations $h_1, \ldots, h_n$) can be found as the components of the eigenvector corresponding to the largest eigenvalue of S. Alternatively the logarithmic least squares method can be applied and the solutions found in the form $h_i = \sqrt[n]{\prod_{j=1}^{n} s_{ij}}$, $i = 1, \ldots, n$. The *consistency condition* for matrices of preference intensities suggested by Saaty is expressed as

$$s_{ik} = s_{ij} \cdot s_{jk}, \text{ for all } i, j, k = 1, 2, \ldots, n. \tag{1}$$

It is well known that using the values $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and their reciprocals in the matrix of preference intensities (see the Saaty's fundamental scale in Table 1), the consistency condition might not be fulfilled for expertly defined matrices of preference intensities, particularly if these are of larger order. In cases when "pure" consistency cannot be reached, Saaty defines the inconsistency index CI based on the spectral radius ($\lambda_{\max}$) of S by $CI = \frac{\lambda_{\max}-n}{n-1}$. For a perfectly consistent matrix $\lambda_{\max} = n$ and hence $CI = 0$. For other matrices Saaty defines the inconsistency ratio $CR = CI/RI_n$, where RI_n is a random inconsistency index computed as an average of inconsistency indices of randomly generated reciprocal matrices of preference intensities of the order n. As long as $CR < 0.1$ the matrix S is considered to be consistent enough.

In this paper we discuss one of the most important parts of the use of AHP (and methods based on matrices of intensities of preferences in general) - the input of preference-intensities. In Section 2 we recall the Saaty's fundamental scale and the linguistic interpretations of its numerical values and point out discrepancies between the numerical and linguistic level in the context of consistency of Saaty's matrix. Section 3 first provides an overview of the currently available calibration methods for the Saaty's scale, in Subsection 3.2 we propose a novel fuzzy preference relation based calibration and in Subsection 3.3 we discuss the new and original calibration methods in the context of a weakened consistency condition. Section 4 provides a discussion and summary of the presented topic.

2 The (fundamental) scale

The elements s_{ij} are estimations of the actual values of h_{ij} and are provided by the decision maker as answers to questions "*What is the intensity of preference of A_i over A_j for you according to the given criterion?*" or alternatively "*How much more do you prefer A_i over A_j?*".

s_{ij}	linguistic labels of the numerical intensities of preferences
1	alternative i is **equally preferred as** alternative j
3	alternative i is **slightly/moderately more preferred than** alternative j
5	alternative i is **strongly more preferred than** alternative j
7	alternative i is **very strongly more preferred than** alternative j
9	alternative i is **extremely/absolutely more preferred than** alternative j
2,4,6,8	correspond with the respective intermediate linguistic meanings obtained by joining the respective two linguistic labels $\mathcal{T}_k$ and $\mathcal{T}_l$ by "between" into the label "between $\mathcal{T}_k$ and $\mathcal{T}_l$"

Table 1: Saaty's scale - 9 numerical values of its elements for expressing preference (or indifference in case of 1) of alternative i over alternative j and their suggested linguistic labels.

We can observe that the questions are not easy to answer - at least not in a precise manner. We can not expect all decision makers to be able to provide answers such as "*A_3 is 3.56 more preferred than A_5*". Even if the decision-maker was able to provide such a precise answer, can we be sure that such a value is reliable? Saaty suggests the scale presented in Table 1 to be used to express preference intensities. Linguistic labels are suggested to represent five elements of the scale. There are several possible modes in which the intensity of preference can be provided - out of all these, two were originally suggested by Saaty - the linguistic mode and the numerical mode (see Table 1, which also summarizes their mutual relation). Huizingh and Vrolijk [3, p. 243] found that under their experimental design "*the numerical* *and the verbal mode are equally able to predict the ranking within the set of alternatives*" (note that they consider ranking, not the actual evaluations of alternatives). The verbal (linguistic) mode, however, results in higher inconsistency according to [3]. They also conclude [3, p. 245] that "using the verbal mode without knowing how people interpret the preference phrases leads to a small loss in decision quality" (again just rankings were considered). But Dyer and Forman [1] clearly warn against the combination of these two nodes in any set of judgements. This seems to suggest, that although both modes can work separately, their relation might not be as easy as it is presented in Table 1. We aim to add more arguments to support the claim that the assignment of linguistic labels to the numerical values of the fundamental scale is arbitrary and even contradicts the consistency condition set by Saaty. It is clear that the appropriateness of a linguistic label for a given preference-intensity (or vice versa the numerical meaning of a given linguistic expression) varies among the decision makers, is context dependent and can even be language specific. Since there is no one-to-one mapping between languages that would maintain meaning, we cannot accept the translations of Saaty's fundamental scale (its linguistic mode) into different languages to work with the same numerical values as meanings. It is questionable, whether the meanings of the translations maintain the same position (and relative position) on the universe $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. It may be that the differences in meaning among different language mutations of AHP are not too big, but not to check them and simply accept a translation of the linguistic scale does not follow any methodological good practice at all. Whether the OR researchers want to accept it or not, there is need for empirical confirmation that we are using the methods correctly - mainly with methods that employ a linguistic mode of description.

If we interpret the numerical values of the fundamental scale {1,2,3,4,5,6,7,8,9} and their reciprocals in the suggested way, that is as describing how many times is one alternative preferred to another one, we need to explain why there is no greater value than 9. There is in fact no natural maximum of the number of times something can be preferred to something else. This problem is even stressed when quantitative criteria are used and the ratio of their values is explicitly larger than 9 (say 10 EUR is ten times more than 1 EUR, but the restricted fundamental scale allows us to use 9 as a maximum value for any element in S). When we now assign a linguistic label "*extremely (or absolutely) preferred to*" to "*9 times more preferred than*", we can confuse the decision makers. What is more, using Table 1 we define "*3 times more preferred*" to be just "*slightly preferred*" which also does not seem to correspond with our intuition. It would seem that any calibration (setting appropriate labels for the numerical values of this scale) will be problematic, unless approached from a context where natural minima and maxima of preference-intensity ratios exist.

If we examine the linguistic mode of Saaty's fundamental scale in the context of his consistency condition (1), we run into further problems. To be compatible with the linguistic labels used in Saaty's scale, the condition should make sense also when we substitute the linguistic labels into it. Let us consider $s_{ik} = 3$ and $s_{kj} = 3$ then based on (1) we need $s_{ij} = 3 \cdot 3 = 9$. If we transform this into the linguistic level, we get if "A_i *is slightly more preferred than* A_k" and "A_k *is slightly more preferred than* A_j" then "A_i *is extremely/absolutely more preferred than* A_j". This is rather counterintuitive - we would expect a much smaller preference between A_i and A_j induced by two slight preferences. The consistency condition (1) is not well defined for the linguistic labels (or the linguistic labels are not well defined). In any cases if the decision maker provides information in linguistic form only and (1) is required, we declare as consistent something that is counterintuitive.

3 Possible solutions

It seems to us that there is no strong correspondence between the linguistic labels and the numbers that are supposed to represent their meaning (other than the fact that the linguistic labels can be ordered in the same way as their numerical counterparts). As was already mentioned, this might not be a big problem, if just one of the levels is used to obtain S - that is if the decision maker provides either just numbers, or just linguistic values. If, however, these two levels are combined and the decision maker has to deal with the fact that e.g. "9 times more = absolutely more", or "3 times more = slightly more" problems can occur due to the possible ambivalence of the assigned meaning. Also the resulting evaluations can be questionable, if the numerical values do not reflect the meaning of the linguistic labels well enough. The control of the decision maker over the process of obtaining evaluations fades in this case. There are at least two problems that need to be addressed - the discrepancy between the linguistic mode and the numerical mode, and the poor performance of linguistic labels in terms of the consistency condition and its "linguistic performance".

3.1 Calibration - achieving correspondence of the two discussed modes

Since the correct meaning of the linguistic labels of the intensities of preferences (e.g. "*strongly more preferred*") can depend on the decision-maker, on the type of decision making situation (context) and on the whole set of linguistic labels chosen for the description of preference intensities (including its cardinality and the choice of linguistic terms), the idea of customizing the meaning of the linguistic terms (expressed as a value from the fundamental scale {1,2,3,4,5,6,7,8,9} and the respective reciprocal values, or simply as values from the interval $[1/9, 9]$) has been addressed by several authors with varying success. Finan and Hurley [2] propose a calibration method that results in a geometric numerical scale providing meanings for the linguistic terms. They use the process of *transitive calibration*, which first specifies the number of the linguistic terms to be assigned numerical values (say four linguistic values: {*equally important, slightly more important, more important, much more important*} $= \{\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3, \mathcal{T}_4\}$; we are looking for their numerical equivalents $\{t_1, t_2, t_3, t_4\}$, $t_1 = 1$ by definition), in the second step they ask the decision-maker to decide the following: if (x is $\mathcal{T}_2$ than y) and (y is $\mathcal{T}_2$ than z), what is the relation of x to z - is it $\mathcal{T}_3$ or $\mathcal{T}_4$? At the end they require the decision maker to specify the meaning of (x is $\mathcal{T}_2$ than y) in terms of "x being k% more important than y". If the decision-maker specified, that (x is $\mathcal{T}_3$ than z), they get $t_2 = 1 + k$, $t_3 = (1 + k)^2$ and $t_4 = (1 + k)^4$. Note that they operate under the consistency condition expressed for the numerical values as $((s_{ij} > 1) \wedge (s_{jk} > 1)) \implies (s_{ik} > \max\{s_{ij}, s_{jk}\})$. This, however, seems to be too restrictive, since the combination of $\mathcal{T}_2$ and $\mathcal{T}_2$ does not need to be perceived as $\mathcal{T}_l$, where $l > 2$. In our opinion $\mathcal{T}_2$ should also be considered as a possible answer (see the concept of weak consistency further in the text).

Ishizaka and Nguyen [4] suggest a calibration procedure for qualitative criteria, that computes the numerical values of linguistically expressed intensities of preferences on the bases of pair-wise comparisons of geometrical figures with different areas (ordered in increasing order with respect to the area). The decision maker is asked to describe linguistically how big the difference between the areas of the given pair of objects is. The judgements are then compared with the real ratios of the areas of the objects and the numerical meaning of the linguistic terms is computed. Although this might seem as a simple and useful idea, there are several methodological issues connected with this solution - 1) there was never a higher ratio of the areas than 1 : 9 used in the study, 2) calibrating a scale for a qualitative criterion on a quantitative universe with different interpretation is very risky (since we know that the meanings of the linguistic terms are context-specific). Paper [4] however presents more evidence for the necessity of addressing the issue of calibration properly.

In both these approaches we can identify the problem of trying to calibrate the meanings of linguistic terms of a variable that might not have a natural maximum on a restricted universe (e.g. $[1/9, 9]$) . This can be bypassed by transforming this restricted universe into a different universe with natural minimum and maximum, with which the decision-maker can work reasonably well.

3.2 Calibration - a fuzzy preference relation approach

It is possible to transform a *multiplicative* (elements interpreted in terms of ratios/multiples) pairwise comparison matrix $S = \{s_{ij}\}_{i,j=1}^n$ into an *additive* (elements interpreted in terms of differences) pairwise comparison matrix $Z = \{z_{ij}\}_{i,j=1}^n$ using the following transformation (see e.g. [6]):

$$z_{ij} = \frac{1}{2}(1 + \log_9 s_{ij}). \tag{2}$$

The resulting matrix Z then carries the same information concerning the preferences of the decision maker. Z is additively reciprocal, that is $z_{ij} = 1 - z_{ji}$, $z_{ij} \in [0, 1]$ and $z_{ii} = 0.5$ for all $i, j = 1, \ldots, n$. As such the matrix Z can be interpreted as a fuzzy relation, its elements z_{ij} representing the degree of preference of A_i over A_j. Obviously, $z_{ij} = 0.5$ is interpreted as indifference between A_i and A_j, $z_{ij} = 1$ is interpreted as A_i is absolutely preferred to A_j, and $z_{ij} = 0$ is interpreted as A_j is absolutely preferred to A_i. The transformation formula (2) can be used in the context of fuzzy pairwise comparison matrices as well (see [5, 9]).

As we already know the meanings of the linguistic labels in the multiplicative case (see Table 1), we can now transform them into the additive representation using (2) and see, whether they "make sense" in the additive case, where a natural minimum and maximum of the degree of preference exists. Figure 1 summarizes the results of the transformation of the meanings of linguistic labels from the multiplicative case into the additive representation. We can see, that at least the meaning of the linguistic label "slightly preferred" seems to be misplaced - it is exactly half the way between indifference and extreme preference. We therefore suggest to construct the meanings of the available linguistic labels to be more intuitive (and thus making the linguistic labels more compatible with the multiplicative consistency condition) by constructing the meanings in the additive model instead. That is to define an appropriate meaning of each of the five linguistic labels used in Saaty's scale as a number from $[0.5, 1]$. Then these values are transformed back to the multiplicative universe (model) and a closest integer value from $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ would be assigned to them. The result of such an approach, when the numerical meanings are chosen to be equidistant in the additive universe $[0.5, 1]$ is summarized in Figure 2. We can see that after this modification even the multiplicative consistency condition seems to work better for the linguistic mode of the fundamental scale - if we again consider that "*A_i is slightly more preferred than A_k*" and "*A_k is slightly more preferred than A_j*" then the consistent result is "*A_i is between strongly and very strongly more preferred than A_j*" (numerically $s_{ik} = 2$ and $s_{kj} = 2$ and therefore $s_{ij} = 2 \cdot 2 = 4$) - this seems to us slightly closer to the intuitive expectation of the aggregated preference than in the original case. The uniform distribution of meanings of the five linguistic labels on $[0.5, 1]$ is just an example to illustrate the proposed approach. The meanings of these terms would have to be set in accordance with the understanding of these linguistic labels by the decision maker.

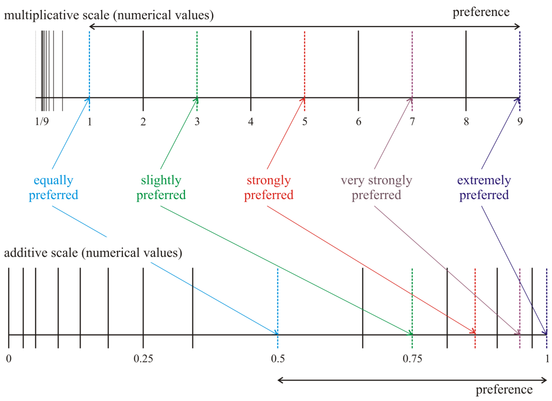


Figure 1: Transformation of the numerical values corresponding with the linguistic labels (in colour) and the intermediate numerical values and their reciprocals from Saaty's multiplicative scale into the values of the additive scale.

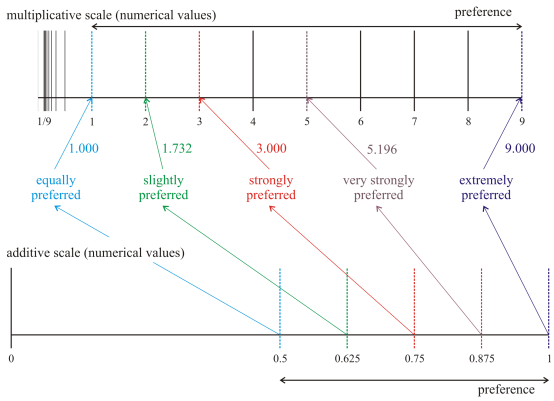


Figure 2: Transformation of the meanings of the linguistic labels of Saaty's scale that are considered to be uniformly distributed on $[0.5, 1]$ in the additive approach back to numerical values of the multiplicative scale. For each linguistic label an exact numerical value of its meaning after transformation is presented and the closest integer is assigned as its meaning in the multiplicative case.

3.3 Revision of consistency requirements

To use the linguistic level for inputs more safely, a weaker consistency condition (see definition 1), that reflects the linguistic labels well has been proposed in [12] and further elaborated in [10].

Definition 1. (Weak consistency condition [10]) Let $S = \{s_{ij}\}_{i,j=1}^{n}$ be a matrix of preference intensities. We say, that S is *weakly consistent*, if for all $i, j, k \in \{1, 2, \dots, n\}$ the following holds:

$$s_{ij} > 1 \wedge s_{jk} > 1 \implies s_{ik} \geq \max\{s_{ij}, s_{jk}\}; \tag{3}$$

$$(s_{ij} = 1 \wedge s_{jk} \geq 1) \vee (s_{ij} \geq 1 \wedge s_{jk} = 1) \implies s_{ik} = \max\{s_{ij}, s_{jk}\}. \tag{4}$$

The properties of this condition are discussed in more details in [9, 10, 11, 12]. It is important to note here, that this condition is reasonable on both the numerical and the linguistic level of description. In situation when "*A_i is slightly more preferred than A_k*" and "*A_k is slightly more preferred than A_j*" we just require A_i to be "*at least slightly more preferred than A_j*". Such a condition is obviously much weaker than (1). It can therefore be seen as a minimum requirement on the consistency of expertly defined matrices of preference intensities.

4 Conclusion

Defining the meanings of the linguistic labels of Saaty's scale in the additive model on a universe with natural minimum and maximum proposed in this paper seems reasonable - the decision maker is asked to simply find a fixed point on the interval $[0.5, 1]$ (that is between indifference and absolute preference) for each linguistic label. This way we are able to assign linguistic description to 5 real numbers representing the strength of a preference (we can also consider the intermediate linguistic terms thus obtaining 9 linguistically interpretable values). We can also consider the real number assigned to a linguistic term as its meaning to be a typical (best) representative of the given linguistic term and use fuzzy sets with kernels containing these typical values and define appropriate fuzzy scales to be able to linguistically label all the possible numerical values from the given scale. The issue of calibration of the linguistic mode of Saaty's scale remains open, but we hope that this paper identifies at least some of the possible paths to its solution.

Acknowledgements

Supported by the grant GA 14-02424S Methods of operations research for decision support under uncertainty of the Grant Agency of the Czech Republic and partially also by the grant IGA PrF 2016 025 of the internal grant agency of Palacký University Olomouc.

References

[1] Dyer, R. F., and Forman, E. H.: *An Analytic Approach to Marketing Decisions*. Prentice Hall International, New Jersey, 1991.

[2] Finan, J. S., and Hurley, W. J.: Transitive calibration of the AHP verbal scale. *European Journal of Operational Research*, **112**, 2, (1999), 367–372.

[3] Huizingh, E. K. R. E., and Vrolijk, H. C. J.: A Comparison of Verbal and Numerical Judgments in the Analytic Hierarchy Process. *Organizational Behavior & Human Decision Processes*, **70**, 3, (1997), 237–247.

[4] Ishizaka, A., and Nguyen, N. H.: Calibrated fuzzy AHP for current bank account selection. *Expert Systems with Applications*, **40**, 9, (2013), 3775–3783.

[5] Krejčí, J.: Additively reciprocal fuzzy pairwise comparison matrices and multiplicative fuzzy priorities. *Soft Computing* (2015), 1–16. DOI 10.1007/s00500-015-2000-2

[6] Ramík, J., and Vlach, M.: Measuring consistency and inconsistency of pair comparison systems. *Kybernetika*, **49** (3), (2013), 465–486.

[7] Saaty, T. L.: A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, **15**, 3, (1977), 234–281.

[8] Saaty, T. L.: *Fundamentals of Decision Making and Priority Theory With the Analytic Hierarchy Process*. RWS Publishers, Pittsburgh, 2000.

[9] Stoklasa, J.: *Linguistic models for decision support*. Lappeenranta University of Technology, Lappeenranta, 2014.

[10] Stoklasa, J., Jandová, V., and Talašová, J.: Weak consistency in Saaty's AHP - evaluating creative work outcomes of Czech Art Colleges Classification of works of art. *Neural Network World*, **23**, 1, (2013), 61–77.

[11] Stoklasa, J., Talášek, T., and Talašová, J.: AHP and weak consistency in the evaluation of works of art - a case study of a large problem. *International Journal of Business Innovation and Research*, **11**, 1, (2016), 60–75.

[12] Talašová, J., and Stoklasa, J.: A model for evaluating creative work outcomes at Czech Art Colleges. In *Proceedings of the 29th International Conference on Mathematical Methods in Economics - part II* (2011), University of Economics, Prague, Faculty of Informatics and Statistics, 698–703.

‘Soft’ consensus in decision-making model using partial goals method

Vojtěch Sukač[1] , Jana Talašová [2] , Jan Stoklasa[3]

Abstract. The paper deals with the method of reaching consensus in group decision-making under fuzziness proposed by Sukač et al. in 2015. This multi-expert (ME) multiple-criteria decision-making (MCDM) model utilizes fuzzy evaluations of absolute type and stems from the partial-goals tree evaluation methodology introduced by Talašová in 2003. The fuzzy evaluations of alternatives are in the form of fuzzy numbers on a universe representing the degree of fulfillment of a given goal and as such they express the fuzzy degrees of fulfillment of the given goal by the respective alternatives for each expert. Each expert is assigned a decision competence expressed by a fuzzy number. The group aggregation of expert evaluations is based on the idea of the ‘soft’ consensus introduced by Kacprzyk and Fedrizzi in 1986. The alternatives which are evaluated “good enough” by a “sufficient amount” of “important” experts are identified and those alternatives form the set of possible solutions. In this paper we briefly recall the steps of the multi-expert ‘soft’ consensus reaching method by Sukač et al. and discuss its performance on practical examples in comparison with one of the most commonly used ME-MCDM approaches based on the weighted average of expert fuzzy evaluations. The center of gravity based ordering is applied to select the best evaluated alternative in this case. This way the paper analyzes the performance of the ‘soft’ consensus based MCDM method in the context of a well known benchmark and points out the added value of the ‘soft’ consensus based MCDM model.

Keywords: Fuzzy, group decision-making, fuzzy weighted average, consensus reaching, soft consensus.

JEL classification: C44
AMS classification: 91B74

1 Introduction

We make lot of decisions every day, in most cases we decide intuitively. Usually this does not require a lot of time to think about nor does it require formal models and decision support. Sometimes, however, we have to make decisions whose consequences are serious (e.g. choosing the best candidate for a job, choosing the best area to live, etc.). In these situations, an appropriate decision-making model should be used. Important and strategic decisions also take into account the view of more parties. In companies strategic decision-making frequently involves more people in order to reach higher objectivity (and also to avoid wrong decisions caused by overlooking some of the important aspects of the decision-making situation). In such a case it is desirable for each of the individuals to have (slightly) different preferences. Sometimes, however, the assessments of multiple experts (decision-makers — DMs) are too different, which makes find an alternative acceptable for everyone difficult or even impossible. Due to this, the negotiation and consensus reaching processes have been studied quite frequently in the literature. In most of the papers dealing with consensus reaching in group decision-making the consensus reaching is based on the aggregation of fuzzy preference relations.

In this paper we recall and analyse the fuzzy group decision-making model proposed in [5, 6] which utilizes fuzzy evaluations of absolute type. These evaluations can be interpreted as the degrees of fulfillment of given goals [7]. Such evaluations are desirable, since they are not dependent on the set of

[1]Palacký University Olomouc, Department of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/12, 77146 Olomouc, vojta.sukac@gmail.com
[2]dtto, jana.talasova@upol.cz
[3]Palacký University Olomouc, Department of Applied Economics, Křížkovského 12, 77180 Olomouc, jan.stoklasa@upol.cz

alternatives and describe the acceptability of the alternatives (this also removes the *"one-eyed is king among the blind" effect* from the evaluation). Group aggregation of evaluations of alternatives is performed by a fuzzy weighted average operation with fuzzy weights (see e.g. [4]). Consensus reaching in the model analyzed in this paper is based on the idea that the choice of the best alternative should be made only among the alternatives that are good enough according to most of the relevant experts. Such pre-selection of alternatives is based on the idea of 'soft' consensus introduced by Kacprzyk and Fedrizzi [1, 2].

The paper is organized as follows. The next section summarizes the definitions of the concepts necessary to introduce the model. In the third section the decision-making model is described in brief. The fourth section contains two examples on which the results and the performance of the 'soft' consensus based model are compared with the standard approach based on the weighted average of expert fuzzy evaluations and the defuzzification by the center of gravity method.

2 Preliminaries

Let U be a nonempty set called *universe*. *Fuzzy set* A on U is determined by its *membership function* $\mu_A(x) : U \to [0,1]$, where $\mu_A(x)$ expresses the *degree of membership* of x in fuzzy set A — from 0 for "x definitely does not belong to A" to 1 for "x definitely belongs to A", through all the intermediate values. The family of all fuzzy sets on the universe U is denoted by $\mathcal{F}(U)$. Let A be a fuzzy set on U and $\alpha \in [0,1]$. The crisp set $A_\alpha = \{x \in U \mid \mu_A(x) \geq \alpha\}$ is called *the α-cut of fuzzy set A. The support of fuzzy set A* is a (crisp) set $\mathrm{Supp}(A) = \{x \in U \mid \mu_A(x) > 0\}$. *The kernel of fuzzy set A* is a (crisp) set $\mathrm{Ker}(A) = \{x \in U \mid \mu_A(x) = 1\}$.

In cases, when the support of A is a discrete set ($\mathrm{Supp}(A) = \{x_1, \ldots, x_k\}$), then the fuzzy set A can be denoted as $A = \{{}^{\mu_A(x_1)}/_{x_1}, \ldots, {}^{\mu_A(x_k)}/_{x_k}\}$. A special type of fuzzy sets whose universe is a subset of $\mathbb{R}$, the so called *fuzzy numbers*, can be defined in the following way. Let $U \subset \mathbb{R}$ be an interval. *A fuzzy number N* is a fuzzy set on the universe U which fulfills the following conditions: a) $\mathrm{Ker}(N) \neq \emptyset$; b) for all $\alpha \in (0,1]$, N_α are closed intervals; c) $\mathrm{Supp}(N)$ is bounded. The family of all fuzzy numbers on U is denoted by $\mathcal{F}_N(U)$.

Each fuzzy number N is determined by $N = \{[\underline{N}(\alpha), \overline{N}(\alpha)]\}_{\alpha\in[0,1]}$, where $\underline{N}(\alpha)$ and $\overline{N}(\alpha)$ is the lower and upper bound of the α-cut of fuzzy number N respectively, for $0 < \alpha \leq 1$ and $[\underline{N}(0), \overline{N}(0)]$ is the closure of the support of N, i.e. $[\underline{N}(0), \overline{N}(0)] = \mathrm{Cl}(\mathrm{Supp}(N))$. A trapezoidal fuzzy number N is determined by ordered quadruple $(n^1, n^2, n^3, n^4) \subset U^4$ of significant values of N satisfying $(n^1, n^4) = \mathrm{Supp}(N)$ and $[n^2, n^3] = \mathrm{Ker}(N)$. The membership function of a trapezoidal fuzzy number N is

$$\mu_N(x) = \begin{cases} \frac{x-n^1}{n^2-n^1} & \text{if } n^1 \leq x < n^2, \\ 1 & \text{if } n^2 \leq x \leq n^3, \\ \frac{n^4-x}{n^4-n^3} & \text{if } n^3 < x \leq n^4, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

N is called a *triangular fuzzy number* if $n^2 = n^3$. Closed real intervals and real numbers can be represented by special cases of trapezoidal fuzzy numbers. See e.g. [3] for more details.

This paper utilizes the linguistic approach to group decision-making problems. *Linguistic variable* [8] is the 5-tuple $(\mathcal{V}, \mathcal{T}(\mathcal{V}), U, G, M)$, where $\mathcal{V}$ is the name of the linguistic variable, $\mathcal{T}(\mathcal{V})$ is the set of its linguistic terms (the values of $\mathcal{V}$), $U \subset \mathbb{R}$ is the universe on which fuzzy numbers expressing meanings of these linguistic terms are defined, G is grammar used for generating of linguistic terms $\mathcal{T}(\mathcal{V})$ and M is a mapping that assigns to each term $\mathcal{C} \in \mathcal{T}(\mathcal{V})$ its meaning $C = M(\mathcal{C})$ (a fuzzy number on U).

3 Description of the model

For the purpose of this paper we consider the group decision-making problem addressed in [6] with a group $\boldsymbol{E} = \{E_1, \ldots, E_p\}$ of $p \geq 2$ individuals (experts) with possibly different competences and a set $\boldsymbol{X} = \{X_1, \ldots, X_n\}$ of $n \geq 2$ alternatives. Experts' competences are given by trapezoidal fuzzy numbers L^k, $k = 1, \ldots, p$ on the interval $[0,1]$, where the endpoints of this continuum have the following interpretations: 0 means an incompetent expert and 1 means a fully competent expert. Each expert provides fuzzy evaluations of each alternative X_i. These fuzzy evaluations H_i^k, $i = 1, \ldots, n$, $k = 1, \ldots, p$ are provided in the form of trapezoidal fuzzy numbers on the interval $[0, 1]$ and express the degree of fulfillment of the given goal by an alternative X_i according to expert E_k (0 is interpreted as goal not being fulfilled at all, 1 means complete fulfillment of the given goal). These evaluations can also be the result of a multiple-criteria assessment of the given alternative by the given expert. As such these evaluations are of absolute type, that is they are not dependent on the set of alternatives and describe the acceptability of the alternatives for each expert.

3.1 The benchmark model — standard approach based on weighted averaging and Center of gravity defuzzification

For each alternative X_i, its fuzzy evaluations H_i^k, provided by all experts are aggregated using the fuzzy weighted average operation with competences L^k as fuzzy weights to obtain group evaluation H_i, $i = 1, \ldots, n$. Then the group fuzzy evaluations of alternatives are defuzzified using the centre of gravity method

$$t_{H_i} = \frac{\int_0^1 x\mu_{H_i}(x)\mathrm{d}x}{\int_0^1 \mu_{H_i}(x)\mathrm{d}x} \tag{2}$$

and the alternatives are sorted in ascending order with respect to their center of gravity. The alternative with the highest centre of gravity is chosen as the best one. In the case of more alternatives with the same or very similar centre of gravity sharing the highest place on the list, all these are the result of decision-making model and any of them can be chosen as the optimal one. This approach is considered as standard approach and the results obtained from this approach will be compared with results obtained from the 'soft' consensus based model.

3.2 The 'soft' consensus-based model

The model proposed in [5] and further developed in [6] also assumes, that for each alternative X_i, its fuzzy evaluations H_i^k are provided by all experts. The subsequent computations are, however, based on the idea that the optimal alternative should be chosen among such alternatives which are good enough according to the meaning of a sufficient amount of important experts. Due to this aim is defined subset of the set of alternatives, which includes only such alternatives which are *good enough according to the given quantity of relevant experts*. This subset is called *the candidate set*. Note, that it can happen that there is no element in the candidate set (i.e. the candidate set is empty). In this case, no alternative is chosen (none of the alternatives is considered good enough — in the consensus sense — by the experts). This constitutes a fundamental difference to the standard approach that there some alternative is always chosen, even if it's completely unsatisfactory. Not to recommend any alternative in cases when none is good enough in the consensus sense is also well in line with the basic idea and purpose of evaluations of absolute type. The model can be summarized in the following steps (the exact procedure of defining the candidate set far exceeds the scope of this paper and its detailed description can be found in [5, 6]):

1. First a linguistic variable $\widehat{\mathcal{A}}$ with the linguistic term set $\{\widehat{\mathcal{A}}_1, \ldots, \widehat{\mathcal{A}}_5\} = \{excellent, good, acceptable, borderline, unacceptable\}$ is introduced to express the level of acceptance of alternatives by experts. Using the values of this linguistic variable and their fuzzy-number meanings a modified set of linguistic terms $\{\mathcal{A}_1, \ldots, \mathcal{A}_5\}$ expressing *"at least* $\widehat{\mathcal{A}}_r$*"*, where $\mathcal{A}_r, \widehat{\mathcal{A}}_r \in \mathcal{F}_N([0,1])$, $r = 1, \ldots, 5$, is defined.

2. The fuzzy sets F_{ri} of *experts suggesting* $\mathcal{A}_r$ *for* X_i, $F_{ri} \in \mathcal{F}(\{E_1, \ldots, E_p\})$, $r = 1, \ldots, 5$, $i = 1, \ldots, n$, are subsequently defined as $F_{ri} = \{{}^{\theta_{ri}^1}/_{E_1}, \ldots, {}^{\theta_{ri}^p}/_{E_p}\}$, where $\theta_{ri}^k = \sup_{x \in [0,1]} \{\min\{\mu_{A_r}(x), \mu_{H_i^k}(x)\}\}$, $k = 1, \ldots, p$.

3. The linguistic quantifier set $\{\widehat{\mathcal{Q}}_1, \ldots, \widehat{\mathcal{Q}}_4\} = \{almost\_all, more\_than\_half, about\_half, minority\}$ is introduced and the meanings of the derived linguistic quantifiers $\{\mathcal{Q}_1, \ldots, \mathcal{Q}_4\}$ expressing *"at least* $\widehat{\mathcal{Q}}_s$*"* are defined respectively by fuzzy numbers on $[0, 1]$. The linguistic term *important expert* is also introduced, with its meaning $B \in \mathcal{F}_N([0, 1])$. The fuzzy set of *important experts* $I \in \mathcal{F}(\{E_1, \ldots, E_p\})$ is defined as $I = \{{}^{\zeta_1}/_{E_1}, \ldots, {}^{\zeta_p}/_{E_p}\}$, where $\zeta_k = \sup_{x \in [0,1]} \{\min\{\mu_B(x), \mu_{L^k}(x)\}\}$, $k = 1, \ldots, p$.

4. The truth value of the statement "the alternative X_i is $\mathcal{A}_r$ (e.g. *at least good*) with respect to the opinion of quantity ($\mathcal{Q}_s$) of important experts (I)" (denoted $\xi_i^{r,s}$) is determined, $i = 1, \ldots, n$, $r = 1, \ldots, 5$, $s = 1, \ldots, 4$ (see [6] for details).

5. For all r, s the set $\Upsilon_{r,s}$ is defined, which includes such alternatives which are $\mathcal{A}_r$ according to $\mathcal{Q}_s$ of important experts, $\Upsilon_{r,s} = \{X_i \in \boldsymbol{X} \mid \xi_i^{r,s} = 1\}$.

6. The order $\mathcal{K}$ in which to scan the sets $\Upsilon_{r,s}$ for non-emptyness is defined, e.g. $\mathcal{K} = \{(1,1), (1,2), \ldots, (2,3)\}$, where each pair of values represents a given combination of r and s. In our case we suggest to check first the set of alternatives considered (*excellent* by *almost all*) then the set of alternatives considered (*excellent* by *majority*), ..., and the last set we check for non-emptyness is the set of alternatives that are (*at least good* by *at least half*). The *candidate set*, which is the first non-empty set $\Upsilon_{r,s}$ according to $\mathcal{K}$ (denoted Υ^*) includes those alternatives, among which the most promising one is to be chosen.
   Note that $\mathcal{K}$ can be defined to include only such sets that are still relevant for the given problem. E.g. looking for a set of alternatives that are at least borderline might not be appropriate if we are not willing to accept borderline alternatives. In this case pairs $(3, s)$ and $(4, s)$ will not be present in $\mathcal{K}$ for any value of s. Similarly we can reflect our requirements of a sufficient amount of experts that agree on the evaluation.

4 Examples — comparison of the 'soft' consensus model with the benchmark

In this section two examples will be presented. Our aim is to show that in situations when there is no significant discrepancy among the evaluations provided by the experts, there is not much difference in the results provided by both methods (see Example 1). However, in situations when there are large differences in evaluations of alternatives among experts, both approaches diverge and provide different results — the standard approach aiming for an "ideal compromise" while the 'soft' consensus model aims for an alternative sufficiently acceptable by a sufficient amount of important experts (see Example 2).

4.1 Example 1 — no discrepancy in the evaluations among the experts

A company has decided to buy a new camera. It has a choice of 8 different cameras in the price category from 10 up to 20 thousands CZK. Their assessments have been provided by 6 experts and it can be found in Table 1 (the four numbers in each row represent the four significant values of the respective trapezoidal fuzzy number evaluation). Experts were assigned competences based on their position in the company and their experience with photographing. Competences are expressed by fuzzy numbers on the interval $[0, 1]$, where 0 is a definitely incompetent expert (his/her involvement is only an effort of an unbiased view) and 1 is a fully competent expert who has either a high position in the company, or he/she is a very experienced photographer. Comptences of experts are as follows: $L^1 = (0.54, 0.62, 0.69, 0.77)$, $L^2 = (0.23, 0.31, 0.38, 0.46)$, $L^3 = (0.69, 0.77, 0.85, 0.92)$, $L^4 = (0.54, 0.62, 0.69, 0.77)$, $L^5 = (0.38, 0.46, 0.54, 0.62)$, $L^6 = (1, 1, 1, 1)$. Effort of the management of the company was to choose such camera that the greatest part of important experts designated as the most suitable.

In the standard approach, group evaluations are computed using fuzzy weighted average operation and alternatives are compared using the center of gravity of these overall group fuzzy evaluations. The centres of gravity of group evaluations are listed in Table 2. It is clear that *camera7* is evaluated the best and the company will buy this camera. In the 'soft' consensus-based model, we compute the acceptability of each alternative by a given quantity of experts and thus the candidate set is defined. The candidate set constitutes of alternatives evaluated as *good or better* by *majority* of the important experts, which contains three alternatives, namely *camera1, camera2, camera7.* Centres of gravity of group evaluations of these 3 alternatives were calculated:

$$t_{H_1} = 0.6355, \qquad t_{H_2} = 0.6629, \qquad t_{H_7} = 0.6942,$$

and the alternative *camera7* was chosen as the optimal one. Note, that the ordering of the alternatives and the centers of gravity are the same, the only difference being that only *camera1, camera2* and *camera7* are considered as "first choice" viable candidates for the best alternative in the 'soft' consensus-based approach.

alternative	Expert 1				Expert 2				Expert 3			
camera1	0.43	0.59	0.76	0.84	0.23	0.36	0.50	0.65	0.20	0.33	0.49	0.64
camera2	0.41	0.57	0.74	0.82	0.26	0.39	0.55	0.70	0.19	0.29	0.47	0.62
camera3	0.26	0.39	0.52	0.64	0.22	0.35	0.48	0.61	0.06	0.14	0.28	0.40
camera4	0.27	0.37	0.55	0.69	0.49	0.64	0.80	0.86	0.08	0.18	0.36	0.51
camera5	0.19	0.32	0.46	0.58	0.19	0.32	0.46	0.59	0.13	0.26	0.41	0.54
camera6	0.52	0.67	0.81	0.91	0.46	0.59	0.72	0.84	0.25	0.40	0.58	0.74
camera7	0.62	0.76	0.92	0.97	0.46	0.59	0.75	0.85	0.31	0.45	0.60	0.73
camera8	0.49	0.61	0.74	0.86	0.37	0.50	0.64	0.77	0.15	0.29	0.45	0.58
alternative	Expert 4				Expert 5				Expert 6			
camera1	0.53	0.67	0.81	0.90	0.53	0.70	0.89	0.95	0.53	0.70	0.87	0.92
camera2	0.65	0.79	0.93	0.97	0.56	0.74	0.92	0.97	0.58	0.74	0.91	0.95
camera3	0.41	0.54	0.67	0.79	0.27	0.40	0.58	0.76	0.22	0.36	0.50	0.64
camera4	0.22	0.37	0.53	0.65	0.15	0.30	0.48	0.64	0.16	0.25	0.43	0.58
camera5	0.27	0.39	0.52	0.64	0.35	0.47	0.63	0.79	0.18	0.30	0.45	0.60
camera6	0.40	0.55	0.71	0.81	0.49	0.67	0.84	0.91	0.30	0.40	0.59	0.72
camera7	0.58	0.71	0.86	0.91	0.69	0.82	0.95	0.98	0.43	0.57	0.75	0.83
camera8	0.37	0.51	0.66	0.79	0.60	0.74	0.89	0.96	0.26	0.35	0.54	0.70

Table 1 Evaluations of eight cameras by each expert.

t_{H_1}	t_{H_2}	t_{H_3}	t_{H_4}	t_{H_5}	t_{H_6}	t_{H_7}	t_{H_8}
0.6355	0.6629	0.4276	0.4080	0.4094	0.5967	0.6942	0.5480

Table 2 Centres of gravity of group evaluations of eight cameras in the benchmark approach.

In this example we have shown that the 'soft' consensus-based model provides the same result as the standard approach in cases when there is not a significant discrepancy between the evaluations of the experts. It can therefore be considered as a substitute tool for the standard approach in these circumstances. In addition, it provides an information about how much are experts satisfied with the result. We know, that *the majority of the important experts finds the given three alternatives good or better* — a linguistic summary of the decision making problem and the strength of the consensus is therefore automatically available as an additional piece of information that is easy to use and understand.

4.2 Example 2 — evaluations of alternatives differ among experts

The family of four looking for a holiday destination. Evaluations of alternatives are presented in Table 3. Comptences of experts (family members) are as follows: $L^1 = (1, 1, 1, 1)$, $L^2 = (0.85, 0.92, 1, 1)$, $L^3 = (0.69, 0.77, 0.85, 0.92)$, $L^4 = (0.69, 0.77, 0.85, 0.92)$.

alternative	Expert 1				Expert 2				Expert 3				Expert 4			
destination1	0.27	0.41	0.57	0.73	0.22	0.33	0.49	0.62	0.08	0.18	0.35	0.51	0.64	0.77	0.90	0.96
destination2	0.16	0.27	0.45	0.62	0.39	0.52	0.67	0.78	0.15	0.24	0.43	0.60	0.64	0.77	0.90	0.96
destination3	0.25	0.37	0.56	0.73	0.23	0.33	0.48	0.61	0.15	0.24	0.43	0.60	0.57	0.71	0.85	0.91
destination4	0.61	0.74	0.88	0.95	0.15	0.25	0.39	0.53	0.10	0.17	0.35	0.50	0.24	0.36	0.49	0.61
destination5	0.20	0.35	0.50	0.63	0.34	0.47	0.61	0.74	0.19	0.29	0.47	0.62	0.29	0.41	0.55	0.69
destination6	0.44	0.60	0.77	0.86	0.14	0.21	0.36	0.51	0.14	0.28	0.43	0.58	0.70	0.83	0.95	0.98
destination7	0.18	0.31	0.47	0.63	0.66	0.79	0.92	0.96	0.07	0.15	0.33	0.51	0.22	0.33	0.48	0.62
destination8	0.17	0.28	0.44	0.58	0.21	0.29	0.45	0.60	0.70	0.84	0.97	0.99	0.22	0.31	0.47	0.63

Table 3 Evaluation of eight destinations by each family member.

In the standard approach, group evaluations are calculated using fuzzy weighted average operation and alternatives are compared by the center of gravity method. The centres of gravity of group evaluation are presented in the top row of Table 4. It's clear from the table that *destination6* is chosen as the best, since it's center of gravity of group evaluation is the highest.

In our consensus-based model, for each destination is computed acceptability by a given quantity of experts and the candidate set is defined. The candidate set contains only one alternative, namely *destination5*, which has been evaluated as *acceptable or better* by *majority* of the important experts. Whereas *destination6* is labelled as *at least good* by *about half* of important experts. Note that in this case the worst alternative according to the standard approach is selected as the optimal one. Let us consider all experts have assigned full competence (that is (1,1,1,1)). From the Table 4 it is clear that according to the standard approach the order of alternatives is slightly different but the consensus-based model again recommends the alternative with the lowest (or one of the lowest) group evaluation.

competence	t_{H_1}	t_{H_2}	t_{H_3}	t_{H_4}	t_{H_5}	t_{H_6}	t_{H_7}	t_{H_8}	optimal
different	0.4975	0.5261	0.4971	0.4732	0.4606	0.5440	0.4864	0.4972	5
fully	0.5019	0.5342	0.5022	0.4596	0.4603	0.5477	0.4781	0.5104	5

Table 4 Centers of gravity of group evaluations of destinations.

5 Conclusion

In this paper we have analyzed the 'soft' consensus model proposed by Sukač et al. [6] and its performance in comparison with the standard ME-MCDM approach based on the (fuzzy) weighted average of partial evaluations and the center of gravity defuzzification method applied to the overall evaluations to obtain the ordering of the alternatives. The first example shows that if experts evaluate alternatives broadly similar, we obtain the same (or similar) results as in the standard ME-MCDM approach. This suggests that in "standard" situations when there is no apparent conflict between the evaluations provided by the experts, both models agree on the conclusion. The 'soft' consensus based approach, however, provides also an additional piece of information concerning the degree of consensus of the evaluators. The second example shows that in the case of different opinions of experts, the result of the 'soft' consensus-based model can be very different from the benchmark approach — even such alternatives that are evaluated very bad according to the standard approach can be recommended as the optimal solution of the problem. The requirement of consensus is taken into account in this case, hence alternatives with ambivalent evaluations cannot be recommended. The 'soft' consensus-based model is not looking for an "ideal compromise" as the standard approach does, but rather for a solution that is sufficiently acceptable for sufficient amount of important experts. It therefore aims closer to a non-conflict solution than the standard approach. This clearly shows that the standard approach is suitable for different problems, when a "compromise" solution is acceptable and the consensus is not required. In situations when consensus (sufficient agreement of a sufficient amount of experts) is of importance the 'soft' consensus model can provide more appropriate solutions than the standard approach.

References

[1] Kacprzyk, J., and Fedrizzi, M.: 'Soft'consensus measures for monitoring real consensus reaching processes under fuzzy preferences. *Control and Cybernetics* **15** (1986), 309–323.

[2] Kacprzyk, J., Fedrizzi, M., and Nurmi, H.: 'Soft'degrees of consensus under fuzzy preferences and fuzzy majorities. In: *Consensus under Fuzziness* (J. Kacprzyk et al., eds.), Springer Science+Business Media, New York, 1997, 55–83.

[3] Klir, G.J., and Yuan, B.: *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall PTR, 1995.

[4] Pavlačka, O.: Modeling uncertain variables of the weighted average operation by fuzzy vectors. *Information Sciences* **181** (2011), 4969–4992.

[5] Stoklasa, J. et al.: Soft consensus model under linguistically labelled evaluations. In: *Proceedings of the 33rd International Conference on Mathematical Methods in Economics 2015*, University of West Bohemia, Plzeň, 2015, 743–748.

[6] Sukač, V., Talašová, J., and Stoklasa, J.: A linguistic fuzzy approach to the consensus reaching in multiple criteria group decision-making problems. *Acta Universitatis Palackianae Olomucensis. Facultas Rerum Naturalium Mathematica* **55** (2016), (in press).

[7] Talašová, J.: *Fuzzy metody vícekriteriálního hodnocení a rozhodování*. Vydavatelství UP, Olomouc, 2003, (in Czech).

[8] Zadeh, L., A.: The concept of a linguistic variable and its application to approximate reasoning-I. *Information Sciences* **8** (1975), 199–249.

Modelling of an asymmetric foreign exchange rate commitment in the Czech economy. DSGE model with constraints.

Karolína Súkupová[1], Osvald Vašíček[2]

Abstract. In this paper we analyse the use of the foreign exchange rate as an instrument of monetary easing. The asymmetric commitment used by the Czech National Bank is modelled as a constraint in a DSGE model. The model we used for our analysis is based on the concept of Justiniano, Preston (2010) and we redesigned it to use nominal exchange rate in a uncovered interest parity (UIP) condition. We estimated the model using a set of the Czech and Euro area data series and then performed a simulation using the method and tools proposed and developed in Holden (2016). Our aim was to show the impact of a long-term use of asymmetric exchange rate commitment into the economy. To answer this question we analyse the simulated trajectories of an endogenous variables that represent the development of economy.

In the first part we described the model structure and we compared the different specifications of UIP condition we used for the analysis. Then we provided a brief characteristic of a method we used to implement constraints into a DSGE model. The final section provides discussion of the obtained results.

Keywords: Zero Lower Bound, unconventional monetary policy, constraints in DSGE models.

JEL classification: C32, E17
AMS classification: 91B51, 91B64

1 Introduction

In the aftermath of the 2008 – 2009 recession, monetary authorities lost their ability to adjust the interest rates to fulfil the monetary policy targets. The persistent low price levels and negative price shocks pushed the economies towards the deflation and the need of the new instruments of the monetary policy restored the discussion among both academia and central bankers.

The possible ways to deal with the liquidity trap were discussed in the early 2000s, for example in Eggertson, Woodford [3]. The papers reacted to the Japanese stagnation and exceptionally low policy rates in the US economy. One sort of the recommendations for the optimal policy is based on the operations that increase the monetary base (quantitative easing, well conducted fiscal stimulation) and it emphasizes the importance of the right management of the public expectations. Some authors tried to find another monetary policy instrument that could directly substitute the policy rates and the nominal exchange rate was natural candidate. The intervention to foreign exchange rate and whole process of escaping from the zero lower bound and deflationary trap is described in Svensson [10].

The Czech National Bank decided to use the foreign exchange rate as an additional instrument of the monetary policy in November of 2013. The bank board reacted to the secondary recession and consequent low inflation rate that was induced both by low domestic demand and low foreign inflation rate. The bank board declared so called asymmetric commitment, bank operates on the foreign exchange rate market to hold the FX rate above some declared value, in this case 27 CZK/EUR.

[1]Masaryk University, Faculty of Economics and Administration, Department of Economics, Lipová 41a, 602 00 Brno, 324246@mail.muni.cz.
[2]Masaryk University, Faculty of Economics and Administration, Department of Economics, Lipová 41a, 602 00 Brno, osvald@econ.muni.cz.

The mechanism of the depreciation pass through is based on the import price channel. The prices of imported goods rise, as the change in FX rate changes the relative prices of domestic and foreign goods. The relatively cheaper domestic goods cause higher domestic demand for domestic goods and that consequently increases the price of domestic production. Overall price level rises, which changes the general price level expectations. The secondary result of the depreciation is the stimulation of domestic production both by higher domestic and foreign demand. This scenario is linked to unexpected change in foreign exchange rate and in the long time period the relative prices should accommodate the depreciation. The aim of this paper is to analyse the impact of the asymmetric commitment, i.e. repeated intervention that holds the nominal exchange rate above declared level.

2 Model

The base of the model that was used for our analysis was derived in [6] and [1]. The structure of the model consists of households, two stages of domestic producers and domestic importers and the monetary authority. Households offer their labour to domestic intermediate goods producers and set their price of labour in a Calvo manner. The intermediate goods producers use the labour as the only input and set their prices in a same way as households. The final goods producers buy intermediate goods from the previous group of firms and process these inputs into the products that are offered to households. The prices of final goods are also staggered. The importers buy products from abroad and sell them to households for prices that are directly linked to foreign price level, but also staggered in a Calvo manner. Monetary authority sets the nominal interest rates following the Taylor type monetary rule that includes both inflation and price level. The decision to include price level targeting was motivated by Holden in [5], who incorporated price level to monetary rule to improve stability of the algorithm. The foreign economy is modelled as a closed version of domestic economy.

The original version of a model included an UIP condition with real exchange rate and a risk premium that depends on a net foreign position of a country. For our purposes we changed the condition and related equations to use nominal exchange rate and the condition has following form

$$r_t - r_t^* = E_t e_{t+1} - e_t - risk_t,$$

where r_t and r_t^* denote domestic and foreign nominal interest rate, e_t denotes a nominal exchange rate, and $risk_t$ is the risk premium that is modelled as

$$risk_t = \mu_{risk,t} + \chi * nfa_t,$$

where $\mu_{risk,t}$ is shock in a risk premium modelled as an AR(1) process and nfa_t denotes a net foreign asset position, parameter χ is a risk premium elasticity.

The second UIP condition specification is following Montoro, Ortiz in [9] and takes form

$$E_t e_{t+1} - e_t = r_t - r_t^* + \gamma\sigma^2(\varpi_t^* + \varpi_t^{cb*}),$$

where ϖ_t^* and ϖ_t^{cb*} denote amount of foreign bond sales and purchases on a FX market from foreign investors and central bank, γ is a risk aversion parameter of FX market dealers and σ denotes a standard deviation of a nominal exchange rate. Capital inflows ϖ_t^* are modelled as AR(1) process. The amount of bonds purchased or sold by central bank in this operation is determined by the equation

$$\varpi_t^{cb*} = \chi_T(e_t - e_T) + \chi_e(e_t - e_{t-1}) + \chi_q q_t + \epsilon_t^{cb}$$

Monetary authority may intervene under the different regimes. The combination of the parameters $(\chi_T = 1, \chi_e = 0, \chi_q = 0)$ implies exchange rate targeting, the combination $(\chi_T = 0, \chi_e = 1, \chi_q = 0)$ means so called "leaning against the wind", i.e. smoothing of exchange rate. The combination $(\chi_T = 0, \chi_e = 0, \chi_q = 1)$ describes the case of intervention to the real exchange rate. The calibration $(\chi_T = 0, \chi_e = 0, \chi_q = 0)$ implies that only unexpected intervention shocks influence the foreign exchange.

To distinguish between two specifications of the model in the text, we call the original version of the model RISK and the model with the UIP condition following [9] is denoted as MODUIP. The parameters of both model specifications were estimated using seven seasonally adjusted time series of the Czech economy and Eurozone from the period between 2000Q1 and 2015Q4. Both domestic and foreign economies were represented by the time series of percentual change in the real output, the inflation rate measured by consumer prices and the nominal interest rate (3M PRIBOR, 3M EURIBOR respectively). The link between the domestic and foreign economy is represented by the nominal CZK/EUR exchange rate depreciation. The estimation of the model was performed with the use of toolbox Dynare, all calculations were performed in the system Matlab.

3 Foreign exchange rate intervention

The foreign exchange rate intervention may be modelled in different manners. First attempts to include exchange rate targeting into the standard DSGE model framework considered FX rate targeting as an additional monetary policy objective. The FX rate was incorporated into the decision function and monetary authority used interest rate movements as a tool to meet its goals. The inflation, output growth and exchange rate are operational targets. This approach cannot be used in case of the Czech intervention, because exchange rate is identified as a tool not a target. Beneš et al. in [2] suggested the model structure with two separated decision functions, one for interest rate and one for the exchange rate. Exchange rate as well as interest rate are considered as monetary policy tools and interact with each other through the UIP condition. In this approach, both decision rules are equivalent. Nevertheless, this approach is not in line with policy of the Czech National Bank. The exchange rate is not considered as a substitute to interest rate, central bank does not adjust FX rate following any decision rule to meet its operational targets.

Montoro, Ortiz constructed a model of the interventions to FX rate based on the FX market approach included in the UIP condition described above. Their approach was not intended to model the use of the FX rate as an monetary policy tool, however it is close to the policy of CNB and managed floating. The approach described above was successfully used for modeling the intervention of CNB by Malovana in [8], nevertheless it has some disadvantages. First, the FX targeting rule is symmetrical, i.e. we are not able to study asymmetric commitment that is recently used by CNB. Second, the changes in the purchase and sale orders of foreign bonds has only limited influence on FX rate. Third, we are not able to use this rule to set target on exact value. We decided to take advantage from the FX market based UIP condition, but we use it only to model unexpected changes in the FX rate. The asymmetric commitment is modeled as a constraint in a model and I use an algorithm from Tom Holden [4] and [5]. The nature of the shadow shock corresponds to expected interventions to FX rate performed by CNB.

4 Nonlinear constraint

Holden and Paetz in [4] introduced and derived a method that is based on so called shadow price shocks. The shadow price shock ϵ_{SP} is incorporated into the equation of the constrained variable and it drives this variable from the negative values back to the zero level. The constrained model equation takes a form

$$x_t = \mu_x + \phi_{-1}y_{t-1} + \phi_0 y_t + \phi_{+1}E_t y_{t+1} - \mu_y(\phi_{-1} + \phi_0 + \phi_{+1}) + \epsilon_{SP}.$$

The interesting feature of the shock is that it is expected. The participants of economy know that the constrained variable cannot violate the bound. In case of the ZLB, the shadow shock may be interpreted as the positive expected monetary innovation, i.e. the interest rate is higher than in unbounded case and that has influence on economy. The constraint imposed on the exchange rate has similar meaning - each time the bound on the nominal exchange rate binds, there is the expected depreciation shock that holds the FX rate above declared value.

The algorithm saves values of ϵ_{SP} and impulse responses of the model variables to all values of this shock. The impulse response to the model's shock ϵ under constraint may be written as $\mathrm{irf}_\epsilon^x = \mu_x + v + \alpha M$, where v denotes the impulse response to this shock without constraint, M is the matrix of the impulse response functions to the values of the shadow price shock and α is the magnitude of these shocks. To find the magnitude α, Holden solves a quadratic programming problem

$$\alpha^* = \underset{\substack{\alpha \geq 0 \\ \mu^* + v^* + M^* \alpha \geq 0}}{\operatorname{argmin}} \{\alpha'(\mu^* + v^*) + \frac{1}{2}\alpha'(M^* + M^{*'})\alpha\}.$$

Holden extended this algorithm into toolbox DynareOBC described in [5]. This toolbox enables us to simulate and estimate models with constraints. The aim of author was to develop easily applicable toolbox and to derive the conditions of existence and uniqueness of solution of the model with non-linear constraints. The unfortunate feature of the tool DynareOBC is its high computational complexity. The algorithm demands to set a assumed horizon of binding constraints and solves the quadratic programming problem for each time period. For larger models this requires to use the computational sources that provide sufficient computational capacity. The Czech academic community may use Metacentrum (https://metavo.metacentrum.cz/).

5 Results

The limited space allows us to show only small fraction of the results, nevertheless we try to choose examples that both illustrate the use of the tool and are relevant for recent economic situation. In the Figure 1, we compare the simulated trajectories of the models with bounded interest rates and different specifications of the UIP condition. The nature of the algorithm enables us to interpret the zero lower bound as a positive monetary shock that struck the economy independently to decisions of the monetary authority.

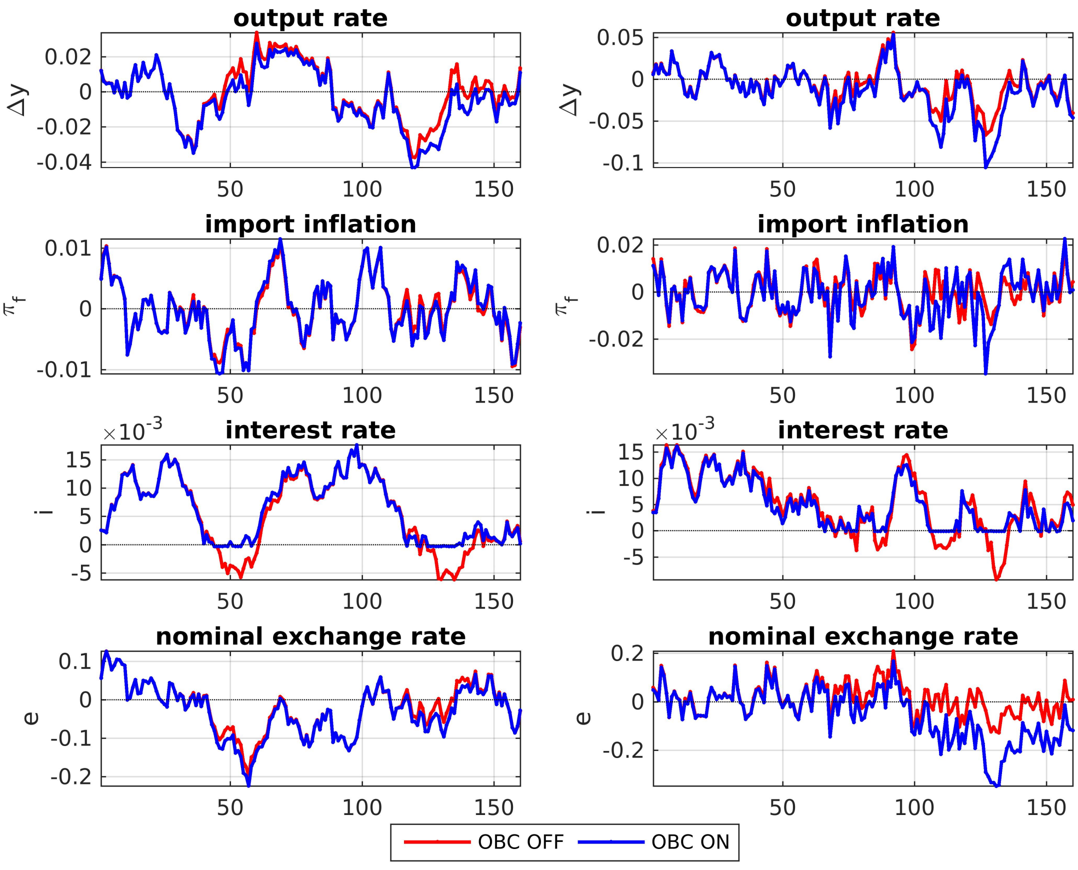


Figure 1 Selected trajectories of a model with constraint on the interest rate. Specification MODUIP in the left column, specification RISK in the right column.

In both versions of the model, constraint on the interest rate pushes the nominal exchange rate towards the greater appreciation. The combination of the FX rate appreciation and the positive shadow monetary shock mutes the output growth and deepens the recessions. There is a noticeable impact to foreign goods (imports) inflation rate - the bound is linked to lower imports inflation rate. These results provide an explanation of the decision of central bank to depreciate the FX rate. Domestic and foreign interest rates reached the zero lower bound in the third quarter of 2012. Domestic economy should be stimulated both by lower interest rate and depreciation of the exchange rate, but the floor that constraints domestic interest rate pushes exchange rate to further appreciation. Central bank decided to intervene

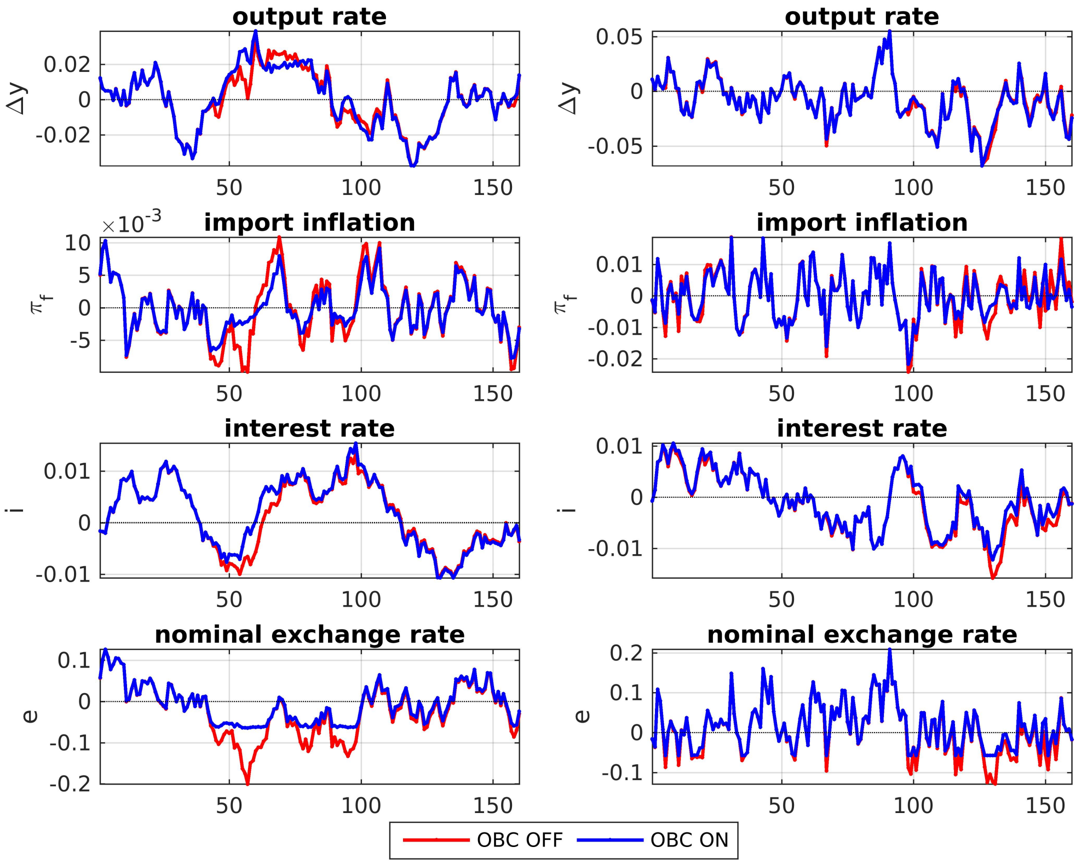


Figure 2 Selected trajectories of a model with constraint on the nominal exchange rate. Specification MODUIP in the left column, specification RISK in the right column.

and depreciated the exchange rate to compensate the influence of the constraint. The main difference between the two specifications of the model is the divergence of the nominal exchange rate when the interest rate is bounded in case of the RISK specification. That is caused by the net foreign position (not depicted) that is deviated by the shadow shock and returns to steady state very slowly.

The simulation of the asymmetric commitment is depicted in the Figure 2. The specification MODUIP shows some interesting results. The nominal interest rates are linked to exchange rate through UIP condition and that causes rise of the interest rates in case of bounded FX rate. The rise of the interest rates may be interpreted also as a reaction of central bank to the output rate, that is positively stimulated by weaker exchange rate. That is decribed in [7] as a consequence of depreciation, but it is not in line with economic reality. The persistence of the negative effect of the interest rate can be seen after period 50, when the exchange rate depreciates and the unbounded case of output growth is higher than the case related to commitment. Another interesting result is the trajectory of imports inflation, that is significantly muted by commitment, especially in the negative range. That suggests that the commitment may be suitable to use when the risk of deflation is caused by foreign factors. The results for the specification RISK are similar, but less significant.

In our further research we would like to continue to use the toolbox DynareOBC to simulate the synergic effect of presence of the Zero Lower Bound on domestic interest rates in combination with other constraints. The primary goal is to combine ZLB and the asymmetric commitment, the next could be the combination of domestic and foreign ZLB, as the Eurozone struggles with zero rates.

6 Conclusion

In the paper we deal with the use of the nominal exchange rate as an unconventional instrument of monetary policy, when the interest are bounded by zero and central bank needs to ease the monetary conditions. The use of a new modelling tool enables us to model the asymmetric commitment as a nonlinear constraint in a model. We compared two different specifications of the UIP condition in a model and the results suggested that the UIP condition based on the FX market solution provides more satisfactory results.

The results showed the impact of the Zero Lower Bound on interest rates into the economy and showed that the bound is equivalent to the monetary restriction. In such case, foreign exchange rate depreciation compensates the impact of a bound. The analysis of the simulated trajectories in case of the bounded foreign exchange rate suggested that the asymmetric commitment may have desired implications to development of the economy, especially used to fight deflationary pressures from foreign sector.

Acknowledgements

This work is supported by funding of specific research at Faculty of Economics and Administration, project MUNI/A/1040/2015. This support is gratefully acknowledged.

Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated.

References

[1] Alpanda, S., Kotz, K. and Woglom, G.: The role of the exchange rate in a New Keynesian DSGE model for the South African economy. *South African Journal of Economics* **78** (2010), 170-191.

[2] Beneš, J. Berg, A. Portillo, R.A. and Vávra, D.: *Modeling Sterilized Interventions and Balance Sheet Effects of Monetary Policy in a New-Keynesian Framework.* Working Paper WP/13/11, International Monetary Fund, 2013.

[3] Eggertson, G.B. and Woodford, M.: The Zero Bound on Interest Rates and Optimal Monetary Policy. *Brookings Papers on Economic Activity* **34**, 1 (2003), 139–235.

[4] Holden, T. and Paetz, M.: *Efficient simulation of DSGE models with inequality constraints* Discussion Paper 1612, School of Economics, University of Surrey, 2012.

[5] Holden, T.: *Existence, uniqueness and computation of solutions to dynamic models with occasionally binding constraints.* Component of the DynareOBC tool. https://github.com/tholden/dynareOBC School of Economics, University of Surrey, 2016.

[6] Justiniano, A. and Preston, B.: *Monetary policy and uncertainty in an empirical small open economy model.* Working Paper Series WP-09-21, Federal Reserve Bank of Chicago, 2010.

[7] Lízal, L. and Schwarz, J.: *Foreign Exchange Interventions as an (Un)Conventional Monetary Policy Tool* (October 2013). BIS Paper No. 73i.

[8] Malovaná, S.: *The Efectiveness of Unconventional Monetary Policy Tools at the Zero Lower Bound: A DSGE Approach.* Karlova Univerzita, Fakulta sociálních studií. Master thesis, 2014.

[9] Montoro, C. and Ortiz, M.: *Foreign Exchange intervention and Monetary policy Design: A Market Microstructure Analysis.* Third Draft, 2014.

[10] Svensson, L.E.O.: The Zero Bound in an Open Economy: A Foolproof Way of Escaping from a Liquidity Trap. *Monetary and Economic Studies* **19**, S1 (2001), 277–312

Long-term unemployment in terms of regions of Slovakia

Kvetoslava Surmanová[1], Marian Reiff[2]

Abstract. The paper focuses on the analysis of long-term unemployment in Slovak regions oriented on the regional disparities at NUTS 3 level, using annual data from a regional database of the Slovak Statistical Office for the period of 2001- 2014. Unemployment as one of the key indicators of the economy, contributes to creating the picture about the level of development and living standards in the country. Since regional disparities in Slovakia are large, this issue is highly topical. In recent years, it is necessary to address not only the total number of unemployed in Slovakia, but to look particularly at the situation in the regions. The highest rates of the long-term unemployment are in the districts of central and eastern Slovakia. Despite the fact, that the reduction of unemployment rate belongs to basic economic policy framework, it is difficult to ensure minimization of this negative phenomenon. We use to analyse the methodology of classical linear model and methodology of seemingly unrelated regressions (SUR model).

Keywords: unemployment, Okun's law, SUR model, region.

JEL Classification: E24, C30
AMS Classification: 62P20

1 Introduction

Unemployment is one of the long-term objectives pursued in the field of economic policy. The necessity to deal with this issue arises from the fact that unemployment plays important role in society and has not only economic, but also social dimension. The current effort of economists to not only understand the problem, but even to find a way of eliminating unemployment, or at least minimize its consequences is therefore justified. The paper is organized in chapters. After introduction in first chapter, the second chapter deals with theoretical basis of the Okun's law. Subsequently, in the third chapter it is briefly point out the regional differences in the selected indicators. The fourth chapter describes the methodological apparatus that is used in the estimation of the Okun's relationship. Estimation results are presented and interpreted in chapter five. In this part of the article we compare the results achieved at the NUTS 3 level with the results obtained by Ordinary least squares method (OLS) on data of Slovak Republic as a whole. The paper closes with concluding remarks.

2 Okun´s Law – theoretical background

Economic theory knows several approaches focused at unemployment in the national economy. Provided examples are Pigou's the theory of unemployment, The Classical Theory of Unemployment, Keynesian Theory of Unemployment, Okun's law theory, Theory of Phillips curve, etc. (for more details see [5], [8], [9], [10]). All conceptions are well known and are supported by academic literature (see e.g. [7], [11] and [13]). In [12] authors attempted to find some simple original dynamic models of the labor market equilibrium in order to derive a Phillips curve. In [4] there is used Okun's law and estimated age and country specific Okun coefficients for five different age cohorts. In [2] there is examined Okun's Law in European countries by distinguishing between the permanent and the transitory effects of output growth upon unemployment and by focusing on the effect of labour market protection policies. In [1] there is examined the validity of Okun‘s Law for V4 countries. In [6] author focused on two questions. The first, is Okun's law a reliable and stable relationship? And the second, is this law a useful for forecasting?

In this paper, we will focus on one of the approaches that is mentioned above, namely Okun's low. The traditional interpretation of Okun's law in terms of macroeconomic theory captures the relationship between economic output and unemployment. It is not only formulation of theoretical economic basis, it is mainly supported by empirical research, which was presented in 1962 by Arthur M. Okun (Okun was an advisor to President Kennedy in the 1960´s). He has focused on the relationship of unemployment and real output. Okun's law captures the negative relationship between these two indicators on the background of a simple linear model. This theory was

[1] University of Economics in Bratislava, Department of Operation Research and Econometrics, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia, surmanova@yahoo.com.
[2] University of Economics in Bratislava, Department of Operation Research and Econometrics, Dolnozemská cesta 1, 852 35 Bratislava, Slovakia, marian.reiff@euba.sk.

gradually modified by other economist authors who have implemented different factors into the original relationship the Okun has not considered.

Although several versions are known, the paper analysis focuses on the incremental (differential) version of Okun's Law. This captures the quarterly change in the unemployment rate in relation to the change of the growth in output, which can be written as follows:

$$\Delta u = \alpha + \beta \cdot gGDP + \varepsilon_t, \tag{1}$$

where $\triangle u = u_t - u_{t-1}$ and $gGDP = \frac{GDP_t - GDP_{t-1}}{GDP_{t-1}}$. Δu is change in the unemployment rate, $gGDP$ is growth of GDP and ε_t is stochastic random error. Parameter β is known as Okun's coefficient and it is assumed to have negative sign (see for example [6] or [8]). Ratio ($-\alpha/\beta$) represents GDP growth rate, consistent with a stable unemployment rate. Above mentioned ratio equally reflects how quickly the economy must grow in order to maintain a constant level of unemployment.

3 Regional disparities at NUTS 3 level

By joining the EU, Slovakia committed itself to the implementation of regional policy. The aim of EU regional policy is reducing disparities between the levels of development of individual regions and mitigating the backwardness of less developed regions of the country. Under the above the focus is on monitoring the disparities between regions in the amount of the GDP and the unemployment rate. Slovakia is at the NUTS 3 level divided into eight regions (BA - Region of Bratislava, TT – Region of Trnava, TN – Region of Trenčín, NT – Region of Nitra, ZI – Region of Žilina, BB – Region of Banská Bystrica, PO – Region of Prešov and KE – Region of Košice), and these territorial units are too heterogeneous and incomparable. And not just in terms of comparison of the level of convergence, as well as regional disparities. Therefore by analyzing them, we can expect significant differences. For this reason, analysis on the regional level provides more detail and richer results compare to analysis conducted at the NUTS 1 – the national level. We used the data from the web page of Statistical Office of the Slovak Republic [14].

We can divide unemployment rate in the regions of Slovakia into 2 groups (see Figure 1). The first group with lower rate of unemployment includes regions BA, TT, TN, ZI a NT. The second group includes east country regions, namely PO, KE a BB region. High value of sample variance (see Table 1) as well as a graph representation of the unemployment rate over time points to the highest variability in NT region (22.64). By comparison with the regions with the lowest sample variance - BA (2.26), it's up to ten times more.

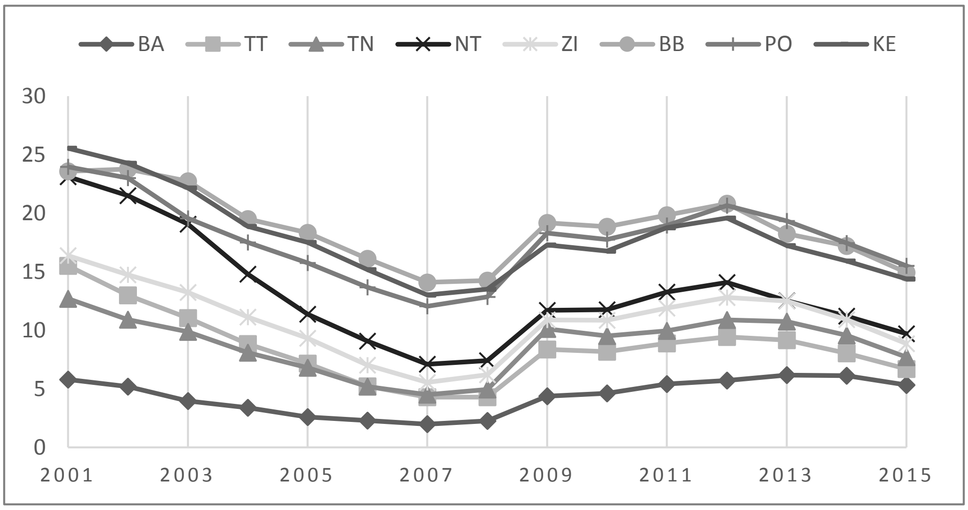


Figure 1 Unemployment rate in region SR in time (in %)

2001 - 2015	U_BA	U_TT	U_TN	U_NT	U_ZI	U_BB	U_PO	U_KE
Median	4.63	8.37	9.56	11.76	10.91	18.86	17.75	17.30
Standard Deviation	1.50	3.03	2.47	4.76	3.04	3.13	3.43	3.69
Sample Variance	2.26	9.20	6.08	22.64	9.26	9.78	11.74	13.61
Range	4.19	11.22	8.20	16.02	10.83	9.67	11.91	12.53
Minimum	1.98	4.29	4.50	7.10	5.55	14.10	12.05	13.02
Maximum	6.17	15.51	12.70	23.12	16.38	23.77	23.96	25.55

Table 1 Descriptive statistics of unemployment rate in time 2001-2015

With regard to regional differences is necessary to mention wages. The amount of wages to some extent determines the need to be employed. A monthly salary varies in different regions, while wages are the highest in the BA region, the region with the lowest wage is PO (see Figure 2).

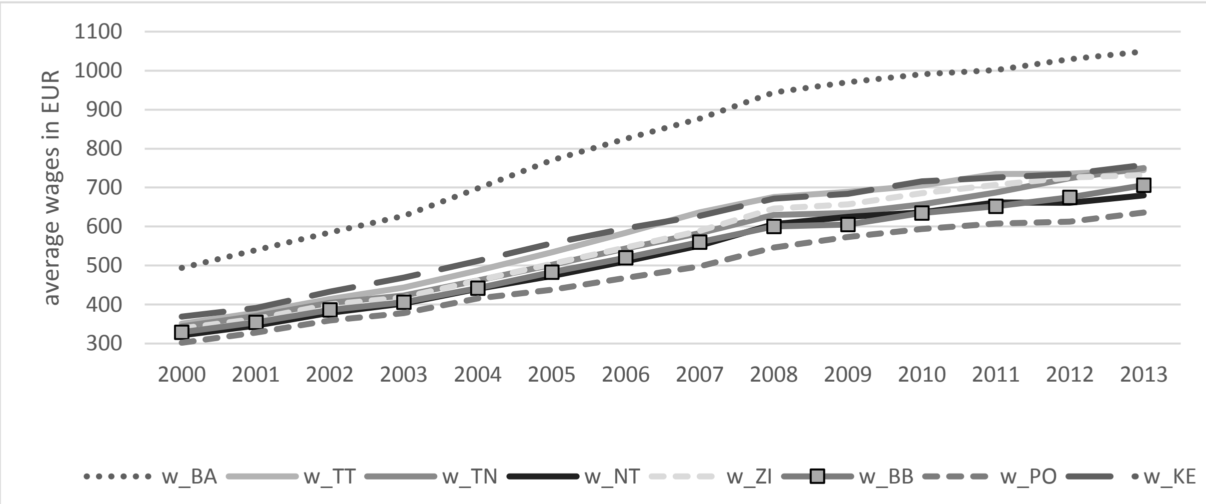


Figure 2 Development average nominal wages in regions in time (in EUR)

For the demonstration of maturity regions it is used quantile display of work productivity (see Figure 3). We have chosen first and last observation available at the regional level, i.e. 2001 and 2013. From Figure 3 it is possible to observe the dynamics in the development of labor productivity in terms of quintile breakdown of regions into five groups within the study period. In two regions (TT and TN), the labor productivity per employed person has declined. In two regions (ZI and BB), however, it has increased. PO region remains weakest regions even after 13 years period, characterized with the lowest productivity of work.

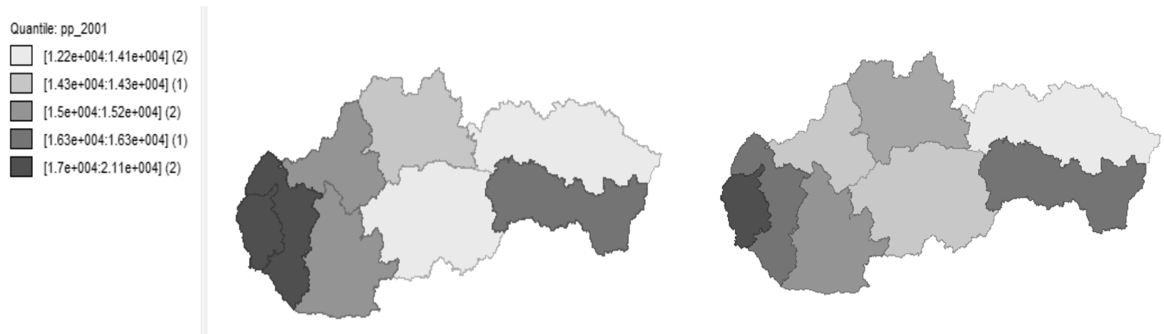


Figure 3 The dynamics of labor productivity in the period 2001 - 2013

4 SUR model

SUR model is created by joining multiple single-equation models in a single multi-equation model, while it may seem that this link is at first sight "Seemingly, unrelated". If there is the correlation between the random component of the individual equations of the same observation, $cov(u_{t1}, u_{t2}) \neq 0$, then the model estimates to seemingly unrelated regressions using the method of estimating model parameters to seemingly unrelated regression (method SUR) (for more details see eg. [3]). Analytically this type of multi- equation model can be written as follows:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \cdot \\ \mathbf{y}_g \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & 0 & \ldots & 0 \\ 0 & \mathbf{X}_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \mathbf{X}_g \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \cdot \\ \boldsymbol{\beta}_g \end{bmatrix} + \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \cdot \\ \mathbf{u}_g \end{bmatrix} \tag{2}$$

where
y is a “blocks” vector of dependent variable with dimension (g × n × 1),
X is a “blocks” matrix of independent variables with dimension (g × n × k),
β is a “blocks” vector parameters with dimension (g × k),
u is a “blocks” vector of stochastic errors with dimension (g × n × 1),
g is a number of cross-section units,
k is a number of determinants
and n is a number of observations in time.
Equation (2) displays a system of equations in which the explanatory variables contained in one equation at the same time do not act in a different equation of the system.

5 Empirical Results

The empirical analysis on regions of Slovakia was realized on data for the period 2001 - 2014. As 14 observations are available for each region, it was used multi- equation model seemingly unrelated regressions - SUR model, as appropriate method for the analysis. This methodology is relatively rarely used, as in the analysis of panel data, which are combination of spatial and time information it is mostly used panel data methodology. The whole analysis was carried out in the software EViews.

In estimation of the differential version of Okun’s law (1) we use data on unemployment rate (in %) and more specifically its first difference. It acts as a determinant of relative indicator of real GDP growth, which we obtained from the conversion of nominal GDP at the regional level and GDP at constant prices for the whole Slovakia.

The above presented theory of Okun’s law and SUR model were used to detect Okun’s coefficient of eight Slovak NUTS 3 regions. We estimate two versions SUR model. The first SUR model with all regions and the second with seven regions, without BA region. The region of BA was excluded because it shows extreme levels of wages. We monitored, whether the exclusion of BA model region causes significant changes. Table 2 contains the results of the regression analysis. Our expectations regarding the impact of excluding BA region of the model, however, did not materialize. Okun’s coefficient (β) nor the required rate of growth of GDP ($-\alpha/\beta$), which should ensure stable unemployment rates in individual models do not differ significantly.

Region	**SUR_8 regions**			**SUR_7 regions (without BA)**		
	β	$-\alpha/\beta$	R^2	β	$-\alpha/\beta$	R^2
BA	-0.04***	9.23	0.32	-	-	-
TT	-0.09***	2.00	0.48	-0.10***	2.62	0.51
TN	-0.11***	5.08	0.44	-0.12***	5.17	0.46
NT	-0.21***	3.04	0.47	-0.19***	2.72	0.45
ZI	-0.13***	5.12	0.47	-0.14***	5.35	0.50
BB	-0.10*	1.47	0.26	-0.09*	1.15	0.25
PO	-0.11***	3.80	0.20	-0.12**	4.11	0.21
KE	-0.22***	2.89	0.59	-0.19***	2.38	0.54
Average	**-0.13**	**4.08**	**0.40**	**-0.12**	**3.36**	**0.42**
Breusch-Pagan test	203.69			158.36		
χ^2 critical	41.34			32.67		
Portmanteau test	63.68			43.67		
χ^2 critical	83.68			66.34		

*** parameter is significant on 1 %, ** significant on 5 %, * significant on 10 % level of significance.

Table 2 The results of the estimation of the Okun’s SUR model

Both estimated versions of models after Breusch-Pagan test validate the hypothesis of correlation of random components of two different equations at the same time *t*. Based on Portmanteau test it was rejected hypothesis of correlation of random components in each of the eight equation formulas, and for two different periods of time.

Values in column ($-\alpha/\beta$) denote what growth rate in each of the regions is need to maintain a stable unemployment rate in the region. It means, by more than 5% of the economy should grow in the BA (9.23%), TN Region (5.08%) and ZI (5.12%) (Model SUR_8). For the second version of the estimate SUR_7, economic growth must be in TN (5.17%) and ZI (5.35%) of the region, results are similar to the SUR_8. The lowest growth, maintaining a stable unemployment regions have BB, TT and KE region. Its value is about 2%. Okun's coefficient itself can be interpreted as follows. According to Okun's it law should be true, if the growth rate of GDP will be in the PO region 4.8% and TN region 6.08%, unemployment in mentioned regions will decrease equally by 0.11%. The low value of the Okun's coefficient in Bratislava region indicates a low sensitivity to GDP growth in order to changes the unemployment. Highest sensitivity of the change in unemployment show results for NT region and KE region.

For further comparison it is estimated the equation (1) with OLS method on aggregate data for the Slovak Republic on basis of annual data for the period 2001-2014. If the growth rate of GDP was zero, the unemployment rate will rise by 1.6795%. To maintain a stable rate of unemployment requires that growth GDP can amount to 5.82%. According to Okun's law it should be true if the GDP growth rate will be 6.82%, the unemployment rate will drop by 0.303%. Linear regression result is captured on Figure 4. The estimated parameters and the model are significant on the 5 % level of significance. In the model is absent autocorrelation of random error and $R^2 = 0.7252$.

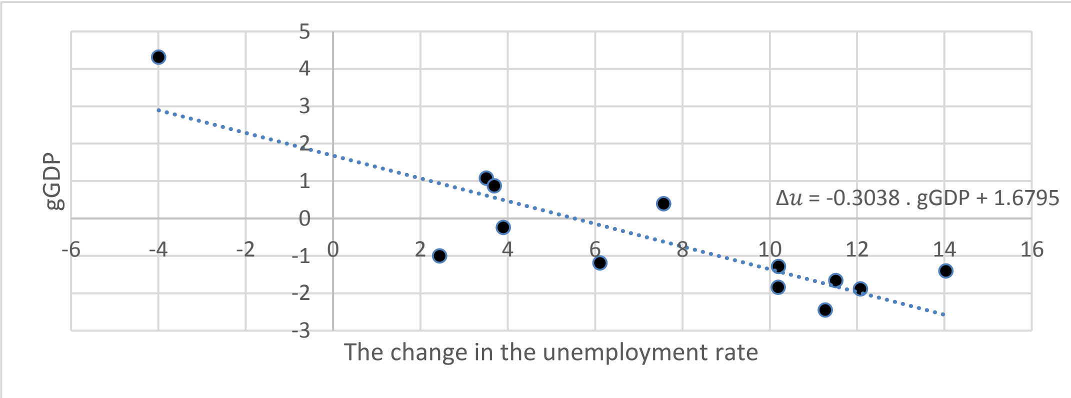


Figure 4 Estimate Okun's law with OLS method for all Slovak Republic

6 Conclusion

In the paper we test the hypothesis of Okun's law in NUTS 3 regions of Slovak republic during the 2001-2014 period using model of seemingly unrelated regressions. Since the analysis at regional level encounters a problem with sufficient and actual database, quantitative analysis has limited capabilities.

In analysing the Okun's law on regional data from Slovakia, we have come in terms of the negative value of the Okun's coefficient to same conclusion as [6], [8] and others. Results in individual regions reveal a significant disparity. There are regions in which unemployment rate reacts more sensitive to economic growth and the regions where the impact is negligible. Results of the analysis indicate that the coefficient with the lowest values is in BA and KE regions and with highest values in NT region, this fact implies that labour market is rigid. Higher values of the Okun's coefficient refer to stronger fluctuations of fixed linking unemployment to fluctuations in product surveyed regions. By comparing the results for the whole country and the results on a regional basis, we can conclude that the regional dimension, by looking at the unemployment rate is eye-catching because the analysis of differences in the regions can serve in large measure to the implementation of proper and targeted regional social policy.

Since the unemployment rate and the impact on other determinants, it would be appropriate to extend the basic model by other information that would lead to increased variability of the model. Equally interesting analysis could be extended to dummy variable that would have corrected the disturbances which occurred as a result of expression of the crisis in 2009.

Acknowledgement

This work was supported by the Grand Agency of Slovak Republic – VEGA grand No. 1/0444/15 "Econometric Analysis of Production Possibilities of the Economy and the Labour Market in Slovakia".

References

[1] Boďa, M. et al.: (A)symetria v Okunovom zákone v štátoch Vyšehradskej skupiny. *Politická ekonomie* **63,** 6 (2015), 741–758.

[2] Economou, A., Psarianos, I. N.: Revisiting Okun´s Law in European Union Countries. *Journal of Economic Studies* **43,** 2 (2016), 275–287.

[3] Greene, W. H.: *Econometric Analysis* (Seventh ed.). Upper Saddle River, Pearson Prentice-Hall, 2012.

[4] Hutengs, O., Stadtmann, G.: Don´t trust anybody over 30: Youth unemployment and Okun´s law in CEE countries. *Bank I Kredyt* **45,** 1 (2014), 1–16.

[5] Keynes, J.: *The General Theory of Employment, Interest and Money*. London, Harcourt, 1936.

[6] Knotek, S. E.: How Useful is Okun´s Law? Kansas City: Federal Reserve Bank of Kansas City, 2007.

[7] Mouhammed, A. H.: Important Theories of Unemployment and Public Policies. *Journal of Applied Business and Economics* **12,** 5 (2011), 100–110.

[8] Okun, A. M.: Potential GNP: Its Measurement and Significance. *American Statistical Association, proceedings of the Business and Economics Statistics Section,* Alexandria, Virginia: American Statistical Association, 1962.

[9] Pigou, A. C.: *The Theory of Unemployment*. London, Macmillan, 1933.

[10] Phillips, A. W.: The Relation between Unemployment and the Rate of Change of Money Wages in The United Kingdom, *Economika* **25,** 100 (1958), 283–299.

[11] Sweezy, P.: Professor Piqou´s Theory of Unemployment. *The Journal of Political Economy* **42,** 6 (1934), 800–811.

[12] Szomolányi, K. at al.: The Relation between the rate of change of money wage and unemployment rates in the frame of the microeconomic dynamics. *Ekonomické rozhľady* **41,** 1 (2012), 44–65.

[13] Trehan, B.: Unemployment and Productivity, *Federal Reserve Bank of San Francisco Economic Letter* **28**, (2001), 1–3.

[14] www.statistics.sk.

The Effect of Terms-of-Trade on Czech Business Cycles: A Theoretical and SVAR Approach

Karol Szomolányi[1], Martin Lukáčik[2], Adriana Lukáčiková[3]

Abstract. Empirical and theoretical studies imply contradictory results about terms-of-trade effect on business cycles. Empirically the terms-of-trade have a negligible effect on the short-term economic fluctuations. But theoretical models – both Keynesian and real business cycle models – predict significant influence of the terms-of-trade on business cycles. To lower a significance of terms-of-trade shocks, theoretical models have been extended by introducing non-tradable goods in the economy. However this modification does not help and the theoretical impact of the terms-of-trade is too high. Using Czech data and MXN theoretical model, we confirm that model predictions of the terms-of-trade impact do not match with observations gathered from the SVAR model estimate of Czech cyclical components of the terms-of-trade trade balance-to-output ratio, output consumption, gross investment and real exchange rate.

Keywords: Terms-of-trade, business cycle, SVAR model, MXN model.

JEL Classification: C44
AMS Classification: 90C15

1 Introduction

Terms-of-trade is theoretically significant source of business cycles and it causes shifts in trade balance. However different theoretical and empirical studies lead to different results of the short-run terms-of-trade impact on output and on trade balance. There are two theoretical effects of terms-of-trade impact on trade balance. Harberger [6] and Laursen and Metzler [8] used traditional Keynesian model to show that trade balance grows with terms-of-trade. On the contrary, dynamic optimizing models of Obstfeld [10] and Svensson and Razin [12] leads to a conclusion that positive effect of terms-of-trade on the trade balance is weaker the more persistent a terms-of-trade shock is. Uribe and Schmitt-Grohé [15] showed that in small open economy real business cycle model (or dynamic stochastic general equilibrium model) with capital costs sufficiently permanent terms-of-trade shocks have negative impact on the trade balance. Empirical studies of Aguirre [1], Broda [4] and Uribe and Schmitt-Grohé [15] surprisingly do not support statistically significant impact of term-of-trade on output in poor and emerging countries. In general authors can confirm an intuition that the more open the economy is the higher effect of terms-on-trade on trade balance is. This result may not be achieved in theoretical general equilibrium models even if non-tradable goods are considered. Uribe and Schmitt-Grohé [15] developed MXN model with importable (M) exportable (X) and non-tradable (N) goods to show that an existence of non-tradable goods “reduce the importance of terms-of-trade shocks.” However authors state that the theoretical model still overestimates the signification of the terms-of-trade impact on the business cycles.

In this paper we confirm this theoretical and empirical mismatch. In the first part of the paper we present and estimate the empirical SVAR model of the terms-of-trade impact on the Czech business cycles to state that the influence of the terms-of-trade is very small. In the second part we present and calibrate MXN model of Uribe and Schmitt-Grohé [15] using Czech observations to state that the theoretical model predicts relatively high impact of the terms-of-trade on business cycles.

2 Empirical Model

First, we used vector autoregressive (VAR) models for our analysis. Every endogenous variable is a function of all lagged endogenous variables in the system in VAR models. See Lütkepohl [9] for more details about them. The mathematical representation of the VAR model of order p is:

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + ... + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{e}_t \tag{1}$$

[1] University of Economics, Department KOVE, Dolnozemská cesta 1, Bratislava, Slovakia, szomolan@euba.sk.
[2] University of Economics, Department KOVE, Dolnozemská cesta 1, Bratislava, Slovakia, lukacik@euba.sk.
[3] University of Economics, Department KOVE, Dolnozemská cesta 1, Bratislava, Slovakia, istvanik@euba.sk.

where $\mathbf{y_t}$ is a k vector of endogenous variables; $\mathbf{A_1}$, $\mathbf{A_2}$, …, $\mathbf{A_p}$ are matrices of coefficients to be estimated; and $\mathbf{e_t}$ is a vector of innovations that may be contemporaneously correlated but are uncorrelated with their own lagged values.

The VAR model (1) can be interpreted as a reduced form model. A structural vector auto-regressive (SVAR) model is structural form of VAR model and is defined as:

$$\mathbf{A}\mathbf{y}_{\mathbf{t}} = \mathbf{B}_1\mathbf{y}_{\mathbf{t-1}} + \mathbf{B}_2\mathbf{y}_{\mathbf{t-2}} + \ldots + \mathbf{B}_{\mathbf{p}}\mathbf{y}_{\mathbf{t-p}} + \mathbf{B}\mathbf{u}_{\mathbf{t}} \tag{2}$$

A SVAR model can be used to identify shocks and trace these out by employing impulse response analysis and forecast error variance decomposition through imposing restrictions on used matrices.

Uribe and Schmitt-Grohé [15] proposed a specification of the SVAR, through which we can determine responses on terms-of-trade impulse:

$$\mathbf{A}\begin{pmatrix} f_t \\ tb_t \\ y_t \\ c_t \\ i_t \\ rer_t \end{pmatrix} = \sum_{i=1}^{p}\mathbf{B}_i\begin{pmatrix} f_{t-i} \\ tb_{t-i} \\ y_{t-i} \\ c_{t-i} \\ i_{t-i} \\ rer_{t-i} \end{pmatrix} + \mathbf{B}\begin{pmatrix} u_t^f \\ u_t^{tb} \\ u_t^y \\ u_t^c \\ u_t^i \\ u_t^{rer} \end{pmatrix} \tag{3}$$

where f is relative cyclical component of the terms of trade, tb is relative cyclical component of the trade balance to output ratio, y is relative cyclical component of output, c is relative cyclical component of consumption, i is relative cyclical component of investment and rer is relative cyclical component of real exchange rate.

The u_t^f, u_t^{ttb}, u_t^y, u_t^c, u_t^i and u_t^{rer} are structural shocks of given variables. We estimated the parameters of the SVAR specification (3) using Amisano and Giannini [3] approach. The class of models may be written as:

$$\mathbf{A}\mathbf{e}_{\mathbf{t}} = \mathbf{B}\mathbf{u}_{\mathbf{t}} \tag{4}$$

The structural innovations $\mathbf{u_t}$ are assumed to be orthonormal, i.e. its covariance matrix is an identity matrix. The assumption of orthonormal innovations imposes the following identifying restrictions on **A** and **B**:

$$\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{e}}\mathbf{A}^{\mathbf{T}} = \mathbf{B}\mathbf{B}^{\mathbf{T}} \tag{5}$$

Noting that the expressions on both sides of (5) are symmetric, this imposes $k(k+1)/2 = 21$ restrictions on the $2k^2 = 72$ unknown elements in **A** and **B**. Therefore, in order to identify **A** and **B**, we need to impose $(3k^2\text{-}k)/2 = 51$ additional restrictions. The matrix **A** of unrestricted specification is a lower triangular matrix with unit diagonal (15 zero and 6 unity restrictions) and matrix **B** is a diagonal matrix (30 zero restrictions) in this just-identified specification. Other tested restrictions are imposed on elements of matrix **A** (matrix of contemporary effects of endogenous variables), which means that specification becomes over-identified and testable.

The selected lag of model (3) is validated by sequential modified likelihood ratio test statistic and information criteria and by the LM test for autocorrelations. Significant values of serial correlation for lower lags could be a reason to increase the lag order of an unrestricted VAR, but this is not our case. We verified the stability of a VAR model (i.e. whether all roots have modulus less than one and lie inside the unit circle). We estimated the parameters of restricted and unrestricted specifications. Using the logarithm of the maximum likelihood functions of both specifications we calculated the likelihood ratio statistics and verified the significance of restrictions. Tests confirmed no immediate impact of terms-of-trade on output, consumption, investment and exchange rate, trade balance on output, consumption and exchange rate and investment on exchange rate. All tests are explained in Lütkepohl [9] for example.

Data for Czech economy are gathered from the Eurostat portal [5]. The responses to the terms-of-trade shock in the Czech Republic are in the Figure 1. As output shock elasticity coefficient is not statistically significant, the improvement in terms-of-trade has no impact on the aggregate activity and the one-quarter delayed output expansion is statistically insignificant. The same result applies to the consumption and real exchange rate. Investment displays a small expansion. On the other hand, the impact of the terms-of-trade shock on trade balance is statistically significant. The 10 % increase in the terms of trade causes a decrease about 2 % in trade balance. The result suggests confirmation of Obstfeld-Svensson-Razin effect rather than Harberger-Laursen-Metzler effect of the terms-of-trade in the Czech Republic.

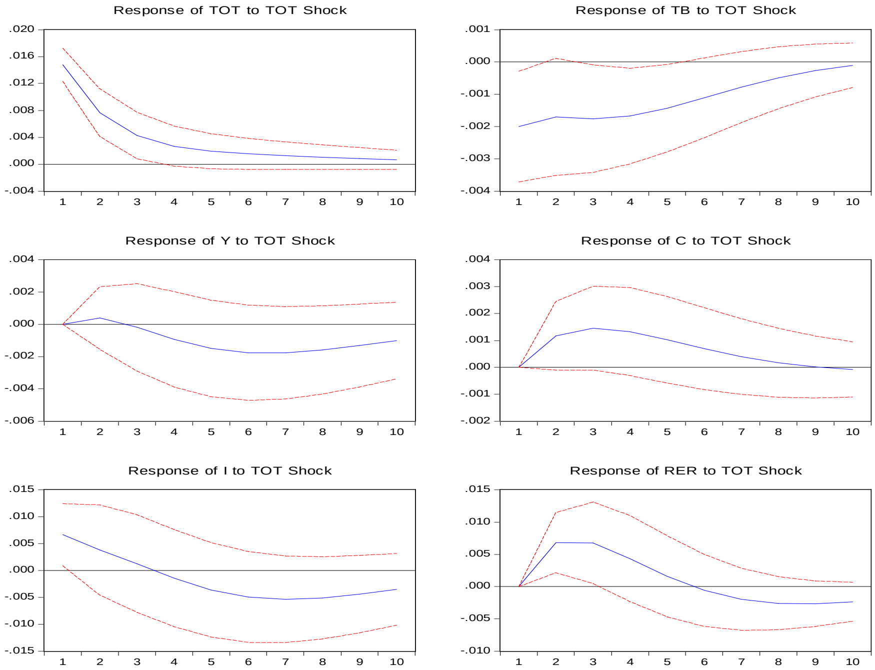


Figure 1 Impulse Response Functions to Terms-of-trade (TOT) Shock in the Czech Republic

3 Theoretical Model

Uribe and Schmitt-Grohé [15] presented the model with import-able (*m*), export-able (*x*) and non-tradable (*n*) sectors. The presence of non-tradable goods should reduce the importance of terms-of-trade shock.

We consider a large number of identical households with preferences described by the utility function:

$$U = E_0 \sum_{t=0}^{\infty} \beta^t \frac{\left[c_t - \frac{\left(h_t^m\right)^{\omega_m}}{\omega_m} - \frac{\left(h_t^x\right)^{\omega_x}}{\omega_x} - \frac{\left(h_t^n\right)^{\omega_n}}{\omega_n} \right]^{1-\sigma} - 1}{1-\sigma} \tag{6}$$

where c_t denotes consumption, for sector $j \in \{m,x,n\}$, h_t^j denotes hours worked in the sector j. Sectoral labour supplies are wealth inelastic and parameters ω_j denote wage elasticity in the sector j. The symbol E_0 denotes the expectations operator conditional on information available in initial period 0. The parameter σ measures the degree of relative risk aversion.

Households maximize the lifetime utility function (6) subject to the budget constraint:

$$\begin{aligned} c_t + i_t^m + i_t^x + i_t^n + \phi_m\left(k_{t+1}^m - k_t^m\right)^2 + \phi_x\left(k_{t+1}^x - k_t^x\right)^2 + \phi_n\left(k_{t+1}^n - k_t^n\right)^2 + p_t^\tau d_t = \\ = \frac{p_{t+1}^\tau d_{t+1}}{1+r_t} + w_t^m h_t^m + w_t^x h_t^x + w_t^n h_t^n + u_t^m k_t^m + u_t^x k_t^x + u_t^n k_t^n \end{aligned} \tag{7}$$

where for sector $j \in \{m,x,n\}$, i_t^j denotes gross investment, k_t^j denotes capital, w_t^j denotes real wage rate and u_t^j is the rental rate of capital in the sector j. Quadratic terms of the budget constraint (7) are capital adjustment costs, where ϕ_j denotes capital adjustment cost parameter in the sector j. The variable p_t^τ denotes the relative price of the tradable composite good in terms of final goods, d_t denotes the stock of debt in period t denominated in units of the tradable composite good and r_t denotes the interest rate on debt held from period t to $t + 1$. Consumption, investment, wages, rental rates, debt, and capital adjustment costs are all in units of final goods.

The capital stocks accumulation is given by:

$$k_{t+1}^j = \left(1-\delta\right) k_t^j + i_t^j;\ \forall j \in \{x, m, n\} \tag{8}$$

where δ denotes constant depreciation rate.

There are 5 types of large number of identical firms in the economy which differ according to their output: firms producing final goods, tradable composite goods, import-able goods, export-able goods and non-tradable goods.

Final goods are produced using non-tradable goods and a composite of tradable goods via CES technology

$$B\left(a_t^{\tau}, a_t^{n}\right) = \left[\chi_{\tau}\left(a_t^{\tau}\right)^{1-\frac{1}{\mu_{\tau n}}} + \left(1-\chi_{\tau}\right)\left(a_t^{n}\right)^{1-\frac{1}{\mu_{\tau n}}}\right]^{\frac{1}{1-\frac{1}{\mu_{\tau n}}}} \tag{9}$$

where a_t^{τ} denotes the tradable composite good and a_t^{n} the non-tradable good, $0 < \chi_{\tau} < 1$ denotes distribution parameter and $\mu_{\tau n} > 0$ is the elasticity of substitution between tradable composite good and non-tradable good.

The tradable composite goods is produced using importable and exportable goods as intermediate inputs via CES technology

$$a_t^{\tau} = A\left(a_t^{m}, a_t^{x}\right) = \left[\chi_{m}\left(a_t^{m}\right)^{1-\frac{1}{\mu_{mx}}} + \left(1-\chi_{m}\right)\left(a_t^{x}\right)^{1-\frac{1}{\mu_{mx}}}\right]^{\frac{1}{1-\frac{1}{\mu_{mx}}}} \tag{10}$$

where a_t^{m} denotes import-able good and a_t^{x} the export-able good, $0 < \chi_m < 1$ denotes distribution parameter and $\mu_{mx} > 0$ denotes the elasticity of substitution between import-able and export-able goods. Import-able, export-able and non-tradable goods are produced with capital and labour via the Cobb-Douglas technologies:

$$y_t^{j} = A^{j}\left(k_t^{j}\right)^{\alpha_j}\left(h_t^{j}\right)^{1-\alpha_j}\ ;\ \forall j \in \left\{x, m, n\right\} \tag{11}$$

where sector $j \in \{m,x,n\}$, y_t^j denotes output and A_t^j denotes total factor productivity in the in sector j.

To ensure a stationary equilibrium process for external debt, we assume that the country interest-rate premium is debt elastic

$$r_t = r^{*} + \psi\left(e^{d_{t+1}-\bar{d}} - 1\right) \tag{12}$$

where r^* denotes the sum of world interest rate and the constant component of the interest-rate premium, the last term of (12) is the debt-elastic component of the country interest-rate premium and we assume the parameter debt-elastic $\psi > 0$.

Model implied terms-of-trade f_t is assumed to follow AR(1) process

$$\log\frac{f_t}{\bar{f}} = \rho\log\frac{f_{t-1}}{\bar{f}} + \pi\varepsilon_t \tag{13}$$

where ε_t is a white noise with mean zero and unit variance, and $\bar{f} > 0$. The serial correlation parameter is $0 < \rho < 1$ and terms-of-trade standard error is $\pi > 0$.

For details of households' and firms' problem first-order conditions, market clearing and competitive equilibrium derivation and definitions see Uribe and Schmitt-Grohé [15].

Calibrating the model we follow Uribe and Schmitt-Grohé [15] process. The calibrated values of the model parameters are in the Table 1 for Czech economy. We assume that values of the parameters σ, δ, r^* in Czech repulic fit with the values commonly used for calibrating small open economies (Uribe and Schmitt-Grohé [15]) We assume that wage elasticity is same in all three sectors. Uribe and Schmitt-Grohé [15] provide a rich discussion with literature references on calibrating the elasticity of substitution between import-able and export-able goods. The values of ω_m, ω_x, ω_m, $\mu_{\tau n}$, α_m, α_x, α_n, are gathered from Ambriško [2]. Considering high-frequently (i.e. quarterly) data it is assumed that $\mu_{mx} = 0.8$ in Czech economy. Calibrating $\bar{f}$, A^m and A^n we adopt values of Uribe and Schmitt-Grohé [15]. The values of terms-of-trade serial correlation, ρ, and standard error, π, correspond to the data characteristics used in empirical models. To calibrate χ_m, χ_τ and A^x we follow a process of Uribe and Schmitt-Grohé [15] and implied moment restrictions of average share of value-added exports in *GDP*, s_x, average trade balance-to-*GDP* ratio, s_{tb}, and average share of non-tradable goods in *GDP*, s_n. Likewise Uribe and Schmitt-Grohé [15] we use OECD Trade in Value-Added (TiVA) [13] and UNCTAD statistical databases [14] to find values of these moment restriction. The values of the rest implied structural parameters, $\bar{d}$ and β come from the values of calibrated ones. We fail to reach a negative reaction of the trade balance to a terms-of-trade shock in the theoretical MXN model to follow empirical facts observed in the Figure 1. After substituting big values for the parameter ψ, the response of the trade balance is negative but very small in absolute value. Therefore we calibrate ϕ_j, $j \in \{m,x,n\}$ and ψ to capture moments observed in the empirical model. Firstly, using SVAR model in the section 2 we estimate the value of investment-terms-of-trade volatility ratio conditional on terms-of-trade shocks to equal approximately 0.45. Secondly, as Uribe and Schmitt-Grohé [15] pointed out, the standard deviation conditional on terms-of-trade shock of investment in the trade sector is 1.5 times as large as its counterpart in the non-traded sector.

calibrated structural parameters			**moment restrictions**		
σ	2.00	Uribe and Schmitt-Grohé [15]	s_n	0.27	UNCTAD [14]
δ	0.10		s_x	0.34	OECD [13]
r^*	0.04		s_{tb}	0.08	
ω_m	2.70	Ambriško [2]	$p^m y^m/(p^x y^x)$	1.00	Uribe and Schmitt-Grohé [15]
ω_x	2.70		$\sigma_{im+ix}/\sigma_{in}$	1.50	
ω_n	2.70		σ_i/σ_{tot}	0.45	Empirical model
$\mu_{\tau n}$	0.76		**implied structural parameter values**		
α_m	0.35		χ_m	0.672	
α_x	0.35		χ_τ	0.979	
α_n	0.25		$\overline{d}$	1.800	
μ_{mx}	0.80	Uribe and Schmitt-Grohé [15]	A^x	1.436	
$\overline{f}$	1.00		β	0.962	
A^m	1.00		ϕ_m	0.0125	
A^n	1.00		ϕ_x	0.021	
π	0.013	Empirical model	ϕ_n	0	
ρ	0.464		ψ	0.0176	

Table 1 Calibration of the MXN Model in the Czech Republic

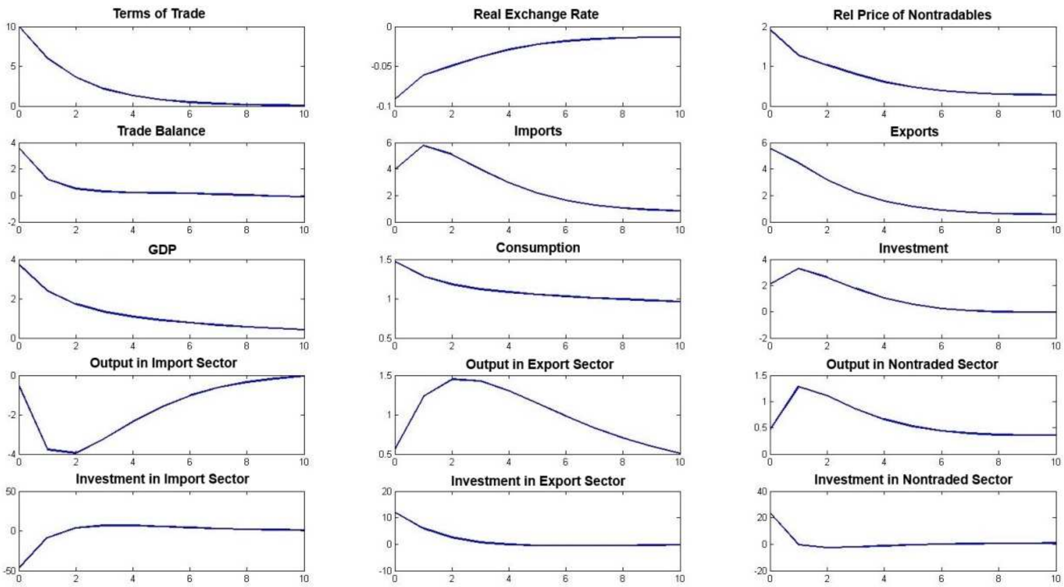


Figure 2 Impulse Response Functions to Terms-of-trade (TOT) Shock in the Czech Republic

In order of finding model equilibrium the first order linear approximation to the nonlinear solution are applied using algorithms created and modified by Klein [7] and Schmitt-Grohé and Uribe [11]. Responses to the terms-of-trade impulses and covariance-variance matrix conditional on the terms-of-trade shock is computed using algorithm of Uribe and Schmitt-Grohé [15].

4 Conclusion

In empirical models we observe small impact of terms-of-trade on business cycles in Czech Republic. In Czech Republic investment reacts immediately, while other aggregates do not change (or they change later mostly as reaction of other variables) on terms-of-trade shock. Terms-of-trade has negative effect on the trade balance in Czech Republic.

However, theoretical model calibrated to suit empirical observations overestimates the influence of terms-of-trade shocks in Czech Republic. Both output and investment rise after terms-of-trade shock realization in Czech Republic. The theoretical falls in real interest rates are overestimated as well. As we already pointed out, we cannot achieve a negative reaction of the trade balance in the theoretical model as it is in the empirical model.

Acknowledgements

Supported by the grant No. 1/0444/15 of the Grant Agency of Slovak Republic - VEGA.

References

[1] Aguirre, E.: *Business Cycles in Emerging Markets and Implications for the Real Exchange Rate*. Ph.D. Dissertation, Columbia University, New York, 2011.
[2] Ambriško, R.: *A Small Open Economy with the Balassa-Samuelson Effect*. Working Paper Series (ISSN 1211-3298), Charles University CERGE EI, Prague, 2015.
[3] Amisano, G., and Giannini, C.: *Topics in Structural VAR Econometrics*, 2nd ed. Springer-Verlag, Berlin, 1997.
[4] Broda, C.: Terms of Trade and Exchange Rate Regimes in Developing Countries. *Journal of International Economics* **63** (2004), 31–58.
[5] EUROSTAT portal. Retrieved from http://ec.europa.eu/eurostat/data/database [Accessed 4th April 2016]
[6] Harberger, A. C.: Currency Depreciation, Income, and the Balance of Trade. *Journal of Political Economy* **58** (1950), 47–60.
[7] Klein, P.: Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model. *Journal of Economic Dynamics and Control* **24** (2000), 1405–23.
[8] Laursen, S., and Metzler, L. A.: Flexible Exchange Rates and the Theory of Employment. *Review of Economics and Statistics* **32** (1950), 281–299.
[9] Lütkepohl, H.: *New Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin, 2005.
[10] Obsfeld, M.: Aggregate Spending and the Terms of Trade: Is There a Laursen-Metzler Effect? *Quarterly Journal of Economics* **97** (1982), 251–270.
[11] Schmitt-Grohé, S. and Uribe, M.: Closing Small Open Economy Models. *Journal of International Economics* **61** (2003), 163–185.
[12] Svensson, L. E. O., and Razin, A.: The Terms of Trade and the Current Account: The Harberger-Laursen-Metzler Effect. *Journal of Political Economy* **91**(1983), 97–125.
[13] Trade in Value Added database (OECD.STAT). Retrieved from: http://www.oecd.org/sti/ind/measuringtradeinvalue-addedanoecd-wtojointinitiative.htm [Accessed 4th March 2016]
[14] UNCTAD database. Retrieved from http://unctadstat.unctad.org/wds/ReportFolders/reportFolders.aspx?IF_ActivePath=P,15912 [Accessed 4th March 2016]
[15] Uribe, M., and Schmitt-Grohé, S.: *Open Economy Macroeconomics*. Columbia University, New York, 2016. Retrieved from http://www.columbia.edu/~mu2166/book/ [Accessed 4th March 2016]

The Dynamic Behaviour of Wonderland Population–Development–Environment Model

Blanka Šedivá [1]

Abstract. The Wonderland Population-Development-Environment Model (PDE) allows to study the interactions between the economic, demographic and environment factors of an idealized world, thereby enabling them to obtain insights transferable to the real world. This model was first introduced by Sanderson in 1994 and now there are several modification of this model. From a mathematical perspective, the PDE model is a system of non-linear differential equations characterized by slow-fast dynamics. This means that some of the system variables vary much faster than others. The existence of speed of dynamical in variables in model implies problems with numerical solution of models. This article concentrates on the numerical solutions of model and on the visualization dynamical behaviour of a four dimensional continuous dynamical system, the Wonderland model. We analyse the behaviour of model for selected part of the parametric space and we showed that the system of four differential equations Wonderland model can generate behaviour typical for chaotic dynamic systems.

Keywords: Dynamical System, Population–Development–Environment Model, Slow–Fast Dynamics.

JEL classification: C63
AMS classification: 34C60, 91B55

1 Introduction

Many models focused on relations relationships between economic factors, demographic and environmental factors have been developed to address climate change policy and these models are used for analysing several scenarios of future development of population from a critical (nightmare) scenario to a dream scenario. Between these two extreme limitations is usually model of the sustainable development expected. Sustainable development, although a widely used phrase and idea, has many different meanings and therefore provokes many different responses. In broad terms, the concept of sustainable development is an attempt to combine growing concerns about a range of environmental issues with socio-economic issues.

Today's economic and civilization development does not imply, in the opinion of many experts, sustainable concept for future. The economic expansion leads to population growth. Population in the world is currently (2016) growing at a rate of around 1.13% per year. The current average population change is estimated at around 80 million per year. Annual growth rate reached its peak in the late 1960s, when it was at 2% and above. The rate of increase has therefore almost halved since its peak of 2.19 percent, which was reached in 1963. The annual growth rate is currently declining and is projected to continue to decline in the coming years. Currently, it is estimated that it will become less than 1% by 2020 and less than 0.5% by 2050. This means that world population will continue to grow in the 21st century, but at a slower rate compared to the recent past. World population has doubled (100% increase) in 40 years from 1959 (3 billion) to 1999 (6 billion). It is now estimated that it will take a further 39 years to increase by another 50%, to become 9 billion by 2038. The latest United Nations projections *http://esa.un.org/unpd/wpp/* indicate that world population will reach 10 billion persons in the year 2056 (six years earlier than previously estimated).

The endless population growth creates more pressure on economic development and the related increasing of economic production puts a strain on the ability of the natural environment to absorb the

[1] University of West Bohemia, Plzeň, Czech Republic, Department of Mathematics, sediva@kma.zcu.cz

high level of pollutants that are created as a part of this economic growth. Therefore, solutions need to be found so that the economies of the world can continue to grow, but not at the expense of the public good. In the world of economics the amount of environmental quality must be considered as limited in supply and therefore is treated as a scarce resource. This is a resource to be protected and the only real efficient way to do it in a market economy is to look at the overall situation of pollution from a benefit-cost perspective.

This article focuses on the numerical solutions and the visualisation dynamical behaviour of the Wonderland model, which can be regarded as one of the simple dynamic models of geographic, economic and environmental interactions. The special feature of Wonderland model is the fact that not all system variables evolve with the same velocity. The velocity varies by two magnitudes and this slow fast dynamics makes solving difficult. The resulting mixture of slow and fast dynamics can lead to unpredictable, catastrophic transition even if all functions in model are deterministic, i.e. no stochastic force are introduced.

2 Sanderson Wonderland model

The first version of Wonderland model was published by Warren Sanderson as a part of IIASA study - The International Institute for Applied System Analysis in 1994 [8]. The original model was written in discrete time. Next, the model has been reformulated in continuous variables and this system has been used to investigate sustainability of economic growth [5],[2], [3] and as a benchmark problem for visualization techniques for higher dimensional dynamical systems [9].

In this parer the continuous version of Wonderland model of economic, demographic and environmental interaction is used. The dynamics in Wonderland is determined by four state variables:

- $x(t)$ - population, demographic variable;
- $y(t)$ - per capital output, economic performance;
- $p(t)$ - pollution per unit of output;
- $z(t)$ - quality of environment.

The state variables for population x and per capita output y can assume all non-negative real values $x, y \in [0, \infty)$, while the stock of natural capital z and the pollution per unit of output p are confined to the unit interval $z, p \in [0, 1]$. The variable z can be interpreted as a level of natural capital. If the natural capital is not polluted at all it takes on the value $z \approx 1$. On the other extreme, when the environmental is so polluted, that it produces the maximum possible damage to human health and to the economy, $z \approx 0$. The value of $p =\approx 1$ represent situation when there is maximal pollution per unit of output and on the other hand $p \approx 0$ implies no pollution per unit of output.

These four state variables evolve according to following set of non-linear, difference equations:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x \cdot n(y, z) \tag{1}$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = y \cdot \left[\gamma - (\gamma + \eta)(1 - z)^{\lambda}\right] \tag{2}$$
$$\frac{\mathrm{d}z}{\mathrm{d}t} = \nu z(1 - z) \left[\mathrm{e}^{\omega(g(z) - f(x,y,z,p)) - 1}\right] \tag{3}$$
$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\chi p \tag{4}$$

where

$c(y,z) = \varphi(1-z)^{\mu} y$	is pollution control,
$\overline{y}(y,z) = y - c(y,z)$	is net per capital output,
$f(x,y,z,p) = pxy - \kappa \left[\frac{e^{\sigma c(y,z)x}}{1+e^{\sigma c(y,z)x}} - \frac{1}{2} \right]$	is pollution flow,
$n(y,z) = b(y,z) - d(y,z)$	is population growth rate, that is calculated as difference between crude birth rate $b(y,z)$ and crude death rate $b(y,z)$.

The complex dynamical model includes 20 parameters, which can be grouped as follows

demographic:	β_1, β_2, β, δ_1, δ_2, δ_3, α, ϑ,
economy:	γ, η, λ,
environmental:	κ, σ, δ, ρ, ω, ν,
environmental policy:	φ, μ, χ.

2.1 Population $\frac{dx}{dt} = x \cdot n(y,z)$

The first differential equation describes population growth which depends endogenously only on per capital output y and the level of natural capital z. In eq. (1), the population growth rate, $n(y,z)$, can be written as the difference between the crude birth rate, $b(y,z)$ (number of births per 1000 population per time at a given time step) and the crude death rate, $d(y,z)$ (number of deaths per 1000 population per time at a given time step):

$$b(y,z) \quad = \quad b(\overline{y}) = \beta_1 \left[\beta_2 - \frac{e^{\beta\,\overline{y}}}{1+e^{\beta\,\overline{y}}} \right] \tag{5}$$

$$d(y,z) \quad = \quad d(\overline{y},z) = \delta_1 \left[\delta_2 - \frac{e^{\alpha\,\overline{y}}}{1+e^{\alpha\,\overline{y}}} \right] \cdot \left[1 + \delta_3(1-z)^{\vartheta} \right] \tag{6}$$

The parameters β_1, β_2, β govern the birth rate, while the parameters δ_1, δ_2, δ_3, α and ϑ govern the death rate. At eq. (5) and (6) it can be seen how both the birth rate and death rate decrease with increases in the per capita output, y. Furthermore, in eq. (6), the death rate is seen to increase when the environment deteriorates, i.e., when z decreases. These effects are in line with recent studies relating population growth with industrial output [1]. The maximum crude birth rate is realised for zero net per capital output $b(\overline{y}=0) = \beta_1(\beta_2 - \frac{1}{2})$ and minimal crude birth rate is the limit situation for unlimited grow of output: $\lim_{\overline{y}\to\infty} b(\overline{y}) = \beta_1(\beta_2 - 1)$.

In the literature [9] there exists also other type of function, that describes the population's growth in form $b_2(\overline{y},z) = \beta_1 \left[\beta_2 - \frac{1}{2}\left(1 + \frac{\beta\,\overline{y}}{1+\beta\,\overline{y}}\right)\right]$. The limit behaviour of this alternative function is the same as for the function previous formulated $b(\overline{y})$.

2.2 Product $\frac{dy}{dt} = y \cdot \left[\gamma - (\gamma+\eta)(1-z)^{\lambda}\right]$

Net per capital output $\overline{y}(y,z)$ is defined as per capital output y minus expenditures on pollution control $c(y,z)$. Pollution control expenditures $c(y,z) = \varphi(1-z)^{\mu} y$ depend on how polluted environmental is and not on the current flow emissions. The availability of natural capital also influence the growth rate of economy as indicated by equation (2). When environmental is totally polluted, i.e. $z \approx 0$, per capital output shrink at the rate $-\eta$, while if environmental is not polluted at all, i.e $z \approx 1$ per capital output increases at rate γ.

2.3 Quality of environmental (natural capital) $\frac{dz}{dt} = \nu z(1-z)\left[e^{\omega(g(z)-f(x,y,z,p))-1}\right]$

The growth of quality of environmental is depended on all state variables. The first part of equation (3) assumes the logistic model for quality of environmental. The speed at which natural capital regenerates is indicated by term $\nu\left[e^{\omega(g(z)-f(x,y,z,p))-1}\right]$.

This term depends positively on the self-cleaning ability of nature, which is described by function $g(z) = \frac{\delta}{\omega} z^{\rho}$, and this term depends negatively on the amount of pollution, which is modelled function

$$f(x,y,z,p) \quad = \quad pxy - \kappa \left[\frac{\mathrm{e}^{\sigma c(y,z)x}}{1 + \mathrm{e}^{\sigma c(y,z)x}} \right] \tag{7}$$

The difference between two flows $g(z) - f(x,y,z,p)$ is the net effect of natural and human forces on environmental. As we can see the model of dynamic behaviour quality of environmental is based on the assumption that the interaction of net effect on environmental is non-linear.

2.4 The pollution flow $\frac{\mathrm{d}p}{\mathrm{d}t} = -\chi p$

Pollution per units of output p exponential falls towards to zero as technology improves and environmental consciousness growth with rate χ.

3 Numerical experiment

The numerical experiments were inspirited by the origin Sanderson's work [8] and special attention was focused on analysis the behaviour the solving of differential system equations for different values of parameters χ, φ and κ. These parameters involve to the governing the expenditures for pollution abatement. The parameter φ from function $c(y,z)$ influences how expenditure increase.

3.1 Slow–fast dynamics

The numerical calculation have been performed using the Matlab R2013b. For numerical solving a stiff solver was used, that is based on an implicit Runge-Kutta formula for solving stiff differential equations with a trapezoidal rule step as its first stage and a backward differentiation formula of order two as its second stage.

The Wonderland model is an example of set differential equations exhibits slow-fast behaviour. One system variable (in this model z) varies much faster then the others (x, y and p). The analyse this phenomenon, we can rewrite the system using a scaling factor ε with $0 < \varepsilon \ll 1$. The scaling factor is called the perturbation parameter in singular perturbation theory [4]. The application of methods of perturbation theory enables us to separate slow and fast components of the system and to analyse both components separately. The reduced system captures the slow dynamics reformulate the equation (3) to form $0 = \nu z(1-z) \left[\mathrm{e}^{\omega(g(z) - f(x,y,z,p)) - 1}\right]$. The layer system captures the fast dynamics, where right sides in equations (1),(2) and (4) are zero and dynamic of system reduced only on the dynamic equation (3) for pollution per unit of output.

3.2 Visualization results and dynamics

Visualisation the behaviour of dynamical systems can provide us a deeper understanding of underlying dynamics. For analysing the model we used the following value of parameters: **demographics factors** $\beta_1 = 0.04$, $\beta_2 = 1.375$, $\beta = 0.16$, $\delta_1 = 0.01$, $\delta_2 = 2.5$, $\delta_3 = 4$, $\alpha = 0.18$, $\vartheta = 15$, **economy factors** $\gamma = 0.02$, $\eta = 0.1$, $\lambda = 2$, **environmental** $\kappa \in (1, 100)$, $\sigma = 0.02$, $\delta = 1$, $\rho = 2$, $\omega = 0.1$, $\nu = 1$ and **environmental police** $\varphi \in (0,1)$, $\mu = 2$, $\chi \in (0.01, 0.02)$.

Our analysis has concentrated on changing three parameters χ, φ and κ while holding all the other parameters fixed. The parameter χ is one of the most responsible parameter for system, this one governs the economic decoupling rate, i.e. the rate at which technological innovation reduce the pollution flow per unit of output. Parameters φ and κ describe policy planning, the parameter φ can be interpreted as a rate at which expenditures increase, so the net per capital output is calculated as $\overline{y}(y,z) = y - \varphi(1-z)^{\mu} y$. Since we require $c < y$, this limit $\varphi < 1$. The parameter κ is included in pollution function $f(x,y,z)$ (7) and determines the effectiveness of expenditures.

The future of Wonderland economy modelled by our system of differential equations is demonstrated for three sets of parameters that describe the three basic scenarios:

(A) Economists' dream scenario $\chi = 0.02$, $\kappa = 20$ and $\varphi = 0.5$. In this case, the per capital output increases over time, population converges toward stationary level of zero growth rate and the pollution flows steadily decreases.

(B) Environmentalists' nightmare scenario $\chi = 0.01$, $\kappa = 20$ and $\varphi = 0.5$. In this case, when parameter χ is reduced from 0.02 to 0.01, the system don't obeys the criteria of sustainability.

(C) Sustainable development scenario $\chi = 0.01$, $\kappa = 70$ and $\varphi = 0.5$ describes the situation when the progress of technology can help to sustainable economic increase.

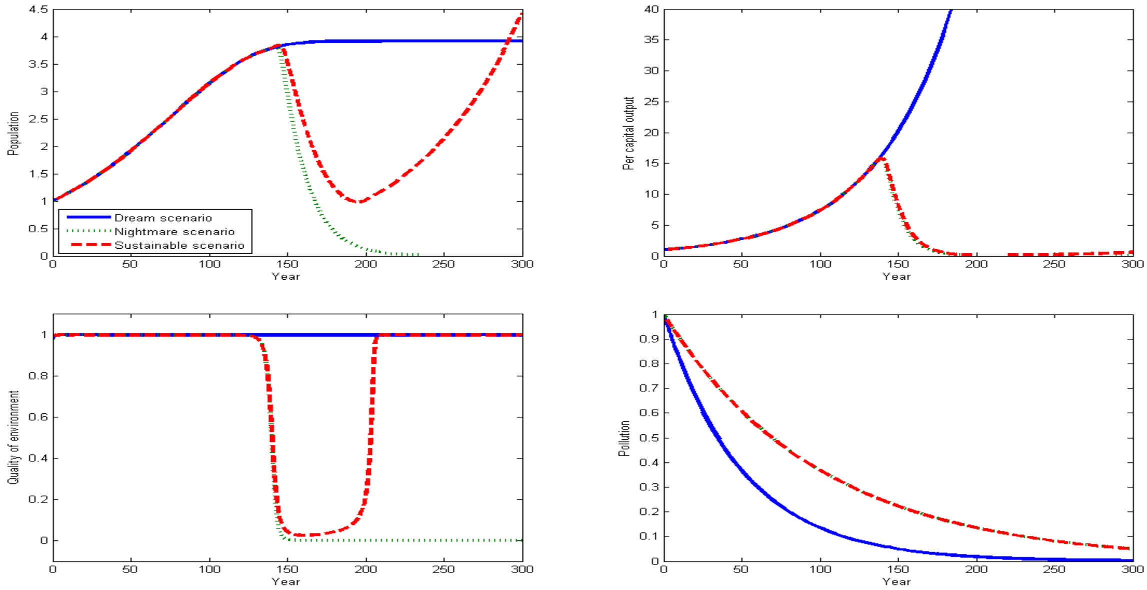


Figure 1 The population $x(t)$, per capital output $y(t)$, quality of environment $z(t)$ and pollution $p(t)$ for scenarios **(A)**, **(B)** and **(C)**

The other numerical experiments has been focused on the stability of the solution. We analyse the behaviour of the solution system for parameter κ from interval 40 to 100 and parameter φ from interval $(0.2, 0.5)$. The modelled value of per capital output at the time $t = 300$ is in the pictures bellow as a scaled coloured graph. There is the region of nightmare scenarios, where per capital output tends for $t \to 0$ to zero (black color) and the region of dream scenarios where per capital output tends to ∞ (white color). As we can see in detail, the border between this two scenarios is not exactly a clear line. This type of behaviour is typical of dynamic system operating in chaotic regime.

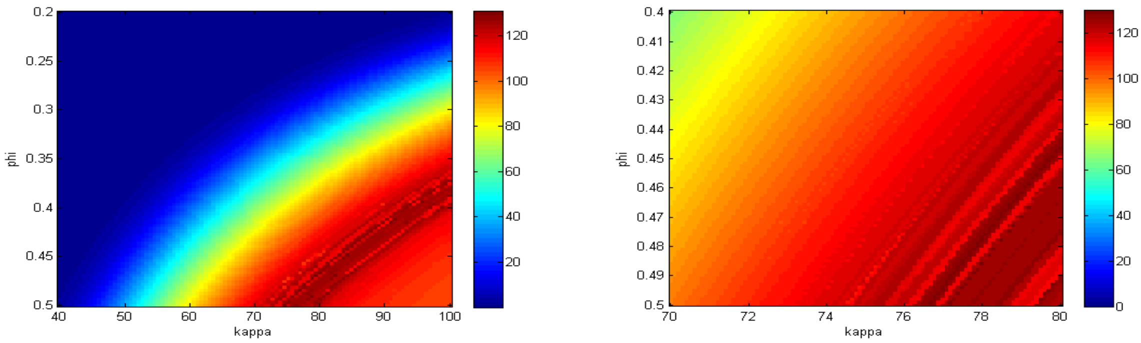


Figure 2 The scaled color graph of per capital output at time $t = 300$ for parametric region $\kappa \in (40, 100) \times \varphi \in (0.2, 0.5)$ (left picture) and detail for parameters $\kappa \in (70, 80) \times \varphi \in (0.4, 0.5)$ (right picture)

4 Conclusion

The behaviour of the four dimensional dynamical system which describe Wonderland Population-Development-Environment Model was presented. The future of the model for three basic scenarios – dream scenario, nightmare scenario and sustainable development scenario – was demonstrated. The behaviour of state variables for several part of parametric space (parameters χ, φ and κ) have been detailed analysed. The detailed comparison of parameters in the Wonderland model with estimation of this parameters based on real data sets are great challenge for our future works.

Of course, Wonderland is a toy model which captures the global features of complete human–environment interaction, but not its a detail. On the other hand though there are no stochastic, exogenous shocks, the future of state variables seems quite unpredictable in some part of parametric space that is typical for chaotic dynamic systems. The next studying of a chaotic behaviour of Wonderland model is the second motivation for our continuing interest in this model.

Acknowledgements

This article was supported by the European Regional Development Fund (ERDF), project "NTIS - New Technologies for the Information Society", European Centre of Excellence, CZ.1.05/1.1.00/02.0090.

References

[1] Bar, M. and Leukhina, O. M.: A model of historical evolution of output and population. *Federal Reserve Bank of Minneapolis, Available at: http://www.stanford.edu/group/SITE/papers2005/Leukhina.05.pdf* (2005).

[2] Guckenheimer, J., Krauskopf, B., Osinga, H. M. and Sandstede, B.: Invariant manifolds and global bifurcations. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25.9 (2015): 097604.

[3] Kohring, G. A.: Avoiding chaos in Wonderland. *Physica A: Statistical Mechanics and its Applications*, 368.1 (2006), pp. 214–224.

[4] Kuehn, C.: Geometric Singular Perturbation Theory. *Multiple Time Scale Dynamics.* Springer International Publishing, 2015, pp. 53–70.

[5] Milik, A., Prskawetz, A., Feichtinger, G. and Sanderson, W. C.: Slow-fast dynamics in Wonderland. *Environmental Modeling & Assessment*, 1.1-2 (1996), pp. 3–17.

[6] Milik, A. and Prskawetz, A.: Slow-fast dynamics in a model of population and resource growth. *Mathematical Population Studies*, 6.2 (1996), pp. 155–169.

[7] Pruyt, E and Kwakkel, J. H.: A bright future for system dynamics: From art to computational science and beyond. In *Proceedings of the 30th International Conference of the System Dynamics Society*, St. Gallen, Switzerland, 22-26 July 2012. System Dynamics Society.

[8] Sanderson, W. C.: Simulation models of demographic, economic, and environmental interactions. *PopulationDevelopmentEnvironment.* Springer Berlin Heidelberg, 1994, pp. 33–71.

[9] Wegenkittl, R., Grller, M. E., & Purgathofer, W.: *A guided tour to wonderland: Visualizing the slow-fast dynamics of an analytical dynamical system.* Technical Report TR-186-2-96-11, Institute of Computer Graphics and Algorithms, Vienna University of Technology, 1996.

Searching for suitable method for clustering the EU regions according to their agricultural characteristics

Ondřej Šimpach[1], Marie Pechrová[2]

Abstract. Not only Common Agricultural Policy, but also other policies at the EU level require unified approach. However, agriculture in various member states as same as in particular regions differs significantly. Therefore, the aim of the paper is to find the appropriate method which will group the most similar regions according to their agricultural characteristics and create appropriate number of groups which would be suitable for the application of the agricultural policy. These results from different methods of HCA are compared and discussed. Particularly single, average, complete, weighted average, median, centroid, and Ward's methods are applied. Also various metrics (Euclidean, Square Euclidean, absolute and maximum-value, Minkowski and Canberra distances, and correlation coefficient and angular separation measures) are used. The data about NUTS II regions of the EU are obtained from Eurostat for the latest available year (mainly 2013). Main indicators for the description of agricultural in particular region were: agricultural income, utilized agricultural area, labour and others. The most suitable well-balanced groups were created by Ward's method with Canberra distance.

Keywords: Clustering, measures, NUTS 2 regions, agricultural policy

JEL Classification: C38, Q18
AMS Classification: 62H30

1 Introduction

Spatial econometrics is an important tool for support of the policy-making decisions. The analyses enable to see the results of taken measures and suggest needed correction. For example research of Smith et al. [12] used spatial econometric techniques to evaluate RDPs in the European Union, at the NUTS 2 level. They focused specifically on labour productivity in the agricultural sector. At first side their results seemed to show that spending within the regional development programs on the competitiveness program (axis 1) had a statistically significant positive relationship with the increase of agricultural labour productivity in southern Europe. However, when their controlled properly for spatial effects (rural versus urban areas), the effect disappeared. "This shows how not taking spatial econometrics into account can lead to erroneous (policy) conclusions" [12]. Similarly Becker et al. [1] examined the effects of EU's structural policy (particularly of the Objective 1 facilitating convergence and cohesion within the EU regions). They observed the development of average annual growth of GDP per capita at purchasing power parity (PPP) during a programming period and average annual employment growth at NUTS 2 and NUTS 3 levels. Becker et al. [1] concluded that "Objective 1 treatment status does not cause immediate effects but takes, in the average programming period and region, at least four years to display growth effects on GDP per capita". Also Palevičienė and Dumčiuvienė [6] used multivariate statistical methods to analyse the EU's NUTS 2 level socio-economic data and to identify the clusters of socio-economic similarity. "The results showed that despite long lasting purposeful structural funds allocations there are still big regional development gaps between European Union member states" [6]. Pechrová and Šimpach [7] searched for the development potential of the NUTS 2 regions using hierarchical cluster analysis (Ward's method and Squared Euclidean distances), the regions were clustered into groups with same characteristics.

"The cluster analysis objective is to find out which objects are similar or dissimilar to each other" [9]. There are various clustering algorithms. Basic division is on hierarchical and non-hierarchical methods. First mentioned group contains agglomerative polythetic approach, two-dimension agglomerative clustering, division monothetic and division polythetic approaches. Non-hierarchical methods can be based on *k*-means algorithm or use fuzzy approach to cluster analysis. "When compared to standard clustering, fuzzy clustering provides more flexible and powerful data representation" [10]. Schäffer et al. [11] proposed new Bayesian approach for quantifying spatial clustering that employs a mixture of gamma distributions to model the squared distance of points to their second

[1] University of Economics Prague, Faculty of Informatics and Statistics, Department of Statistics and Probability, W. Churchill sq. 4, 130 67 Prague 3, Czech Republic, ondrej.simpach@vse.cz.
[2] Institute of Agricultural Economics and Information, Modelling of the Impacts of Agricultural Policy, Mánesova 1453/75, 120 00 Prague 2, Czech Republic, pechrova.marie@uzei.cz.

nearest neighbours. Ritter et al. [8] proposed a new method for autonomously finding clusters in spatial data belonging to the nearest neighbour approaches. “It is a repetitive technique which produces changing averages and deviations of nearest neighbour distance parameters and results in a final set of clusters” [8].

2 Data and methods

The aim of the paper is to find the appropriate method which will group the most similar regions according to their agricultural characteristics and create balanced groups. Firstly, the data are introduced, than the hierarchical cluster analysis methods and metrics used in the article are described. The data matrix was obtained from Eurostat [5] for the latest available years (mainly 2013). The data were not available for all EU’s regions (only for 2014). All NUTS 2 regions from Germany, Belgium and Slovenia were missing. Unfortunately, very important agricultural indicators such as the herds’ sizes and types, structure and acreage of crops were not available for majority of regions. Hence, they were not included into the analysis. The data were checked whether there is the correlation between them. Based on pairwise Pearson correlation coefficients (the values higher than 0.9), some of indicators were excluded from the analysis. In Table 1 we present the descriptive characteristics only for those variables which were chosen to be used in the next step. The agricultural holdings created an output in average height of 1 533 thousand CZK in year 2013. They used around 706 thous. ha of agricultural area and 411 thous. ha of arable land on average. The average share of arable land was 56% on average, but it differed across the EU. Similarly the share of agricultural land was higher (81%), but varied across the EU regions. Indicators related to the number of workers were highly correlated. Therefore, only the number of sole holders working on the farm was used. There were 45 898 of them on average. High standard deviation points on higher variability in the data. Sometimes it is even twice as high (in case of the number of sole holders working on the farm). The presence of variability in the data can negatively influence the results of clustering methods. Zero values stand for region Inner London with no agricultural production and land.

Used variables	**Mean**	**Std. Dev.**	**Min**	**Max**
Agricultural output (million EUR)	1533	1593	0	9 390
Utilized agricultural area (ha)	706 038	752 940	0	5 295 680
Arable land (ha)	410 724	474 818	0	3 371 340
Share of arable land	56%	27%	0%	99%
Share of agricultural land	81%	19%	0%	100%
Sole holders working on the farm (pers.)	45 898	99 096	0	749 260

Table 1 Descriptive characteristics of data from Eurostat (2016); own elaboration

As the variables are in different units standardized. Consequently, hierarchical cluster analysis was applied. At first, the resemblance matrix was calculated. Choice of similarity or dissimilarity measure depends on the type of variables (nominal, ordinal, ratio, interval, and binary). We utilized Euclidean, squared Euclidean, absolute-value, and maximum-value distances, and Minkowski distance with p argument, Minkowski distance with p argument raised to power (2), Canberra distance, correlation coefficient similarity measure, and angular separation similarity measure. Euclidean and squared Euclidean distances are based on Pythagoras theorem. Euclidean distance between two data points (X_i and Y_i) is calculated as the square root of the sum of the squares of the differences between corresponding values (1).

$$d = \sqrt{\sum_{i=1}^{n} \left(X_i - Y_i\right)^2} \tag{1}$$

(It is similar to Minkowski distance with argument p = 2.) The Euclidean Squared distance metric uses the same equation (1) without the square root. As a result, clustering with the Euclidean Squared distance metric is faster. Calculation of absolute value distance (i.e. Manhattan distance) is similar to Minkowski distance with argument p = 1. Maximum-value distance (Czebyshev) is calculated as Minkowski with $p \to \infty$. Minkowski distance with other arguments is calculated as (2).

$$d = \left(\sum_{i=0}^{n-1} \left| X_i - Y_i \right|^p \right)^{1/p} \tag{2}$$

It is often used when variables are measured on ratio scales with an absolute zero value. Its disadvantage is that even a few outliers with high values bias the result. Canberra metric is a dissimilarity coefficient defined in interval a_{jk} = < 0;1 >, where a_{jk} = 0.0 means maximum similarity when objects j and k are identical. “Each term in the sum is scaled between 0.0 and 1.0 equalizing the contribution of each attribute to overall similarity” [9]. Multiplier $1/n$ averages the n proportions (3):

$$a_{jk} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|X_{ij} - X_{ik}\right|}{\left(X_{ij} + X_{ik}\right)}. \tag{3}$$

Correlation coefficient similarity measure is defined in interval r_{jk} = < -1;1 >, where r_{jk} = 1.0 represents maximum correlation between the variables. Its advantage is that it can be applied on non-normalized data and its accuracy of the score increases. The calculation is as follows (4):

$$r_{jk} = \frac{\sum_{i=1}^{n} X_{ij}X_{ik} - \frac{1}{n}\sum_{i=1}^{n} X_{ij}\sum_{i=1}^{n} X_{ik}}{\sqrt{\left(\sum_{i=1}^{n} X_{ij}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} X_{ij}\right)^2\right)\left(\sum_{i=1}^{n} X_{ik}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} X_{kj}\right)^2\right)}}. \tag{4}$$

Angular separation similarity measure represents cosine vectors between two angles (5). It is defined in interval s_{jk} = < -1;1 >, where higher value indicates the similarity. The cosine of 0° is 1, for any other angle is < 1.

$$s_{jk} = \frac{\sum_{i=1}^{n} X_{ij}X_{ik}}{\left(\sum_{i=1}^{n} X_{ij}^2 \sum_{r=1}^{n} X_{ir}^2\right)^{1/2}} \tag{5}$$

Both above stated distances place anti-correlated objects maximally far apart (see e.g. [14]). Extended discussion on similarity measures can be found e.g. in [3].

Joining clusters by dissimilarity coefficients two clusters identified with the smallest dissimilarity coefficient values are merged. When using similarity coefficient, two clusters identified with the largest similarity coefficient value are joined. The article uses different clustering methods: single linkage (SLINK), average linkage method (i.e. unweighted pair-group method using arithmetic averages, UPMGA), complete linkage (CLINK), weighted average linkage, median linkage, centroid linkage, and Ward's linkage. Single linkage method finds the two most similar spanning objects in different clusters. SLINK tents to produce compacted trees. It is useful only when clusters are obviously separated. When objects are close to each other, SLINK tends to create long chain-like clusters that can have a relatively large distance separating observations at either end of the chain. Average linkage method defines the similarity between two clusters as the arithmetic average of the similarities between the objects in one cluster and objects in the other cluster. It usually produces trees which are between extremes of those created by single or complete linkage. It also tends to give higher values of the cophenetic correlation coefficient. "On average UPGMA produces less distortion in transforming the similarities between objects into a tree" [9]. It can also have a weighted form. In complete linkage clustering method the similarity between any two clusters is determined by the similarity of their two most dissimilar spanning objects. SLINK tends to produce clusters with similar diameters and extended trees. It is useful when the objects form naturally separated clusters. However, the results can be sensitive to outliers. Median linkage belongs to the averaging techniques, but uses medians instead of arithmetic means. This enables to mitigate the effect of possible outliers in the data. With the median linkage method, the distance between two clusters is the median distance between an observation in one cluster and an observation in the other cluster. Centroid method determines the distance of clusters by the distance of their centres (calculated as averages of real values). Its weighted form is useful when different size of clusters is expected. It usually utilizes the squared Euclidean distance metrics. Ward's method [15] merges the clusters with minimal within-cluster sum of squared deviations from objects to centroids. The distances of objects are again usually measured by squared Euclidean distance. It tends to create relatively small clusters because of the squared differences, but with similar numbers of observations. However, it is sensitive to outliers. Another disadvantage is that the distance between clusters calculated at one step of clustering is dependent on the distance calculated in previous step.

The clustering was cut when the number of clusters was 5 to create reasonable number of groups for policy treatment. The calculations were done in Stata 11.2 where above stated metrics and methods are available.

3 Results and discussion

For the policy making purposes, it is important that the clusters contain similar number of regions. We created 5 clusters by each method. Usual disadvantage of single linkage method – the tendency to chaining – appeared also in our application. Regardless the distance metrics created four clusters with 1 or 2 regions and 1 cluster (number 5) with others. Therefore, it will not be further taken into account. Average linkage method produced also groups where majority of regions was included in one cluster (mostly 1 or 2) and the remaining clusters included only few regions. Canberra distance rearranged the regions differently (majority of them was in 5th cluster, than in 1st). Correlation coefficient and angular separation similarity measures crated more balanced clusters in case of using average linkage clustering method. Complete linkage with majority of used metrics was also not optimal. Almost all regions were clustered to the 1st group. Using Canberra distance, majority of regions was included in 3rd group. It may be due to the presence of outliers in the data. Correlation coefficient and angular separation similarity measures created the most balanced clusters for complete linkage method. Weighted average linkage grouped to 3rd cluster more regions than its unweighted variant. Canberra distance in this case put more regions in 3rd cluster while in unweighted case it was the 5th regions. The number of regions in each cluster was balanced in case of using correlation coefficient and angular separation similarity measures. Median linkage put almost all regions into the 1st cluster. Only using angular separation similarity measure the 3rd cluster emerged as the biggest. Despite that it should mitigate the influence of the outliers to some extent, this was not the case. Probably London region with minimal values in almost all variables represented a problem to this method. Similar problem is with centroid linkage. All regions are in the 1st cluster, only with correlation coefficient and angular separation similarity measures the clusters are more balanced.

Despite that Ward's linkage is normally used with squared Euclidean distance, more feasible seems absolute and maximum values metrics (see Table 2). While the first mention creates one cluster with only three regions (cluster 5), the number of regions grouped by the other stated methods is more equal; but maximum-value method put only 10 regions in the fourth cluster. Therefore, Canberra distance seems to be optimal in terms of the number of regions in each cluster. Minkowski distance with *p* argument provided similar results to Euclidean distance and Minkowski distance with *p* argument raised to power to squared Euclidean distance in some cases as the value of parameter *p* was 2. Correlation coefficient and angular separation similarity created the most suitable groups. As the first one is only a special case of the latter one (angular separation standardized by centring the coordinates on its mean value), we may suggest using the angular separation similarity measure. Canberra distance enabled to create relatively well balanced groups in the last category. Despite that it has certain disadvantage with higher number of variables, this is not our case and we can recommend its usage.

Distance	**Cluster's number**				
	1	**2**	**3**	**4**	**5**
Euclidean distance	62	3	58	73	18
squared Euclidean distance	98	22	59	32	3
absolute-value distance	24	22	63	68	37
maximum-value distance	79	57	29	10	39
Minkowski distance with *p* argument	62	3	58	73	18
Minkowski distance with *p* argument raised to power	98	22	59	32	3
Canberra distance	71	42	23	45	33
correlation coefficient similarity measure	49	31	64	46	24
angular separation similarity measure	46	27	57	61	23

Table 2 Number of regions in clusters using Ward's linkage method and various metrics; own elaboration

As Ward's (and average) method is according to [4] the most suitable in majority of cases, we suggest combining it with Canberra distance method. Analogical results for UPGMA and SLINK methods were achieved by Bouguettaya et al. [2]. They compared distance measure functions and found out that Canberra distance seems better than the Euclidean counterpart. „With no noticeable difference in computational cost, correlation achieved by the Canberra method is consistently higher than the correlation obtained by the Euclidean method on a same data set with either UPGMA or SLINK" [2]. In our case, this approach grouped regions with minimal values in almost all variables to the fifth cluster and with maximal values to the first cluster (see Table 3).

Cluster	**Agric. output (million EUR)**	**Utilized agric. area (ha)**	**Arable land (ha)**	**Share of arable land**	**Share of agric. land**	**Sole holders working on the farm (pers.)**
1	2 827	1 474 755	871 046	62%	89%	96 704
2	1 189	510 578	224 192	45%	86%	43 615
3	1 198	560 681	473 554	83%	61%	21 287
4	827	160 985	79 111	51%	94%	6 579
5	382	145 464	66 148	44%	54%	10 266

Table 3 Characteristics of clusters created by Ward's method using Canberra distance; own elaboration

Regions grouped to the fifth cluster produced the less agricultural output with the less acreage of UAA and arable land. It contained mainly regions from France, Italy, and Poland. Regions in first cluster produce the highest agricultural output, utilize the most agricultural area and arable land in absolute values and had the most sole holders working on the farm. It included e.g. regions from Austria. The results of Ward's clustering method using different metrics are presented at Figure 1 in Appendix.

4 Conclusion

Agriculture in European Union varies in particular regions. As Common Agricultural Policy requires unified approach, the aim of the paper was to find the appropriate method which will group the most similar regions according to their agricultural characteristics and create appropriate number of groups which would be suitable for the application of the agricultural policy. The results from different methods of hierarchical cluster analysis showed that finding a suitable method for clustering regions is not an easy task. Each clustering method or distance metric provided different results (see also comparison in [13]). Besides, majority of them proved to be sensitive to outliers. Mostly the best results were provided by correlation coefficient and angular separation similarity measures. Surprisingly, often used Ward's linkage method with squared Euclidean distances did not provide well balanced groups. Normally this combination of methods tends to create relatively small clusters because of the squared differences, but with similar numbers of observations which is important for the application in the area of agricultural policy-making. Therefore, we recommend using Ward's method when clustering EU regions for policy purposes using rather Canberra distance. This measure provided well balanced groups with clearly different characteristics which enable to formulate appropriate political measures.

Acknowledgements

The research was supported by the Czech Science Foundation project no. P402/12/G097 DYME – "Dynamic Models in Economics" and also by Thematic task no. 4107/2016 of the Institute of Agricultural Economics and Information.

References

[1] Becker, S. O., Egger, P. H., von Ehrlich, M.: Going NUTS: The effect of EU Structural Funds on regional performance. *Journal of Public Economics* **94** (2010), 578–590.

[2] Bouguettaya, A., Yub, Q., Liub, X., Zhouc, X., Songa, A.: Efficient agglomerative hierarchical clustering. *Expert Systems with Applications* **42** (2015), 2785–2797.

[3] Cha, S. H.: Comprehensive Survey on Distance / Similarity Measures between Probability Density Functions. *International Journal of Mathematical Models and Methods in Applied Sciences* **4** (2007), 300–307.

[4] Dunn, G., Everitt, B. S.: *An introduction to mathematical taxonomy*. Cambridge University Press, 1982.

[5] Eurostat: Agriculture, forestry and fisheries. [on-line] Available at: http://ec.europa.eu/eurostat/data/database [cit. 2016-02-20].

[6] Palevičienė, A., Dumčiuvienė, D.: Socio-Economic Diversity of European Regions: Finding the Impact for Regional Performance. *Procedia Economics and Finance* **23** (2015), 1096–1101.

[7] Pechrová, M., Šimpach, O.: The Development Potential of the Regions of the EU. In: *Region in the Development of Society*. Mendel University in Brno, Brno, (2013), 322–335.

[8] Ritter, G. X., Nieves-Vázquez, J.-A., Urcid, G.: A simple statistics-based nearest neighbor cluster detection algorithm. *Pattern Recognition* **48** (2015), 918–932.

[9] Romesburg, H. C.: *Cluster Analysis For Researchers*. Lulu Press, North Carolina, 2004.

[10] Rovetta, S., Masulli, F.: Visual stability analysis for model selection in graded possibilistic clustering. *Information Sciences* **279** (2014), 37–51.

[11] Schäffer et al.: A Bayesian mixture model to quantify parameters of spatial clustering. *Computational Statistics and Data Analysis* **92** (2015), 163–176.

[12] Smith, M. J., van Leeuwen, E. S., Florax, R. J. G. M., de Groot H. L. F.: Rural development funding and agricultural labour productivity: A spatial analysis of the European Union at the NUTS2 level. *Ecological Indicators* **59** (2015), 6–18.

[13] Šimpach, O.: Application of Cluster Analysis on the Demographic Development of Municipalities in the Districts of Liberecky Region. In: *7th International Days of Statistics and Economics*. Slaný: Melandrium, (2013), 1390–1399.

[14] van Dongen, S. Enright, A.J.: Metric distances derived from cosine similarity and Pearson and Spearman correlations. [on-line] Available at: http://arxiv.org/abs/1208.3145 [cit. 2016-04-14].

[15] Ward, J. H.: Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58** (1963), 236–244.

Appendix

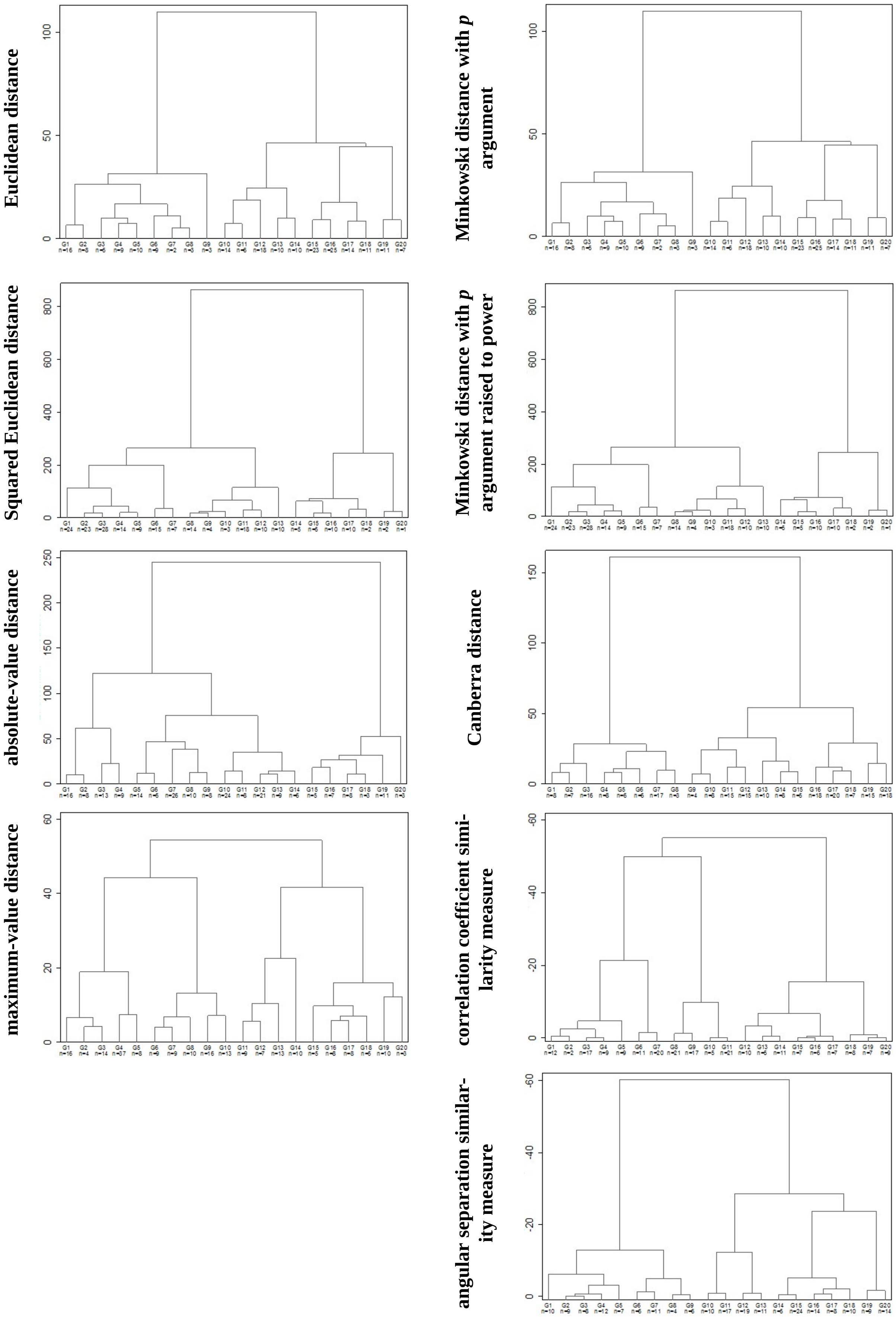


Figure 1 Dendrograms for Ward’s clustering method and various distance metrics; own elaboration

A computer-based VBA model for optimal spare parts logistics management in a manufacturing maintenance system

Eva Šlaichová[1], Márcio Rodrigues[2]

Abstract. Spare parts management in maintenance systems has become increasingly challenging in major global corporations, since it's hard to predict exactly when a piece of equipment fails and needs repair or replacement of its parts. Service time length, lead times, spare parts costs, the cost of non-production and service efficiency modelled with operations research tools and mathematical methods guide this multi-criteria decision-making process. Moreover, not always predictive and preventive maintenances are able to provide the adequate manufacturing management information about which equipment would need proper assistance and historical data of corrective maintenance hadn't revealed a periodic pattern of failure modes for accurate demand forecasting.

This paper proposes a conceptual model using VBA language that links timelines of preventive and corrective maintenances of equipments with stock levels and quality indicators, for a non-periodic time series for a more accurate management of optimal inventory levels and demand forecasting for spare parts in manufacturing maintenance systems. All tests and results acquired with actual corporate data as inputs are also presented.

Keywords: spare parts, maintenance, operations research, logistics.

JEL classification: C44, C53
AMS classification: 90B05

1 Introduction

The proper management of spare parts in maintenance systems has become very important in major companies, due to a continuous search for increasing plant efficiency, reduce unproductive times and most relevant, costs. It depends on a combination of many factors and cooperation of many key departments, not only setting up correctly the values of initial stocks and periodic purchase orders of such parts for each equipment but also combined with proper cost control to let plant management wisely decide which types of spare parts are suitable to stock or to purchase on-demand.

Key questions should be periodically asked and answered during this continuous process: Which are my key spare parts? How are their inventories? Keep or don't keep safety stocks? Which are the lead times of their suppliers? Is it wise to use Third-Party-Logistics (3PL)? How long are the production downtimes due to spare parts replacement or machinery repairs? And the most important: what's the impact on the final customer?

Additionally, pattern recognition to predict failure modes in industrial machines is directly linked with predictive and preventive maintenance plans of such equipments, through detailed analysis of intrinsic parameters of the process and the machines themselves, such as pressure variations, temperature ranges , tension, state of the components, calibration, electrical, hydraulic drives, sensors, impellers, thermostats, actuators, thermocouples, among many others and also associated with databases of historical data of non-conformities.

[1]Technical University of Liberec/Faculty of Economics, Department of Business Administration and Management, Studentska 1402/2 Liberec 1, 461 17 Czech Republic, eva.slaichova@tul.cz
[2]Universidade Federal da Bahia - UFBA/Escola Politécnica da UFBA, Departamento de Engenharia Mecânica, R. Prof. Aristides Novis, 2 - Federação, Salvador - BA, 40210-630, Brazil, marciovtr@gmail.com

The main goal of this paper is to propose a conceptual model for spare parts logistics in maintenance systems, which outputs demand and cost forecast key informations to help companies make wise and fast decisions regarding accurate spare parts purchases, according to corrective and preventive maintenance plans of main and bottleneck machines, time saving plus managing more efficiently their stocks and reduce costs.

The methodology used during the model development and this paper contribution includes the following activities:

- a wide literature review and consolidation of main concepts: maintenance types, Total Productive Maintenance, spare parts classification, inventory and demand forecasting methods;
- mathematical and economic model definitions, with the help of operations research tools and economic concepts;
- survey creation and submission for corporate data acquisition;
- real corporate data treatment and analysis;
- VBA programming and development of user interface;
- tests and empirical improvements.

This paper is divided in five sections. The first section is a brief introduction to spare parts logistics management and an overview of the main contribution, as described above. Section 2 presents the literature research regarding relevant inventory management methods applied to spare parts. Section 3 describes the main contribution of this paper, with both mathematical, economic and computational approaches of the proposed model. Section 4 discusses some of the results acquired using real corporate data and future improvement opportunities for the model and the topic. The last section resumes the main content discussed on the previous sections of this paper.

2 Literature Review

All over the past years many authors have developed and applied their own methods and models for proper spare parts management in maintenance systems. Aiming an accurate management of spare parts, two main aspects should be considered: demand forecasting and inventory policies. The next two subsections detail relevant used methods in major companies all over the past years.

2.1 Demand Forecasting methods

Relevant demand forecasting methods applied in major industries are shown as follows labelled by authors:

- Krever (2005): mean and variance of demand during lead time. Single Demand Approach (SDA) (Different than Periodic Demand Approach PDA). Variables of demand as a function of time, where

$$D(t) = (Quantity\,Ordered\,During\,Lead\,Time, Time\,Intervals\,Between\,Demands, Individual\,Lead\,Time) \tag{1}$$

- Croston (1972): an alternative method which separates the estimation of intervals between demands of the amounts demanded in each occurrence; [3]
- Willemain, Smart and Schwarz (2004): bootstrapping techniques for intermittent demands - assess the demand distribution during lead-time; [3]

2.2 Inventory methods applied to spare parts

The inventory models for products with high and independent demand is a consolidated area in operations management. From the original work by Harris (1913) on Economic Order Quantity (EOQ), different inventory models have been developed, including the following classic models: [2]

- Continuous Review (R, Q): in this model the inventory is continuously monitored and, when the level reached the order point "R" a lot size of "Q" (economic order quantity) is placed;
- Periodic Review (T, S): in this case, orders are placed in fixed interval of time "T", in an amount to replace the inventory position to the maximum inventory level "S";
- Base Stock (B): here, at each withdrawal from inventory, an order of the same amount is made for replacing the baseline, keeping the inventory position (inventory on hand plus open orders) constant and equals to "B";
- Periodic Review Model: maximum inventory demand distribution with accumulated historical information. Popovic (1987) developed a periodic review model in which the maximum inventory are determined based on a demand distribution estimated a priori and adjusted a posteriori using Bayesian inference with accumulated historical information. Initially, the demand rate (λ) is considered constant and, then, an alternative model is presented for time varying demand rate, given by
$$\Lambda(t) = (k+1)\lambda.t.k \tag{2}$$
- Petrovic's (1990): Needs to consider subjective aspects such as quality failure modes. Petrovic claims that the decisions in the spare parts management need to consider subjective aspects in addition to the traditional data for costs, lead times and demand. They developed an expert system to manage inventory with the premise that the distribution of the time between failures of a component follows the exponential distribution (therefore did not consider the cases of early mortality and aging/wearing). Users estimate the failure rates of components (or classify them in a subjective scale) and then answer questions about repairability, repair time, cost and criticality level of components. After inputting and processing data, the user gets the list of recommended lot sizes and the total expected inventory cost; [4]
- Aronis' (2004): used the Bayesian model to forecast the demands and determine the single parameter B of a base stock model for spare parts. Estimates and failure history of similar items are used a priori to fit the proper distribution of the demand and then the levels of required inventories (base level) are calculated; [5]
- Johnston and Boylan's (1996): Compared forecasting made through exponential smoothing to the ones made through Crostons method, and concluded that the latter is superior when the average interval between demands is greater than 1.25 time periods (time bucket). Syntetos and Boylan (2001) pointed out a bias in the original Croston's model and proposed a correction that gave rise to the SBA (Syntetos-Boylan Approximation) model; [6];[7]
- Krever's Intermittent demand (2005): computes the mean and variance of demand during the lead time. In their approach, known as Single Demand Approach (SDA), as opposed to the more traditional Periodic Demand Approach (PDA) with time buckets, three random variables are used: amounts demanded during the lead-time, time intervals between demand occurrences, and the lead-time itself. [8];[9]

In the following section is shown the proposed model and concepts to achieve the main contribution of this paper.

3 Conception of the proposed model

The basic concept of widely-used Croston's method for demand forecasting, consistent of separating estimation of intervals between demands of the amounts demanded in each occurrence was used as a starting point for the proposed model.

The main difference is that the timeline for demand events is a combination of both preventive maintenance date series for an specific equipment and spare part type and unpredicted corrective maintenance cumulative events time series to adjust previous amount demanded before next preventive maintenance scheduled event, as shown in Figure 1 below:

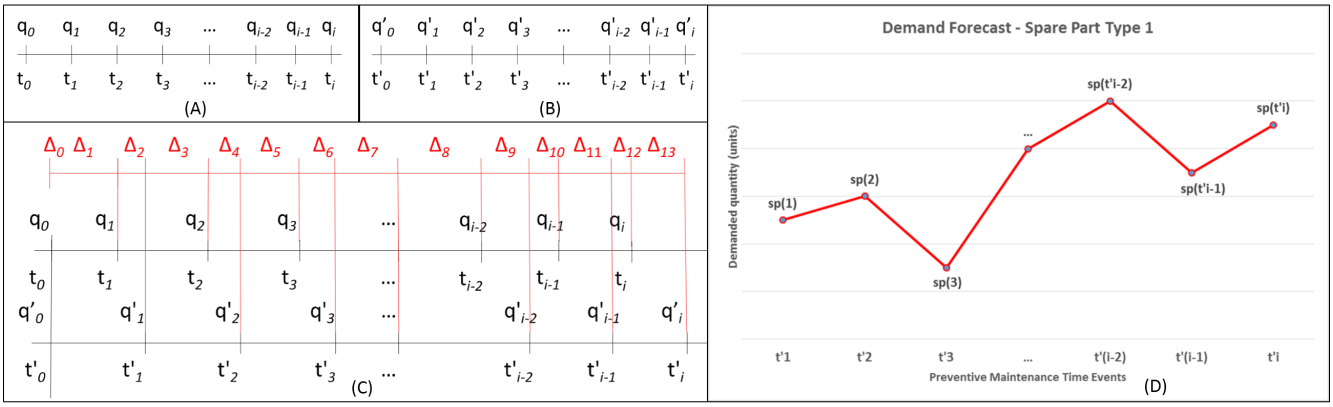


Figure 1: Preventive (A), Corrective (B) and Combined (C) maintenance. Conceptual demand forecast chart for Spare Part Type 1 (D)

As a direct implication of the unpredicted corrective maintenance events, once there is a new event the demand forecast needs to be adapted with new quantities. The following subsection describes the mathematical development of this time series.

3.1 Mathematical development of the spare parts inventory model

As represented in Figure 1 (A) and (B), each timeline diagram is a group of (t_i, q_i) and (t'_i, q'_i) pairs, whereas t_i and t'_i represent planned preventive maintenance dates and unpredicted corrective maintenance events and q_i and q'_i amount of spare parts planned/needed, respectively. Figure 1 (C) represents a combined timeline where (A)'s and (B)'s maintenance events are sorted in order of occurrence. An additional auxiliary variable Δ_i is used to represent time intervals between first (t_i) and subsequent time event $(t_{i+1}$ or $t'_{i+1})$ and this last one can be whether preventive or corrective.

For timeline consolidation, it is necessary that there must be a last time event, here addressed as T, which represents the last preventive maintenance time event for this machine that will require this spare part or in case the demand forecast is estimated for a lower range, this brings the restriction turned into the equation

$$\sum_{i=0}^{n} \Delta_i = \sum_{i=0}^{n} |(t_i - t'_{i-1})| \therefore \sum_{i=0}^{n} |(t_i - t'_{i-1})| \leq T \tag{3}$$

The amounts of spare parts demanded q_i and q'_i for preventive and corrective maintenance time events, respectively, and stock levels (in the following equations addressed as variable sp) for demand forecast are used to formulate the conceptual demand forecast amounts of spare parts. The initial stock levels for each type of spare part (sp_0) is assumed in this paper as an input data from the company and is chosen according to its own stock policies and limitations (warehouse capacity or size, inventory cost upper budget limit, supplier lead times, internal or external logistics):

$$sp(t_i) = sp(t_{i-1}) + sp_e(t_i - t_{i-1}) + r_q(i) + sp_0 \tag{4}$$

(4) represents the demand forecast as a function of preventive maintenance time series. As an conceptual output for the model, Figure 1 (D) presented a chart of the proposed theoretical conceptual demand forecast function.

3.2 Economic development of the spare parts inventory model

As previously mentioned in Section 1, the main goal of proposed model is to help companies to make smarter and quicker decisions for spare parts management, which leads to the main economic goal: costs reduction. From mathematical model presented on previous subsection, the initial stock levels once considered as given information were obtained using ABC curve of spare parts unitary costs and ABC curve of spare parts' suppliers costs from lower lead times for process critical machines. The decision to make is which demand forecast scenario delivers the minimum total cost, defined by the objective function $Z = f(q_i, q_i')$, where

$$Z_{min} = f(q_i, q_i', k_q) = min \begin{cases} \sum_{i=0}^{n} k_q(q_i + q_i'), & \text{when } q_i \neq q_i' \\ \int_{i=0}^{n} k_q . q_i \, di, & \text{when } q_i = q_i' \end{cases} \tag{5}$$

As a direct implication of (5), it's also possible to determine the total spare parts inventory cost forecast, as (5) is defined for each machine and each type of spare part. Considering conceptually a group of 10 machines (auxiliary variable a) and 15 different types of spare parts suitable to all 10 machines (auxiliary variable b) comes a new objective function $Y = f(Z)$, where

$$Y_{min} = f(q_i, q_i', k_{qab}) = min \begin{cases} \sum_{a=0}^{10} \sum_{b=0}^{15} \sum_{i=0}^{n} k_{qab}(q_i + q_i'), & \text{when } q_i \neq q_i' \\ \int_{a=0}^{10} \int_{b=0}^{15} \int_{i=0}^{n} k_{qab} . q_i \, da \, db \, di, & \text{when } q_i = q_i' \end{cases} \tag{6}$$

3.3 Computer-Based VBA model

Moreover, using VBA programming language, $Solver^{®}$ tool and $NeuroXL^{®}$ neural network add-in for $Microsoft^{®}Excel^{®}$ multicriteria decision ended up with the proposed spare parts logistics model user interface. It contains traceability information regarding each type of spare parts, quality non-conformity historical data, supplier lead times and all input functions allowing the users to generate reports.

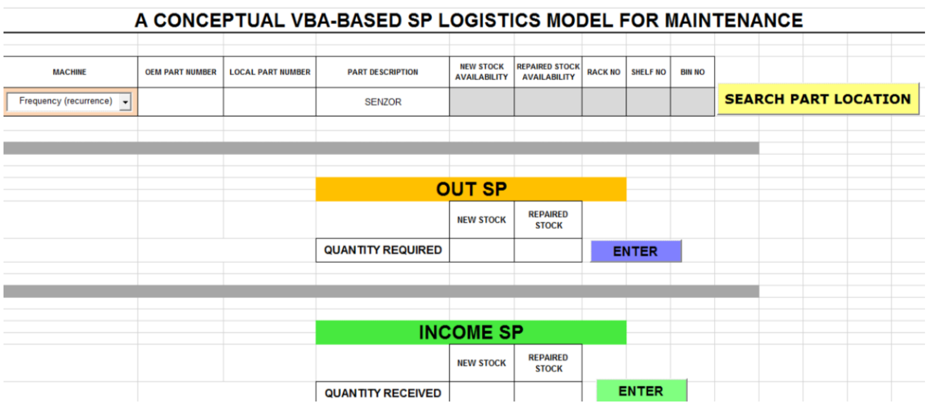
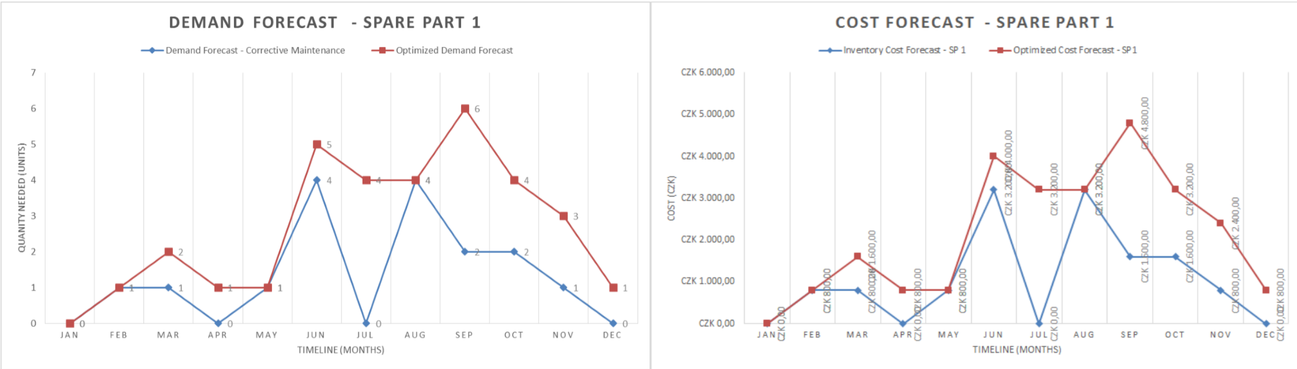


Figure 2: Model user interface

In the next section results regarding real data processing and analysis are presented.

4 Results and discussion

After tests and debugging, reports from proposed model provide optimized demand and cost information for each type of spare part in preventive maintenance timeline, as shown in Figures 3 (E) and (F) respectively, also total inventory cost for chosen plant and the amount of spare parts needed and many other criteria and filters can be applied to it, which proves to be a powerful tool in time-cost savings and can be widely used in complex spare parts logistics maintenance systems.



(E) (F)

Figure 3: Demand (E) and cost (F) forecast charts using company data

5 Conclusion

This paper intended to present a conceptual model for spare parts management in maintenance systems, showing its mathematical, economic and computational interfaces and working principles. Confirms how demanding and urgent the seek for optimization and cost reduction in industrial processes are nowadays. Future works of this paper are possible, where results can be compared with existing methods for spare parts management and also opportunities for its improvement.

References

[1] Rêgo, J., and Mesquita, M.: Spare parts inventory control: a literature review. In: *Production Journal '11*, vol. 21, issue 4, São Paulo, 2011, 656-666.

[2] Waller, D.: *Operations Management: A Supply Chain Approach*. Thomson Learning, London, 2003.

[3] Bacchetti, A. and Saccani, N.: Spare parts classification and demand forecasting for stock control: Investigating the gap between research and practice, *Omega* **40** (2012), 722–737.

[4] Rego, J. S. and de Mesquita, M. A.: Demand forecasting and inventory control: A simulation study on automotive spare parts, *International Journal of Production Economics* **161** (2015), 1–16.

[5] Martnez-Jurado, P. J. and Moyano-Fuentes, J.: Lean Management, Supply Chain Management and Sustainability: A Literature Review, *Journal of Cleaner Production* **85** (2014), 134–150.

[6] Bacchetti, A. and Saccani, N.: Spare parts classification and inventory management: a case study In: *Syntetos Salford Business School Working Paper (WP) Serie* (2010), 1–36.

[7] Gan, S., Zhisheng Zhang, Y. and Jinfei, S.: Joint optimization of maintenance, buffer, and spare parts for a production system, *Applied Mathematical Modelling* **39** (2015), 6032–6042.

[8] Jouni, P. and Huiskonen, J. and Pirttila, T: Improving global spare parts distribution chain performance through part categorization: a case study, *International Journal Production Economics* **133** (2011), 164–171.

[9] Behfard, S., van der Heijden, M. C. and Zijm, W. H., M.: Last time buy and repair decisions for spare parts distributions on the line, *European Journal of Operational Research* **244** (2015), 498–510.

Optimization of the Transportation Plan with a Multicriterial Programming Method

Eva Šlaichová[1], Eva Štichhauerová[2], Magdalena Zbránková[3]

Abstract. Environmental policy and corporate social responsibility have become very important criteria in considering of performance of the enterprises. Transportation is the second largest source of greenhouse gas emissions, in particular CO_2 and N_2O, in the EU, and road transport is so far the largest emission source within the transport sector. This article deals with the of best route choice decisions on vehicle energy consumption and emission rates for different vehicle types using multicriterial linear programming method. The purpose was to find the optimal transportation plan so as to simultaneously minimize the fuel consumption and the total CO_2 emission. The ecologic-economical model is tested on the real data of the selected logistical enterprise. The paper also presents general applicability of the multicriterial model. From the model results it cannot be implied that the amount of emissions produced was directly proportional to fuel consumption, rather the contrary.

Keywords: emission reduction, mathematical modeling, linear programming, operation analysis, sustainable development

JEL Classification: C02, Q52
AMS Classification: 90C05

Introduction

Greenhouse gas emissions and their component CO_2 have a significant negative effect on human health and negatively effects also the economy of the country. The Czech Republic deals with this urgent problem as many other countries in Europe and all over the world. With an increasing worldwide concern for the environment, logistics providers and freight carriers have started paying more attention to the negative externalities of their operations [4]. The question is whether between the amount of produced emissions and fuel consumption is a direct proportion. If so, it would be possible to prove that the environmental and economic aspects are important during the provision of transport services. In a number of, mainly small or medium-sized businesses, the environmental performance is a result of economic management choices, with little or no regard to their environmental impact. Furthermore companies are unlikely to invest in environmental issues unless they come under economic or legal pressure [8]. There is also an agreement that strategic decisions should have a larger impact on emissions than operative decisions. There is, however, a disagreement on what specific decisions have the largest impact, and what those decisions really will lead to regarding environmental impact [3].

In the frame of these issues it is possible to use also methods of the operation research, e.g. adaptive evolution algorithm [7]. However, linear programming seems to be one of the most efficient techniques because it can be used to solve problems dealing also with environmental and ecological topics [9]. The authors of the article decided to use two simple assignment models with different objective functions. The next step was to compare the results with the multi-objective (multi-criteria) linear programming with the weighted sum of these objectives. This article presents the results of the case study from the real logistic company (with modified data) focused on comparison of two models, model optimizing the transportation plan while minimizing emissions of CO_2 and the same model minimizing fuel costs.

[1] Faculty of Economics, Technical University of Liberec, Department of Business Administration and Management, Studentská 1402/2, Liberec 1, Czech Republic, eva.slaichova@tul.cz.
[2] Faculty of Economics, Technical University of Liberec, Department of Business Administration and Management, Studentská 1402/2, Liberec 1, Czech Republic, eva.stichhauerova@tul.cz.
[3] Faculty of Economics, Technical University of Liberec, Department of Business Administration and Management, Studentská 1402/2, Liberec 1, Czech Republic, magdalena.zbrankova@tul.cz.

1 Literature Review

Over the past two decades, many organizations have taken steps to integrate the principles of sustainability into their long and short-term decision-making [1]. At the local and regional levels, a significant portion of freight transportation is carried out by trucks, which emit a large amount of pollutants. In the transportation sector emissions are dominated by CO_2 emissions from burning fossil duele [4]. These cause atmospheric changes and climate disruptions which are harmful to the natural and built environments and pose health risks. Environmental issues and the inclusion of environmental strategies in strategic thinking are interesting subjects of investigation [13]. Already in 1995 Wu and Dunn discussed logistical decisions in light of their environmental impact [10]. Regarding transportation, they highlight three areas with high-environmental impacts, the construction of transport networks related to infrastructure and inventory location, the operation of vehicles, including fuel choice and the disposal of vehicles.

As Oberhofer claims, more efficient and reduced use of vehicles, fuel switching, renewable electricity and fuels and the use of hydrogen and biomass as transportation fuels should be considered as a means of stabilization [8]. On the other hand, Wu and Dunn point out that companies must re-evaluate where facilities are located, whom they cooperate with, what technology is used and the whole logistics structure [10]. According to the lean management principles environmental friendly logistics structures are characterised by fewer movements, less handling, shorter transportation distances, more direct shipping routes and better utilisation. However, the business main objective is not the protection of the environment and therefore companies strive from within to balance ecology and economy [11]. Therefore as Ghobadian (2006) declares, where cost savings can be generated from green practices, green logistics initiatives tend to be applied in the freight transport sector [5]. As Demir points out the minimization of the total travelled distance has been accepted as the most important objective in the field of vehicle routing and freight transportation [4]. The number, quality and the availability of the truck models with lower emmisions have increased considerably recently.

Most studies in the field of green transportation have focused on a limited number of factors, mainly vehicle load and speed. Mainly vehicle speed is shown to be quite important, of course the crucial is to make the drivers drive with minimum fuel consumption for a given routing plan. Light vehicles should be preferred over medium duty and heavy vehicles. Medium should be also preferred to heavy vehicles if possible. Another point is that positive road gradient leads to an increase in fuel consumption. This should be taken into account especially in connection with GIS software which can nowadays provide information of the road gradient. Besides CO_2 emissions pollutants, noise, accidents and environmental damage and other traffic externalities could be examined at the local and regional levels. A number of studies focus mainly on the routing aspect of green logistics [4]. However, other problems linked to routing presents former reductions in emissions. For example, facility location is concerned with locating a set of depots so as to maximize the profit generated by providing service to a set of customers. The relocation of a depot may lead to reductions in CO_2 emissions [6]. Driver working hours should be considered, not only because of labour law requirements, but as well as for the acknowledgement of the health hazards arising from demanding workload of some transportation plans. As Zhang claims, to enhance the interaction between companies and investors for higher environmental goals economic incentives to business-led voluntary initiatives should be used [12].

2 Optimization of the transportation plan while minimizing emissions of CO_2

This part of the contribution contains information about applied methodology, about the structure of the model, and the description of data used in the case study.

2.1 Data and methodology

This contribution presents the results of the case study on the example of the logistic company. The aim is to compare three models. A model optimizing the transportation plan while minimizing emissions of CO_2 and the same model minimizing fuel costs. The third model is based on multi-criteria evaluation of both factors. The data were modified at the request of the company. For a research and data collection, interview method was used.

Linear programming was applied to set both models because all the constraints can be described by the linear equations and inequations. Calculations were processed by the software tool Solver provided by Microsoft Excel 2007.

2.2 The multi-objective assignment problem – the structure of the model

The goal is to minimize logistics companies produced a total volume of CO_2 emissions per day, while minimizing the costs of daily fuel consumption when assigning trucks to routes from the main depot to the target depot. The prerequisite is that one day is realized transport on all routes just once.

The **multi-objective assignment problem** can be formulated as follows:

Minimize
$$z_1 = \sum_{i=1}^{m} \sum_{j=1}^{n} e_{ij} * x_{ij} \tag{1}$$

$$z_2 = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} * x_{ij} \tag{2}$$

$$x_{ij} \in \{0;1\} \tag{3}$$

$$\sum_{j=1}^{n} x_{ij} \leq 1 \tag{4}$$

$$\sum_{i=1}^{m} x_{ij} = 1 \tag{5}$$

$$i = 1, \ldots, m; j = 1, \ldots, n; m \geq n$$

where (1) the objective function to determine the total CO_2 emissions produced per day, and (2) the objective function to determine the total cost of the fuel consumption per day, the two criteria are of the minimization type.

E_ij structural factor is the amount of CO_2 emissions produced by the i-th vehicle on the route from the main depot to depot j-th and back (on the j-*th* line). Structural coefficient c_{ij} is the cost of fuel consumption i-th vehicle to the j-th route.

Logistics company has a number of m trucks with them it provides transport n routes, the number of vehicles exceeds the number of available routes. Theoretically there are $m * n$ possible combinations of vehicle and route.

When assigning cars to routes so that each route was secured by just one car and each car was assigned to no more than one route. An obligatory condition (3) indicates that the variable x_{ij} can have only two values; it equals 1 if a car is assigned to depot j and 0 otherwise.

Assuming that the number of vehicles in the fleet exceeds the number of routes that need to be serviced. For fleet logistics company at maximum one car can be assigned to a route between the capital and the target depot. Capacity of the fleet can not be exceeded. Thus, no more than one of the variables x_{ij} must be for the i-th car equal to 1, and this must apply to all vehicles. Line constraints are given in equation (4).

Each route between the depot and the depot target must be implemented in one car, at least one vehicle must be assigned to the route. Thus, just one of the variables x_{ij} must be for the j-th route is equal to 1, and this must apply to all route. Column constraints are given by equation array (5).

After application of the principle of aggregation of the objective function to the objective function (1) and (2) and the transformation of the original values of structural coefficients e_{ij} and $c_{i,j}$ the problem can be solved as a normal linear programming (6) under the conditions (3), (4) and (5):

Maximize
$$z = \sum_{i=1}^{m} \sum_{j=1}^{n} e_{ij}^{T} * x_{ij} * v_1 + \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}^{T} * x_{ij} * v_2 \tag{6}$$

$$e_{ij}^{T} \in \langle 0;1 \rangle$$

$$c_{ij}^{T} \in \langle 0;1 \rangle$$

where the weight v_1 has the relative importance of the criteria of the total volume of CO_2 emissions and weight v_2 indicates the relative importance of the criteria of the cost of the overall fuel consumption coefficients and the transformed values of the original structural coefficients e_{ij} and c_{ij}.

2.3 The application of the model on the real data

Necessary information about the composition of the fleet logistics companies are contained in Table 1. This is a truck with a maximum load between 15–19 m^3. Limitation in the case of vehicles B13 is the maximum length of the route without refueling (CNG), suggesting that it can be assigned only to routes within a distance of 400 km.

Type of the truck*	Identification of the truck B_i	Year of production	Emissions in CO_2 [g/km]	Consumption of fuel [l/km]*
Renault Master 2.8 DTI	B1	1999	287	0.108
Fiat Ducato Maxi 2.8 JTD	B2	2002	235	0.072
Mercedes-Benz Sprinter 308 CDI	B3	2004	282	0.107
Fiat Ducato 33 Multijet 120 L3H2	B4	2007	218	0.082
Mercedes-Benz Sprinter 315 CDI	B5	2009	252	0.104
Fiat Ducato 2.3 JTD 17H L4H3 Maxi	B6	2012	225	0.085
Renault Master 2.8 DTi	B7	2012	215	0.081
Iveco Daily 35S15V Euro 5	B8	2013	187	0.094
Mercedes-Benz Sprinter 319 CDI	B9	2014	205	0.078
Mercedes-Benz Sprinter 313 D Blue Efficiency +	B10	2014	195	0.075
Mercedes-Benz 319 CDI	B11	2014	207	0.080
VW LT35 - MAXI JUMBO	B12	2003	254	0.103
Fiat Ducato 3.0 Natural Power CNG L2H2	B13	2011	239	0.088**

* estimated combined consumption in v litres per 1 km

** consumption of gas CNG in kg per 1 km

Table 1 Volume of emmisions CO_2 and consumption of fuel regarding the type of the truck

Final Depot	Route A_j	Length of the route [km]
Ústí nad Labem	A1	214
Olomouc	A2	526
Hradec Králové	A3	250
Ostrava	A4	708
Brno	A5	376
Plzeň	A6	224
České Budějovice	A7	268
Liberec	A8	242
Chodov (KV)	A9	324
Jihlava	A10	226
Velké Přílepy + Praha 10*	A11	126 (tj. 92+34)

* two routes are secured by one vehicle

Table 2 Length of the route between main depot in Modletice and final depots outward journey included

Structural factors e_{ij} of the objective function z_1 (1) were obtained by multiplying the volume of CO_2 emissions i-th vehicle g/km (see Table 1) and the path length between the depot and the j-th depot in km (see Table 2).

Due to the fact that the fleet is made up of both diesel trucks, as well as one-consuming CNG, consumption was converted into monetary terms. For this purpose, the price of 1 liter of diesel was set at CZK 32 and the price of 1 kg of CNG gas CZK 26.

Structural coefficients c_{ij} from the objective function z_2 (2) were obtained by multiplying the monetary value of consumption fuel i-th vehicle CZK/km (see Table 1) and the length of the j-th km route (see Table 2).

An exception in the matrix of prices are coefficients $e_{13,2}$ and $e_{13,4.}$ To avoid assigning vehicles B13 routes to relevant depots Olomouc and Ostrava, distant from the main depot Modletice more than 400 km, the authors (Burkard et al., 2012) recommend to replace the original value with an extraordinary high value – there 999,999 (g or CZK). In the case of multi-criteria evaluation which work with the transformed values after conversion, additional equating coefficients are indigenous to the structural coefficients range from 0 to 1.

In the case study, the question of determining the exact weights of the criteria has not been addressed. Optimization was performed for the situation when two criteria are considered as important (e.g. the application of multi-criteria evaluation during the same values scales, where $v_1 = v_2 = 0.5$), and in extreme situations where it is important to only one or only the second criterion (e.g. monocriterial optimization, where $v_1 = 1$ and $v_2 = 0$, and vice versa).

3 Discussion of the results

The total cost of fuel consumption per day in a situation that the vehicles are assigned to routes with respect to the requirement for a minimum total volume of produced CO_2 is CZK 9,329, which is 4.29 % (CZK 384) more than cost-optimal combination.

Total daily volume of CO_2 emissions in a situation where the company respects the requirement for minimum consumption costs, is 766,062 g, which is about 3.77 % (27,800 g) than the combination of vehicles minimizing CO_2 emissions. If the vehicles are assigned to routes based on current optimization of the two criteria, the total daily volume is 749,190 grams of CO_2 emissions and total cost of consumption of fuel is CZK 9,031 per day.

Destination Depot	Route	Cars assigned to minimize:		
		both objectives together	**total CO_2 emissions per day**	**total cost of the consumption per day**
Ústí nad Labem	A1	**B12**	B5	**B12**
Olomouc	A2	**B9**	B10	B10
Hradec Králové	A3	**B7**	B4	**B7**
Ostrava	A4	**B10**	B8	B2
Brno	A5	**B11**	B9	B13
Plzeň	A6	**B6**	B13	B8
České Budějovice	A7	**B13**	B7	B11
Liberec	A8	**B4**	B6	**B4**
Chodov (KV)	A9	**B2**	B11	B9
Jihlava	A10	**B8**	B2	B6
Velké Přílepy + Praha 10	A11	**B5**	B12	**B5**
(unused car)		**B1**	**B1**	**B1**
(unused car)		**B3**	**B3**	**B3**
The total volume of CO_2 emissions per day		**749,190 g**	**738,262 g**	766,062 g
The total cost for fuel consumption per day		**CZK 9,031**	CZK 9,329	**CZK 8,945**

Table 3 Comparison of the models

The result of this article is an optimization of the transportation plan regarding important side effects, namely CO_2 emissions and fuel consumption. The contributions of this article are description of the model including the factors of the fuel consumption and CO_2 emissions into existing planning methods for vehicle routing, introduction of multicriterial integer programming formulations that minimizes a fuel consumption. The possible limitation of the research is that the model does not include labor costs and emission costs are not expressed as a function of load, speed and other parameters.

Results of the model analysis based on real data came to the following important conclusion. Within minimizing emissions from the cost perspectives also other factors should be taken into account, e.g. labor costs. Another point is that use of fewer vehicles generally implies a lower fuel consumption and a higher utilization rate of the vehicle capacity. Heavy vehicles typically generate more emissions at low speed than medium vehicles. Medium vehicles produce significantly low emissions under the same operational conditions, so it indicates that more low-emitting vehicles should be assigned to lower-speed routes.

As it can be also concluded, significant improvements in air quality and energy consumption are achievable by educating drivers [2]. Economical driving in addition to fuel economy carries even further: it is possible to calculate the savings of brakes, tires, clutch, etc. Safe and defensive driving style, which is linked to the driver's ability to anticipate and deal with crisis situations, eventually to avoid them, can translate even in lower accident rates and other related benefits, e.g. insurance, liability. It follows that drivers should learn the rules of economic and defensive driving and should be encouraged to respect them. In the case of achieving fuel savings, financial award can be offered to them in a certain amount for a certain periods.

From the model results it cannot be implied that the amount of emissions produced was directly proportional to fuel consumption, rather the contrary. Although according to the results, between low emissions and costs can be find a positive relationship. If the company prefer vehicles that have lower emissions for longer transport routes, it is possible to reduce overall emissions, thereby reducing the carbon. As indicated by the results, a vehicle that was not deployed to transportation were not meeting the requirements of EURO II and III Those vehicles were produced before 2004 and their emissions are among the highest (more than 280 g per kilometer). At the same time, these vehicles have the highest consumption (Table 1), so that even when selecting the vehicles to minimize the total cost of the fuel consumption, these vehicles would not be put into operation as well. From this perspective, it is worth considering a further economic evaluation of whether it would be economically and environmentally efficient to replace the "outdated" new vehicles that fulfill stricter emission standards for pollutants EURO VI.

Conclusion

Road freight transportation is essential for the economic development, but it emissions of CO_2 are harmful to the environment and to human health. For many logistic companies the planning of transportation activities has mainly focused on cost minimization. However by minimizing high emitting driving behavior, air quality can be improved significantly. The article has compared models that have been developed in order to optimize the fuel consumption and greenhouse gas emissions associated with road freight transportation. When comparing the modeled results data from the selected logistic company were used.

References

[1] Ahi, P.: An analysis of metrics used to measure performance in green and sustainable supply Chin. *Journal of Cleaner Production* **86** (2015), 360–377.
[2] Ahn, K. and Raka, H.: The effects of route choice decisions on vehicle energy consumption and emissions. *Transportation Research, Part D* **13,** 3 (2008), 151–167.
[3] Aronsson, H. and Brodin, M. H.: The environmental impact of changing logistics structures. *International Journal of Logistics Management* **1**, 3 (2006), 394–415.
[4] Demir, E., et al.: A review of recent research on green road freight transportation. *European Journal of Operational Research* **237**, 3 (2014), 775–793.
[5] Ghobadian, A.: In search of the drivers of high growth in manufacturing SMEs. *Technovation* **26**, 1 (2006), 30–41.
[6] Harris, I. and Naim, M.: Assessing the Impact of Cost Optimization Based on Infrastructure Modelling on CO_2 Emissions. *International Journal of Production Economics* **131**, 1 (2011), 313–321.
[7] Koblasa, F. and Manlig, F.: Application of Adaptive Evolution Algorithm on real-world Flexible Job Shop Scheduling Problems. In: *Proceedings of the 32nd International conference on Mathematical Methods in Economics* (Talašová, J., Stoklasa, J., Talášek, T., eds.). Olomouc, 2014, 425–430.
[8] Oberhofer, P., et al.: Sustainable development in the transport sector: Influencing environmental behaviour and performance. *Business Strategy and the Environment* **22**, 6 (2013), 374–383.
[9] Štichhauerová, E., et al.: Use of linear programming method to constructing a model for reduction of emission in a selected company. In: *Proceedings of the 32nd International conference on Mathematical Methods in Economics.* (Martinčík, D., Ircingová, J., Janeček, P., eds.). Cheb, 2015, 805–811.
[10] Wu, H. and Dunn, S.: Environmentally responsible logistics systems. *International Journal of Physical Distribution & Logistics Management* **25**, 2 (1995), 20–38.
[11] Wygonik, E. and Goodchild, A.: Evaluating CO_2 emissions, cost, and service quality trade-offs in an urban delivery system case study. *IATSS Research* **35** 1 (2011), 7–15.
[12] Zhang, Y. R.: Analyzing the Promoting Factors for Adopting Green Logistics Practices: A Case Study of Road Freight Industry in Nanjing, China. *Procedia - Social and Behavioral Science* **125** (2014), 432–444.
[13] Žabkar, V., et al.: Environmental Strategy: A Typology Of Companies Based On Managerial Perceptions Of Customers' Environmental Activeness And Deterrents. *E+M Economics and Management* **16**, 3 (2013), 57–74.

Similarity of Stock Market Series Analyzed in Hilbert Space

Jakub Šnor[1]

Abstract. Time series from stock markets exhibit various types of fluctuations and can be studied as samples from unknown n-dimensional distribution. The main question behind the paper is how to recognize similar stocks or similar time periods of their history. Whenever current stock segment is similar to another former one of the same stock or the other, the investment strategy should be also similar. Similarity measure is built in Hilbert space of infinite dimension as follows.
Parzen estimate of probability density function was used to obtain scalar products of stock segment pairs. This product and corresponding metrics were expressed analytically as double sum over segment items to enable the use of Principal Component Analysis, Cluster Analysis and Self-Organized Maps for the description of stock similarities and differences. Mathematical formulation of similarity task and analytic calculations are results of my original research work.
General principle is demonstrated on real stocks of leading IT companies in the period 2009-2015. Both PCA and SOM enable to visualize the similarities in infinite Hilbert space. All the calculations were performed in the Matlab environment.

Keywords: Hilbert space, self-organization, Parzen estimate, stock market, time series, market similarities.

JEL classification: G11
AMS classification: 91B84

1 Introduction

Every decision process is based on knowledge of actual state of given object. Object similarities can help to improve the quality of decisions. In the particular case of stock market operations, the time series of price history are good source of supporting information. Using sliding windows we can convert the problem to investigation of statistical sample similarities. Resulting task is also complex but it should be solved using Parzen estimate [7] of probability density function and data processing in Hilbert space [13]. My original research work consists of novel method how to calculate the dot product of probability density function estimates which is based on statistical sample analysis. It is useful not only for distance and similarity measurement but also for self-organization. This paper is continuation of my previous work [10] which was focused on Kernel Principal Component Analysis (KPCA), Cluster Analysis (CA) and Self-Organized Mapping (SOM) in Hilbert space of infinite dimension.

2 Sample Description

Let $M \in \mathbb{N}$ be the length of a sequence and $a_n > 0$ the value of the sequence for $n = 1, \ldots, M$. Then the sequence is traditionally described as $\{a_n\}_{n=1}^{M}$. Sequence can be also described by logarithmic differences of neighboring values defined as

$$z_k = \ln \frac{a_{k+1}}{a_k} \tag{1}$$

for $k = 1, \ldots, M-1$.

[1] Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Břehová 7, Praha 1, 115 19, Czech Republic, snorjaku@fjfi.cvut.cz

The sequence can be also cut to smaller sub-sequences. Then we can apply sliding window of given length l to $\{z_k\}_{k=1}^{M-1}$ and obtain vectors $\mathbf{x}_k = (z_k, \ldots, z_{k+l-1})$ for $k = 1, \ldots, M-l$. Segment can be described as statistical sample $\mathcal{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ where $N \in \mathbb{N}$, $N = M - l$.

In this paper, we analyze stocks of different IT companies. The main goal is to compare the stock states of various titles. The subject of comparison are two statistical samples x and y. We can compare:

- two segments of different length of the same stock;
- two segments of different length of different stocks;
- two segments of different period of the same stock;
- two segments of different period of different stocks.

The comparison of two segments of different stocks may be used to identify stocks which values are developing with the similar rate. Analyzing two segments of different length may serve to discover two stocks that are behaving similarly but with different speed (for example mobile industry is growing way faster than was computer industry one decade ago). Studying stock in different time period will most probably discover stocks with some seasonal development or it may help with prediction how one stock will develop based on the previous development of the same stock or different one. As a special case we may analyze the full period of two different stocks as well.

3 Parzen Estimate in Hilbert Space

Let $N \in \mathbb{N}$ be number of observations and $h > 0$ be the bandwidth parameter. Let $||\mathbf{x}||$ be Euclidean norm of sample $\mathbf{x}$ of dimension $n \in \mathbb{N}$. Then Parzen Estimate [7] is defined as

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} h^n N} \sum_{k=1}^{N} \exp\left(-\frac{||\mathbf{x} - \mathbf{x}_k||^2}{2h^2}\right) \tag{2}$$

This estimate is consistent for

$$\lim_{N\to\infty} h = 0 \tag{3}$$

$$\lim_{N\to\infty} Nh^n = +\infty \tag{4}$$

Traditionally, the parameter h is chosen according to the number of observations N. We set

$$h = h_0 N^{-\alpha} \tag{5}$$

for $h_0 > 0$ and $\alpha \in (0, n^{-1})$.

The investigation in Hilbert Space is based on dot product existence. In our case we calculate dot product of probability distribution functions as

$$(\mathrm{f}|\mathrm{g}) = \int_{\mathbb{R}^n} \mathrm{f}(\mathbf{x})\mathrm{g}(\mathbf{x})\mathrm{d}\mathbf{x}. \tag{6}$$

Using Parzen estimate from samples $\mathbf{x}$, $\mathbf{y}$ of density f, g with length N_f, N_g we obtain

$$(\mathrm{f}|\mathrm{g}) = \frac{1}{(2\pi)^n h_\mathrm{f} h_\mathrm{g} N_\mathrm{f} N_\mathrm{g}} \sum_{k=1}^{N_\mathrm{f}} \sum_{j=1}^{N_\mathrm{g}} \int_{\mathbb{R}^n} \exp\left(-\frac{||\mathbf{x} - \mathbf{x}_k||^2}{2\sigma^2}\right) \mathrm{d}\mathbf{x} \tag{7}$$

where $h_\mathrm{f} = h_0 N_\mathrm{f}^{-\alpha}$, $h_\mathrm{g} = h_0 N_\mathrm{g}^{-\alpha}$. After the integration in $\mathbb{R}^n$ resulting formula is obtained as

$$(\mathrm{f}|\mathrm{g}) = \frac{1}{(2\pi)^{n/2} \sigma^n N_\mathrm{f} N_\mathrm{g}} \sum_{k=1}^{N_\mathrm{f}} \sum_{j=1}^{N_\mathrm{g}} \exp\left(-\frac{||\mathbf{x}_k - \mathbf{y}_j||^2}{2\sigma^2}\right) \tag{8}$$

where

$$\sigma = h_0 \sqrt{N_{\mathrm{f}}^{-2\alpha} + N_{\mathrm{g}}^{-2\alpha}} \tag{9}$$

In this case $\alpha = 1/2n$ is recommended in the middle of its acceptable range. The parameter h_0 will be subject of investigation. The product $(\mathrm{f}|\mathrm{g})$ of densities $\mathrm{f}(\mathbf{x})$, $\mathrm{g}(\mathbf{x})$ in infinite dimensional Hilbert Space will be directly used for PCA calculations and adequate metrics [10]

$$\mathrm{d}(\mathrm{f}, \mathrm{g}) = \sqrt{(\mathrm{f}|\mathrm{f}) + (\mathrm{g}|\mathrm{g}) - 2(\mathrm{f}|\mathrm{g})} \tag{10}$$

is used as distance in cluster analysis and SOM learning.

4 Cluster Analysis in Hilbert Space

Let $M, H \in \mathbb{N}$ be number of patterns from Hilbert space and given number of clusters. Pattern set is then $\mathcal{S} = \{x_k \in \mathcal{H} : k = 1, \ldots, M\}$. Traditional *batch clustering* [3] is driven by *partition vector* [4] $\mathbf{p} = (p_1, \ldots p_M) \in \{1, \ldots, H\}^M$. Cluster $\mathcal{C}_j \subset \mathcal{S}$ is defined as $\mathcal{C}_j = \{x_k \in \mathcal{S} : p_k = j\}$. The clusters $\mathcal{C}_1, \ldots, \mathcal{C}_H$ represent *partition* of $\mathcal{S}$ as trivial to verify. For non-empty cluster $\mathcal{C}_j$ we define weight for pattern k as $w_k = 1/\mathrm{card}(\mathcal{C}_j)$. *Partition quality* $Q(\mathbf{p})$ is defined [12] as

$$Q(\mathbf{p}) = \sum_{j=1}^{H} \sum_{k \in \mathcal{C}_j} ||x_k - t_j||^2 \tag{11}$$

where t_j is centroid of $\mathcal{C}_j$. After small rearrangement we directly obtain

$$Q(\mathbf{p}) = \sum_{j=1}^{H} \mathrm{card}(\mathcal{C}_j) \sum_{k \in \mathcal{C}_j} w_{j,k} ||x_k - t_j||^2 = \sum_{j=1}^{H} q_j \tag{12}$$

where q_j is quality of cluster $\mathcal{C}_j$ calculated according [10] as

$$q_j = \mathrm{card}(\mathcal{C}_j) \sum_{k=2}^{\mathrm{card}(\mathcal{C}_j)} \sum_{l=1}^{k-1} w_{j,k} w_{j,l} ||x_k - x_l||^2. \tag{13}$$

General aim of batch cluster analysis is to obtain *optimum clustering (partition)* driven by partition vector $\mathbf{p}_{\mathrm{opt}} \in \operatorname{argmin} Q(\mathbf{p})$. According to [1], optimum clustering is NP-hard task in Euclidean space which is special case of finite Hilbert space. Therefore, optimum clustering in Hilbert space is also NP-hard one. In this case, trivial defense against resignation is the application of any optimization heuristics (K-means [12], Simulated Annealing (SA) [2], Fast Simulated Annealing (FSA) [9], Integer Cuckoo Search (ICS) [5] based on integer Lévy flights [11], Steepest Descent (SD) [8] or Random Descent (RD) [6].

5 Self-Organized Mapping in Hilbert Space

SOM structure and learning strategies were widely discussed in [3]. We are focused on batch learning and our main motivation is in adapting the main principles of SOM learning in metric space [4] to general Hilbert space. Let $\mathcal{N}$ be a non-empty set of vertices (nodes). Let

$$\mathcal{R} \subset \binom{\mathcal{N}}{2} \tag{14}$$

be a set of edges (relations). Then the pair $\mathcal{G} = \langle \mathcal{N}, \mathcal{R} \rangle$ is an undirected connected graph which typically represents the topology of SOM [3]. The number of output nodes $H = \mathrm{card}(\mathcal{N})$. Let $d^* : \mathcal{N} \times \mathcal{N} \to \mathbb{N}_0$ be a vertex distance in the graph of SOM. Then $\mathcal{M} = \langle \mathcal{N}, d^* \rangle$ is a metric space over the SOM. The maximum value of the vertex distance over $\mathcal{N} \times \mathcal{N}$ is called *diameter* D of the SOM. The SOM in Hilbert space is a function $\mathrm{SOM} : \mathcal{S} \to \mathcal{N}$, while the *SOM learning* is an algorithm of the SOM function design according to the set of patterns $\mathcal{S}$ [4].

Batch learning of SOM in Hilbert space $\mathcal{H}$ is also driven by vector $\mathbf{p} = (p_1, \ldots p_M) \in \{1, \ldots, H\}^M$ which places patterns from $\mathcal{S} \subset \mathcal{H}$ into nodes from $\mathcal{N}$ of given SOM graph. Novel penalization strategy is based on weighted centroids around individual SOM nodes and corresponding centroid quality. When the pattern is placed into investigated node, its weight is maximum possible but when it is placed in node neighborhood the weight is function of d^* according to formula

$$w_{i,k} = \frac{\chi(\mathrm{d}^*(i, p_{x_k}))}{\sum_{j=1}^{M} \chi(\mathrm{d}^*(i, p_{x_j}))} \tag{15}$$

where characteristics $\chi : \{0, \ldots, D\} \to \mathbb{R}^+$ is non-increasing function satisfying $\chi(0) = 1, \chi(D) \in (0, 1)$. In the case of batch SOM learning [3], only direct neighbors are involved in centroid calculations with full weight. Therefore, $\chi(d^*) = 1$ for $d^* \leq 1$ and $\chi(d^*) = 0$ otherwise.

The other characteristics are also useful as:

- *rectangular* with characteristics $\chi(d^*) = \mathrm{I}(d^* \leq R)$,
- *Gaussian* with characteristics $\chi(d^*) = \exp\left(-\frac{1}{2}\left(\frac{d^*}{R}\right)^2\right)$,
- *Exponential* with characteristics $\chi(d^*) = \exp\left(-\frac{d^*}{R}\right)$,
- *Cauchian* with characteristics $\chi(d^*) = \left(1 + \left(\frac{d^*}{R}\right)^2\right)^{-1}$

where $R > 0$ is learning radius. Gaussian characteristics is frequently used in traditional Kohonen SOM learning and should be preferred also in this method.

As seen from (15) the weights $w_{i,k}$ are well defined and positive satisfying

$$\sum_{k=1}^{M} w_{i,k} = 1, \text{ for all } i = 1, \ldots H \tag{16}$$

which represents node views to given data set $\mathcal{S}$. SOM quality is then defined as a sum of node view penalizations according to formula

$$\mathrm{S}(\mathbf{p}) = \sum_{i=1}^{H} s_i \tag{17}$$

where s_i is "cluster" quality with weights $w_{i,k}$ for given node i according to (13). The main difference between cluster analysis and SOM in Hilbert space is in interaction among patterns from various nodes. SOM partition is formally the same as clustering partition but the influence of any pattern overcome its node limitations. Therefore, SOM learning is just integer minimization of $\mathrm{S}(\mathbf{p})$ for given pattern set $\mathcal{S}$.

This task is similar to optimum clustering. General integer minimization heuristics (SA, FSA, ICS, SD, RD) can be employed again. Alternative way is to modify K-means heuristic without necessity of node centroid calculations. Novel heuristics of SOM learning in Hilbert space begins with random partition $\mathbf{p}$. The main loop performs SOM partition revisions till penalization $\mathrm{S}(\mathbf{p})$ decreases as follows: For every pattern $x_k \in \mathcal{S}$ and every node i we calculate $\mathrm{d}(x_k, t_i)$ where t_i is hidden centroid of $\mathcal{C}_i$ around node i. New pattern position in SOM is then given by formula $p_k \in \mathrm{argmin}_{i=1,\ldots,H}\, \mathrm{d}(x_k, t_i)$.

6 Application to Stock Indexes

The proposed method was tested on data from stock markets[1] since 2$^{\text{nd}}$ January 2009 to 22$^{\text{nd}}$ March 2016 of 10 popular IT companies: Adobe Systems (ADO), Apple Computer (APP), Amazon.com (AMA), Advanced Micro Devices Inc. (AMD), Autodesk Inc. (AUT), Cisco Systems Inc. (CIS), Computer Sciences Corporation (CSC), Intel (INT), Microsoft (MIC) and Yahoo! (YAH). Each time series contained 1824 values. For each two neighboring values was computed logarithmic difference according to (1) reducing its length by one and then the sliding window of length 10 was applied. After this data preprocessing, the scalar product of each time series pair was computed.

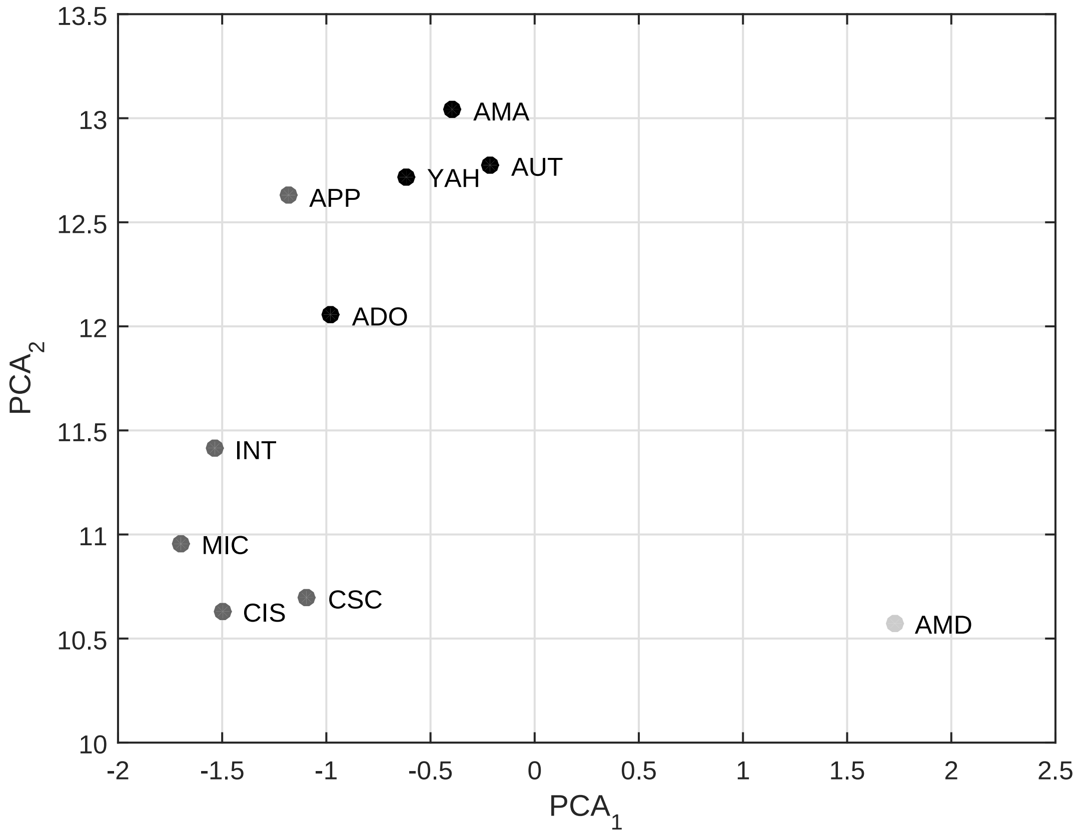


Figure 1 Kernel Principal Component Analysis and Cluster Analysis markers of stocks

Kernel Principal Component Analysis with kernel $K(x, y) = (x|y)$ was used as referential method and its results were mapped into two-dimensional space. Two dimensions were chosen for easy comparison with planar SOM. The results of 2D KPCA are depicted on Fig. 1.

Cluster analysis into three clusters was used as the second referential method. Steepest descent heuristic with random initial point was used for finding the minimum of (12). Stocks ADO, AMA, AUT, and YAH formed the first cluster performed as black dots in Fig. 1. The second cluster consists of stocks APP, CIS, CSC, INT and MIC as dark grey dots. The third cluster contained only stock AMD as light grey point in the same image.

The stock market data was finally self-organized by presented algorithm and compared with referential methods. Due to small number of stock titles, hexagonal topology of SOM with seven nodes and traditional Gaussian repulsion function with the radii $R = 1.5$ and $R = 7.5$ were applied to the same data. The steepest descent heuristic was again used for the minimization (17). The results are depicted on Fig. 2. As seen, smaller repulsion brought results similar to KPCA, while higher repulsion formed structure similar to CA. Therefore, the novel method is able to make compromise between two traditional analytic approaches.

[1] http://www.wessa.net/stocksdata.wasp

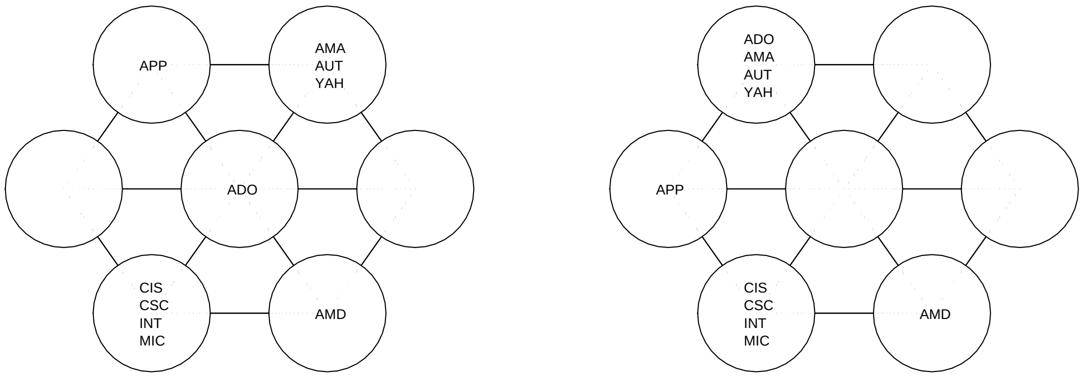


Figure 2 Self-Organized Mapping for $R = 1.5$ (left) and $R = 7.5$ (right)

7 Conclusion

Novel method of SOM in Hilbert space was used for self-organization of statistical samples. This technique enables to analyze stock market states as demonstrated on IT company stocks. The novel method is a compromise between Principal Component Analysis and Cluster Analysis. The proposed method together with hexagonal topology of SOM is a general tool for stock market analysis but the choice of pre-processing parameters l, σ and SOM parameters R, H is essential and will be subject of future research.

References

[1] Garey, M. R., and Johnson, D. S.: *Computers and intractability: a guide to the theory of NP-completness*. W.H. Freeman and Company, New York, 1997.

[2] Kirkpatrick, S., Gellat C. D., and Vecchi, M. P.: Optimization by Simulated Annealing. *Science* **4598** (1983), 671-680.

[3] Kohonen, T.: *Self-Organizing Maps*. Springer Verlag, New York, 1995.

[4] Kukal, J.: SOM in Metric Space. *Neural Network World*. Prague, 2004, 469-488.

[5] Kukal, J., Mojzeš M., Tran, Q. V., and Boštík, J.: Integer Cuckoo Search. In: *Proceedings of 18th International Confeence on Soft Computing MENDEL 2012*. VUT Press, Brno, 2012, 298-303.

[6] Nesterov, Y.: Efficiency of Coordinate Descent Methods on Huge-Scale Optimization Problems. *SIAM Journal on Optimization* **2** (2012), 341-362.

[7] Parzen, E.: On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics* **3** (1962), 1065-1076.

[8] Snyman, J. A.: *Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms*. Springer, New York, 2005.

[9] Szu, H., and Hartley, R.: Fast simulated annealing. *Physics Letters A* **3-4** (1987), 157-162.

[10] Šnor, J., and Kukal, J.: Self-organized mapping in Hilbert Space. *Kybernetika* (submitted). Prague, 2015.

[11] Viswanathan G.M.: Lévy flights and Superdiffusion in the Context of Biological Encounters and Random Searchers. *Physics of Life Reviews* **3** (2008), 133-150.

[12] Xu, R., and Wunsch, D. C.: *Clustering*. IEEE Press, Piscataway, NJ, 2009.

[13] Young, N.: *An introduction to Hilbert space*. Cambridge University Press, New York, 1988.

The role of distance and similarity in Bonissone's linguistic approximation method – a numerical study

Tomáš Talášek[1], Jan Stoklasa[2], Jana Talašová[3]

Abstract. Linguistic approximation is a common way of translating the outputs of mathematical models in the expressions in common language. These can then be presented to decision makers who have difficulties with interpretations of numerical outputs of formal models as an easy-to-understand alternative. Linguistic approximation is a tool to stress, modify or effectively convey meaning. As such it is an important yet neglected area of research in management science and decision support.

During the last forty years a large number of different methods for linguistic approximation were proposed. In this paper we investigate in detail the linguistic approximation method proposed by Bonissone (1979). We focus on its performance under different "fit" measures in its second step - we consider various distance and similarity measures of fuzzy sets to choose the most appropriate linguistic approximation. We conduct a numerical study of the performance of this linguistic approximation method, present its results and discuss the impact of a particular choice of a "fit" measure.

Keywords: Linguistic approximation, two-step method, fuzzy number, distance, similarity.

JEL classification: C44
AMS classification: 90B50, 91B06, 91B74

1 Introduction

In practical applications of decision support models that employ fuzzy sets it is often necessary to be able to assign a linguistic label (from predefined linguistic scale) to a fuzzy set (usually obtained as an output of some mathematical model). This process is called *linguistic approximation*. The main reason for applying linguistic approximation is to "translate" (abstract/formal) mathematical objects into the common language (recent research also suggests that the ideas of linguistic approximation can be used e.g. for ordering purposes - see [7]). This way the outputs of mathematical models can become easier to understand and use for the decision-makers. The process of linguistic approximation involves the selection of the best fitting linguistic term from a predefined term set as a representative of the given mathematical object (fuzzy set). Obviously, since the set of linguistic terms is finite (and usually contains only a few linguistic terms), the process distorts the actual output of the mathematical model to some extent (add or decrease uncertainty, shift the meaning in the given context etc. - hence *approximation*). The key to a successful linguistic approximation is to find an appropriate tradeoff between understandability and loss (distortion) of information (see e.g. [10, 6]). Linguistic approximation relies in many cases on distance and similarity of fuzzy sets, on the subsethood and the differences in relevant features of the output to be approximated and the meaning of its approximating linguistic term.

In this paper we focus on the Bonissone's two-step method for linguistic approximation [1], since it combines the idea of semantic similarity with the requirement of the closeness of meaning. In the first step, the method preselects a given amount of linguistic terms, that embody the semantic best fit (based on a specified set of features). In the second step the linguistic term whose meaning is the closest based on some distance/similarity measure is selected. We investigate the role of distance/similarity measure in the

[1]Palacký University, Olomouc, Křížkovského 8, Olomouc, Czech Republic and Lappeenranta University of Technology, Skinnarilankatu 34, Lappeenranta, Finland, tomas.talasek@upol.cz
[2]Palacký University, Olomouc, Křížkovského 8, Olomouc, Czech Republic, jan.stoklasa@upol.cz
[3]Palacký University, Olomouc, Křížkovského 8, Olomouc, Czech Republic, jana.talasova@upol.cz

second step of this method and on a numerical study we compare the performance of the Bhattacharyya distance suggested by Bonissone with the dissemblance index distance measure and two fuzzy similarity measures. Based on the results of a numerical study we analyze what features employed in the first step (namely position and uncertainty - the same features employed by Wenstøp [9] in his method) are emphasized and which are distorted by each of the distance/similarity measures. This way the paper strives to contribute to the scarce body of research on good practices in linguistic approximation.

2 Preliminaries

Let U be a nonempty set (the universe of discourse). A *fuzzy set* A on U is defined by the mapping $A : U \to [0, 1]$. For each $x \in U$ the value $A(x)$ is called a *membership degree* of the element x in the fuzzy set A and $A(.)$ is called a *membership function* of the fuzzy set A. $\mathrm{Ker}(A) = \{x \in U | A(x) = 1\}$ denotes a *kernel* of A, $A_\alpha = \{x \in U | A(x) \geq \alpha\}$ denotes an *α-cut* of A for any $\alpha \in [0, 1]$, $\mathrm{Supp}(A) = \{x \in U | A(x) > 0\}$ denotes a *support* of A.

A fuzzy number is a fuzzy set A on the set of real numbers which satisfies the following conditions: (1) $\mathrm{Ker}(A) \neq \emptyset$ (A is *normal*); (2) A_α are closed intervals for all $\alpha \in (0, 1]$ (this implies A is *unimodal*); (3) $\mathrm{Supp}(A)$ is bounded. A family of all fuzzy numbers on U is denoted by $\mathcal{F}_N(U)$. A fuzzy number A is said to be defined on [a,b], if $\mathrm{Supp}(A)$ is a subset of an interval $[a, b]$. Real numbers $a_1 \leq a_2 \leq a_3 \leq a_4$ are called *significant values* of the fuzzy number A if $[a_1, a_4] = \mathrm{Cl}(\mathrm{Supp}(A))$ and $[a_2, a_3] = \mathrm{Ker}(A)$, where $\mathrm{Cl}(\mathrm{Supp}(A))$ denotes a closure of $\mathrm{Supp}(A)$. Each fuzzy number A is determined by $A = \left\{[\underline{a}(\alpha), \overline{a}(\alpha)]\right\}_{\alpha\in[0,1]}$, where $\underline{a}(\alpha)$ and $\overline{a}(\alpha)$ is the lower and upper bound of the α-cut of fuzzy number A respectively, $\forall\alpha \in (0, 1]$, and the closure of the support of A $\mathrm{Cl}(\mathrm{Supp}(A)) = [\underline{a}(0), \overline{a}(0)]$. A *union* of two fuzzy sets A and B on U is a fuzzy set $(A \cup B)$ on U defined as follows: $(A \cup B)(x) = \min\{1, A(x) + B(x)\}$, $\forall x \in U$.

The fuzzy number A is called *linear* if its membership function is linear on $[a_1, a_2]$ and $[a_3, a_4]$; for such fuzzy numbers we will use a simplified notation $A = (a_1, a_2, a_3, a_4)$. A linear fuzzy number A is said to be *trapezoidal* if $a_2 \neq a_3$ and *triangular* if $a_2 = a_3$. We will denote triangular fuzzy numbers by ordered triplet $A = (a_1, a_2, a_4)$. More details on fuzzy numbers and computations with them can be found for example in [2].

Let A be a fuzzy number on $[a, b]$ for which $a_1 \neq a_4$. Then A could be described by several real number characteristics, such as *cardinality*: $\mathrm{Card}(A) = \int_{[a,b]} A(x)dx$; *center of gravity*: $\mathrm{COG}(A) = \int_{[a,b]} xA(x)dx/\mathrm{Card}(A)$; *fuzziness*: $\mathrm{Fuzz}(A) = \int_{[a,b]} \mathrm{S}(A(x))dx$, where $S(y) = -y \ln(y) - (1 - y) \ln(1 - y)$ and *skewness*: $\mathrm{Skew}(A) = \int_{[a,b]}(x - \mathrm{COG}(A))^3 A(x)dx$.

A *fuzzy scale* on $[a, b]$ is defined as a set of fuzzy numbers $T_1, T_2, \ldots, T_s$ on [a,b], that form a Ruspini fuzzy partition (see [5]) of the interval $[a, b]$, i.e. for all $x \in [a, b]$ it holds that $\sum_{i=1}^{s} T_i(x) = 1$, and the T's are indexed according to their ordering. A *linguistic variable* ([11]) is defined as a quintuple $(\mathcal{V}, \mathcal{T}(\mathcal{V}), X, G, M)$, where $\mathcal{V}$ is a name of the variable, $\mathcal{T}(\mathcal{V})$ is a set of its linguistic values (terms), X is an universe on which the meanings of the linguistic values are defined, G is an syntactic rule for generating the values of $\mathcal{V}$ and M is a semantic rule which to every linguistic value $\mathcal{A} \in \mathcal{T}(\mathcal{V})$ assigns its meaning $A = M(\mathcal{A})$ which is usually a fuzzy number on X. Linguistic variable $(\mathcal{V}, \mathcal{T}(\mathcal{V}), X, G, M)$ is called a *linguistic scale* on $[a, b]$ if $X = [a, b]$, $\mathcal{T}(\mathcal{V}) = \{\mathcal{T}_1, \ldots, \mathcal{T}_s\}$ and $M(\mathcal{T}_i) = T_i, i = 1, \ldots, s$ form a fuzzy scale on $[a, b]$. Terms $\mathcal{T}_i, i = 1, \ldots, s$ are called *elementary terms*. Linguistic scale on $[a, b]$ is called *extended linguistic scale*, if besides elementary terms contains also *delivered* terms in the form $\mathcal{T}_i$ to $\mathcal{T}_j$ where $i < j$, $i, j \in \{1, \ldots, n\}$ and $M(\mathcal{T}_i \text{ to } \mathcal{T}_j) = T_i \cup T_{i+1} \cup \cdots \cup T_j$.

3 Bonissone's two step method for linguistic approximation

Bonissone's two step approach for linguistic approximation [1] was proposed in 1979. In contrast to the majority of linguistic approximation approaches, Bonissone suggested to split the process into two steps – in the first step the set of suitable linguistic terms for the approximation of a given fuzzy number is found (this "pre-selection step" is done based on the semantic similarity), then in the second step the most appropriate term for the linguistic approximation is found from this set of suitable linguistic terms.

In the pre-selection step the set $\mathcal{P} = \{\mathcal{T}_{p_1}, \ldots, \mathcal{T}_{p_k}\}$ of $k\,(k \leq s)$ suitable linguistic terms from $\mathcal{T}(\mathcal{V})$

is formed in the way that the meaning of these pre-selected terms are similar to the fuzzy set O (an output of a mathematical model to be approximated) with respect to four characteristics (cardinality, center of gravity, fuzziness and skewness). These characteristics are assumed to capture the semantic value of a fuzzy set used to model the meaning of a linguistic term. The semantic value of a fuzzy set on a given universe can thus be represented by a quadruple of real numbers (values of 4 features in four-dimensional space). Let A be a fuzzy set on $[a, b]$. Then the respective characteristic quadruple is denoted as (a^1, a^2, a^3, a^4) where $a^1 = \text{Card}(A)$, $a^2 = \text{COG}(A)$, $a^3 = \text{Fuzz}(A)$ and $a^4 = \text{Skew}(A)$.

Let the fuzzy set O on $[a, b]$ be an output of a mathematical model that needs to be linguistically approximated by one linguistic term from the set $\mathcal{T}(\mathcal{V}) = \{\mathcal{T}_1, \ldots, \mathcal{T}_s\}$. $\mathcal{T}(\mathcal{V})$ is a linguistic term set of a linguistic variable $(\mathcal{V}, \mathcal{T}(\mathcal{V}), [a, b], G, M)$, such that $T_i = M(\mathcal{T}_i)$, $i = 1, \ldots, s$ are fuzzy numbers on $[a, b]$. Linguistic terms $\{\mathcal{T}_1, \ldots, \mathcal{T}_s\}$ are ordered with respect to the distance of their characteristic quadruples from the characteristic quadruple of O. The ordered set $\mathcal{N} = (\mathcal{T}_{p_1}, \ldots, \mathcal{T}_{p_s})$ is thus obtained, such that $d(T_{p_1}, O) \leq d(T_{p_2}, O) \leq \cdots \leq d(T_{p_s}, O)$ where

$$d(T_{p_i}, O) = \sum_{j=1}^{4} w_j |t^j_{p_i} - o^j|, \; i = 1 \ldots, s, \tag{1}$$

and w_j, $j = 1, \ldots, 4$ are normalized real weights (i.e. $\sum_{j=1}^{4} w_j = 1, \; w_j \geq 0, \; j = 1, \ldots, 4$). The choice of weights is usually left with the user of the model and some features could be even optional. Wenstøp [9] for example proposed (in his method for linguistic approximation) to use only two features – cardinality (uncertainty) and center of gravity (position). First k linguistic terms (the parameter k is specified by the decision maker) from the ordered set $\mathcal{N}$ are stored in the set $\mathcal{P}$ and the pre-selection step is finished.

In the second step, the linguistic approximation $\mathcal{T}_O \in \mathcal{P}$ of the fuzzy set O is computed. The fuzzy set $T_O = M(\mathcal{T}_O)$ is computed as

$$T_O = \arg \min_{\mathcal{T}_{p_i} \in \mathcal{P}} d_1(T_{p_i}, O) \tag{2}$$

using the modified Bhattacharyya distance:

$$d_1(A, B) = \left[1 - \int_U (A^*(x) \cdot B^*(x))^{1/2} dx\right]^{1/2}, \tag{3}$$

where $A^*(x) = A(x)/\text{Card}(A(x))$ and $B^*(x) = B(x)/\text{Card}(B(x))$. This way the linguistic term $\mathcal{T}_O$ is found as the closest linguistic approximation among the pre-selected linguistic terms.

The Bhattacharyya distance (3) can be substituted by different distances or similarity measures[1] of fuzzy numbers – this step will, however, modify the behaviour of the linguistic approximation method. In the next section the following distance and similarity measures of fuzzy numbers are considered:

- A *dissemblance index* (introduced by Kaufman and Gupta [4]) of fuzzy numbers A and B is defined by the formula

$$d_2(A, B) = \int_0^1 |\underline{a}(\alpha) - \underline{b}(\alpha)| + |\overline{a}(\alpha) - \overline{b}(\alpha)| \, d\alpha, \tag{4}$$

- A *similarity measure* (introduced by Wei and Chen in [8]) of fuzzy numbers A and B is defined by the formula

$$s_1(A, B) = \left(1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4}\right) \cdot \frac{\min\{Pe(A), Pe(B)\} + \min\{\text{hgt}(A), \text{hgt}(B)\}}{\max\{Pe(A), Pe(B)\} + \max\{\text{hgt}(A), \text{hgt}(B)\}}, \tag{5}$$

  where $Pe(A) = \sqrt{(a_1 - a_2)^2 + (\text{hgt}(A))^2} + \sqrt{(a_3 - a_4)^2 + (\text{hgt}(A))^2} + (a_3 - a_2) + (a_4 - a_1)$, $Pe(B)$ is defined analogically

- A *similarity measure* (introduced by Hejazi and Doostparast in [3]) of fuzzy numbers A and B can be defined by the formula

$$s_2(A, B) = \left(1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4}\right) \cdot \frac{\min\{Pe(A), Pe(B)\}}{\max\{Pe(A), Pe(B)\}} \cdot \frac{\min\{Ar(A), Ar(B)\} + \min\{\text{hgt}(A), \text{hgt}(B)\}}{\max\{Ar(A), Ar(B)\} + \max\{\text{hgt}(A), \text{hgt}(B)\}}, \tag{6}$$

  where $Ar(A) = \frac{1}{2}\text{hgt}(A)(a_3 - a_2 + a_4 - a_1)$, $Ar(B)$ is defined analogically and $Pe(A)$ and $Pe(B)$ are computed identically as in the previous method.

[1] In the case of similarity measure the *arg min* function in the formula (2) must be changed to *arg max* function.

4 Numerical experiment

We restrict ourselves for the purpose of this paper to the linguistic approximation of triangular fuzzy numbers. We assess the performance of the distance measures d_1 and d_2 and of the two similarity measures s_1 and s_2 in the context of Bonisonne's linguistic approximation method by the following numerical experiment. We randomly generate 100 000 triangular fuzzy numbers on $[0,1]$ to be linguistically approximated (denoted $\{O_1,\ldots,O_{100000}\}$) and compute their cardinalities $\{o^1_1,\ldots,o^1_{100000}\}$ and their centers of gravity $\{o^2_1,\ldots,o^2_{100000}\}$. We assume that for all these generated fuzzy numbers the hypothetical result of the first phase of Bonisonne's method is the set of linguistic terms $\mathcal{P}=\{\mathcal{T}_{p_1},\ldots,\mathcal{T}_{p_k}\}$ - in our numerical study this set is the linguistic term set of an extended linguistic scale constructed from a uniform Ruspini fuzzy partition of the universe $[0,1]$ with 5 triangular fuzzy numbers. This way we obtain the linguistic term set $\mathcal{P}=\{\mathcal{T}_{p_1},\ldots,\mathcal{T}_{p_{15}}\}$, the meanings of these linguistic terms are $\{T_1,\ldots,T_{15}\}$, with cardinalities $\{t^1_1,\ldots,t^1_{15}\}$ and centers of gravity $\{t^2_1,\ldots,t^2_{15}\}$. Using each distance and similarity measure we find the linguistic approximation of each generated output applying the second step of Bonisonne's method - this way we obtain $\{\mathcal{T}^{d_1}_{O_1},\ldots,\mathcal{T}^{d_1}_{O_{100000}}\}$, $\{\mathcal{T}^{d_2}_{O_1},\ldots,\mathcal{T}^{d_2}_{O_{100000}}\}$, $\{\mathcal{T}^{s_1}_{O_1},\ldots,\mathcal{T}^{s_1}_{O_{100000}}\}$ and $\{\mathcal{T}^{s_2}_{O_1},\ldots,\mathcal{T}^{s_2}_{O_{100000}}\}$ as the linguistic approximations of the generated triangular fuzzy numbers using d_1,d_2,s_1 and s_2 respectively.

Figure 1 plots the cardinalities of the approximated fuzzy numbers (horizontal axis) against the cardinality of the meaning of the respective linguistic approximation for all the distance/similarity measures. We can clearly see from the plots, that all four measures provide linguistic approximations with both higher cardinality (points above the main diagonal) and with lower cardinality. It, however, seems, that higher cardinality case is more frequent (points in the left upper corner of the plots). This can be reasonable, since even in common language we tend to use super-categories to generalize the meaning. In all the methods it is possible to also get a linguistic approximation with a lower cardinality (i.e. the meaning of the linguistic approximation is less uncertain than the original output of the model). Note, that since we have generated triangular fuzzy numbers on $[0,1]$, the maximum possible cardinality of any generated fuzzy number was 0.5. The Bhattacharyya distance is the only one from the investigated measures, that provides very highly uncertain approximations. This behavior could be tolerated only if the reason for the addition of uncertainty is the tendency of the measure to achieve a linguistic approximation that is more general than the approximated fuzzy set. Table 1 summarizes in how many cases the kernel of the resulting linguistic approximation is a superset of the kernel of the approximated results - in these cases the "typical representatives" of the output are also the "typical representatives" of the approximated linguistic term. We can see that Bhattacharyya distance focuses on this aspect much more than the other investigated methods.

The situation for the centers of gravity is summarized analogically in Figure 2. Here the desired state can be no presence of a systematic bias of the approximation. This corresponds with the points being close to the main diagonal in the respective plot, or evenly distributed to the left and to the right. We can see that with respect to this requirement the Bhattacharyya distance performs rather well. Both similarities perform in most cases in the following way: i) in case of lower centers of gravity of the approximated result they shift the center of gravity of the meaning of the linguistic approximation lower than the original center of gravity of the approximated results, ii) in case of higher centers of gravity of the approximated result they shift the center of gravity of the meaning of the linguistic approximation higher than the original center of gravity of the approximated results. Similarities seem to have an amplifying effect on the center of gravity - shifting the center of gravity to the endpoints of the universe. This can be a desirable property in cases, when such an amplification of meaning is needed.

k	d_1	d_2	s_1	s_2
$\frac{\mathrm{Card}\{O_i \mid \mathrm{Ker}(O_i)\subseteq \mathrm{Ker}(T^k_{O_i}), i=1,\ldots,100000\}}{100000}$	0.2868	0.1559	0.1778	0.1632

Table 1: The relative count of cases when the $\mathrm{Ker}(O_i)\subseteq \mathrm{Ker}(T^k_{O_i})$, $k\in\{d_1,d_2,s_1,ss_2\}$ out of the given 100 000.

5 Conclusion

In the paper we have investigated the role of different distance and similarity measures of fuzzy numbers in the second step of Bonissone's linguistic approximation method. We focused on the cardinality and

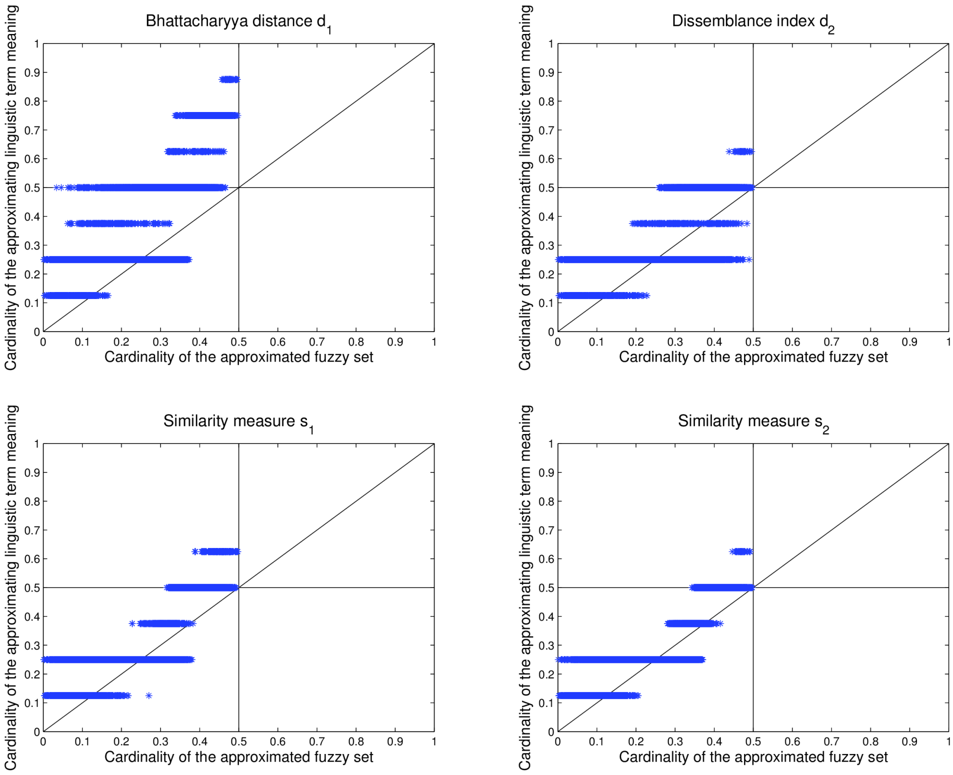


Figure 1: Comparison of cardinalities of the approximated outputs (horizontal axis) and the meanings of their linguistic approximations (vertical axis) for d_1, d_2, s_1 and s_2.

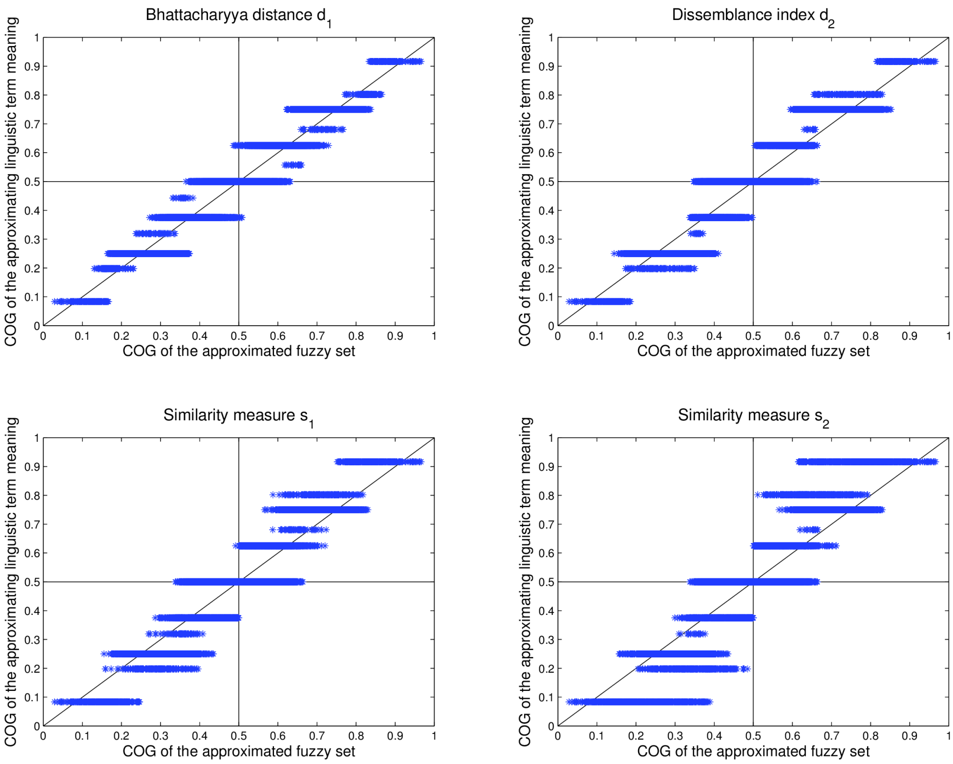


Figure 2: Comparison of centers of gravity of the approximated outputs (horizontal axis) and the meanings of their linguistic approximations (vertical axis) for d_1, d_2, s_1 and s_2.

center of gravity characteristics of fuzzy numbers. We have performed a numerical experiment which investigated the differences between the chosen characteristics of randomly generated triangular fuzzy numbers on the interval $[0, 1]$ and the characteristics of the meanings of their linguistic approximations computed by Bonissone's method. This served as a basis for the analysis of the performance of two different distance measures and two similarity measures of fuzzy numbers in the linguistic approximation context.

The results of the numerical experiment suggest, that the Bhattacharyya distance tends to provide more uncertain approximations than the other methods and is more likely to provide approximating linguistic term that "catch" the typical representatives (the kernel of the approximating linguistic term meaning is a superset to the kernel of the approximated fuzzy number). Both presented similarity methods have amplifying effect on the center of gravity – they shift the center of gravity of the approximating linguistic term meaning to the endpoints of the universe.

Acknowledgements

Supported by the grant by the grant GA 14-02424S Methods of operations research for decision support under uncertainty of the Grant Agency of the Czech Republic and partially also by the grant IGA PrF 2016 025 of the internal grant agency of the Palacký University, Olomouc.

References

[1] Bonissone, P. P.: A pattern recognition approach to the problem of linguistic approximation in system analysis. In: *Proceedings of the IEEE International Conference on Cybernetics and Society*, 1979, 793–798.

[2] Dubois, D., and Prade, H.: *Fuzzy sets and systems: theory and applications*, Academic Press, 1980.

[3] Hejazi, S. R., Doostparast, A., and Hosseini, S. M.: An improved fuzzy risk analysis based on a new similarity measures of generalized fuzzy numbers. *Expert Systems with Applications*, **38**, 8 (2011), 9179–9185.

[4] Kaufman, A., and Gupta, M. M.: *Introduction to Fuzzy Arithmetic*, Van Nostrand Reinhold, New York, 1985.

[5] Ruspini, E.: A New Approach to Clustering. *Information and Control*, **15** (1969), 22–32.

[6] Stoklasa, J.: *Linguistic models for decision support.* Lappeenranta University of Technology, Lappeenranta, 2014.

[7] Talášek, T., Stoklasa, J., Collan, M., and Luukka, P.: Ordering of Fuzzy Numbers through Linguistic Approximation Based on Bonissone's Two Step Method. In: *Proceedings of the 16th IEEE International Symposium on Computational Intelligence and Informatics.* Budapest, 2015, 285–290.

[8] Wei, S. H., and Chen, S. M.: A new approach for fuzzy risk analysis based on similarity measures of generalized fuzzy numbers. *Expert Systems with Applications*, **36**, 1 (2009), 589–598.

[9] Wenstøp, F.: Quantitative analysis with linguistic values. *Fuzzy Sets and Systems*, **4**, 2 (1980), 99–115.

[10] Yager, R. R.: On the retranslation process in Zadeh's paradigm of computing with words. *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics*, **34**, 2 (2004), 1184–1195.

[11] Zadeh, L. A.: The concept of a linguistic variable and its application to approximate reasoning I, II, III. *Information Sciences*, **8** (1975), 199–257, 301–357, **9** (1975), 43–80.

Registry of Artistic Performance - the final state of the evaluation model

Jana Talašová[1], Jan Stoklasa[2], Pavel Holeček[3], Tomáš Talášek[4]

Abstract. The amendment to the Law on Higher Education Institutions adopted in 2016 means the enactment of the Registry of Artistic Performance (RAP) and its transfer under the administration of the Ministry of Education, Youth and Sports. In this context it was necessary to finalize the mathematical model for the evaluation of the outcomes of artistic creative work stored in RAP. The results of the model – scores assigned to the outcomes – are used for the distribution of a part of the funding from the state budget among universities in the Czech Republic. For the purpose of the evaluation, the creative work outcomes are classified into categories based on three criteria: significance of the outcome, its extent, and its institutional reception. These categories have been assigned scores by the AHP. In time modifications to the Saaty's matrix have proven to be necessary. The algorithm for multi-expert classification of the creative-work outcomes has been improved in the process. The description of the development and successive modifications of the evaluation model are the focus of this paper.

Keywords: Registry of Artistic Performance, creative-work outcomes, multiple criteria evaluation, AHP, group evaluation.

JEL Classification: C44
AMS Classification: 90B50

1 The Registry of Artistic Performance

The Registry of Artistic Performance (RAP) is a web application for storing the information on the outcomes of artistic creative work (denoted OACW or outcomes further in the text) produced by the universities in the Czech Republic. Since 2013, the application has been performing this task. Another purpose of the application is the evaluation of the registered outcomes: on the basis of expert evaluations, the outcomes are divided into categories with a pre-defined point gain. The sum of the points obtained by each individual university in the past 5 years then represents one of the indicators used for the funding of these universities from the state budget. In 2016, the amendment to the Law on Higher Education Institutions, in which the RAP is explicitly stated, was adopted. As soon as the Law on Higher Education Institutions becomes effective (September 1, 2016), RAP will enter a fully professional mode under the administration of the Ministry of Education, Youth and Sports. It was therefore necessary to finish the tuning of the evaluation model based on the results of the most recent analyses of the development and performance of RAP and the respective evaluation methodology.

A quantitative evaluation of OACW (i.e. pieces of art and artistic performances) represents a task comparable in its complexity to the evaluation of the research and development (R&D) results. While various scientometric tools [1] have been developed for the quantitative evaluation of the R&D outcomes since 1950s, there was only a very limited possibility to reuse some previously developed procedures in case of OACW evaluation.

The evaluation model used in the RAP has undergone significant development since its introduction in 2010 [6] till its current state. The development was motivated primarily by the endeavor to ensure the maximum objectivity of the evaluation of OACW and the comparability of the evaluations among various fields of artistic production. This paper describes the final state of the evaluation model at the time of the transition of RAP into a fully professional mode and the development of the evaluation model.

From the mathematical point of view, the methodology of the evaluation of OACW, which is used in the RAP, is comprised of two models – (1) a multiple-criteria evaluation model based on the Saaty's method, which

[1] Palacký University Olomouc/Faculty of Science, Department of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/12, 771 46 Olomouc, jana.talasova@upol.cz.
[2] Palacký University Olomouc/Faculty of Arts, Department of Applied Economics, Kateřínská 17, 772 00 Olomouc, jan.stoklasa@upol.cz.
[3] Palacký University Olomouc/Faculty of Science, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 241/27, 783 71, Olomouc – Holice, pavel.holecek@upol.cz.
[4] Palacký University Olomouc/Faculty of Science, Department of Mathematical Analysis and Applications of Mathematics, 17. listopadu 1192/12, 771 46 Olomouc, tomas.talasek@upol.cz.

has been used to derive the scores of the individual categories of the artistic work, and (2) a group decision-making model, which is used to assign specific OACW into the individual categories. Another separate challenge was to design the categories of OACW so that comparisons across different fields of artistic production are possible.

2 Categorization of OACW and the scores of categories

For the purpose of the registration and evaluation of the creative- work outcomes in the RAP, the whole area of artistic production is divided into 7 fields: *architecture*, *design*, *film*, *fine arts*, *literature*, *music*, and *theatre*.

OACW (i.e. pieces of art and artistic performances) stored in the RAP database are evaluated according to the following three criteria:

1. relevance or significance,
2. extent,
3. institutional reception.

The first mentioned criterion is the most important one; it reflects the quality of the piece of art or artistic performance. The second criterion has been included into the evaluation because the financial subsidy for a given university should be similar when it produces a lesser number of extensive (large) pieces of art and when it produces a greater number of smaller pieces of art, provided that the evaluation of the quality of the outcomes is similar. The third criterion represents, to some extent, the confirmation of the evaluation according to the first criterion (very significant pieces of art are usually presented in respected institutions – although not exclusively). Statistically speaking, there is a dependency between the evaluation according to the first criterion and the third one; on the other hand, the expert evaluation according to the first criterion is done independently of the value of the third criterion – the significance is assessed independently of the institutional reception.

For the criterion *relevance or significance of the outcome*, four levels of the evaluation are considered:

- an outcome of crucial significance and originality (level “A”),
- an outcome containing numerous important innovations (level “B”),
- an outcome pushing forward modern trends (level “C”),
- other artistic outcomes (level “D”).

For the criterion *extent* of the outcome, three evaluation levels have been defined:

- large (level “K”),
- medium (level “L”),
- limited (level “M”).

And finally, according to the criterion *institutional reception*, creative-work outcomes created and reflected in the following context are distinguished:

- international context (level “X”),
- national context (level “Y”),
- regional context (level “Z”).

Each outcome is assigned a triplet of the letters – for example, AKX, BKY, or CLZ – representing the resulting category of the outcome. In total, there are 27 categories with nonzero scores. Concerning the categories beginning with the letter “D”, they contain *outcomes without any RAP point gain*, regardless of the values of the other two criteria (for the purposes of the evaluation, we denote all such categories simply as a “D” category). If the outcome does not belong to any of the defined categories according to one or more of the criteria, or if the outcome does not meet the formal requirements, it is assigned into the category *rejected* (denoted as “R”).

The decision on the *significance* of the outcome (levels “A”, “B”, “C”, or “D”) is based exclusively on the evaluations provided by experts. Originally the *extent* of the outcome was also decided based on the opinion of experts from the given field of arts. The meanings of the terms *large*, *medium* and *limited,* however, depend on the context (on the particular field of artistic production). That is why in the context of the extent criterion in each field of the artistic production, general types of outcomes with a pre-assigned level of extent have been defined. The assignment of the outcome into one of the *extent* categories is thus done by selecting one of these general outcome types. Concerning the evaluation of the outcome’s *institutional reception,* lists of institutions corresponding with the possible levels of this criterion (X, Y, Z) are defined and annually actualized based on the opinions of experts.

For the determination of scores of each of the categories (AKX,…,CMZ) in the RAP, Saaty's pair-wise comparison method described in [3, 4] has been used. The method uses an evaluation scale with verbal descriptors to express the intensities of preferences between the pairs of categories. The verbal level of this method was utilized, because it was necessary to obtain the information on preferences from the experts from the field of art. The mathematicians in the project team paid a great attention to the organization and facilitation of the whole process of gathering the expert data – to ensure that the process of data input was as natural for the experts as possible. The aim was to obtain an objective reflection of the experts' preferences among the categories.

The representatives of each field of artistic production (the so called guarantors of the fields) were first instructed to specify one example of an outcome for each of the categories with nonzero point gain in their field. The experts could use these examples for the comparisons in the process of determination of scores of categories and, therefore, they did not have to compare the abstract categories. The whole process of determining the preferences was divided into two steps. First, the experts used the pair-wise comparison method to determine the ordering of the categories according to their significance. Subsequently, they determined the preference intensities (the elements of Saaty's matrix) for these ordered categories.

Because of the size of the used Saaty's matrix (27x27), it was difficult to achieve its sufficient consistency. In this context, the notion of a weak consistency of a Saaty's matrix has been applied ([5], [6]).The weak consistency condition is very easy to check in cases when the categories are ordered in the descending order according to their significance (the values in the rows of the Saaty's matrix are non-decreasing from left to right; the values in columns are non-decreasing from below to the top). The individual elements of the pair-wise comparison matrix and, subsequently, also Saaty's matrix were the result of the consensus among the team of experts representing all individual fields of the artistic production; the process was also influenced by the mathematicians who pointed out minor inconsistencies in the preferences. The eigenvector method was used to derive the relative evaluations of the categories from the Saaty's matrix. The vector was then scaled so that the highest possible category (AKX) has the same score (305 points) as the research and development result with the highest evaluation according to the Methodology of research and development results evaluation valid in the Czech Republic at the time when the model has been designed. The initial scores (used for the evaluation in RAP in 2010 - 2015) are listed in Table 1.

Category	Relevance or significance	Extent	Institutional reception	Initial scores	Adjusted scores
AKX	crucial significance	large	international	305	305
AKY	crucial significance	large	national	259	242
AKZ	crucial significance	large	regional	210	98
ALX	crucial significance	medium	international	191	220
AMX	crucial significance	limited	international	174	188
ALY	crucial significance	medium	national	138	162
ALZ	crucial significance	medium	regional	127	85
BKX	containing numerous important innovations	large	international	117	137
AMY	crucial significance	limited	national	97	114
AMZ	crucial significance	limited	regional	90	76
BKY	containing numerous important innovations	large	national	79	68
BKZ	containing numerous important innovations	large	regional	66	38
BLX	containing numerous important innovations	medium	international	62	62
BMX	containing numerous important innovations	limited	international	48	56
BLY	containing numerous important innovations	medium	national	44	49
BLZ	containing numerous important innovations	medium	regional	40	32
BMY	containing numerous important innovations	limited	national	37	42
BMZ	containing numerous important innovations	limited	regional	31	28
CKX	pushing forward modern trends	large	international	26	24
CLX	pushing forward modern trends	medium	international	24	22
CKY	pushing forward modern trends	large	national	19	20
CKZ	pushing forward modern trends	large	regional	17	12
CMX	pushing forward modern trends	limited	international	16	17
CLY	pushing forward modern trends	medium	national	12	15
CLZ	pushing forward modern trends	medium	regional	10	10
CMY	pushing forward modern trends	limited	national	9	13
CMZ	pushing forward modern trends	limited	regional	8	8

Table 1 The comparison of the initial and the adjusted scores for the individual categories.

Thorough analyses of the data stored in RAP are carried out annually after the input phase is completed. The analyses aim to: (a) facilitate the comparability among the individual fields of artistic production, (b) increase the objectivity of the evaluation, (c) identify possible problems of various kinds that might be encountered in the process of gathering and evaluating the outputs. The results of the analyses performed in 2013 initiated significant changes of the types of outputs registered in the RAP – whole pieces of art, whose extent can differ in various segments significantly were replaced by individual artistic performances in some fields of artistic production (such as *theatre*). Primarily, possible general types of outcomes according to the *extent* criterion were predefined for each of the field and these were assigned an appropriate K or L or M level.

The analyses also discovered that outputs, whose significance differs considerably from their institutional reception (A.Z, B.Z), occur in the data more frequently than it was expected. Such cases were assumed to be theoretically possible – e.g. if the first presentation of an outcome was at the end of the year, the experts could recognize its high quality, but the output did not manage to receive the institutional reception yet (the reception evaluation can be increased in the following years, which would increase the output's score in the RAP automatically). Such real life examples of outputs from the individual fields were used by the experts for setting the preferences intensities of these categories in the Saaty's matrix for the calculation of the scores of categories. Long-term observation showed, however, that the assignment of the A.Z category often represents an overestimation of the output; the high expert evaluation is not always supported by a higher institutional reception in the whole monitored 5 year period. That is why a new determination of the significance order of categories and a revision of the elements of Saaty's matrix was performed in 2015. The categories of outputs with only regional *reception* (the triplet of letters ending with Z) were moved towards the end of the group of categories with the given *significance* level (A, B, or C). Afterwards, the elements of Saaty's matrix were again defined; the original expertly defined preferences intensities of the other categories served as an important basis of this process. The resulting weakly-consistent Saaty's matrix that was used for the calculations is depicted in Figure 1. The new adjusted scores of categories are listed in Table 1 together with the initial scores. The new scores were again standardized so that score of the highest category (AKX) was 305 points. These modifications of Saaty's matrix increased also its consistency measured by the CR index (see [3] for the definition).

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
		AKX	AKY	ALX	AMX	ALY	BKX	AMY	AKZ	ALZ	AMZ	BKY	BLX	BMX	BLY	BMY	BKZ	BLZ	BMZ	CKX	CLX	CKY	CMX	CLY	CMY	CKZ	CLZ	CMZ
1	AKX	1	3	3	5	5	5	5	7	7	7	7	7	7	7	7	7	7	7	9	9	9	9	9	9	9	9	9
2	AKY	1/3	1	3	3	3	5	5	5	5	5	5	5	5	7	7	7	7	7	7	7	7	7	9	9	9	9	9
3	ALX	1/3	1/3	1	3	3	5	5	5	5	5	5	5	5	5	7	7	7	7	7	7	7	7	9	9	9	9	9
4	AMX	1/5	1/3	1/3	1	3	3	3	5	5	5	5	5	5	5	7	7	7	7	7	7	7	7	7	7	9	9	9
5	ALY	1/5	1/3	1/3	1/3	1	3	3	3	3	5	5	5	5	5	5	5	7	7	7	7	7	7	7	7	7	9	9
6	BKX	1/5	1/5	1/5	1/3	1/3	1	3	3	3	3	5	5	5	5	5	5	5	5	7	7	7	7	7	7	7	9	9
7	AMY	1/5	1/5	1/5	1/3	1/3	1/3	1	3	3	3	3	3	3	5	5	5	5	5	5	5	5	7	7	7	7	7	9
8	AKZ	1/7	1/5	1/5	1/5	1/3	1/3	1/3	1	3	3	3	3	3	3	5	5	5	5	5	5	5	5	7	7	7	7	7
9	ALZ	1/7	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1	3	3	3	3	3	3	3	5	5	5	5	5	5	5	5	7	7	7
10	AMZ	1/7	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1	3	3	3	3	3	3	5	5	5	5	5	5	5	5	5	7	7
11	BKY	1/7	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1	3	3	3	3	3	5	5	5	5	5	5	5	5	5	7	7
12	BLX	1/7	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	3	5	5	5	5	5	5	5	5	7	7
13	BMX	1/7	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	5	5	5	5	5	5	5	5	5	7
14	BLY	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	3	5	5	5	5	5	5	5	7
15	BMY	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	3	5	5	5	5	5	5	7
16	BKZ	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	3	5	5	5	5	5	7
17	BLZ	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	3	5	5	5	5	7
18	BMZ	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1	3	3	3	3	3	5	5	5	5
19	CKX	1/9	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	3	5	5	5
20	CLX	1/9	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	3	5	5
21	CKY	1/9	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	5	5
22	CMX	1/9	1/7	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1	3	3	3	3	5
23	CLY	1/9	1/9	1/9	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1	3	3	3	5
24	CMY	1/9	1/9	1/9	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1	3	3	3
25	CKZ	1/9	1/9	1/9	1/9	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1/3	1	3	3
26	CLZ	1/9	1/9	1/9	1/9	1/9	1/9	1/7	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1/3	1	3
27	CMZ	1/9	1/9	1/9	1/9	1/9	1/9	1/9	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/7	1/5	1/5	1/5	1/5	1/5	1/5	1/3	1/3	1/3	1

Figure 1 The new Saaty's matrix of preference intensities for the individual categories.

3 Assignment of the OACW into the individual categories

The model used for assigning OACW stored in the RAP into the individual categories of the artistic production is, from the mathematical point of view, a group decision-making model.

The original model used until 2013 was based on the following principle: The initial assignment of the outcome into one of the categories was proposed by the university that has registered this outcome into the RAP. Subsequently, two independent experts proposed their evaluations for the outcome. If a majority agreement has been achieved according to the criteria *significance* and *extent* (the value of the criterion *institutional reception* has been given by assignment of the institutions into one of the levels X, Y, Z, which has been already known at the moment of the evaluation of the outcome by the experts), the overall evaluation has been determined. If the majority agreement was not achieved, the outcome has been discussed in the Arbitration committee that has consisted of the representatives of the individual universities. The Arbitration committee arbitrated also the outcomes which have been proposed to rejection by at least one of the experts and also the outcomes where the university appealed against their stated evaluations.

In 2013, it was decided that the initial evaluation of the outcome proposed by the university should not enter the calculation of the resulting evaluation any longer. From 2013 this initial evaluation is assessed by the guarantor of the respective field of art. This way, the comparability of the evaluation within each field of artistic production is ensured. The guarantor's evaluation and the evaluations of the other two independent (randomly assigned) experts form the basis for the resulting evaluation calculation.

The analyses of the stored data performed in 2015 resulted in other modification proposals of the algorithm for assigning the OACW into the individual categories of the artistic production. The goal was to create a deterministic algorithm that would make it possible to remove the Arbitration committee from the decision process. The basic rule is that if the majority agreement has not been achieved, the RAP application would select another independent expert, whose task it would be (for each of the criteria) to select one of the evaluations proposed by the previous experts (including the possibility to *reject* the outcome). Taking into account the number of levels, which have been defined for the individual criteria, the requirement that the new expert should "incline" to the opinion of one of the previous ones does not represent a problem for the expert.

The practical realization of the proposed rule differs among the three criteria. The list of institutions assigned according to the *institutional reception* into the categories X, Y, or Z are updated every year before the beginning of the period for gathering the data; after the data are gathered (before the expert evaluation is performed) the new institutions registered into the RAP application are also assigned into one of the categories. The assignment is given by the opinion of the guarantor of the respective field and two independent experts – if the majority agreement has not been reached, another assigned expert assesses the output by inclining to the opinion of one of the previous experts (however, this case is quite improbable). All the undecided cases that could in the original version of the methodology end up before the Arbitration committee are thus eliminated.

In case of the *extent* criterion, the assignment of the corresponding level (K, L, or M) is given automatically by the type of the outcome, which is selected from a list when an output is being registered into the RAP. However, because of the continuous progress in the art, new kinds of OACW emerge. That is why a possibility to evaluate an outcome's *extent* by the experts as *large (K)*, *medium (L)*, or *limited (M)* has been preserved for exceptional cases. The same rule as in case of the other criteria is followed – the consensus of at least two of the experts is required.

The main emphasis in the expert evaluation of the outcomes stored in the RAP is put on the *significance* criterion. Let us note that if the university assigns some of their outcomes into the D category (which does not result in any point gain), such outcome does not proceed further into the evaluation process - this way, the need for expert evaluation for insignificant outcomes is reduced. The target state of the model is that even the D category outputs are assessed by the guarantor of the given field of artistic production concerning the minimum requirements for an output to be stored in RAP. The guarantor and the independent experts can select any of the four levels of the *significance* criterion. If the guarantor and both of the independent experts (randomly assigned to the output by the RAP application) select different evaluations, another expert selected by the RAP application would assess the output by inclining to the opinion of one of the previous experts. This way the overall evaluation of the outcome according to this criterion would be provided.

A special case occurs if any of the evaluators proposes the rejection of the outcome. In this case, the same general rules are applied as in the previous cases – the output is rejected if at least two of the evaluators proposed its rejection. The main difference is that the reason for the rejection does not need to be connected with any of the three criteria; rejection can be also caused by formal flaws of the record entered into the RAP itself. There is a significant difference between the D category (with no point evaluation) and the situations when an outcome is rejected; the rejected results are expected to be subjected to sanctions for incorrectly registered outcomes.

4 Conclusion

The mathematical model of evaluation used in the Registry of artistic performance was being developed during the years 2010 to 2016, with changes and modifications inspired by the findings of the annual analyses of the data stored in the RAP. The model's authors were striving to keep the mathematical model as simple and comprehensible as possible. One goal of the performed modifications in the evaluation model was to improve the outcomes' comparability among various fields of artistic production (the solution comprised a modification of the types of the registered outcomes and clear assignment of extent categories K, L, M to these types). Another goal was to increase the evaluation objectivity (implemented by decreasing the point scores of questionable categories A.Z and B.Z, and by introducing the rule that agreement of at least two external evaluators is required in any evaluation performed in the RAP application).

Although various experiments with significantly innovated versions of the model (see e.g. [2]) were performed according to the requests of external authorities, it has always turned out that the originally designed conception of the solution is more suitable for the given purpose.

The scores of the outputs from the past 5 years are used in the Czech Republic for the distribution of a part of the funding from the state budget for the *pedagogical activities* of the universities for the following year, similarly to the scores obtained by these universities for their activities in the research and development. However, it would be desirable to take the scores for the artistic creative-work into account also in the division of the funding for *creative-work* of the Czech universities, where only the points for the research and development have been used so far.

Acknowledgements

The research has been supported by the grant GA14-02424S Methods of operations research for decision support under uncertainty of the Grant Agency of the Czech Republic and partially also by the grant IGA PrF 2016 025 of the internal grant agency of the Palacký University, Olomouc.

References

[1] Garfield, E.: Citation indexes for science: A new dimension in documentation through association of ideas. *Science* **122**, 3159 (1955), 108–111.

[2] Jandová, V., Stoklasa, J., Talašová, J.: Modification of the AHP based model for evaluating artistic production of Czech colleges. *Proceedings of the 32nd International Conference on Mathematical Methods in Economics 2014*, Palacký University, Olomouc, (2014), 372–377.

[3] Saaty, T.L.: *The fundamentals of decision making and priority theory with the Analytic Hierarchy Process*. Vol. VI of the AHP Series, RWS Publ., 2000.

[4] Saaty, T. L.: Decision making with the analytic hierarchy process. *International Journal of Services Sciences*. **1**, (2008), 83–98.

[5] Stoklasa, J., Jandová, V., Talašová, J.: Weak consistency in Saaty's AHP – evaluating creative work outcomes of Czech Art Colleges, *Neural network world*. **23**, 1 (2013), 61–77.

[6] Talašová, J., Stoklasa, J.: A model for evaluating creative work outcomes at Czech Art Colleges. *Proceedings of the 29th International Conference on Mathematical Methods in Economics 2011 – part II*, Praha, (2011), 698–703.

Multicriteria coalitional game with choice from payoffs

Michaela Tichá[1]

Abstract. Multicriteria games are model situations in which at least one player has more than one criterion. If we consider that every player is able to assign weights to the criteria, then the situation is simple. In the case of noncooperative games, the game is reduced to a one-criterion classic game. However, in coalitional games one problem remains if the players have different weights: We cannot simplify the vector payoff function to a scalar payoff function. We can analyze the contribution of each player, but it is difficult to analyze the contribution of coalitions. We use the known formulation of linear model that split the gain of the coalition by use of bargaining theory - egalitarian solution. Then we generalize the model for the case when players and coalitions have possibility to choose from more payoffs. We use the binary variables and formulate a mixed integer programming problem. Finally we shall determine the final coalitional structure.

Keywords: multicriteria games, game theory, cooperative games, multiple multicriteria games.

JEL classification: C72
AMS classification: 91A10

1 Introduction

The multicriteria coalitional game is a class of cooperative games with vector pay-off functions. The players create coalitions depending on what kind of benefit they obtain. Knowing the preferences of the players we can compare the utility of the players. The utility of a player is the weighted sum of the individual criteria. We can easily determine the utility of a player when he is alone. However, in the case of (at least) two-member coalition we cannot easily determine the benefit of the coalition because the players usually have different preferences (weights). We use the egalitarian solution for allocation of the gain among players. We maximized the extra utility for the individual players in the case of egalitarian solution. The condition was that neither of the players attains lower utility, compared to being in coalition with a smaller number of members. This concept is taken from [1]. We use this concept to solve the problem where players can choose from more vector payoffs.

2 Problem formulation

Here we focus on the type of game where the gain depends on several options from which players can choose. Players have several criteria and several payoffs to choose, they may vote among them. To better clarify this issue, provide the following example. Consider that a coalition of $\{\alpha, \beta\}$ can choose between vectors $(1, 3)$ or $(5, 2)$. In the classic monocriterial game this selection is not logical because a player would always choose a higher gain. However, in multicriterial situation, the final choice is not obvious, because one solution does not dominate the second solution. If we know the preferences (weights) of the players, we can determine what gain the player chose, if he were in a one-member coalition. But even then we can not determine the selected gain of the more-member coalition because players may have different preferences.

[1]University of Economics, Prague, Department of Econometrics, W. Churchill Sq. 1938/4, 130 67 Prague 3, xticm11@vse.cz

Definition 1. *Multicriteria coalitional game with choice* is (N, v), where $N = \{1, 2, \ldots, n\}$ is the set of players, $n \in \mathbb{N}$ represents the count of players. Nonempty subset $S \subseteq N$ of the set of players is called coalition. The characteristic function of the game

$$v : S \to \mathbb{R}^{m \times l}$$

assigns the gain $v(S) \in \mathbb{R}^{m \times l}$ to any coalition $S \subseteq N$, $m \in \mathbb{N}$ represents the count of the criteria (the count of all the criteria of players), $l \in \mathbb{N}$ represents the count of variants of payoffs, that players can choose. We denote $v^{\tau\sigma}(S)$ - the gain in the criterion τ while choosing the variant σ:

$$v(S) = \begin{pmatrix} v^{11}(S) & v^{12}(S) & \ldots & v^{1l}(S) \\ v^{21}(S) & v^{22}(S) & \ldots & v^{2l}(S) \\ \vdots & & & \\ v^{m1}(S) & v^{m2}(S) & \ldots & v^{ml}(S) \end{pmatrix}.$$

The characteristic function determines the total payoff of the coalition in each criterion. We assume that $v^{\tau\sigma}(S) \geq 0 \quad \forall \tau \in \{1, 2, \ldots, m\}, \forall \sigma \in \{1, 2, \ldots, l\}$ fo all the coalitions S.

Let assume for example $l = 3$, $m = 2$, which means that every coalitione S can choose from three payoff vectors, for instance $(2, 6)$ or $(4, 2)$ or $(1, 9)$. It is only on the decision of the players from the coalition to decide what vector they prefer. Let consider further the known preferences of the players.

Definition 2. *Multicriteria coalitional game with choice and known preferences* is a triplet (N, v, W), where $N = \{1, 2, \ldots, n\}$ is the set of players, $n \in \mathbb{N}$ represents the count of players. Nonempty subset $S \subseteq N$ of the set of players is called coalition. The characteristic function of the game

$$v : S \to \mathbb{R}^{m \times l}$$

assigns the gain $v(S) \in \mathbb{R}^{m \times l}$ to any coalition $S \subseteq N$. $m \in \mathbb{N}$ represents the count of the criteria (the count of all the criteria of players), $l \in \mathbb{N}$ represents the count of variants of payoffs, that players can choose. We denote $v^{\tau\sigma}(S)$ - the gain in the criterion τ while choosing the variant σ. Let further

$$W = \begin{pmatrix} w_{11} & w_{12} & \ldots & w_{1m} \\ w_{21} & w_{22} & \ldots & w_{2m} \\ \vdots & & & \\ w_{n1} & w_{n2} & \ldots & w_{nm} \end{pmatrix}$$

is the matrix of preferences of the players, $w_{a\tau}$ represents the weight of the criterion τ of the player a.

$$w_{a\tau} \in [0, 1] \qquad a \in N, \tau \in \{1, 2 \ldots, m\},$$
$$\sum_{\tau=1}^{m} w_{a\tau} = 1 \qquad \forall a \in N.$$

Definition 3. *An allocation* $X(S) \in \mathbb{R}^{k \times m}, k = |S|$ is a payoff matrix whose line represents the payoff for each player from the coalition S when variant σ is selected. $x_{a\tau}(S)$ represents the payoff of the player a in the criterion τ in the coalition S by the variant σ.

3 Solution

Assuming known preferences of players, we can easily determine which payoff the player gives priority if he is in one-member coalition. He prefers the payoff that gives the highest utility. Unfortunately it can not be easily determined for coalitions with more than one member with different preferences. Individual members of the coalition may prefer a different payoff vector. We will show the approach that gives a complete coalition structure in multicriterial coalitional game with choice with known preferences.

We proceed from the smallest coalition to the grand coalition. We determine for each coalition which option would be chosen and how the gain is split among the players. We consider that the gain is split by egalitarian approach.

1) $|S| = 1$
For the player $a \in N$, who is alone, his gain is given because it is not divided among several players. So his utility is given, too, it is easy to compute:

$$u_a(\{a\}) = \max_{\sigma \in \{1,2,\dots,l\}} \left(w_{a1} v^{1\sigma}(\{a\}) + w_{a2} v^{2\sigma}(\{a\}) + \ldots + w_{am} v^{m\sigma}(\{a\}) \right).$$

2) $|S| = 2$
In the case of two-member coalition we solve the following program based on [1]. The program is extended to mixed integer programming, binary variables y_σ are added. When the variant σ is chosen, then $y_\sigma = 1$, otherwise $y_\sigma = 0$.

$$\begin{array}{rcll}
\max\limits_{x_a, x_{a'}, D} & & D & \\
u_a(\{a,a'\}) - u_a(\{a\}) & \geq & D & \\
u_{a'}(\{a,a'\}) - u_{a'}(\{a'\}) & \geq & D & \\
x_{a\tau} & \geq & 0 & \forall \tau \in \{1,2,\dots,m\} \\
x_{a'\tau} & \geq & 0 & \forall \tau \in \{1,2,\dots,m\} \\
x_{a\tau} + x_{a'\tau} & = & \sum\limits_{\sigma=1}^{l} y_\sigma v^{\tau\sigma}(S) & \forall \tau \in \{1,2,\dots,m\} \\
u_a(\{a,a'\}) & \geq & u_a(\{a\}) + \varepsilon & \\
u_{a'}(\{a,a'\}) & \geq & u_{a'}(\{a'\}) + \varepsilon & \\
\sum\limits_{\sigma=1}^{l} y_\sigma & = & 1 & \\
y_\sigma & & \text{binary} \quad \forall \sigma \in \{1,2,\dots,l\}, &
\end{array}$$

where

$$u_a(\{a,a'\}) = w_{a1} x_{a1}(\{a,a'\}) + w_{a2} x_{a2}(\{a,a'\}) + \ldots + w_{am} x_{am}(\{a,a'\}),$$
$$u_{a'}(\{a,a'\}) = w_{a'1} x_{a'1}(\{a,a'\}) + w_{a'2} x_{a'2}(\{a,a'\}) + \ldots + w_{a'm} x_{a'm}(\{a,a'\}).$$

This is what we assign to each coalition S, which does not arise because it is not profitable for any of players:

$$x_{a\tau}(S) := 0 \qquad \forall a \in S, \forall \tau \in 1,2,\dots m.$$

3) $|S| = z$; $z = 3, \dots, n$
We use the same procedure for $|S| = 3$, then for $|S| = 4$ etc. up to $|S| = n$. For every coalition S we solve the following program:

$$\begin{array}{rcll}
\max\limits_{x_a, a \in S; D} & & D & \\
u_a(\{S\}) - u_a(\{a\}) & \geq & D & \forall a \in S \\
x_{a\tau} & \geq & 0 & \forall a \in S,\ \forall \tau \in \{1,2,\dots,m\} \\
\sum\limits_{a \in S} x_{a\tau} & = & \sum\limits_{\sigma=1}^{l} y_\sigma v^{\tau\sigma}(S) & \forall \tau \in \{1,2,\dots,m\} \\
u_a(S) & \geq & u_a(S') + \varepsilon & \forall a \in S; \forall S' \in \{S'; a \in S' \wedge |S'| \leq |S|\} \\
\sum\limits_{\sigma=1}^{l} y_\sigma & = & 1 & \\
y_\sigma & & \text{binary} \quad \forall \sigma \in \{1,2,\dots,l\}, &
\end{array}$$

where
$$u_i(S) = \sum_{\tau=1}^{m} w_{i\tau} x_{i\tau}(S).$$

For all coalitions S, that shall not ares because they are not profitable for any of players, we assign

$$x_{a\tau}(S) := 0 \qquad\qquad \forall a \in S, \forall \tau \in 1, 2, \ldots m.$$

Subsequently we solved $\sum_{i=2}^{n} \binom{n}{i}$ programs. It is the mixed integer programming problem. When we know the allocation $X(S)$ for all coalitions, we can determine the final coalitional structure. We will build on stable coalitions.

Definition 4. We say that the coalition S is stable if every player in the coalition has higher utility by at least ε versus the coalitions in which he is a member and if there is no coalition with the same utility and fewer members.

In one-criterial case it is reduced to a standard definition of coalition stability, and it is therefore its natural generalization.

It should also be noted that if a grand coalition (the coalition of all players) is stable, then it is the solution. Such a case cannot occur in which there would exist division of a grand coalition with all conditions satisfied, while there would concurrently exist another coalition with the highest overall utility because the condition for the division of the grand coalition is that every player must have higher utility from it than from being in any other coalition.

If the grand coalition is not stable, we choose a coalition with the highest value of the objective function, i.e., that where the players obtain the highest value of the extra utility. Furthermore, we select the remaining players' coalition with the highest value of the objective function and thus we proceed until we get a coalition structure of the game. Like in classical one-criterion cooperative game such a case may occur that there are more coalitions with the highest utility. In this case, there may be multiple solutions and the coalition structure is not uniquely determined.

Example 1. Let us consider the game of three players α, β, γ with two criteria and two variants with the characteristic function shown in the Table 1 and the preferences of players shown in the Table 2.

	$v(\{\alpha\})$	$v(\{\beta\})$	$v(\{\gamma\})$	$v(\{\alpha,\beta\})$	$v(\{\alpha,\gamma\})$	$v(\{\beta,\gamma\})$	$v(\{\alpha,\beta,\gamma\})$
1^{st} variant	(1;2)	(3;5)	(7;2)	(7;7)	(7;8)	(12;11)	(17;15)
2^{nd} variant	(2;1)	(4;4)	(5;6)	(6;8)	(8;7)	(11;12)	(16;16)

Table 1: Characteristic function of players α, β, γ with choice

	player α	player β	player γ
1^{st} criterion	0.5	0.2	0
2^{nd} criterion	0.5	0.8	1

Table 2: Preferences of players α, β, γ with choice

We consider $\varepsilon = 0,1$. ε is the amount which a player must obtain by entering into a coalition to be willing to cooperate and not to prefer a coalition with a smaller number of members in which he has more power. This parameter is important because otherwise the solution where only one player would get all the extra gain is possible. We proceed from the coalition with the least number of members to the grand coalition.

1) $|S| = 1$
   In the first step, we determine the utility of one-member coalitions, which is the maximum utility

from all the variants:

$$
\begin{aligned}
u_\alpha(\{\alpha\}) &= \max_{\sigma\in\{1,2\}} \left(w_{\alpha 1}v^{1\sigma}(\{\alpha\}) + w_{\alpha 2}v^{2\sigma}(\{\alpha\})\right) = \max(0.5\cdot 1 + 0.5\cdot 2; 0.5\cdot 2 + 0.5\cdot 1) = 1.5,\\
u_\beta(\{\beta\}) &= \max_{\sigma\in\{1,2\}} \left(w_{\beta 1}v^{1\sigma}(\{\beta\}) + w_{\beta 2}v^{2\sigma}(\{\beta\})\right) = \max(0.2\cdot 3 + 0.8\cdot 5; 0.2\cdot 4 + 0.8\cdot 4) = 4.6,\\
u_\gamma(\{\gamma\}) &= \max_{\sigma\in\{1,2\}} \left(w_{\gamma 1}v^{1\sigma}(\{\gamma\}) + w_{\gamma 2}v^{2\sigma}(\{\gamma\})\right) = \max(0\cdot 7 + 1\cdot 2; 5\cdot 0 + 1\cdot 6) = 6.
\end{aligned}
$$

2) $|S| = 2$
In the second step we find the distribution of two-member coalitions. First we solve the program for $S = \{\alpha, \beta\}$:

$$
\begin{aligned}
\max_{x_\alpha, x_\beta, D} \quad & D\\
u_\alpha(\{\alpha,\beta\}) - u_\alpha(\{\alpha\}) &\geq D\\
u_\beta(\{\alpha,\beta\}) - u_\beta(\{\beta\}) &\geq D\\
x_{\alpha 1} &\geq 0\\
x_{\alpha 2} &\geq 0\\
x_{\beta 1} &\geq 0\\
x_{\beta 2} &\geq 0\\
x_{\alpha 1} + x_{\beta 1} &= y_1 v^{11}(\{\alpha,\beta\}) + y_2 v^{12}(\{\alpha,\beta\})\\
x_{\alpha 2} + x_{\beta 2} &= y_1 v^{21}(\{\alpha,\beta\}) + y_2 v^{22}(\{\alpha,\beta\})\\
u_\alpha(\{\alpha,\beta\}) &\geq u_\alpha(\{\alpha\}) + \varepsilon\\
u_\beta(\{\alpha,\beta\}) &\geq u_\beta(\{\beta\}) + \varepsilon\\
y_1 + y_2 &= 1\\
y_1, y_2 \quad & \text{binary,}
\end{aligned}
$$

where

$$
u_\alpha(\{\alpha,\beta\}) = w_{\alpha 1}x_{\alpha 1}(\{\alpha,\beta\}) + w_{\alpha 2}x_{\alpha 2}(\{\alpha,\beta\}),
$$
$$
u_\beta(\{\alpha,\beta\}) = w_{\beta 1}x_{\beta 1}(\{\alpha,\beta\}) + w_{\beta 2}x_{\beta 2}(\{\alpha,\beta\}).
$$

We used methods for mixed integer programming problem and find optimal solution, see Table 3. Then we solve a similar program for coalitions $\{\alpha, \gamma\}$ and $\{\beta, \gamma\}$. Both coalitions are profitable, the programs have feasible and optimal solution shown in the Table 3.

coalition	$\{\alpha,\beta\}$	$\{\alpha,\beta\}$	$\{\alpha,\gamma\}$	$\{\alpha,\gamma\}$	$\{\beta,\gamma\}$	$\{\beta,\gamma\}$
player	α	β	α	γ	β	γ
1^{st} criterion	6	0	7	0	11	0
2^{nd} criterion	0.23	7.77	0	8	4.67	7.33
utility	3.12	6.22	3.5	8	5.93	7.33
total utility		9.34		11.5		13.26
utility from cooperation $(2\cdot D)$		3.24		4		2.66
y_1		0		1		0
y_2		1		0		1

Table 3: Solution of the multicriteria game with choice for coalitions $\{\alpha,\beta\}$, $\{\alpha,\gamma\}$ and $\{\beta,\gamma\}$

Coalition $\{\alpha, \beta\}$ prefers the second variant as we see in the Table 3 ($y_2 = 1$). Coalition $\{\alpha, \gamma\}$ would choose the first variant ($y_1 = 1$) and coalition $\{\beta, \gamma\}$ prefers the second variant.

3) Finally we determine the allocation of the grand coalition by the following program

$$
\begin{array}{rcl}
\max\limits_{x_\alpha, x_\beta, x_\gamma, D} & & D \\
u_a(\{\alpha,\beta,\gamma\}) - u_a(\{a\}) & \geq & D \qquad \forall a \in \{\alpha,\beta,\gamma\} \\
x_{a\tau} & \geq & 0 \qquad \forall a \in \{\alpha,\beta,\gamma\}, \quad \forall \tau \in \{1,2\} \\
\sum\limits_{a \in \{\alpha,\beta,\gamma\}} x_{a\tau} & = & \sum\limits_{\sigma=1}^{2} y_\sigma v^{\tau\sigma}(\{\alpha,\beta,\gamma\}) \qquad \forall \tau \in \{1,2\} \\
u_a(\{\alpha,\beta,\gamma\}) & \geq & u_a(S') + \varepsilon \qquad \forall S' \in \Omega \\
y_1 + y_2 & = & 1 \\
y_1, y_2 & & \text{binary,}
\end{array}
$$

where $\Omega = \{\{\alpha,\beta\};\{\beta,\gamma\};\{\alpha,\gamma\};\{\alpha\};\{\beta\};\{\gamma\}\}$.
For the solution for the grand coalition see Table 4.

	player α	player β	player γ
1^{st} criterion	8.45	7.55	0
2^{nd} criterion	0	7.27	8.73
utility	4.23	7.33	8.73
total utility			20.29
utility from cooperation $(3 \cdot D)$			8.18
y_1			0
y_2			1

Table 4: Solution of the multicriteria game with choice for coalition $\{\alpha,\beta,\gamma\}$

There exists a feasible solution for a grand coalition, it is also a solution with the greatest utility from cooperation. The grand coaliton is stable, players prefer the second variant ($y_2 = 1$). Therefore the coalition structure of the game is

$$(\{\alpha,\beta,\gamma\}).$$

4 Conclusion

In the paper we formulated multicriteria game with choice. This is a new situation in the game where players may have a possibility to choose their payoff vectors from two or more variants. In one-criterial game, this situation would be meaningless, a player or a coalition would always choose a greater payoff. However, if the payoff function is a vector, it is not clear that the payoff vector is greater. We considered known preferences, used utility theory and solve this problem using mixed integer programming problem. At the end we illustrated the solution on example - multicriteria game of three players.

Acknowledgements

The article is supported by internal grant agency VŠE IGA 54/2015.

References

[1] Tichá, M., Dlouhý, M.: Multicriteria coalitional games with known individual weights, *European Journal of Operational Research*, submitted.

Measuring Comovements by Univariate and Multivariate Regression Quantiles: Evidence from Europe

Petra Tomanová[1]

Abstract. The aim of this paper is to measure changes in dependencies among returns on equity indices for European countries in tranquil periods against crisis periods and to investigate their asymmetries in the lower and upper tail of their distributions. The approach is based on a conditional probability that a random variable is lower than a given quantile while another random variable is also lower than its corresponding quantile. Time-varying conditional quantiles are modeled by the regression quantiles. In addition to the univariate conditional autoregressive models, the vector autoregressive extension is considered. In the second step, the conditional probability is estimated through the OLS regression. The results document a significant increase in European equity return comovements in bear markets during the crisis in 1990s and 2000s. The explicit controlling for the high volatility days does not appear to have an impact on the main findings.

Keywords: CAViaR, codependence, contagion, regression quantiles.

JEL classification: C22, C32, G15
AMS classification: 91G70

1 Introduction

Measuring comovements among equity markets has become an important issue in particular for policy makers who are concerned with the stability of the financial system. Due to the dependencies between equity returns, a loss in one market can be accompanied by a loss in another market. It is essential to analyze the dependencies properly since an underestimating of high correlations may cause even more severe underestimating of potential losses which may lead to a weakened stability of the financial system. However, estimating equity market correlations in financial crisis is a difficult exercise and misleading results have often been reported in the past since there is often a spurious relationship between correlation and volatility [7], [8]. An increase in financial market comovements due to the transmission of crises across countries is referred to as the "contagion" in financial literature.

From extensive reviews dealing with a contagion such as [4], [5] and [9] it is evident that the problem could generally be split into three basic categories: modeling first and second moments of returns, estimating the probability of coexceedance and methods based on copula theory. The econometric framework in terms of time-varying regression quantiles, which is used in this paper, has several advantages with respect to those methodologies. The problem of spurious correlation measures is not an issue here since the regression quantiles are robust to heteroskedasticity and outliers. Moreover, it allows to identify and measure asymmetries in comovement in the upper and lower tails of the distributions. Finally, no assumptions on neither the joint distributions nor the marginal distributions of the variables are needed since the regression quantile is a semi-parametric technique [2].

This approach of measuring comovements was introduced by Cappiello et al. [2]. Since then, several applications have been published. Zagaglia et al. [11] investigated the impact of the turmoil on the relation between gold and the U.S. dollar. Beine et al. [1] combined the approach with a panel data analysis to investigate global integration. Cappiello et al. [3] applied the approach on euro area equity returns and tested whether comovements changed between them after the introduction of the euro currency.

[1]University of Economics, Prague, Faculty of Informatics and Statistics, Department of Econometrics, W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic, petra.tomanova@vse.cz.

2 Estimating comovements

The methodology of Cappiello et al. [2] can be split into three separate stages: modeling the time-varying quantiles by regression quantiles, estimating average probabilities of comovements by OLS regression and testing the changes in financial comovements between tranquil and crisis periods. In this paper a generalization for n-tuple of random variables is employed.

2.1 Modeling time-varying quantiles

Let y_{it} be a set-up with n random variables, $\{y_{it} : \ i = 1, \ldots, n\}$, at time t. Then the conditional time-varying quantile $q_{i,j,t}$ at time t for a given confidence level $\theta_j \in (0, 1)$ and for a random variable y_{it} conditional on $\mathcal{F}_{t-1}$ is defined as

$$\Pr[y_{it} \leq q_{i,j,t} | \mathcal{F}_{t-1}] = \theta_j.$$

Let $q_{it}(\boldsymbol{\beta}_{ij})$ be an empirical specification for the $q_{i,j,t}$ where $\boldsymbol{\beta}_{ij}$ is the p-vector of parameters. The optimization problem for an estimator $\hat{\boldsymbol{\beta}}_{ij,T}$ of unknown vector of parameters $\boldsymbol{\beta}_{ij}$ is defined as

$$\min_{\boldsymbol{\beta}_{ij}} \frac{1}{T} \sum_{t=1}^{T} \rho_{\theta_j}(y_{it} - q_{it}(\boldsymbol{\beta}_{ij})) = \frac{1}{T} \sum_{t=1}^{T} (\theta_j - \mathbb{1}[y_{it} \leq q_{it}(\boldsymbol{\beta}_{ij})]) \times (y_{it} - q_{it}(\boldsymbol{\beta}_{ij})),$$

where $\rho_\theta(e) = e\psi_\theta(e), \ \psi_\theta(e) = \theta - \mathbb{1}[e \leq 0]$.

For the purpose of modeling individual quantiles the methodology of Engle et al. [6] is adopted. The authors proposed a conditional autoregressive value at risk by the regression quantiles (CAViaR) models which were primarily constructed to estimate the value at risk (VaR). Since VaR is defined as a particular quantile of future portfolio values conditional on current available information, $\Pr[y_t < VaR_t | \mathcal{F}_{t-1}] = \theta$, the problem of finding the VaR_t is clearly equivalent to estimating the time-varying conditional quantiles. The generic CAViaR model of Engle et al. [6] with dummy variable for crisis can be specified as

$$q_{it}(\boldsymbol{\beta}_{ij}) = \beta_{0,ij} + \sum_{k=1}^{v} \beta_{k,ij} q_{i,t-k}(\boldsymbol{\beta}_{ij}) + \sum_{l=1}^{w} \beta_{l,ij} f(\mathbf{x}_{i,t-l}) + \beta_{p-1,ij} S_{1,t},$$

where $f(\mathbf{x}_{i,t-l})$ is a function of a finite number of lagged values of observable variables. The vector of parameters, $\boldsymbol{\beta}_{ij} = [\beta_{0,ij} \quad \beta_{1,ij} \quad \ldots \quad \beta_{p-1,ij}]'$, is p-dimensional, in particular $p = 1 + v + w + 1$. The autoregressive terms $\beta_{k,ij} q_{i,t-k}(\boldsymbol{\beta}_{ij}), \ k = 1, \ldots, v$, ensure that the quantile changes smoothly over time and the $f(\mathbf{x}_{i,t-l})$ links $q_{it}(\boldsymbol{\beta}_{ij})$ to observable variables that belong to the information set. The dummy variable $S_{1,t}$ identifies the crisis periods.

Furthermore, a simple version of vector autoregressive (VAR) extension to the quantile models for two variables, $y_{it}, \ i = 1, 2$ [10] is adopted:

$$\begin{aligned} q_{1t}(\boldsymbol{\beta}_j) = \quad & b_{11,j} S_{1,t} + \mathbf{x}_t' \boldsymbol{\beta}_{1j} + b_{12,j} q_{1,t-1}(\boldsymbol{\beta}_j) + b_{13,j} q_{2,t-1}(\boldsymbol{\beta}_j), \\ q_{2t}(\boldsymbol{\beta}_j) = \quad & b_{21,j} S_{1,t} + \mathbf{x}_t' \boldsymbol{\beta}_{2j} + b_{22,j} q_{1,t-1}(\boldsymbol{\beta}_j) + b_{23,j} q_{2,t-1}(\boldsymbol{\beta}_j), \end{aligned} \tag{1}$$

where $\boldsymbol{\beta}_{1j} = [b_{10,j} \ b_{14,j} \ b_{15,j} \ldots]'$, $\boldsymbol{\beta}_{2j} = [b_{20,j} \ b_{24,j} \ b_{25,j} \ldots]'$ and $\mathbf{x}_t = [1 \ x_{1t} \ x_{2t} \ldots]'$ denotes predictors belonging to $\mathcal{F}_{t-1}$. Matrix of parameters $\boldsymbol{\beta}_j$ is estimated by the quasi-maximum likelihood (QML) method. Once the QML estimator $\hat{q}_{i,j,t} = q_{it}(\hat{\boldsymbol{\beta}}_j)$ is obtained, conditional quantile functions $\hat{q}_{i,j,t}$ can be computed.

2.2 Estimation of conditional probability of comovement

The approach of measuring comovements is based on an estimation of probability that a random variable y_{it} at time t falls below a conditional quantile, given that other random variables $\{y_{kt} : k = 1, \ldots, n \wedge k \neq i\}$ are also bellow their corresponding quantiles. Let $\overline{F}_{j,T} \equiv T^{-1} \sum_{t=1}^{T} F_t(q_{1,j,t}, \ldots, q_{n,j,t})$ be an average probability that all $\{y_{it} : i = 1, \ldots, n\}$ random variables fall below their quantiles over a given time period, where $F_t(q_{1j}, \ldots, q_{nj})$ is a cumulative distribution function for the set of quantiles $(q_{1,j,t}, \ldots, q_{n,j,t})$. The conditional quantiles are estimated univariately and multivariately by regression quantiles. Then the indicator variables are constructed in such way that they are equal to one if the realized random variable is lower than the conditional quantile and zero otherwise. To obtain the average probability of comovement among $y_{1t}, \ldots, y_{nt}$ the following linear regression model is constructed and estimated:

$$I_{1t}(\hat{\boldsymbol{\beta}}_{1j,T}) \times I_{2t}(\hat{\boldsymbol{\beta}}_{2j,T}) \times \cdots \times I_{nt}(\hat{\boldsymbol{\beta}}_{nj,T}) = \boldsymbol{W}_t \boldsymbol{\alpha}_j^0 + \epsilon_t, \quad j = 1, \ldots, m, \tag{2}$$

where $I_{it}(\boldsymbol{\beta}_{ij}) \equiv \mathbb{1}[y_{it} \leq q_{it}(\boldsymbol{\beta}_{ij})]$, $i = 1, \ldots, n$, $\boldsymbol{W}_t \equiv [1\ \boldsymbol{S}_t]$ and $\boldsymbol{\alpha}_j^0$ is an s-vector of unknown parameters. The $\boldsymbol{S}_t$ denotes an $s-1$ row vector of time dummies. Let $\hat{\boldsymbol{\alpha}}_{j,T}$ be the OLS estimator of (2), $\hat{\alpha}_{lj,T}$ be the $(l+1)$-th element of $\hat{\boldsymbol{\alpha}}_{j,T}$ for $l = 0, 1, \ldots, s-1$ and S_{lt} represent the l-th element of $\boldsymbol{S}_t$. Row vector $\boldsymbol{S}_t^{(-l)}$ denotes the vector $\boldsymbol{S}_t$ from which the l-th element is removed and $\mathbf{0}$ is a zero vector of required dimension. Then dummies $\boldsymbol{S}_t$ are defined as $\{S_{lt} = 1,\ \boldsymbol{S}_t^{(-l)} = \mathbf{0}\}_{t=1}^T$ [2].

Cappiello et al. [2] provided a set of conditions under which the estimated intercept $\hat{\alpha}_{0j,T}$ of (2) converges in probability to the average probability of comovement in the period of $t \in \{t : \boldsymbol{S}_t = \mathbf{0}\}$ and the sum of estimated parameters $\hat{\alpha}_{0j,T} + \hat{\alpha}_{lj,T}$ converges to the average probability of comovement in the period corresponding to the dummy.

2.3 Testing changes in financial comovements

For the confidence level θ_j let

$$\begin{aligned}
F_t^-(\theta_j) &\equiv \theta_j^{-1} \Pr(y_{1t} \leq q_{1t}(\boldsymbol{\beta}_{1j}), \ldots, y_{nt} \leq q_{nt}(\boldsymbol{\beta}_{nj})), \\
F_t^+(\theta_j) &\equiv (1-\theta_j)^{-1} \Pr(y_{1t} \geq q_{1t}(\boldsymbol{\beta}_{1j}), \ldots, y_{nt} \geq q_{nt}(\boldsymbol{\beta}_{nj})).
\end{aligned}$$

Likelihood $p_t(\theta_j)$ of a tail event at time t for any subset Γ_t of random variables $\{y_{it} : i = 1, \ldots, n\}$, given that a tail event occurred for $\{y_{it} : i = 1, \ldots, n \ \wedge \ y_{it} \notin \Gamma_t\}$ is defined as

$$p_t(\theta_j) \equiv \begin{cases} F_t^-(\theta_j) & \text{if } \theta_j \leq 0.5 \\ F_t^+(\theta_j) & \text{if } \theta_j > 0.5 \end{cases}.$$

For the purpose of this study the realization of random variable y_{it} is a return of an equity index for i-th country at time t and the first dummy variable $S_{1,t}$ from the regression (2) identifies crisis times. Then the probability of comovement in tranquil times and the probability of comovement in crisis times is defined as

$$\begin{aligned}
\overline{p}_0(\theta_j) &\equiv C_0^{-1} \sum_{t \in \{t : \boldsymbol{S}_t = \mathbf{0}\}} p_t(\theta_j), \\
\overline{p}_1(\theta_j) &\equiv C_1^{-1} \sum_{t \in \left\{t : S_{1,t} = 1, \boldsymbol{S}_t^{(-1)} = \mathbf{0}\right\}} p_t(\theta_j),
\end{aligned}$$

respectively, where C_0 and C_1 is a number of observations during tranquil and crisis period, respectively. Capiello et al. [2] applies the following rigorous joint test for an increase in comovements:

$$\hat{\delta}(\underline{\theta}, \overline{\theta}) = (\#\theta)^{-1} \sum_{\theta_j \in [\underline{\theta}, \overline{\theta}]} [\overline{p}_1(\theta_j) - \overline{p}_0(\theta_j)] = (\#\theta)^{-1} \sum_{\theta_j \in [\underline{\theta}, \overline{\theta}]} \hat{\alpha}_{1,j},$$

where $\#\theta$ is a number of addends in the sum and $\hat{\alpha}_{1,j}$ is the OLS estimate of $\alpha_{1,j}$ in (2). When the null hypothesis $\mathrm{H}_0 : \hat{\delta}(\underline{\theta}, \overline{\theta}) = 0$ is rejected in favor of the alternative hypothesis $\mathrm{H}_1 : \hat{\delta}(\underline{\theta}, \overline{\theta}) \neq 0$, the comovements do change between tranquil and crisis periods.

3 Data

Returns on Morgan Stanley Capital International (MSCI) world indices, that are used in this paper, are market-value-weighted, do not include dividends and are denominated in the local currency. The sample includes prices of 16 European countries from December 31, 1987 to April 30, 2015. The sample period spans over 7 131 days which results in 7 130 continuously compounded log returns for each of the European countries. However, the final sample includes less observations since data have been adjusted for non-simultaneous closures. The same data cleaning procedure as in [2] is employed. It takes into an account the fact that national holidays and administrative closures do not fully coincide. For each pair of countries a new data set is created. It includes only equity returns for days on which both markets were opened that day and had been opened the day before. After the data cleaning the number of valid daily returns varies from 6 193 (the Greece – Portugal country pair) to 6 738 (the Germany – Netherlands country pair). A next important aspect of the data is the cross-country synchronicity of daily returns.

Descriptive statistics and sample characteristics for each country were investigated. The presence of financial crises in the sample results in extreme returns and strong fluctuation. Common characteristics indicate leptokurtic distribution of returns and the null hypothesis of the Jarque-Bera test, i.e of normally distributed returns, is rejected at the 1% significance level. The presence of autocorrelation is confirmed by the Ljung-Box test and the presence of a unit root in the return series is readily rejected by the Augmented Dickey-Fuller test at the 1% significance level. Hence the methodology is suitable for analyzing the data sample.

Cappiello et al. [2] determined 7 crisis periods, 3 of them have been taken from the paper of Forbes et al. [7]: Tequila crisis – November 1, 1994 to March 31, 1995; Asian crisis – June 2, 1997 to December 31, 1997; Russian crises – August 3, 1998 to December 31, 1998; Argentinean crisis – March 26, 2001 to May 15, 2001; the US sub-prime crisis – February 15, 2007 to March 30, 2007; Lehman bankruptcy – September 1, 2008 to October 31, 2008; turbulences in the euro area – August 1, 2011 to September 30, 2011. The consequent period of September 4, 2012 – April 30, 2015 represents rather a tranquil period of time, thus further crisis periods are not defined. The crisis sample includes 530 equity returns. After the data cleaning procedure is applied, the crisis sample sizes vary from 466 observations (the Ireland – Portugal country pair) to 510 observations (the Netherlands – Norway country pair).

In the data for European countries a clear evidence of higher correlations over turbulent times than over tranquil times is found. The highest increase of correlation indicates the Denmark – Portugal country pair, for which the excess correlation between tranquil and crisis periods is equal to 0.29. The weakest effect is present in returns of four important European economies – Belgium, Germany, France and the Netherlands. For all bivariate combinations of 16 European countries the average correlation is approximately 0.54 and 0.70 over tranquil days and over days of turbulence respectively. Forbes et al. [7] pointed out that such differences are caused by the heteroskedasticity issue rather than by an increased dependence among country-specific returns during the crises. Time-varying regression quantiles are generally robust to heteroskedasticity and hence an improper inference due to the spurious correlation is avoided.

4 Results and discussion

Ten different univariate CAViaR models are considered for the purpose of this study: adaptive model with $G = 10$, symmetric absolute value, asymmetric slope of Engle et al. [6], non-linear specification of Cappiello et al. [2] and their modifications. In addition to univariate models, the multivariate model (1) where $\mathbf{x}_t = [1\ |y_{1,t-1}|\ |y_{2,t-1}|]'$ is estimated. Each model is estimated for 19 quantile probabilities θ_j ranging from 5% to 95% using a quasi-Newton algorithm with a cubic line search procedure.

The selection of the "best" univariate models is based on the in-sample Dynamic Quantile test of Engle et al. [6]. In addition, the following criteria are taken into an account: computational time, occurrence of quantile-crossing problem, number of parameters and their significance. Quantile-crossing problem refers to the situation when the monotonicity assumption $q_{it}(\boldsymbol{\beta}_{i1}) < q_{it}(\boldsymbol{\beta}_{i2}) < \cdots < q_{it}(\boldsymbol{\beta}_{im-1}) < q_{it}(\boldsymbol{\beta}_{im})$ implied by $0 < \theta_1 < \theta_2 \cdots < \theta_{m-1} < \theta_m < 1$ is violated. Further, the sample is split into two sub-periods of January 1, 1988 – September 3, 2012 and September 4, 2012 – April 30, 2015 to assess the univariate and the multivariate models by an out-of-sample DQ test. Similarly as in the case of the univariate models, this comparison takes into account some additional criteria: correct fraction of expectation, computational time and quantile-crossing problem.

Based on the aforementioned criteria, the following CAViaR model is selected as the most suitable model for European countries:

$$q_{it}(\boldsymbol{\beta}_{ij}) = \beta_{0,ij} + \beta_{1,ij} q_{i,t-1}(\boldsymbol{\beta}_{ij}) + \beta_{2,ij} y_{i,t-1} + \beta_{3,ij} y_{i,t-2} + \beta_{4,ij} |y_{i,t-1}| + \beta_{5,ij} S_{1,t}.$$

In the pivotal analysis, all 120 European country pairs are analyzed for the period of January 1, 1988 – April 30, 2015. Four null hypotheses $\hat{\delta}(\underline{\theta}, \overline{\theta}) = 0$ of different intervals are tested for each European country pair. The results are reported in the Figure 1 where the t statistics are visualized. They are computed from the average of $\hat{\alpha}_{1,j}$ over θ_j, i.e. $\hat{\delta}(\underline{\theta}, \overline{\theta}) = (\#\theta)^{-1} \sum_{\theta_j \in [\underline{\theta}, \overline{\theta}]} \hat{\alpha}_{1,j}$, and the associated standard errors. The t statistic provide a joint test for changes in comovements between two countries. The blue color indicates significant changes in comovements between tranquil and crisis periods and the red color indicates insignificant ones. Matrix on the left represents results for the lower tail of distributions: upper triangle represents results for $\hat{\delta}(0.05, 0.25)$ and the lower triangle for $\hat{\delta}(0.25, 0.50)$. And similarly, matrix on the right represents results for the upper tail of distribution: upper triangle represents results for $\hat{\delta}(0.50, 0.75)$ and the lower triangle for $\hat{\delta}(0.75, 0.95)$.

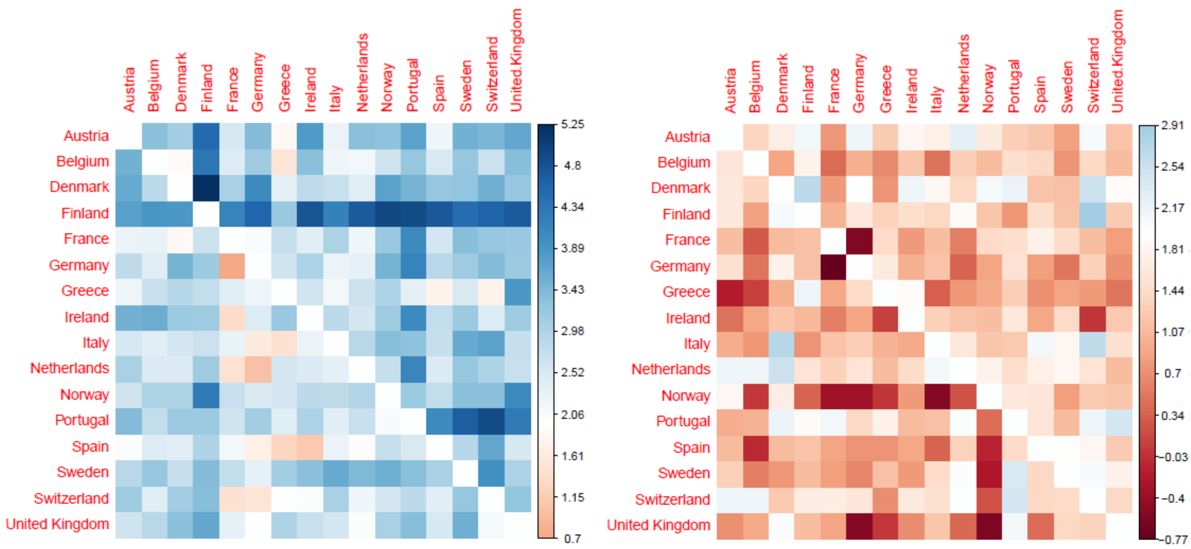


Figure 1 Results from the joint test for changes in comovements (t statistic)

The null hypothesis for $\hat{\delta}(0.05, 0.25)$ is not rejected only for 5 out of 120 country pairs at the 5% significance level, namely Austria – Greece, Belgium – Denmark, Belgium – Greece, Greece – Spain and Greece – Switzerland: the comovements do not change significantly between tranquil and crisis periods for these country pairs. However, for the 96% majority of country pairs, the probabilities of comovements for $\theta_j \in [0.05, 0.25]$ are significantly higher during the crisis times. In case of $\hat{\delta}(0.25, 0.50)$, the differences in comovements are significant for 88% of all country pairs. The null hypothesis $\hat{\delta}(0.25, 0.50) = 0$ is not rejected mostly for country pairs containing France, Germany and Spain. However, for the upper tail of distributions the findings are opposite in sense that for the majority, 86.7% and 85.0% of country pairs, changes in comovements are not significant in the case of $\hat{\delta}(0.50, 0.75)$ and $\hat{\delta}(0.75, 0.95)$, respectively. This evidence clearly suggest that for the most country pairs the comovements change significantly between tranquil and crisis periods in the lower half of distribution but not in the upper half (using the 5% significance level). These findings are also verified by joint tests considering two interval $\hat{\delta}(0.05, 0.50)$ and $\hat{\delta}(0.50, 0.95)$.

To check for the robustness of the estimates, the analyzes have been repeated for two different time spans and controlled for different levels of volatilities. In addition, the analyzes have been repeated for country triplets. Insignificance in changes in comovements in the upper tail of distributions are even more pronounced when a shorter time span of January 2, 1995 – April 30, 2015 is used. For $\hat{\delta}(0.50, 0.75)$ the changes are significant only for 1 out of 120 country pairs, Denmark – Switzerland, and for $\hat{\delta}(0.75, 0.95)$ the changes are significant only for 6 country pairs. Results from the shortest period of June 3, 2002 – April 30, 2015 are quite clear: there is no significant change in probabilities of comovement for all combinations of 16 European countries in the upper tail of distributions except for 3 country pairs although for the lower tail the changes are significant for 96% of country pairs. Moreover, when n is set to 3 and the longest sample period is considered, the results indicate significant changes in the lower tail of the distribution for all country triplets. Upper tail changes are not significant for vast majority of the country triplets, as in the previous case. Finally, dummy variables for different levels of volatilities in financial markets are used as control variables. Let the control dummy $S_{2,t} \equiv \mathbb{1}\left[\sigma^2_{\mathrm{EWMA},t} > q_{0.90}(\sigma^2_{\mathrm{EWMA},t})\right]$, where $\sigma^2_{\mathrm{EWMA},t}$ is an average return on both stock markets forming the country pair computed as the exponentially weighted moving average (EWMA) with decay coefficient equal to $\lambda = 0.97$. Then let $q_{0.90}(\sigma^2_{\mathrm{EWMA},t})$ be its 90% unconditional quantile. Inclusion of the control dummy $S_{2,t}$ into models does not have an impact on the main findings.

Overall, the results confirmed that the distributions of returns for European markets are characterized by strong asymmetries. These asymmetries cannot be detected by a simple correlations or estimates of adjusted correlation coefficients of Forbes et al. [7].

5 Conclusion

The methodology of measuring changes in comovements was applied on 16 European equity markets. Results for the period of January 1, 1988 to April 30, 2015 show that comovements in equity returns across European markets tend to increase significantly in turbulent times against tranquil times in the lower tail of the distribution. However, for vast majority of markets the differences are not significant in the upper tail of the distribution. These results for upper tail contradict the findings of Capiello et al.[2] who analyzed the Latin American countries, since the authors documented significant changes of comovements in crisis times for the whole distribution of the returns. Moreover, the changes in comovements in the upper tail of distributions tend to be insignificant even more for the shorter periods of January 2, 1995 – April 30, 2015 and June 3, 2002 – April 30, 2015. On the other hand, the changes in comovements tend to remain significant for the lower tail of the distributions. This asymmetric dependence documents a significant increase in European equity return comovements in bear markets during the crisis of 1990s and 2000s and insignificant changes in bull markets.

Acknowledgements

This work was supported by the research grant VŠE IGA F4/73/2016, Faculty of Informatics and Statistics, University of Economics, Prague and by the Pigeon. This support is gratefully acknowledged.

References

[1] Beine, M., Cosma, A., and Vermeulen, R.: The dark side of global integration: Increasing tail dependence, *Journal of Banking & Finance* **34** (2010), 184–192.

[2] Cappiello, L., Gérard, B., Kadareja, A., and Manganelli, S.: The impact of the euro on equity markets, *Journal of Financial Econometrics* **12** (2014), 645–678.

[3] Cappiello, L., Kadareja, A., and Manganelli, S.: Measuring comovements by regression quantiles, *Journal of Financial and Quantitative Analysis* **45** (2010), 473–502.

[4] De Bandt, O., and Hartmann, P.: Systemic risk: a survey, *ECB working paper* (2000).

[5] Dungey, M., Fry, R., González-Hermosillo, B., and Martin, V., L.: Empirical modelling of contagion: a review of methodologies, *Quantitative finance* **5** (2005), 9–24.

[6] Engle, R., F., and Manganelli, S.: CAViaR: Conditional autoregressive value at risk by regression quantiles, *Journal of Business & Economic Statistics* **22** (2004), 367–381.

[7] Forbes, K., J., and Rigobon, R.: No contagion, only interdependence: measuring stock market comovements, *The journal of Finance* **57** (2002), 2223–2261.

[8] Longin, F., and Solnik, B.: Extreme correlation of international equity markets, *The journal of finance* **56** (2001), 649–676.

[9] Pericoli, M., and Sbracia, M.: A primer on financial contagion, *Journal of Economic Surveys* **17** (2003), 571–608.

[10] White, H., Kim, T., and Manganelli, S.: VAR for VaR: Measuring tail dependence using multivariate regression quantiles, *Journal of Econometrics* **187** (2015), 169–188.

[11] Zagaglia, P., and Marzo, M.: Gold and the US dollar: tales from the turmoil, *Quantitative Finance* **13** (2013), 571–582.

The Use of Cluster Analysis for Development of Categorical Factors in Exploratory Study: Facts and Findings from the Field Research

Sigitas Vaitkevicius [1], Vladimira Hovorkova Valentova [2], Evelina Zakareviciute [3]

Abstract. This paper describes the development of categorical latent variables and their use for the typology development. The main idea of this modeling was to investigate the possibility of using the classification methods instead of factor analysis for development of the final latent variable which can cumulatively explain the set of primary indicators. The modeling is based on the empirical findings from the online retail consumers' behavior study. Selected data allowed confirm the statement that even the small data sets using the classification data analysis methods can display the significant Ecological Validity. The modeling was performed in three steps. First, the number of primary indicators was reduced using the factor analysis. Based on it several latent variables were created. Second, the k-mean cluster analysis instead of secondary factor analysis was used for development of three cluster variables that represent six clusters in total. Third, all three variables were used for development of the final latent variable which is categorical and represents the eight theoretically possible and five empirically confirmed categories.

Keywords: Categorical variables, latent classes, typology, online retail

JEL Classification: C18, C38, C93
AMS Classification: 62H25, 62H30

1 Introduction

Factor analysis and cluster analysis already have been discussed widely – e. g. in [3], [6], [10], [13], [14]. Factor analysis until now mostly have been used for development of latent variables which describe variability among observed correlated variables in terms of a potentially lower number of unobserved variables called factors. An advantage of the factor analysis is definitely its feature to reduce the number of observed variables connecting them into latent variables that cumulatively explain the correlated content of observed variables. Looking from the point of the interpretation logic, correlation of observed variables is logically significant. If it is not, the development of the factor would be logically unexplainable. In this cases science of methodology already proposed the alternative research methods. It can be cluster analysis because it explores the types of data relations instead of their correlations [6], [10]. *Cluster analysis* is focused on the partitioning of similar objects into meaningful classes when both the number of classes and the composition of the classes are to be determined [5], [10].

There are the studies that allow answer to the questions when and what type of classification design should be used in order to get the correct model of the clusters [4]. For example, in model-based clustering it is assumed that the objects under study are generated by a mixture of probability distributions with one component corresponding to each class. When the attributes of objects are continuous, cluster analysis is sometimes called *latent profile analysis* [8], [11], [1], [17]. When the attributes are categorical, cluster analysis is sometimes called *latent class analysis (LCA)* [11], [9], [1], [15]. There is also cluster analysis of *mixed-mode data* [5] where some attributes are continuous while others are categorical.

This paper describes the situation when the complete verification of the questionnaire is not possible using the factor analysis due to the uncorrelated latent factors. It is based on the real findings from the study that is used as the case which represents the possible research scenario. It combines three of four types of the latent

[1] Kaunas University of Technology, Department of Management, Donelaicio g. 20 – 405, Kaunas, Lithuania, sigitas.vaitkevicius@gmail.com

[2] Technical University of Liberec, Department of Economic Statistics, Studentska 2, Liberec, vladimira.valentova@tul.cz

[3] Kaunas University of Technology, Department of Management, Donelaicio g. 20 – 405, Kaunas, Lithuania, evelina.zakareviciute@gmail.com

variable model. They are in application order: factor analysis, latent profile analysis, and latent class analysis [7]. This example was selected because it explicates the situation when the small number of observations is meaningful and statistically significant. The possibility of use of the factor analysis, the latent profile analysis and the latent class analysis with the small number of observations is also discussed in scientific literature [12]. Just in this paper all this knowledge is combined together in order to answer to the questions if it is possible to verify the questionnaire using the classification design instead of correlative one and how rational and significant it may be when it is used with the small number of observations.

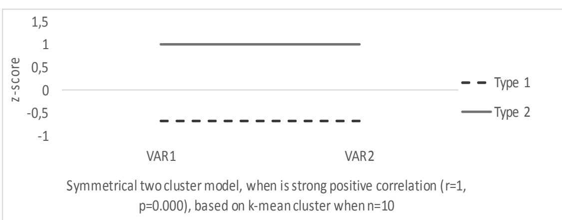
2 Assumptions that Lead to the Use of Classification Methods instead of Correlation Methods for Verification of the Questionnaires

Latent variable model depends on two main components in its structure: *manifest variable* that describes the type of initial variable and *latent variable* which represents the factor variable. The manifest variable and the latent variable depend on its nature and they can be continuous or categorical [2, p. 145]. In the literature presented variable model shows that a factor can be created no matter what type of the variable is. There is just one question that remains unanswered: if it is possible to develop the data analysis design which allow mixing of continuous and categorical types of variables in order to verify the questionnaire. Next, in this paper there is presented the discussion that ground the conditions when it is possible to use correlative methods and when to use the classification methods for the verification.

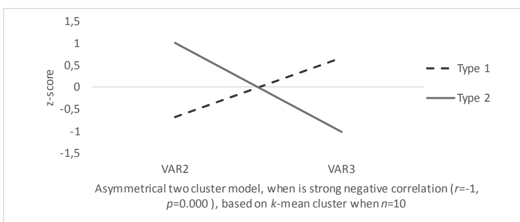
Comparing the factor analysis and latent trait analysis, the latent variables are treated as continuous normally distributed variables, and in latent profile analysis and latent class analysis as from a multinomial distribution [7]. The manifest variables in factor analysis and latent profile analysis are continuous and in most cases their conditional distribution given the latent variables is assumed to be normal. In latent trait analysis and latent class analysis the manifest variables are discrete. These variables could be dichotomous, ordinal or nominal variables [7]. Their conditional distributions are assumed to be binomial or multinomial. Indeed, it is quite different situation that also gives a rise to the question about the possibility of combination of the ways of latent variable development in order to extract one single variable which would verify the consistency of the questionnaire.

The answer to the particular questions becomes possible comparing the behavior of the data. The variables of such analysis are not comparable but the pattern of their behavior can be compared especially if it is enriched by graphs and some statistics. For the illustrative purpose of relationship between the results got by application of correlative and classification methods, the linear diagrams were used together with the main rates from the factor, reliability and k-mean cluster analysis. For illustration, three theoretically possible variables were created – see Figure 1 (d). They are created in a way where VAR1 is filled with random numbers that range from 1 to 7, VAR2 is created by adding 3 to each VAR1 number except the last one, where is added 2.9. It is done because if VAR2 variable is created which would correlate with VAR1 equally by 1 it would not be possible to estimate factor coefficients because the correlation matrix is not positive definite. The correction does not change the essence just allow conducting the comparative analysis. The VAR3 is reversed VAR1. VAR4 is created by repeating VAR1 two times. It means that the values of VAR4 are 1, 5, 2, 7, 3, 4, 5, 7, 1, 2, 1, 5, 2, 7, 3, 4, 5, 7, 1, 2. Firstly, because it would correspond to VAR2 presented as the first 10 items in VAR5 and secondly, it would correspond to VAR3 presented as the second 10 items in VAR5. It means that VAR5 assumes these values: 4, 8, 5, 10, 6, 7, 8, 10, 4, 4,9 ,7, 3, 6, 1, 5, 4, 3, 1, 7, 6. In this way, we work with variables which correlate positively by 1 (VAR1 and VAR2); negatively by -1 (VAR1 and VAR3 (absolutely), or VAR2 and VAR3 (approximately)); and has zero (0) correlation (VAR4 and VAR5) (see Table 1).

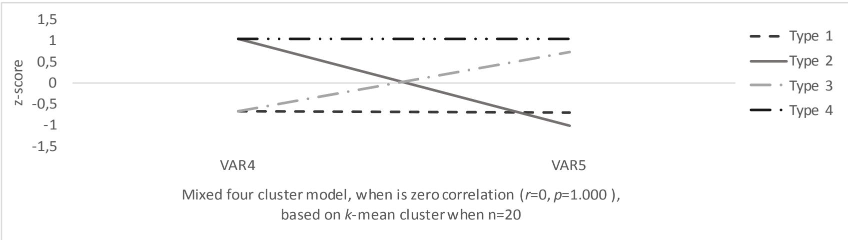
The results of comparative analysis are presented in Figure 1 for the k-mean cluster analysis and in the Table 1 for the Factor and Reliability analysis. The comparative analysis has shown that the Symmetrical two cluster model (a) presented in Figure 1 completely corresponds with the positive strong correlation and can be used as a substitute for the latent variable developed on the base of the factor analysis which is characterized by factor scores equal to 1.000 and by Cronbach's Alpha also equal to 1.000.



(a)



(b)



	VAR1	VAR2	VAR3
1	1.00	4.00	7.00
2	5.00	8.00	3.00
3	2.00	5.00	6.00
4	7.00	10.00	1.00
5	3.00	6.00	5.00
6	4.00	7.00	4.00
7	5.00	8.00	3.00
8	7.00	10.00	1.00
9	1.00	4.00	7.00
10	2.00	4.90	6.00

(c) (d)

Figure 1 Illustrative Data (d) and Their Comparative Models

The Asymmetrical two cluster model (b) presented in Figure 1 completely corresponds to the negative strong correlation and can be used as a substitute for the latent variable developed on the base of the factor analysis which is characterized by factor scores equal to 1.000 and -1.000. In this case the Cronbach's Alpha shows a negative average covariance among items.

Varia-bles	Correlation	Sig. (1-tailed)	**KMO and Bartlett's Test**				**Total Variance Explained. Extraction Sums of Squared Loadings**		
			KMO	*Bartlett's Test*					
				Approx. Chi-square	df	Sig.	% of Variance	Factor Matrix: Factor Score	Cronbach's Alpha
VAR1	*1.000*	*0.000*	*0.500*	*64.658*	*1*	*0.000*	*99.986*	*1.000*	*1.000*
VAR2								*1.000*	
VAR2	*-1.000*	*0.000*	*0.500*	*64.658*	*1*	*0.000*	*99.986*	*1.000*	*-20564.444(a)*
VAR3								*-1.000*	
VAR4	*0.02*	*0.497*	*0.500*	*0.000*	*1*	*0.995*	*0.081*	*0.028*	*0.003*
VAR5								*0.028*	

Extraction Method: Alpha Factoring
(a) The value is negative due to a negative average covariance among items.
1 factor extracted. 1iteration required.

Table 1 Correlation, Factor analysis and Reliability Analysis Coefficients for Theoretical Variables

The Mixed four cluster model (c) displayed in Figure 1 completely correspond to the zero correlation and cannot be used as a substitute for any latent variable developed on the base of the factor analysis. In this case the Cronbach's Alpha is equal to 0.003 which means that no continuous factor is possible. Typically, in this situation the researchers that use factor analysis for questionnaires verification have statement that future development of latent variables is impossible. Moreover, the extraction of one single variable which should verify the consistency of the questionnaire also has to be terminated.

The reason can be seen in the Mixed four cluster model (c). There is the alternative option of usage the Latent profile analysis for development of latent variable. In this case the main difference is that the latent variable is categorical and if it is final variable it means no problem. But if it is in the middle of the verification process, it will need more advanced verification design based on multi-method analysis. In the next chapter an empirical example of the multi-method verification design is presented.

3 Methods and Design of the Research

As an empirical example the Lithuanian on-line retail industry and the internet shops managers´ and owners´ perception about the demand formation motives in the internet shopping have been studied. For this research the questionnaire technique was selected. The correspondent survey method in presence of the researcher has been used. The discussion with the interviewee was used for depth description of the content but not strictly as a tool for the attitude formation.

For this research the questionnaire consisted of 112 questions was used. It was specially designed for this research. 39 questions were used as primary indicators for extraction of latent variables. The 14 indicators together with 13 latent variables have been used for development of the categorical variables. All other questions were designed as the demographical and contextual variables that were used for the description of the interviewee´s characteristics and for description of the attractiveness of the internet shopping. The researched data have been processed using SPSS 20.0 and Microsoft Excel 2010. The license holders are Kaunas University of Technology in Lithuania and Technical University of Liberec in the Czech Republic.

The most of the Cronbach's Alpha values exceed the 0.700 except only one – the "shopping quality" (0.352). It consists of two primary indicators which factor score is 0.626.

Inter-Item Correlations, Extraction Sums of Squared Loadings and Corrected Item-Total Correlation are also high which shows together with Cronbach's Alpha that latent variables are very homogenous according to the extent of their common content. Relatively high homogeneity of latent assertions shows the factor score but in this case it has been observed the higher abstraction rate for the some variables (when the factor score was less

than 0.700) that are included into latent variables: „Functionality“ and „Increasing in accessibility“[4]. Because of this quality their presence in the latent variables has been treated as a feature which only partially related to the content of the latent variable. The primary variables which factor score exceeded 0.700 are interpreted as essential features describing the latent variable.
The data analysis has been made in this order:

- At first, the contextual analysis of each question has been performed. During it the answers to the questions has been compared to the comments from the discussion with the interviewee.
- Next, the “Encoding the real authorial intention using the Phenomenological Hermeneutic system” logical structure for interpretation of the findings has been used [16]. It has been used for the theoretical modelling of the statement truth which was needed for development of theoretical **economic demand formation (EDF) motives** model.
- **Factor analysis.** For reduction of the number of primary items the Factor analysis using Alpha Factoring method and Varimax axis rotation has been selected. The obtained individual latent variables additionally have been tested using the Reliability analysis.
- **Latent profile analysis.** Next, the typology of latent variables has been modelled using the k-mean cluster method by selecting the smallest meaningful number of groups. It was needed to develop the Typology of the economic demand formation (EDF) motives in Lithuanian online retail industry. The typological modelling was performed in two stages. Firstly, it was used for an attribute structure consolidation on the categorical level. The problem of the significance of the results has been resolved transforming the latent variable values to the z-scale.
  The conventional statistical practice is applied for interpretation of the distance on the z-scale. If distance between the points of two clusters was equal or more than one standard deviation, it was interpreted that opinions differ significantly.
  Three cluster models were developed. First one, Demand inclination and growth interferences model (Q1) describes the factors that interfere in the demand formation. The two types of the expert notion were extracted. First one covers the notion of the three experts who think that the growth of online stores trade interferes the complex effect of three latent variables: problems in process of shopping, lack of information, and discrepancy in price and quality. The second type of the expert notion which covers the notion of the four experts is based on the statement that the growth of online stores trade interferes in the lack of immediate communication and the variables: lack of information and the problems in the process of shopping have no impact on the demand inclination.

  The second Demand emergence stimulation model (Q2) also diversifies experts into two groups. First one is based on four expert opinions that the accessibility, time and information of the client about the product are critical to stimulate the demand formation. The other three experts identified the low price as a demand formation stimulation tool.

  The third one, the Demand growth acceleration and growth incentives model (Q3) explored that some e-stores (n=5) are focused on the clients who look for lower price (Price oriented shopping) and others (n=2) offers for Ease oriented shopping client.
  Comparing the clusters depending on the demand emergence stimulation model and those that depend on the Demand growth acceleration and growth incentives model it was noticed that some of the experts see the price as a factor which stimulate the emergence of the demand based on the attraction of the new customers while others see the price as a factor that accelerate the growth of the demand by exploiting the already existing customers.
- **Latent class analysis.** At the last stage, the typological modelling using the Q1, Q2 and Q3 categorical variables was performed. The Two step cluster applying distance measure - Log-likelihood, and selecting number of clusters - fixed number (5) was used. The biggest empirically possible group number was selected for this modelling. The main idea for this model is to make a realistic-empirical model based on expert´s evaluation which would be closest to the theoretical-mathematical model describing idealistic model of typology of the economic demand formation (EDF) motives in Lithuanian online retail industry.

[4] Working with the different size of data sets it should be noticed that the indicators which have the factor score close to the 0.500 comparing to the ones that tend to be close to the 0.999 are much broader by their content. For example, on the scale of 'Openness to changes' the one of nine primary indicators 'risk acceptance' is too broad by its content for the latent variable. It is because the risk can be as an indicator for rational openness to the forthcoming changes and as an indicator for irrational activity like to risk driving the car and so far. Because of such content its factor score for the latent variable was close to 0.500 (Personnel management center (2007) Strategical individual competencies. KTU Faculty of Economics and Management, Personnel management center [CD catalogue], ISBN 978-9955-25-549-9)

4 Discussion on Results

As a result, five empirical types were determined. They were interpreted qualitatively and the final interpretation of the demand formation motives is presented in Table 2. The difference comparing the use of factor analysis for extraction of one single variable which verifies the consistency of the questionnaire is in the interpretation. Using the factor analysis, it is just enough to state that single variable is extracted which shows the consistency of the questionnaire. In this case, no matter that also by using the factor analysis one single variable was extracted, the consistence is presented by the final set of categories which each of them is meaningful itself and can be interpreted separately as well as together. It means that in formativeness of such variable is grater comparing to the factor variable. Moreover, the variable gives the possibility to compare the respondent groups by the behavior pattern that is presented in the inner structure of the variable. For example, from Table 2 it can be seen that respondents which belong to the first group are similar to the second group by the Q1, by the Q2 are similar to the three and four groups, and by the Q3 are similar to the groups four and five. Dissimilarities are also seen. It gives not only the knowledge that consistency exists but also explains what consistency means in a particular case.

Another advantage is the possibility to reduce the complete interpretation to the Key Concept. It is useful developing the typology. In Table 2 is presented the example of such a reduction. In this case, the development of such a typology allows discover the strategies used by the internet shops owners that have led them into success in their business. Therefore, the use of classification methods for verification was not limited only to the evaluation of consistence but also made possible the discovery of business qualities that are important for the better understanding of business nature.

Typology of the economic demand formation (EDF) motives						
		Model of economic motives				
		Deliberate reasonable-price shopping model, n=2	**Clear and easy lower-price shopping model, n=1**	**Fast quality shopping model, n=1**	**Fast and informed reasonable-price shopping model, n=1**	**Fast low-price shopping model, n=2**
Result		**1**	**2**	**3**	**4**	**5**
Demand growth deceleration	Economic motives	Lack of clearness and fairness	Lack of clearness and fairness	Slowness	Slowness	Slowness
Stimulation of demand emergence		Accessibility and information	Low price	Accessibility and information	Accessibility and information	Low price
Demand growth acceleration		Price	Ease	Ease	Price	Price

Table 2 Classification Model of Empirically Explored Economic Demand Formation Motives

Methods selected for this study allowed not only explore the existing patterns but also to explore findings which have not been found across the researched population yet. The categories that identify these three patterns are presented in Table 3 and show the patterns of economic motives that have not been found used as a key success factor for a company success achievement between the Top sellers in Lithuanian on-line retail industry. It can mean that these models have not been enough economic and ideologically attractive for sellers as a way to develop business and strive for profit. It is also the discovery and does not matter that it was not found in empirical data.

Typology of the economic demand formation (EDF) motives				
		Model of economic motives		
		Clear and easy shopping model, n=0	**Clear, lower-price shopping model, n=0**	**Cheap and simple shopping model, n=0**
Result		**1**	**2**	**3**
The strategy for interferences removal *(Demand inclination model)*	Economic motives	Reduction the lack of clearness and fairness	Reduction the lack of clearness and fairness	Immediate communication
The strategy for attraction of the new customers *(Model of demand emergence stimulation)*		Accessibility and information	Low price	Low price
The strategy for retention of existing customers *(Demand growth acceleration model)*		Ease	Price	Ease

Table 3 Typology of the Economic Demand Formation (EDF) Motives, Empirically unexplored Categories

5 Conclusion

In summary, it can be said that beside the simple verification of consistency, the use of classification design for it in addition revealed the complete model of customer's economic motives perceived by managers and online-retail business developers and allow the identification of those that were the exceptionally important for online-retail business development and success at Lithuanian market. It demonstrates the richness of the classification methods when they are used as a continuation of a factor analysis and not only in the situations when the factor analysis is not possible but also when it can be replaced in order to get more information about the inner consistency of the final latent variable.

Finally, it is possible to state that the proposed methodology was useful and allowed objectively identify the differences in the expert opinion that was needed for the empirical description of the economic formation motives of the researched phenomenon.

References

[1] Bartholomew, D. J., and Knott, M. *Latent variable models and factor analysis*. Kendall's Library of Statistics 7. London: Arnold, 1999.

[2] Bartholomew, D. J., Steel, F., Moustaki, I., and Galbraith J. I. *The Analysis and Interpretation of Multivariate Data for Social Scientists*. Chapman & Hall/CRC, New York, 2002.

[3] Child, D. *Essentials of Factor Analysis*. Continuum International Publishing Group, London, 2006.

[4] Creswell, J. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Sage Publications, Thousand Oaks, 2007.

[5] Everitt, B. S. *Cluster Analysis*. Edward Arnold Publishing, New York, 1993.

[6] Everitt, B. S., Landau, S., Leese, M., and Stahl, D. *Cluster Analysis*. John Wiley & Sons, Ltd, London, 2011.

[7] Everitt, B. S. *An Introduction to Latent Variables Models*. Chapman & Hall, New York, 1984.

[8] Gibson, W. A. Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika* **24**, 3 (1959), 229-252.

[9] Goodman, L. A. Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* **61**, 2 (1974), 215-231. doi:10.1093/biomet/61.2.215.

[10] Kaufman, L., and Rousseeuw, P. J. *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons, Inc., New Jersey, 2005.

[11] Lazarsfeld, P. F., and Henry, N.W. *Latent Structure Analysis*. Boston: Houghton Mill, 1968.

[12] Pearson, H. R. Recommended Sample Size for Conducting Exploratory Factor Analysis on Dichotomous Data. *Journal of Modern Applied Statistical Methods* **9**, 2 (2010), 359-368.

[13] Řezanková, H., Húsek, D., and Snášel, V. *Shluková analýza dat*. Professional Publishing, Praha, 2007.

[14] Stankovičová, I., and Vojtková, M. *Viacrozmerné štatistické metódy s aplikáciami*. Iura Edition, spol. s r. o., Bratislava, 2007.

[15] Uebersax, J. S. Probit Latent Class Analysis with Dichotomous or Ordered Category Measures: Conditional Independence/Dependence Models. *Applied Psychological Measurement* **23**, 4 (1999), 283-297.

[16] Vaitkevicius, S. Rethinking the Applicability of Hermeneutic Systems for Coding and Statistical Analysis of Authorial Intentions in Economics. *Inzinerine Ekonomika-Engineering Economics*, **24**, 5 (2013), 415-423.

[17] Vermunt, J. K., and Magidson, J. Latent class cluster analysis. In Hagenaars, J. A. and McCutcheon A. L. (eds.), *Advances in latent class analysis*. Cambridge University Press, 2002.

Analysis of Income of EU Residents Using Finite Mixtures of Regression Models

Kristýna Vaňkátová[1], Eva Fišerová[2]

Abstract. In situations where the classical linear regression is inapplicable due to a heterogeneity of the population, mixtures of regression models are a popular choice. The method acquires parameters estimates by modelling the mixture conditional distribution of the response given the explanatory variables. The mixture distribution is given by the weighted sum over all components and, in the end, an individual regression models, one for each component, can be estimated simultaneously. The estimation of parameters is done via the expectation–maximization (EM) algorithm, a widely applicable algorithm for computing maximum likelihood estimates from incomplete data. Recently, mixture models are used more and more in various fields, including biology, medicine, genetics and the economics. In this paper, income of residents of EU countries is explored, in further detail we analyse the relationship between an annual old age pension and income of people over 65 years. While the data show evident heterogeneity, the mixture regression approach is required.

Keywords: mixture regression models, linear regression, EM algorithm, income, old age pension

JEL classification: C11, C38, C51, C52, E01, E24
AMS classification: 62J05, 62H30, 62F10, 62P20

1 Introduction

Many macroeconomic factors characterizing an economic situation of a country are linked. In this paper, the influence of a social protection on an income of EU countries citizens in retirement age is evaluated. Consequently, two factors representing each side of the relationship were chosen - an annual income of EU countries inhabitants, as a specification of living standard, and old age pension expenditures of given country as a measure of social protection. Focusing on this relation we may proceed to a solution within regression analysis, where the function describing the general dependency (over all countries) is of main interest. Potentially, when the homogeneity is not accomplished, clusters of data should be examined. We approach the problem for people in an age category of 65 years plus. Even basic examination of the data indicates that in this particular case the data comes from at least two latent classes. Hence, the finite mixtures of regression models (FMRM) approach is applied to estimate parameters of each regression model (also called component).

While the classical linear regression is based on an assumption of a homogeneous population, in some situations we suspect that there are several heterogeneous groups in the population that a sample represents. FMRM are designed to model this unobserved heterogeneity using a maximum likelihood methodology as presented for instance in [1], [2], [4] and [5]. An extensive review of finite mixture models is given in [10]. Finite mixture models with a fixed number of components are usually estimated with the expectation-maximization (EM) algorithm [3] within a maximum likelihood framework and with MCMC sampling within a Bayesian framework. In addition, concomitant variable models for the component weights provide the possibility to partition the data into the mixture components through other variables [6]. This extension can provide both more precise parameter estimates and better components identification. Applications of mixtures of regression models can be found in various fields of statistical

[1]Palacký University in Olomouc, Faculty of Science, Department of Mathematical Analysis and Applications of Mathematics, 17. listopadu 12, 771 46 Olomouc, Czech Republic, kristyna.vankatova@upol.cz
[2]Palacký University in Olomouc, Faculty of Science, Department of Mathematical Analysis and Applications of Mathematics, 17. listopadu 12, 771 46 Olomouc, Czech Republic, eva.fiserova@upol.cz

applications such as climatology, biology, economics, medicine and genetics; see e.g. [7] and [8].

2 Mixtures of regression models

Finite mixtures of regression models are frequently used in applications to model unobserved heterogeneity. The sample can be clustered into classes by modelling the conditional distribution $Y_j|\mathbf{x}_j$, where Y_j denotes the response variable and $\mathbf{x}_j$ means the vector of explanatory variables (possibly including an intercept) for jth subject [1]. Then the individual regression models for the classes can be estimated simultaneously. Suppose that each observation $(y_j, \mathbf{x}_j)$ belongs to one of c classes. The mixture of linear regression models is given as follows

$$Y_j = \begin{cases} \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_1 + \varepsilon_{1j} & \text{with probability } \pi_1, \\ \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_2 + \varepsilon_{2j} & \text{with probability } \pi_2, \\ & \vdots \\ \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_c + \varepsilon_{cj} & \text{with probability } \pi_c, \end{cases}$$

where $\boldsymbol{\beta}_i$ denotes the p-dimensional vector of unknown regression parameters for the ith component, ε_{ij} are random errors with normal distribution $\mathrm{N}(0, \sigma_i^2)$, and σ_i^2 is the unknown error variance for the ith component. Probabilities $(\pi_1, \ldots, \pi_c)$ are mixing proportions, i.e., $\pi_i > 0$ and $\sum_{i=1}^{c} \pi_i = 1$, and c is the number of the mixture components.

The response variable Y_j is assumed to be distributed as a finite sum or mixture of conditional univariate normal densities ϕ with the expectation $\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_i$, and the dispersion σ_i^2, $i = 1, \ldots, c$. Accounting for the mixture structure, the conditional density of $Y_j|\mathbf{x}_j$ is

$$f(y_j|\mathbf{x}_j, \boldsymbol{\Psi}) = \sum_{i=1}^{c} \pi_i \phi(y_j|\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_i, \sigma_i^2) = \sum_{i=1}^{c} \pi_i \, (2\pi\sigma_i^2)^{-\frac{1}{2}} \exp\left\{ \frac{-(y_j - \mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_i)^2}{2\sigma_i^2} \right\},$$

where the symbol $\boldsymbol{\Psi}$ denotes the vector of all unknown parameters

$$\boldsymbol{\Psi} = (\pi_1, \ldots, \pi_c, (\boldsymbol{\beta}_1^{\mathrm{T}}, \sigma_1^2), \ldots, (\boldsymbol{\beta}_c^{\mathrm{T}}, \sigma_c^2))^{\mathrm{T}}.$$

2.1 Parameter estimation

Parameter estimation in the mixtures of linear regression models have mainly been studied from a likelihood point of view. For a fixed number of classes c, the parameters $\boldsymbol{\Psi}$ are estimated by maximizing the log–likelihood function

$$\log L(\boldsymbol{\Psi}|\mathbf{x}_1, \ldots, \mathbf{x}_n, y_1, \ldots, y_n) = \sum_{j=1}^{n} \log \Big(\sum_{i=1}^{c} \pi_i \phi(y_j|\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_i, \sigma_i^2) \Big).$$

In this particular case, the observations can be viewed as an incomplete data. The idea here is to think of the data as consisting of triples $(\mathbf{x}_j, y_j, \mathbf{z}_j)$, where $\mathbf{z}_j$ is the unobserved vector indicator that specifies the mixture component from which the observation y_j is drawn [10]. The unobserved component memberships z_{ij} of the observations are treated as missing values and the data are augmented by estimates of the component memberships, i.e. the estimated a-posteriori probabilities τ_{ij}. Thus, any jth observation can be assigned to the ith cluster (using Bayes rule) via the estimated probability:

$$\tau_{ij} = \frac{\pi_i \phi(y_j|\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_i, \sigma_i^2)}{\sum_{h=1}^{c} \pi_h \phi(y_j|\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_h, \sigma_h^2)}.$$

Since mixing proportions sum to unity, $\sum_{i=1}^{c} \pi_i = 1$, the log–likelihood function can be optimize in a sense of Lagrange multipliers method. The maximum likelihood estimates are found by initially forming an augmented log–likelihood function to reflect $\sum_{i=1}^{c} \pi_i = 1$ constraint. In order to obtain stationary equations, we compute the first order partial derivatives of the augmented log–likelihood function and equate them to zero. Afterwards, it is a matter of few simple modifications to acquire a new system of equations obviously corresponding to stationary equations of another optimization problem, that is maximization of

$$\log L_c(\boldsymbol{\Psi}) = \sum_{i=1}^{c}\sum_{j=1}^{n} \tau_{ij} \log \phi(y_j|\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_i, \sigma_i^2). \tag{1}$$

Since in the likelihood (1) only the estimated a-posteriori probabilities τ_{ij} are considered instead of unobservable values z_{ij}, (1) is called expected complete log–likelihood. This particular structure gainfully lends itself to the development of a two-stage EM algorithm.

The number of components is chosen by comparing BIC or AIC of various models, each with a different number of components.

2.2 EM algorithm

The EM algorithm [3] is an iterative procedure which alternates between an Expectation step and a Maximization step. The EM algorithm works on the expected complete log–likelihood and exploits the fact that the expected likelihood is in general easier to maximize than the original likelihood. In the E-step, a-posteriori probabilities τ_{ij} are estimated. Consequently, expected complete log–likelihood is maximized in the M-step and vector of unknown parameters $\boldsymbol{\Psi}$ is updated. The EM algorithm has been shown to increase the likelihood in each step and hence to converge for bounded likelihoods. The $(k+1)$th iteration of EM algorithm can be summarized as follows:

E-step: Given the observed data $\mathbf{y}$ and current parameter estimates $\widehat{\boldsymbol{\Psi}}^{(k)}$ in the kth iteration, replace the missing data z_{ij} by the estimated a-posteriori probabilities:

$$\widehat{\tau}_{ij}^{(k)} = \frac{\widehat{\pi}_i^{(k)}\phi(y_j|\mathbf{x}_j^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_i^{(k)}, \widehat{\sigma}_i^{2(k)})}{\sum_{h=1}^{c}\widehat{\pi}_h^{(k)}\phi(y_j|\mathbf{x}_j^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_h^{(k)}, \widehat{\sigma}_h^{2(k)})}. \tag{2}$$

M-step: Given the estimates $\widehat{\tau}_{ij}^{(k)}$ for the a-posteriori probabilities τ_{ij} (which are functions of $\widehat{\boldsymbol{\Psi}}^{(k)}$), obtain new estimates $\widehat{\boldsymbol{\Psi}}^{(k+1)}$ of the parameters by maximizing expected complete log–likelihood

$$Q(\boldsymbol{\Psi}, \widehat{\boldsymbol{\Psi}}^{(k)}) = \sum_{j=1}^{n}\sum_{i=1}^{c} \widehat{\tau}_{ij}^{(k)} \log \phi(y_j|\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_i, \sigma_i^2). \tag{3}$$

This maximization is equivalent to solving the weighted least squares problem, where the vector $\mathbf{y} = (y_1, \ldots, y_n)^T$ of observations and the design matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T$ are each weighted by $\widehat{\tau}_{ij}^{1/2\,(k)}$. Thus, the entire set of $\widehat{\boldsymbol{\beta}}_i^{(k+1)}$ is derived by performing c separate weighted least-squares analyses. In the same spirit, we estimate $\widehat{\sigma}_i^{2\,(k+1)}$ and, lastly, we update the estimates of probability π_i using

$$\widehat{\pi}_i^{(k+1)} = \frac{\sum_{j=1}^{n}\widehat{\tau}_{ij}^{(k)}}{n} \qquad \text{for} \quad i = 1, \ldots, c.$$

Initial values of regression parameters are based on a random division of observations into c components, i.e. on initial $\widehat{\tau}_{ij}^{(0)}$ probabilities, where for each observation y_j only one of these c probabilities equals to 1 and other ones are set to zero. The EM algorithm is stopped when the (relative) change of log–likelihood is smaller than a chosen tolerance.

3 Mixtures of regression models with concomitant variables

The mixture of regression models is assumed to consist of c components where each component follows a parametric distribution. Each component has assigned a weight which indicates the a-priori probability for an observation to come from this component and the mixture distribution is given by the weighted sum over the c components. If the weights depend on further variables, these are referred to as concomitant variables [6]. We consider FMRM with concomitant variables of the form of

$$f(y_j|\mathbf{x}_j, \boldsymbol{\omega}_j) = \sum_{i=1}^{c} \pi_i(\boldsymbol{\omega}_j, \boldsymbol{\alpha}_i)\phi(y_j|\mathbf{x}_j^{\mathrm{T}}\boldsymbol{\beta}_i, \sigma_i^2),$$

where $\boldsymbol{\omega}_j$ denotes the vector of concomitant variables for the jth observation. In this case, the vector of all unknown parameters is given by

$$\boldsymbol{\Psi} = ((\boldsymbol{\alpha}_1^{\mathrm{T}}, \boldsymbol{\beta}_1^{\mathrm{T}}, \sigma_1^2), \ldots, (\boldsymbol{\alpha}_c^{\mathrm{T}}, \boldsymbol{\beta}_c^{\mathrm{T}}, \sigma_c^2))^{\mathrm{T}}.$$

The component weights π_i need to satisfy

$$\sum_{i=1}^{c} \pi_i(\boldsymbol{\omega}_j, \boldsymbol{\alpha}_i) = 1 \quad \text{and} \quad \pi_i(\boldsymbol{\omega}_j, \boldsymbol{\alpha}_i) > 0 \qquad \text{for} \qquad i = 1, \ldots, c, \tag{4}$$

where $\boldsymbol{\alpha}_i$ are the parameters of the concomitant variable model.

Different concomitant variable models are possible to determine the component weights. The mapping function only has to fulfill condition (4). In the following, we assume a multinomial logit model for the π_i given by

$$\pi_i(\boldsymbol{\omega}_j, \boldsymbol{\alpha}_i) = \frac{\exp^{\boldsymbol{\omega}_j^{\mathrm{T}} \boldsymbol{\alpha}_i}}{\sum_{h=1}^{c} \exp^{\boldsymbol{\omega}_j^{\mathrm{T}} \boldsymbol{\alpha}_h}} \qquad \text{for} \qquad i = 1, \ldots, c,$$

with $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_c)$ and $\boldsymbol{\alpha}_1 \equiv \mathbf{0}$.

In this case the expected complete log–likelihood can be derived in similar way like in the previous section and, as a result, the EM algorithm for FMRM with concomitant variables is the following:
E-step: Given the observed data $\mathbf{y}$ and current parameter estimates $\widehat{\boldsymbol{\Psi}}^{(k)}$ in the kth iteration, replace the missing data z_{ij} by the estimated a-posteriori probabilities τ_{ij}:

$$\widehat{\tau}_{ij}^{(k)} = \frac{\pi_i(\boldsymbol{\omega}_j, \widehat{\boldsymbol{\alpha}}_i^{(k)}) \phi(y_j | \mathbf{x}_j^{\mathrm{T}} \widehat{\boldsymbol{\beta}}_i^{(k)}, \widehat{\sigma}_i^{2(k)})}{\sum_{h=1}^{c} \pi_h(\boldsymbol{\omega}_j, \widehat{\boldsymbol{\alpha}}_i^{(k)}) \phi(y_j | \mathbf{x}_j^{\mathrm{T}} \widehat{\boldsymbol{\beta}}_i^{(k)}, \widehat{\sigma}_i^{2(k)})}.$$

M-step: Given the estimates $\widehat{\tau}_{ij}^{(k)}$ for the a-posteriori probabilities τ_{ij} (which are functions of $\widehat{\boldsymbol{\Psi}}^{(k)}$), obtain new estimates $\widehat{\boldsymbol{\Psi}}^{(k+1)}$ of the parameters $\boldsymbol{\Psi}$ by maximizing

$$Q(\boldsymbol{\Psi}, \widehat{\boldsymbol{\Psi}}^{(k)}) = Q_1(\boldsymbol{\beta}_i, \sigma_i^2, i = 1, \ldots, c\,; \widehat{\boldsymbol{\Psi}}^{(k)}) + Q_2(\boldsymbol{\alpha}, \widehat{\boldsymbol{\Psi}}^{(k)}),$$

where

$$Q_1(\boldsymbol{\beta}_i, \sigma_i^2, i = 1, \ldots, c\,; \widehat{\boldsymbol{\Psi}}^{(k)}) = \sum_{j=1}^{n} \sum_{i=1}^{c} \widehat{\tau}_{ij}^{(k)} \log\left(\phi(y_j | \mathbf{x}_j^{\mathrm{T}} \boldsymbol{\beta}_i, \sigma_i^2)\right)$$

and

$$Q_2(\boldsymbol{\alpha}, \widehat{\boldsymbol{\Psi}}^{(k)}) = \sum_{j=1}^{n} \sum_{i=1}^{c} \widehat{\tau}_{ij}^{(k)} \log\left(\pi_i(\boldsymbol{\omega}_j, \boldsymbol{\alpha}_i)\right).$$

Formulas Q_1 and Q_2 can be maximized separately. The maximization of Q_1 gives new estimates $\widehat{\boldsymbol{\beta}}_i^{(k+1)}, \widehat{\sigma}_i^{2(k+1)}$, $i = 1, \ldots, c$, and the maximization of Q_2 gives $\widehat{\boldsymbol{\alpha}}^{(k+1)}$. Q_1 is maximized using the weighted ML estimation of linear models and Q_2 by means of the weighted ML estimation of multinomial logit models.

The treatment regarding a choice of a number of components and starting estimates of regression parameters is analogous to FMRM.

4 Analysis of income of EU residents

Here we focus on modelling the relationship between the mean of equivalised net income of residents and old age pension (OAP) government expenditures. The main task is to find a general dependency over all countries for people over 65 years. The data come from the Eurostat database and they capture last 15 years of development considering given indicators. Each observation in a dataset stands for one state of EU in one year and both variables are measured in Euros. OAP is standardized by computing OAP per inhabitant.

As the first step in our analysis of income, the contribution of time factor to given model was analysed. After careful consideration, time was not included in the model neither as a predictor or a concomitant variable. Using time as a predictor resulted in a model, where most of time levels were insignificant, moreover, on closer inspection, both OAP and mean income have very similar time development. Meaning, adding time to the model is, in fact, excessive and does not carry any new information. Further, based on an assumption that impact of OAP on income could be one or more years delayed, several new models were created. However, none of them performed better than the original one. Therefore, only a simple relationship of income as a response variable and OAP as an explanatory variable is used in a final study. Also, still in terms of ordinary linear regression, we could extend our model by 27 additional explanatory variables representing levels of factor variable GEO, variable describing country of observation origin, together with interaction terms (GEO*OAP). These models provide reasonably quality fit, with some reservations (over-parametrization and/or incorrectly fitted trend in some countries), but they lack to capture the general relationship we are interested in. We have a desire to find the clusters of countries that share the same development of studied relationship. That is the idea behind applying FMRM.

The statistical software R [11] contain several extension packages for estimation of mixture regression models. The results of our study are built on **flexmix** package, introduced in [9]. The package implements a framework for maximum likelihood estimation with the EM algorithm. The main focus is on finite mixtures of regression models and it allows for multiple independent responses and repeated measurements.

Firstly, a suitable number of components must be determined. In order to do so, FMRM was estimated for mixture consisting of 1 to 4 components and all FMRM models were evaluated based on BIC. This information criterion suggests the two component mixture is desirable.

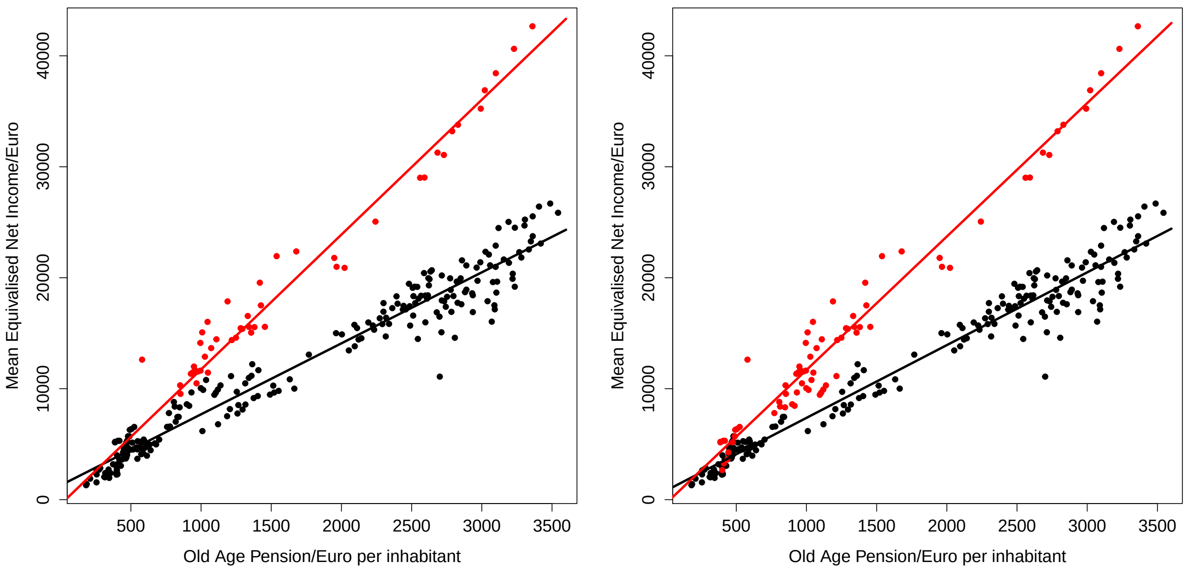


Figure 1 Fitted FMRM without (left) and with concomitant variable (right).

Estimation via FMRM resulted to the following two regression lines (Figure 1, left)

$$\widehat{Income} = 1295.7 + 6.4 * OAP, \quad \widehat{Income} = -420.0 + 12.2 * OAP. \tag{5}$$

The structure of countries included in the first or second component suggests that GEO variable has a strong effect on classification of each observation. By providing GEO as a concomitant variable, the models estimates (5) are updated to (Figure 1, right)

$$\widehat{Income} = 815.6 + 6.6 * OAP, \quad \widehat{Income} = -337.4 + 12.0 * OAP \tag{6}$$

It is visible that not only regression lines are slightly different, but the structure of clusters varies as well. The upper (red) cluster is mainly containing observations from Ireland, Spain, Croatia, Cyprus, Slovenia, Slovakia, Malta and Luxembourg. The rest of EU countries lie in the second, vertically lower, cluster (black). The only regression parameter estimated as insignificantly different from zero is the intercept of the upper component. The accuracy of estimates slightly increases visibly once the concomitant variable is added to a model. For an estimate of intercept standard deviation declines from 517.57 (231.46 for the second component) to 440.57 (210.42), for slope estimate the deviation reduces from 0.30 (0.11) to 0.27 (0.10).

5 Discussion

Income, as an indicator of life quality, is frequently discussed and analysed in economic studies. For people older than 65 years, income appears to be highly dependent on old age pension that given state provide. In order to evaluate this relationship in best possible manner, the FMRM were applied. The results indicate that data follow two different simple regression models, first including most of the EU countries, where the slope is considerably smaller. The second cluster contains mostly observations from 8 specific countries (Ireland, Spain, Croatia, Cyprus, Slovenia, Slovakia, Malta, Luxembourg), where income is growing almost twice as fast as in the first cluster. There may be many reasons behind this behaviour, from different retirement age or significantly varying age structure of a population, to a general tendency of maintaining job even in a retirement age.

Acknowledgements

The authors thankfully acknowledge the support by the grant IGA_PrF_2016_025 of the Internal Grant Agency of Palacký University Olomouc.

References

[1] Bengalia, T., Chauveau, D., Hunter, D. R., Young, D. S.: Mixtools: An R Package for Analyzing Finite Mixture Models, *Journal of Statistical Software* **32**, 6 (2009), 1–29.

[2] De Veaux, R. D.: Mixtures of Linear Regressions, *Computational Statistics & Data Analysis* **8**, 3 (1989), 227–245.

[3] Dempster, A. P., Laird, N. M., Rubin, D. B.: Maximum Likelihood from Incomplete Data Via the EM-Algorithm, *Journal of the Royal Statistical Society, Series B* **39** (1977), 1–38.

[4] Desarbo, W. S., Cron, W. L.: A Maximum Likelihood Methodology for Clusterwise Linear Regression, *Journal of Classification* **5**, 2 (1988), 249–282.

[5] Faria, S., Soromenho, G.: Fitting Mixtures of Linear Regressions, *Journal of Statistical Computation and Simulation* **80**, 2 (2010), 201–225.

[6] Grün, B., Leisch. F.: FlexMix Version 2: Finite Mixtures with Concomitant Variables and Varying and Constant Parameters, *Journal of Statistical Software* **28**, 4 (2008), 1–35.

[7] Grün, B., Scharl, T., Leisch, F.: Modelling time course gene expression data with finite mixtures of linear additive models, *Bioinformatics* **28**, 2 (2012), 222–228.

[8] Hamel, S., Yoccoz, N. G., Gaillard, J. M.: Assessing variation in life-history tactics within a population using mixture regression models: a practical guide for evolutionary ecologists, *Biological Reviews* (2016).

[9] Leisch, F.: FlexMix: A General Framework for Finite Mixture Models and Latent Class Regression in R, *Journal of Statistical Software* **11**, 8 (2004), 1–18.

[10] McLachlan, G., Peel, D.: *Finite Mixture Models.* John Wiley & Sons, New York, 2000.

[11] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, 2016.

Co-movement of Tourism Activity of V4 countries

Petra Vašaničová[1], Eva Litavcová[2], Štefan Lyócsa[3]

Abstract. We study the dependence of tourism activity within four Central and Eastern European Countries (CEE): Poland, Czech Republic, Slovak Republic, and Hungary. Using monthly data of the nights spent at tourist accommodation establishments by residents and non-residents (total) in V4 countries from January 2003 to December 2015 we show, that there exists statistically significant short-term (up to 6 months), mid-term (up to 12 months) and long-term (up to 24 months) co-movement between V4 countries. Maximum long-term co-movement (ρ = 0.88) of tourism activity was found between Poland and Slovak Republic, while minimum short-term co-movement was found between Hungary and Slovak Republic (ρ = 0.29). In general, we can say that lower dependencies between tourism activity were found for short-term co-movements, while larger for long-term co-movement. The results have implications for tourism policies, as they might shed light on whether co-operation between CEE countries is a viable strategy for supporting tourism activities.

Keywords: tourism, nights spent, Central and Eastern Europe, co-movement, long-run correlations

JEL Classification: C32, L83
AMS Classification: 62-07

1 Introduction

When attracting tourist, should neighboring countries cooperate or compete? Using monthly data of the nights spent at tourist accommodation establishments by residents and non-residents in V4 countries from January 2003 to December 2015 we show, that there exists statistically significant short-term (up to 6 months), mid-term (up to 12 months) and long-term (up to 24 months) co-movement between V4 countries. Maximum long-term co-movement (ρ = 0.88) of tourism activity was found between Poland and Slovak Republic, while minimum short-term co-movement was found between Hungary and Slovak Republic (ρ = 0.29). In general, we can say that lower dependencies between tourism activity were found for short-term co-movements, while larger for long-term co-movement.

Most of the existing literature is centered around the tourism-growth nexus (see in: [2], [3], [7], [9]). Bayramoğlu and Ari [2], provided strong unidirectional evidence for causality from the expenditures of foreign tourists who visited Greece to the growth of Greece. Further on, Pérez-Rodríguez, et al. [9], found significant, asymmetric and positive dependence between tourism and GDP growth rates for United Kingdom, Spain and Croatia. Mishra et al. [7], provides evidence of long-run unidirectional causality from tourism activity to economic growth of the India. Finally, Chou [3], in his study applies the bootstrap panel Granger causality approach to test whether domestic tourism spending promotes economic growth. Using a sample of 10 transition countries the results were rather mixed.

Another strand of the literature focuses on interactions of tourism with other economic sectors. For example, Zeren et al. [11], studied the dependence between finance, tourism and advertising in Turkey. Karadžova and Dičevska [4], looked at interactions between the development of the financial system and tourism development in the case of Republic of Macedonia. Serquet and Rebetez [10] examined whether there exists an impact of high temperature values at lower elevation on the number of overnight stays made by domestic tourists in the Swiss Alps and they demonstrated a correlation between domestic tourist demand and hot summer temperatures in the Swiss Alps.

Interestingly, to the best of our knowledge, there are only two studies covering co-movement of tourism activity. Lord and Moore [5] analyzed co-movement in tourist arrivals in the 21 Caribbean destinations for a period from 1977 to 2002. They investigated if tourism can encourage the goal of regional integration. The empirical

[1] University of Presov, Faculty of Management, Konstantinova 16, Presov, Slovakia, petra.vasanicova@smail.unipo.sk.
[2] University of Presov, Faculty of Management, Konstantinova 16, Presov, Slovakia, eva.litavcova@unipo.sk.
[3] University of Presov, Faculty of Management, Konstantinova 16, Presov, Slovakia, University of Economics in Bratislava, Institute of Economics and Management, Dolnozemská cesta 1, Slovakia, stefan.lyocsa@gmail.com.

results presented in the paper provided no evidence for either absolute or conditional convergence in tourism penetration to the Caribbean. This finding is confirmed using both pairwise and panel unit root tests that allow for cross-sectional dependence. Although there are differences in the rate of tourism dependence in the Caribbean, the industry has strengthened regional growth convergence over the past 25 years. The similarities in tourism cycles across the region also suggest that a joint promotion project for Caribbean region could yield significant gains. Mérida [6] contributed to the ability to understand the relationship between the tourist arrivals to Spain from nine developed countries (Belgium, France, Germany, Italy, Netherlands, Portugal, Sweden, United Kingdom, United States and Spain) and the business cycles of both the source countries and Spain. The results from co-movement analysis revealed different patterns for each country. For example, arrivals from France, Germany, Portugal and Sweden seem to be more correlated with the movement of the Spanish business cycle than with their own economic cycle while in other countries (such as Italy or the United Kingdom) both aspects seem to be equally correlated with the tourist outflows to Spain, although for both of them the evolution of their own inner economies seem to be more relevant.

We contribute to the existing literature in two significant ways. First, we fill the gap in the literature on co-movement of tourism activity. Second, existing studies usually cover countries, where tourism is a significant part of the total production of the economy. Our study covers new markets, the CEE countries, where tourism plays only a minor role of the total economy. Still, we have found strong evidence for co-movement between tourism activities of these countries.

The remainder of the paper is structured as follows. Section 2 presents the data. Section 3 describes the methodology. In Section 4 we discuss results and conclude.

2 Data

Our study examines long-run correlation between the nights spent at tourist accommodation establishments (such as: hotels, holiday and other short-stay accommodation, camping grounds, recreational vehicle park and trailer parks) by residents and non- residents (total) in V4 countries, namely: Slovak Republic (SK), Czech Republic (CZ), Poland (PL) and Hungary (HU). Our dataset consists of monthly data from January 2003 to December 2015. Data were obtained from Eurostat database.

Due to the significant seasonal pattern in the data, we have first removed the seasonal component via annualization of the monthly data series. For example, our first observation corresponds to the December 2003 and it is the sum of total monthly nights spent from January 2003 to December 2003. Next observation corresponds to January 2004 and corresponds to the sum of total monthly nights spent from February 2003 to January 2004, etc. The resulting time series are deseasonalized and are displayed on Fig. 1.

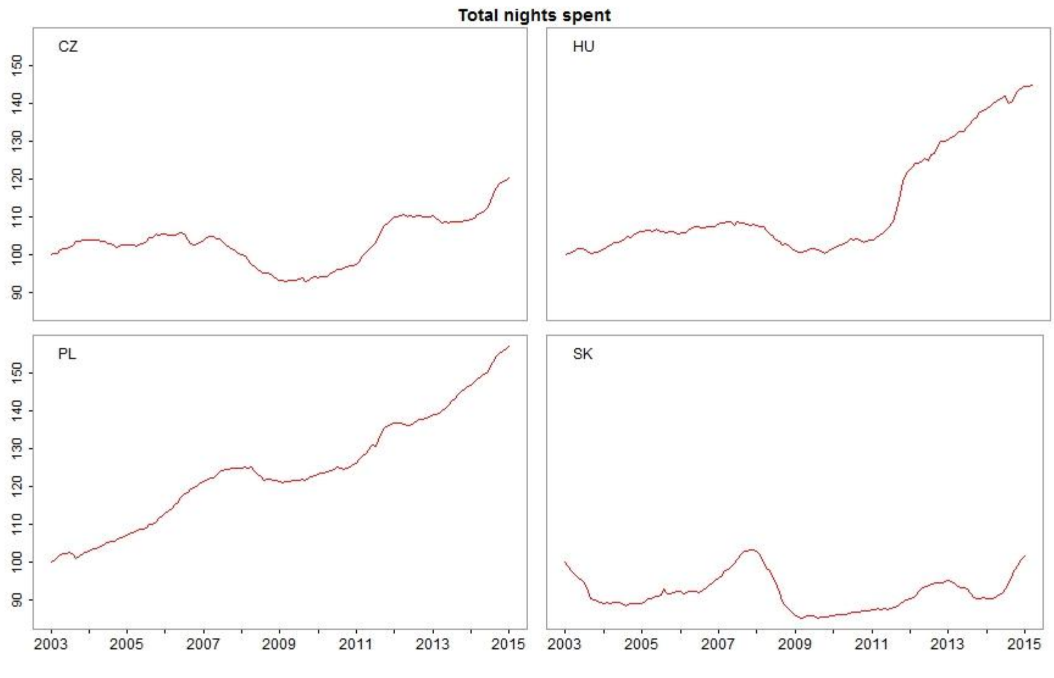


Figure 1 Total nights spent at tourist accommodation establishment of V4 countries

Note: CZ, HU, PL, and SK denotes the Czech Republic, Hungary, Poland, and Slovak Republic.

3 Methodology

In our paper we decided to use long-run correlation coefficient ρ_{xy} , which was presented in [8]. It is a non-parametric approach, which has several advantages, among others it is a model free approach and it leads to a consistent estimate of the variance-covariance matrix even in the presence of autocorrelation and heteroskedasticity. The long-run covariance matrix $\mathbf{\Omega}$ of the process $\mathbf{Z}_t = [y_t, x_t]^{\mathrm{T}}$ is defined as:

$$\hat{\mathbf{\Omega}}_T = \begin{bmatrix} \omega_{yy} & \omega_{xy} \\ \omega_{xy} & \omega_{xx} \end{bmatrix} = \lim_{T\to\infty} T^{-1} \sum_{i=1}^{T} \sum_{j=1}^{T} E\left(\mathbf{Z}_i \mathbf{Z}_j^{\mathrm{T}}\right) \tag{1}$$

where, for a given sample size T:

$$\mathbf{\Omega}_T = \sum_{j=-T}^{T} \Gamma_T\left(j\right) \tag{2}$$

where

$$\Gamma_T\left(j\right) = \begin{Bmatrix} \dfrac{1}{T} \displaystyle\sum_{t=j+1}^{T} \left[Z_t Z_{t-j}\right], & j \geq 0 \\ \dfrac{1}{T} \displaystyle\sum_{t=-j+1}^{T} \left[Z_{t+j} Z_t\right], & j < 0 \end{Bmatrix}, t = 1,2,...,T \tag{3}$$

In line with the study of Andrews [1] the following estimator is employed:

$$\hat{\mathbf{\Omega}}_T = \sum_{j=-T+1}^{T-1} k\left(\frac{j}{M}\right) \hat{\Gamma}_T\left(j\right) \tag{4}$$

where $k(.)$ is a real-valued kernel and M the bandwidth parameter. The kernel function is described in [1] and together with the bandwidth parameter it determines the scheme (and the weight) with which past cross-covariances enter the estimation of the long-run covariance matrix. In our empirical work, we have used the Quandratic spectral kernel weighting function:

$$k_{QS}(x) = \frac{25}{12\pi^2 x^2} \left(\frac{\sin(6\pi\, x/5)}{6\pi\, x/5} - \cos(6\pi\, x/5) \right) \tag{5}$$

Given the variance-covariance matrix, the long-run correlation is estimated as:

$$\hat{\rho}_{xy} \equiv \frac{\hat{\omega}_{xy}}{\sqrt{\hat{\omega}_{xx} \hat{\omega}_{yy}}} \tag{6}$$

Finally, Panopoulou et al. [8] have provided the necessary distributional assumption for the significance test of the long-run correlation coefficient. More specifically, the following test statistics follows the normal distribution under the null hypothesis:

$$\sqrt{\frac{T}{M}} \left(\hat{\rho}_{xy} - \rho_{xy}\right) \sim N\left(0, \left(1 - \rho_{xy}^2\right)^2\right) \tag{7}$$

The long-run correlation was calculated for different bandwidth parameters M. More specifically, if we set the bandwidth parameter to M = 1,2,..,6 the resulting correlation takes into account dependencies across previous or past six months (from one series to another and vice versa). Therefore, such correlations were considered to represent *short-term dependence*. Similarly, if the bandwidth parameter was set to M = 7,8,...,12 the resulting correlation represents the *mid-term dependence* and finally, with M = 13,14,…, the resulting correlation represents the *long-term dependence*.

4 Results and Conclusion

Our results are reported in Table 1. Looking at the significances, it appears clear that almost all relationships are statistically significant. Moreover, the relationships are clearly also economically meaningful, as the correlations range from $\rho = 0.29$ to $\rho = 0.88$.

	Short-term co-movement							**Medium-term co-movement**				
Country/ *M*	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
CZ - HU	0.57^{d}	0.58^{d}	0.59^{d}	0.60^{d}	0.61^{d}	0.63^{d}	0.64^{d}	0.65^{d}	0.67^{d}	0.68^{c}	0.69^{c}	0.71^{c}
CZ - PL	0.54^{d}	0.58^{d}	0.59^{d}	0.60^{d}	0.62^{d}	0.63^{d}	0.64^{d}	0.65^{d}	0.66^{d}	0.67^{c}	0.68^{c}	0.69^{c}
CZ - SK	0.35^{d}	0.39^{d}	0.42^{d}	0.43^{d}	0.44^{c}	0.45^{c}	0.46^{c}	0.47^{c}	0.48^{b}	0.49^{b}	0.49^{b}	0.50^{b}
HU - PL	0.60^{d}	0.66^{d}	$\mathbf{0.67^{d}}$	0.67^{d}	0.67^{d}	0.66^{d}	0.65^{d}	0.65^{d}	0.64^{c}	0.64^{c}	0.63^{c}	0.62^{c}
HU - SK	0.29^{d}	0.33^{d}	0.34^{c}	0.35^{c}	0.35^{b}	0.36^{b}	0.38^{b}	0.39^{a}	0.40^{a}	0.42^{a}	0.43^{a}	0.45^{a}
PL - SK	0.47^{d}	0.51^{d}	0.54^{d}	0.56^{d}	0.58^{d}	0.59^{d}	0.60^{d}	0.61^{d}	0.62^{c}	0.64^{c}	0.66^{c}	0.68^{c}
					Long-term co-movement							
Country/ *M*	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
CZ - HU	0.73^{c}	0.74^{c}	0.75^{c}	0.75^{c}	0.76^{c}	0.77^{c}	0.78^{c}	0.79^{c}	0.80^{c}	0.80^{c}	0.81^{c}	$\mathbf{0.82^{c}}$
CZ - PL	0.70^{c}	0.70^{c}	0.70^{c}	$\mathbf{0.70^{c}}$	0.70^{c}	0.70^{c}	0.69^{b}	0.69^{b}	0.69^{b}	0.69^{b}	0.69^{b}	0.68^{b}
CZ - SK	0.50^{b}	0.51^{a}	0.51^{a}	0.52^{a}	0.53^{a}	0.55^{a}	0.57^{a}	0.58^{a}	$\mathbf{0.58^{a}}$	0.57^{a}	0.56^{a}	0.55^{a}
HU - PL	0.62^{c}	0.61^{b}	0.61^{b}	0.60^{b}	0.60^{b}	0.60^{b}	0.61^{b}	0.61^{b}	0.62^{a}	0.62^{a}	0.63^{a}	0.64^{a}
HU - SK	0.46^{a}	0.47^{a}	0.48^{a}	0.49^{a}	0.50^{a}	0.51^{a}	0.51^{a}	0.52^{a}	0.52^{a}	0.52^{a}	0.52^{a}	$\mathbf{0.52^{a}}$
PL - SK	0.70^{c}	0.73^{c}	0.75^{c}	0.78^{c}	0.82^{c}	0.85^{c}	0.87^{c}	0.88^{c}	0.88^{c}	$\mathbf{0.88^{c}}$	0.88^{c}	0.88^{c}

Table 1 Co-movement between tourism activities of V4 countries.

Note: CZ, HU, PL, and SK denotes the Czech Republic, Hungary, Poland, and Slovak Republic; a, b, c, d denotes 10%, 5%, 1%, and 0.1% significance level; *M* denotes the bandwidth parameter. Bold values represent the highest correlations for a given pair of countries.

Summary statistics of co-movements between tourism activities of V4 countries is listed in Table 2. According to this table, maximum long-term co-movement ($\rho = 0.88$) of tourism activity was found between Poland and Slovak Republic at the bandwidth parameter $M = 22$, while minimum short-term co-movement ($\rho = 0.29$) was found between Hungary and Slovak Republic at the bandwidth parameter $M = 1$. Maximum mean value of co-movement of tourism activity ($\rho_{mean} = 0.71$) was found between Poland and Slovak Republic. Minimum mean value of co-movement ($\rho_{mean} = 0.44$) of tourism activity was found between Hungary and Slovak Republic.

Country	**mean**	**max^{M}**	**min^{M}**
CZ - HU	0.70	0.82^{24}	0.57^{1}
CZ - PL	0.66	0.70^{16}	0.54^{1}
CZ - SK	0.50	0.58^{21}	0.35^{1}
HU - PL	0.63	0.67^{03}	0.60^{1}
HU - SK	0.44	0.52^{24}	$\mathbf{0.29^{1}}$
PL - SK	**0.71**	$\mathbf{0.88^{22}}$	0.47^{1}

Table 2 Summary statistics of co-movement between tourism activities of V4 countries.

Note: CZ, HU, PL, and SK denotes the Czech Republic, Hungary, Poland, and Slovak Republic; *M* denotes the bandwidth parameter. Bold values denote extreme values across all pairs of countries.

In general, we can observe that lower dependencies between tourism activity were found for short-term co-movements, while larger for long-term co-movement. Although, short-term co-movements were observed to have lower values, they were still statistically significant.

To conclude, these results may be helpful for a deeper understanding of the tourism industry in V4 countries. Although, the tourism plays only a minor role of the total economy in these V4 countries, we found strong evidence of co-movement of tourism activity. It appears, that tourist agencies in V4 might find it useful to co-operate when promoting their countries abroad. One readily available explanation for high dependence of tourism activity might be, that the co-movement only manifest the co-movement of real economies (the business cycle). This however, remains to be tested in our subsequent works.

Acknowledgements

Supported by the grant No. 1/0392/15 of the Slovak Grant Agency, by the grant KEGA 037PU-4/2014 of the Cultural and Educational Grant Agency of the Slovak Republic, and by the grant VEGA 1/0412/17 of the Scientific Grant Agency of Ministry of Education SR and Slovak Academy of Sciences.

References

[1] Andrews, D. W. K.: Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* **59** (1991), 817-858.

[2] Bayramoğlu, T., Y. O. Ari.: The relationship between tourism and economic growth in greece economy: a time series analysis. *Computational Methods in Social Sciences* **3** (2015), 89-93.

[3] Chou, M. .C.: Does tourism development promote economic growth in transition countries? A panel data analysis. *Economic Modelling* **33** (2013), 226-232.

[4] Karadžova, V., S. Dičevska.: Interactions between financial system development and tourism development: conditions in Republic of Macedonia. *Sustainable Tourism: Socio-Cultural, Environmental and Economics Impact* (2011), 169-186.

[5] Lorde, T., W. Moore.: Co-movement in tourist arrivals in the Caribbean. *Tourism Economics* **14** (2008), 631-643.

[6] Mérida, A. L.: Tourism and economic growth: causality, persistence and asymmetry. Universidad de Huelva, Huelva, 2015.

[7] Mishra, P. K., H. B. Rout, S. S. Mohapatra.: Causality between Tourism and Economic Growth: Empirical Evidence from India. *European Journal of Social Sciences* **18** (2011), 518-527.

[8] Panopoulou, E., Pittis, N., Kalyvitis, S.: Looking far in the past: revisiting the growth-returns nexus with non-parametric test. *Empirical Economics* **38** (2010), 743-766.

[9] Pérez-Rodríguez, J. V., F. Ledesma-Rodríguez, M. Santana-Gallego.: Testing dependence between GDP and tourism's growth rates. *Tourism Management* **48** (2015), 268-282.

[10] Serquet, G., M. Rebetez.: Relationship between tourism demand in the Swiss Alps and hot summer air temperatures associated with climate change. *Climatic Change* **108** (2011), 291-300.

[11] Zeren, F., M. Koç, F. Konuk.: Interaction between finance, tourism and advertising: evidence from Turkey. *Tourism and Hospitality Management* **20** (2014), 185-193.

On newsvendor model with optimal selling price

Petr Volf[1]

Abstract. The contribution studies the problem of optimizing the inventory amount to achieve maximal profit, under circumstances of uncertain, random demand. This matter, known also as the newsvendor problem, has various variants. After reviewing two basic formulations, we concentrate to the model when the vendor is able to change selling price. Naturally, increased price leads to smaller demand, and vice versa. To describe such an effect, we assume a nonlinear regression-like model of dependence of demand on the price, with an inflexion point at a 'base', customary price. This is in fact the main innovation proposed in the present contribution, as standardly linear or exponential demand curves are considered. We then derive optimal solution concerning both the inventory level and selling price, maximizing either expected profit or chosen quantile of profit distribution. Solutions are illustrated with the aid of simple numerical examples.

Keywords: stochastic optimization, newsvendor model, quantile optimization, demand function.

JEL classification: C41, J64
AMS classification: 62N02, 62P25

1 Introduction

The newsvendor model is a mathematical formulation of the problem how to determine optimal inventory level to be sold in circumstances of uncertain demand. It is named by an analogy with the situation of a newspaper vendor who must decide how many copies of the paper to purchase, knowing that unsold copies will become worthless. The simplest case is characterized by fixed both purchase and selling prices, without any additional costs. In fact it serves as a basic block for building more sophisticated and realistic model variants [3]. The present paper starts by description of such a simple instance. Then, in next parts, some generalizations are provided. First, a model is recalled considering additional costs for stocking unsold items or for not meeting the demand. Then, the main part deals with the case allowing the vendor to change the selling price. We propose a nonlinear, cubic regression-like model for dependence of demand on price, discuss such a choice, and derive rules for simultaneous assessing both optimal stock amount and optimal price in order to maximize chosen characteristics of (random) profit. Namely, we take into account either expected profit or certain profit distribution quantile. We prefer to deal with quantiles, as they are more universal characteristics of distribution (especially of non-symmetric one), simultaneously the calculation of optimal values in considered models is feasible enough. The presentation of models and optimal solutions is accompanied by simple illustrative examples.

1.1 Basic formulation of the newsvendor problem

Let D be a (nonnegative) random variable with known probability distribution representing the demand (of units of certain commodity), further let each unit be sold for price q and purchased for price c, ($c < q$), S be the number of units stocked, i.e. purchased to be sold. In the simplest case c and q are fixed, we search for optimal stocking quantity S which maximizes the profit

$$Z = q \cdot \min(S, D) - c \cdot S. \tag{1}$$

As both D and therefore Z are random, we have a choice which criterion to maximize. It is well known and can be shown easily that to get $\max_S EZ$, optimal solution is $S = S_E = ((q-c)/q)$-quantile of distribution

[1]FP TU Liberec, Department of Applied Mathematics, Univerzitní nám. 1, Liberec, Czech Republic, petr.volf@tul.cz

of D, while for maximization of the α-quantile of Z the optimal stock equals $S = S_\alpha = \alpha$-quantile of D (naturally, in case $q \le c$ the optimal decision is $S = 0$). See for instance [2],[3],[5]. Computation of resulting distribution of profit Z from (1), hence also of achieved values of EZ and quantiles of Z, $Q_Z(\alpha)$, is then rather easy task. Namely, denote f, F, Q the density, distribution function and quantile function of random demand D, further $\bar{F} = 1 - F$, then from (1) it follows that for given S

$$EZ = \int_0^S (qx - cS)f(x)dx + (q-c)S\bar{F}(S) = q\int_0^S \bar{F}(x)dx - cS,$$

after integration by parts. As regards quantiles of Z as function of S, they equal

$$Q_Z(\alpha) = (q-c)S \text{ for } S \le Q(\alpha), \quad Q_Z(\alpha) = qQ(\alpha) - cS \text{ for } S > Q(\alpha).$$

Then maximum of it, which is attained at $S_\alpha = Q(\alpha)$, equals $Q(\alpha)\cdot(q-c)$. This solution can also be interpreted as providing the stock quantity that guarantees that the demand will be satisfied with probability at least α. It is possible and interesting to study dependence of distribution of Z and values of its characteristics, including its variance, on chosen stock amount S. This is actually the purpose of the following example. We have selected the normal distribution for D, $D \sim N(\mu = 150, \sigma = 20)$ (which is practically positive with $P \sim 1$), prices were set to $c = 10$, $q = 15$. Then 0.1, 0.25, 0.5, 0.75, and 0.9 quantiles of this normal distribution equal 124.3690, 136.5102, 150.0000, 163.4898, and 175.6310, respectively. It means that these values are also levels of stock S_α maximizing corresponding α-quantiles of profit Z. Further, the ratio $(q-c)/q = 1/3$, hence the optimal stock level maximizing EZ equals $S_E = 141.3855$, corresponding to 1/3 quantile of distribution of D. Figure 1 shows dependence of considered 5 quantiles of profit Z on S. It contains also the graph of mean values EZ and curves $EZ - 2\sigma_Z$, $EZ + 2\sigma_Z$, where σ_Z denotes standard deviation of Z. Though the value $S_E = 141.3855$ guarantees maximal possible expected profit, simultaneously it is seen that the variance (and therefore uncertainty) of profit really achieved is rather large. However, let us consider just slightly sub-optimal S, for instance $S = 130$. Expected profit is then just slightly smaller, approximately from 600 decreasing to 580, simultaneously standard deviation of distribution of Z is smaller significantly (approximately twice). It is also seen how for even smaller levels of stock (less than 120, say) the distribution of possible profit becomes to be rather narrow, but, naturally, all considered distribution characteristics are already lower.

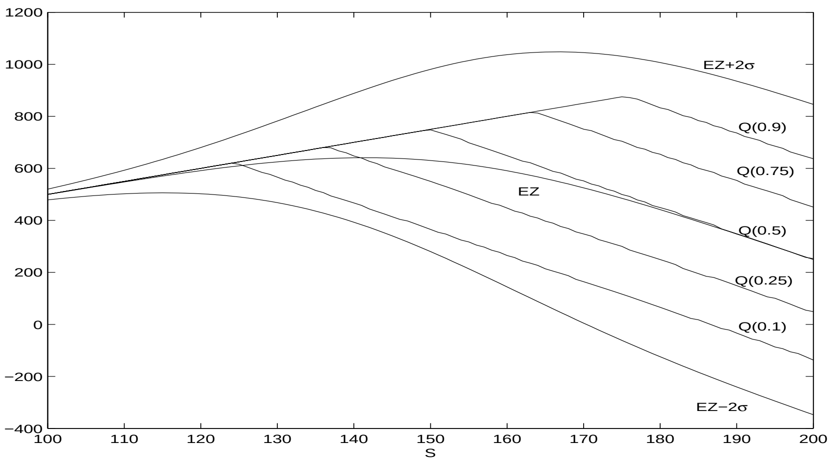


Figure 1 Dependence of characteristics of profit Z on inventory amount S in model (1).

1.2 Generalization with additional costs

The model from above is naturally just a basis for more general cases, see overviews in [1],[3]. One of generalizations can consider a set of additional costs, namely p = penalty (per unit) for non-satisfied demand, and h = cost of stocking each non-sold unit. Moreover, let s be a rest inventory from the previous period which enters the present period as a part of S. Then we search for optimal amount $S - s$ to be purchased. If all costs are fixed and demand is random, then again there is no problem to derive (at least numerically) the distribution of profit Z, where

$$Z = q\cdot\min(S, D) - c\cdot(S - s) - p\cdot\max(D - S, 0) - h\cdot\max(S - D, 0), \text{ hence}$$

$$Z = D(q+h) - S(c+h) + sc \text{ for } S > D, \quad Z = S(q-c+p) - Dp + sc \text{ for } S \leq D, \tag{2}$$

and to find S maximizing either expected profit EZ or certain quantile of Z. Namely, maximum of EZ is attained at $S_E = (q+p-c)/(q+p+h)$-quantile of distribution of demand D, while optimal S maximizing a quantile of Z has to be computed numerically on the basis of distribution (2). Figure 2 illustrates the model. The main characteristics are the same as in Figure 1, moreover we set $h = p = 3$, $s = 0$. Then $S_E = Q(0.381) = 143.94$, optimal S maximizing selected quantile curves can be assessed numerically and are visible from the graph. In the example, under given values of parameters, they are higher than in model (1), while corresponding distribution of profit is shifted down, due additional costs.

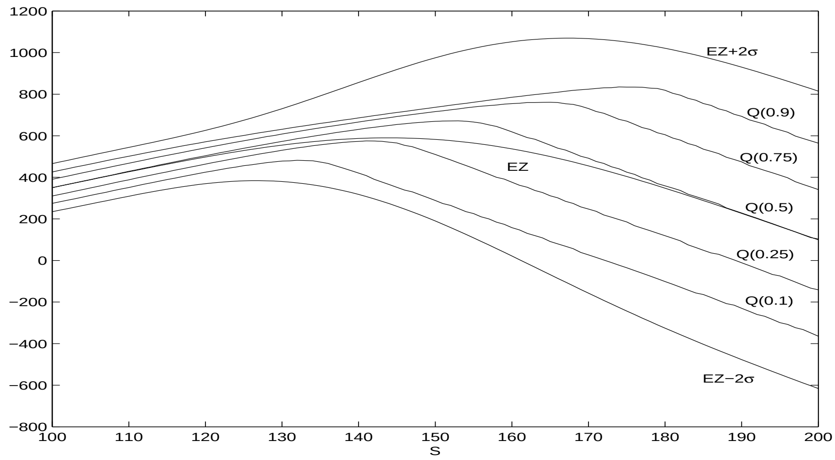


Figure 2 Dependence of characteristics of profit Z on inventory amount S in model (2).

2 Demand dependent on price

In the sequel we shall consider yet another generalization, allowing in the framework of basic model (1) for changes of selling price q. That is why we shall need a relation between price and demand. Thus, the aim is then to find both optimal stock amount and optimal selling price. Inspired for instance by [3] we propose several regression-like models, additive or multiplicative. We shall present basic results with linear demand functions, however later we shall prefer nonlinear trend curves and discuss their use and meaning. Taking into account that, as we have seen, the quantiles of profit distribution are easier to compute, and, mainly, that quantiles are in many cases more reliable characteristics of non-symmetric distributions (for discussion on quantile optimization see [2] or [4]), we shall concentrate to optimization of profit quantiles. Even simple examples in preceding section show how a few quantiles can nicely characterize the main part of profit distribution.

2.1 Additive model

Let as assume that the demand depends on selling price in the following manner

$$D(q) = d(q) + \varepsilon. \tag{3}$$

where $d(q)$ is a trend function, non-increasing, smooth (e.g. with first derivative), ε stands for a random variable, centered and independent of q. In fact, we should assume nonnegative function $d(q)$, to give it a practical sense. Let us denote $F(x), Q(\alpha)$ the distribution function and quantiles of ε, respectively, $\bar{F} = 1-F$. Then, for fixed q, the distribution function of demand $D(q)$ is $F_q(x) = F(x-d(q))$ and its quantiles are $Q_q(\alpha) = Q(\alpha) + d(q)$. The most standard choice of ε is the normal distribution $N(0, \sigma^2)$. In this case, however, we have to truncate it at large values of q in order to keep demand nonnegative. Further, still assuming fixed q, we are able to find optimal stocks maximizing EZ, $S_E(q) = Q((q-c)/q) + d(q)$, or maximizing $Q_z(\alpha)$, which equals $S_\alpha(q) = Q(\alpha) + d(q)$. In the second case we are also able to express achieved maximal value, namely

$$\hat{Q_Z}(\alpha, q) = S_\alpha(q) \cdot (q-c) = [Q(\alpha) + d(q)] \cdot (q-c). \tag{4}$$

Now, the task is to find q maximizing (4). It means to solve equation

$$\frac{d\hat{Q}_Z(\alpha,q)}{dq} = Q(\alpha) + d(q) + d'(q)\cdot(q-c) = 0. \tag{5}$$

Example with linear trend. Let $d(q) = a+bq$, $b<0$, ($d(q)=0$ for $q > -a/b$). From (5) the equation $Q(\alpha)+a+2bq-bc=0$ follows, which is solved at $q = c/2 - (Q(\alpha)+a)/2b$. If we assume that demand at c is positive, i.e. $-a/b > c$, then optimal $q > c$. As regards maximization of expected profit, for fixed q its maximum equals

$$EZ(q) = \int_0^{S_E(q)} \bar{F}(x - d(q))dx - c\cdot S_E(q). \tag{6}$$

It is seen that here the problem of assessing optimal q could be solved just numerically.

2.2 Log-additive model

We can as well speak on multiplicative model for demand, its logarithm (we use natural logarithms) leads to

$$\log D(q) = l(q) + \varepsilon. \tag{7}$$

This model seems to be more convenient than the additive one, guaranteeing positive demand for all q. Again, $l(q)$ is assumed non-increasing. When, standardly, $\varepsilon \sim N(0,\sigma^2)$, $D(q)$ has the log-normal distribution and its variance increases with decreasing price q. Naturally, we could consider an explicit model of dependence of variance on q, however this is beyond scope of the present study. Denote again by $F, \bar{F}, Q$ corresponding characteristics of distribution of ε, then characteristics of distribution of demand (at fixed q) are $F_q(x) = F(\ln x - l(q))$, $Q_q(\alpha) = \exp(Q(\alpha)+l(q))$. Hence we can easily express stock amount maximizing either EZ or a quantile of Z, under given q, namely: $S_E(q) = \exp[Q((q-c)/q)+l(q)]$, $S_\alpha(q) = \exp[Q(\alpha)+l(q)]$. Taking into account the form of optimal EZ (6), we see that the next step, maximization over price q, is not easy and can be attempted by a convenient numerical procedure. On the other hand, the form of optimal quantiles of Z is simpler and in certain instances can be solved by direct computation. At a fixed q, maximal α-quantile of profit Z equals $\hat{Q}_Z(\alpha,q) = \exp[Q(\alpha)+l(q)]\cdot(q-c)$. Equation for assessing optimal q then reads

$$\frac{d\hat{Q}_Z(\alpha,q)}{dq} = \exp[Q(\alpha)+l(q)]\cdot[l'(q)\cdot(q-c)+1] = 0. \tag{8}$$

Example with log-linear trend. Assume now, similarly as above, that $l(q) = a+bq$, $b<0$. Then, searching for optimal q, we obtain the following equation

$$\frac{d\hat{Q}_Z(\alpha,q)}{dq} = \exp[Q(\alpha)+a+bq]\cdot(b(q-c)+1) = 0.$$

It is solved at $q = c - 1/b$, which is larger than c. Notice an interesting feature that optimal q does not depend on α.

3 Nonlinear demand function

Linear (or exponential) functions of dependence of demand on price are the most commonly used models, however I feel that for our purposes other models could be more appropriate. The idea is the following: Let us imagine that there exists a base price, or customary price q_0 for given commodity, and that its small changes actually do not change the demand. However, larger alteration can have already much stronger influence on demand. For instance, under too high price customers start to search for substitution, either from another source or by a similar product. Similarly, low price attracts new buyers. Thus, I guess, a suitable dependence function should be a curve with an inflexion point and with derivative close to zero at q_0, smoothly increasing to left and decreasing to right (for larger price). For instance like a cubic polynomial with conveniently chosen parameters. Figure 3 shows examples of such functions, they are specified below. The purpose of examples is just to illustrate the approach and method of solution. Naturally, there is a rather rich choice of similar functions, for instance an inversion to S-curve (of Gompertz) or functions composed from more parts.

3.1 Example of additive cubic trend

I propose to choose function $d(q) = a + b(q - q_0)^3$, with $b < 0$, possessing all required properties. Namely, its derivative at q_0 equals zero, $a = d(q_0)$, q_0 is also its inflexion point. Optimal values are then obtained via (4) and (5). For a numerical example, let us fix $c = 0.7, b = -0.5, q_0 = 1$, hence also $a = 1$, further let the standard deviation of normal random variable ε be $\sigma = 0.1$, and $\alpha = 0.5$, which means that we wish to optimize median of profit Z. Corresponding cubic demand function $d(q)$ and optimal values (w.r. to S) $\hat{Q_Z}(\alpha, q)$, i.e. medians of profit maximized for fixed q, are displayed in right subplots of Figure 3. Optimal values of price q, stock amount S and achieved median of profit Z then equal

$$q_{opt} = 1.73, \;\; S_{opt} = 0.8055, \;\; Q_{Z,opt}(0.5) = 0.8297.$$

Naturally, if $q \le c$, 'optimal' $S(q) = 0$.

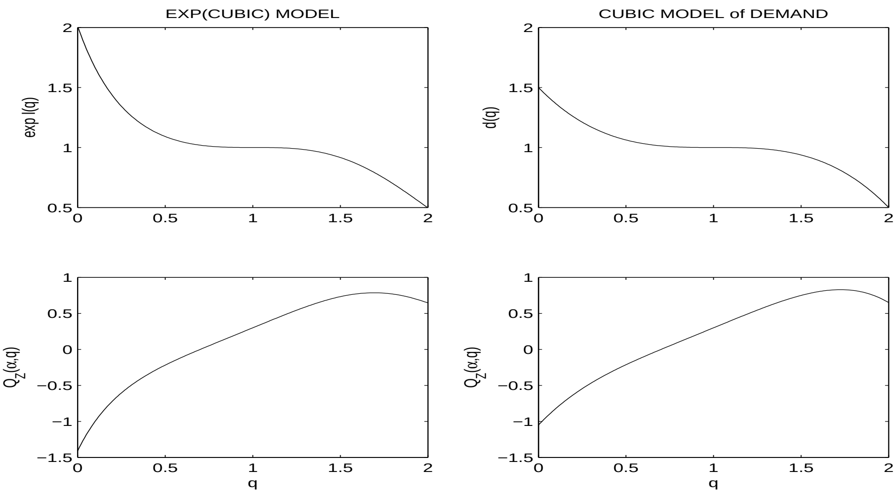


Figure 3 Exponential-cubic trend and cubic trend for demand (first row), optimal values of medians of profit Z as functions of q (second row).

3.2 Example of log-additive cubic trend

Now it is assumed that the log of demand has a cubic trend, $l(q) = a + b(q - q_0)^3$. It means that the trend function of demand equals $d(q) = \exp(l(q))$. Again, 1-st and 2-nd derivatives of both functions are zero at q_0, conditions $d(q_0) = d_0$ implies that $a = \log(d_0)$. Optimal solutions maximizing α-quantile of Z are obtained in accordance with rules derived above, including expression (8).

In a numerical example, we fixed $c = 0.7, b = -0.7, q_0 = 1$, also $d(q_0) = 1$, hence here $a = 0$, further $\sigma = 0.1$, and $\alpha = 0.5$ as above. Left subplots of Figure 3 display function $d(q) = \exp(l(q))$ (above) and corresponding $\hat{Q_Z}(\alpha, q)$ (below). Achieved optimal values of price q, stock amount S and median of profit Z are equal to

$$q_{opt} = 1.69, \;\; S_{opt} = 0.7946, \;\; Q_{Z,opt}(0.5) = 0.7866.$$

We can say that both models offer comparable solution. Optimal recommended price is higher than the base price q_0, under q_{opt} the profit is higher in spite of lower demand.

In order to compare models and their performance, we present here also numerical results from linear and log-linear models. In the former, trend function $d(q) = a + bq$ had parameters $b = -1$, $a = 1 - bq_0 = 2$, as we again fixed $q_0 = 1$, and required $d(q_0) = 1$. However, in these models q_0 is just one of points on demand curve, without the meaning of a "base point" (like above). Further, let again $c = 0.7$, $\sigma = 0.1$ and $\alpha = 0.5$. Functions $d(q)$ and $\hat{Q_Z}(\alpha, q)$, i.e. medians of Z optimized w.r. to S for fixed q, are displayed in Figure 4, right subplots. Optimal values of price q, stock amount S and median of profit Z were equal to

$$q_{opt} = 1.35, \;\; S_{opt} = 0.65, \;\; Q_{Z,opt}(0.5) = 0.4225.$$

For example of log-linear trend $l(q) = a + bq$ we have selected $b = -1$, $a = -b = 1$, then again $\exp(l(1)) = 1$. Functions $\exp(l(q))$ and $\hat{Q_Z}(\alpha, q)$ are displayed in left subplots of Figure 4. In this case the optimal

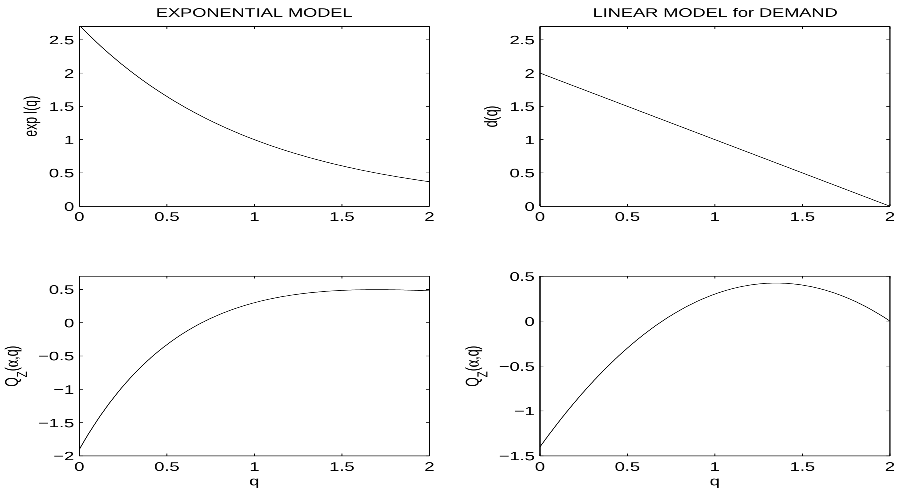


Figure 4 Exponential trend and linear trend for demand (first row), optimal values of medians of profit Z as functions of q (second row).

values of price q, stock amount S and median of profit Z equaled

$$q_{opt} = 1.70, \;\; S_{opt} = 0.4966, \;\; Q_{Z,opt}(0.5) = 0.4966.$$

4 Conclusion

In the present paper several basic variants of the newsvendor model have been reviewed. The main contribution consisted in proposing the solution in the case with simultaneous control of inventory amount and selling price, under nonlinear (here cubic) demand curve. The selection of function form was justified by introducing the term "base price" as inflexion point of the curve. Then, the price optimal from the vendor point of view is, as a rule, higher, which can be interpreted as a base price plus a premium, [3]. A question arises how an appropriate demand function should be estimated. The main source of information should be, I believe, an experience, based on data collection and analysis, extrapolation of extracted relations, as well as on expert knowledge and opinion. Another variant of the model can consider also adaptation of price during selling period. As it has been already said, rather simple models studied here can then serve as building blocks for construction of dynamical multi-period models.

References

[1] Gallego, G.: *IEOR 4000 Production Management Lecture 7.* Columbia University, 1995. Retrieved 30.03.2016 from http://www.columbia.edu/ gmg2/4000/pdf/lect_07.pdf

[2] Kim, J. H., and Powell, W.: *Quantile optimization for heavy-tailed distributions using asymmetric signum functions.* Working Paper, Princeton University, 2011. Retrieved 12.01.2016 from http://castlelab.princeton.edu/Papers/

[3] Petruzzi, N. C., and Dada, M.: Pricing and the Newsvendor Problem: A Review with Extensions, *Operations Research* **47** (1999), 183–194.

[4] Volf, P.: On quantile optimization problem with censored data. In:*Proceedings of the 31st International Conference on Mathematical Methods in Economics 2013*, College of Polytechnics, Jihlava, 2013, 1004–1009.

[5] Wikipedia: *Newsvendor model.* Retrieved 30.03.2016 from https://en.wikipedia.org/wiki/Newsvendor_model

Approximate Transition Density Estimation of the Stochastic Cusp Model

Jan Voříšek [1]

Abstract. Stochastic cusp model is defined by stochastic differential equation with cubic drift. Its stationary density allows for skewness, different tail shapes and bimodality. There are two stable equilibria in bimodality case and movement from one equilibrium to another is interpreted as a crash. Qualitative properties of the cusp model were employed to model crashes on financial markets, however, practical applications of the model employed the stationary distribution, which does not take into account the serial dependence between observations.

Because closed-form solution of the transition density is not known, one has to use approximate technique to estimate transition density. This paper extends approximate maximum likelihood method, which relies on the closed-form expansion of the transition density, to incorporate time-varying parameters of the drift function to be driven by market fundamentals. A measure to predict endogenous crashes of the model is proposed using transition density estimates.

Empirical example estimates Iceland Krona depreciation with respect to the British Pound in the year 2001 using differential of interbank interest rates as a market fundamental.

Keywords: multimodal distributions, stochastic cusp model, approximate transition density.

JEL classification: C46
AMS classification: 62F10

1 Introduction

Stationary density of stochastic cusp model belongs to the class of generalized normal distributions. Since it has four parameters, its flexible enough to allow for skewness, kurtosis and bimodality (Cobb et al. [4]). Modes of stationary density correspond to the stable equilibria of differential equation with cubic polynomial. In bimodality case there are two stable equilibria attracting the process and one unstable one between them, which repulses the process and corresponds to the antimode of stationary density. Movement from one stable equilibrium to another, which is defined by switching from one side of unstable equilibrium to another, can be viewed as an intrinsic crash in the system. This potentially large shift towards another stable equilibrium level is considerable advantage of the cusp model in describing certain systems over traditionally used mean reverting linear models, which has just one stable attracting equilibrium. Another advantage is a random walk behavior (under certain parametrization) in the middle of the domain, which was found appropriate for interest rate modeling by Aït-Sahalia [1].

However, complexity of cusp model brings some disadvantages, along with the lower research interest, which results into the lack of recent theoretical results. The major disadvantage lies in non-existence of closed-form solution of the transition density. There are several different approaches to overcome this obstacle: Euler approximation, simulation based methods, binomial approximations, numerical solution of Kolmogorow equations, and Hermite expansions. The last method, proposed by Aït-Sahalia [2] and [3], gives unlike the other methods closed-form approximation of the transition density, that converges

[1] Academy of Sciences of the Czech Republic, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, CZ-182 08 Prague 8, Czech Republic
Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University Prague, Sokolovská 83, CZ-186 75 Prague 8, Czech Republic, vorisek@karlin.mff.cuni.cz

to the true likelihood function. This methods allows to consistently estimate parameters of the diffusion processes.

Approach, which utilizes time-varying parameters, but only stationary density of stochastic cusp model were theoretically derived by Creedy and Martin [5] and empirically tested by the same authors [6] for the USD/GBP exchange rate. There were attempts to model currency crises by stationary cusp density (e.g. [7], [10]), some of them neglecting serial dependence completely, some reflecting it in variability of estimates. Goal of this paper is to provide an estimation technique, which reflects the time dependence, by extension of approximate transition density method to include exogenous variables, which drive parameters of the model. As an illustration example of proposed methodology serves Icelandic Krona deprecitation with respect to British Pound, given interest rate differential as an explanatory variable.

The rest of this paper is organized as follows. Section 2 recalls a method, which approximates transition density function. Estimation of stochastic cusp model with time-varying parameters and a measure to access potential of crash is introduced in Section 3. Section 4 presents an illustration example of proposed methodology using daily ISK/GBP exchange rate for the period from July 1, 1999 to December 31, 2004. Section 5 summarizes the results and concludes.

2 Approximate transition density function

Transition density of a general diffusion process

$$dX_t = \mu(X_t;\theta)dt + \sigma(X_t;\theta)dW_t, \tag{1}$$

does not have an analytical form. There are different techniques, how it can be approximated including Euler approximation, simulation-based methods or a numerical solution of forward Kolmogorov equation (see [8]). Aït-Sahalia ([3]) proposed a method based on Hermite polynomials to derive closed-form expansion for the transition density. Assume following *assumptions* :

1. The functions $\mu(x;\theta)$ and $\sigma(x;\theta)$ are infinitely differentiable in x, and three times continuously differentiable in θ, for all $x \in D_X = (-\infty, \infty)$ and $\theta \in \Theta$.
2. There exists a constant ζ such that $\sigma(x;\theta) > \zeta > 0$ for all $x \in D_X$ and $\theta \in \Theta$.
3. For all $\theta \in \Theta$, $\mu(x;\theta)$ and its derivatives with respect to x and θ have at most polynomial growth near the boundaries and
$$\lim_{x\to\pm\infty} -\frac{1}{2}\left(\mu^2(x;\theta) + \frac{\partial\mu(x;\theta)}{\partial x}\right) < \infty,$$

and a *transformation* of original process X into a new process Y as

$$Y \equiv \gamma(X;\theta) = \int^X \frac{1}{\sigma(u;\theta)}\mathrm{d}u.$$

By applying Itô's Lemma, Y has unit diffusion

$$\mathrm{d}Y_t = \mu_Y(Y_t;\theta)\mathrm{d}t + \mathrm{d}W_t,$$

where the drift is given by

$$\mu_Y(y;\theta) = \frac{\mu(\gamma^{-1}(y;\theta);\theta)}{\sigma(\gamma^{-1}(y;\theta);\theta)} - \frac{1}{2}\frac{\partial\sigma(x;\theta)}{\partial x}|_{x=\gamma^{-1}(y;\theta)}.$$

Than approximation of the log-transition density of y given initial values y_0 and a time step Δ is given by:

$$l_Y^{(K)}(y|y_0,\Delta;\theta) = -\frac{1}{2}\log(2\pi\Delta) - \frac{C^{(-1)}(y|y_0;\theta)}{\Delta} + \sum_{k=0}^{K} C^{(k)}(y|y_0;\theta)\frac{\Delta^k}{k!} \tag{2}$$

where K is order of expansion connected to the power of Δ and coefficients $C^{(k)}(y|y_0;\theta)$ can be calculated by substitution of proposed solution (2) into forward and backward Kolmogorov equations. Given $l_Y^{(K)}$, the expression for $l_X^{(K)}$ is given by the Jacobian formula

$$l_X^{(K)}(x|x_0,\Delta) = -\frac{1}{2}\log(2\pi\Delta\sigma^2(x)) - \frac{(\gamma(x)-\gamma(x_0))^2}{2\Delta} + \sum_{k=0}^{K} C^{(k)}(\gamma(x)|\gamma(x_0);\theta)\frac{\Delta^k}{k!}. \tag{3}$$

3 Stochastic model cusp

The univariate stochastic cusp model can be characterized by nonlinear diffusion process for the variable of interest

$$dX_t = \left(\alpha + \beta\,\frac{X_t-\lambda}{\sigma} - \left(\frac{X_t-\lambda}{\sigma}\right)^3\right)\frac{1}{2\sigma}dt + \sigma_w dW_t, \tag{4}$$

where W_t is standard Brownian motion, $\sigma > 0$ and $\sigma_w > 0$. Stationary density can be expressed analytically

$$f_s(x;\theta) = \eta_s(\theta)\exp\left[\alpha\,\frac{x-\lambda}{\sigma} + \frac{\beta}{2}\left(\frac{x-\lambda}{\sigma}\right)^2 - \frac{1}{4}\left(\frac{x-\lambda}{\sigma}\right)^4\right], \tag{5}$$

where $\eta_s(\theta)$ is normalizing constant.

For identifying bimodality of the stationary probability density function (5) serves statistic called Cardan's discriminant

$$\delta_C = \left(\frac{\alpha}{2}\right)^2 - \left(\frac{\beta}{3}\right)^3. \tag{6}$$

The parameters α (*asymetry*) and β (*bifurcation*) are invariant with respect to changes in λ (*location*) and σ (*scale*), as is δ_C, and they have following approximate interpretations [4]. If $\delta_C \geq 0$ then α measures skewness and β kurtosis, while $\delta_C < 0$ then α indicates the relative height of the two modes and β their relative separations.

3.1 Time-varying parameters

Following for example Creedy and Martin [6] or Fernandes [7], one can allow parameters α and β from (4) to be time varying:

$$dx_t = \left(\alpha(\xi_t) + \beta(\xi_t)\,\frac{x_t-\lambda}{\sigma} - \left(\frac{x_t-\lambda}{\sigma}\right)^3\right)\frac{1}{2\sigma}dt + \sigma_w dW_t, \tag{7}$$

where ξ_t is a vector of market fundamentals that are strictly exogenous with respect to x_t. Using approximate transition density (3) log-likelihood of the observed values x_i, ξ_i is given by:

$$\begin{aligned} ll^{(2)} = \sum_{i=1}^{n-1}\Bigg[&-\frac{(z_{i+1}-z_i)^2}{2\Delta_\sigma} + \frac{\alpha_i}{2}\frac{z_{i+1}-z_i}{\sigma_w} + \frac{\beta_i}{4}\frac{z_{i+1}^2-z_i^2}{\sigma_w} - \frac{z_{i+1}^4-z_i^4}{8\sigma_w} + \\ &+\frac{\Delta_\sigma}{\sigma_w^2}\left(-\frac{\alpha_i^2}{8} + \frac{\alpha_i}{16}\left(\kappa_i^3 - 2\beta\kappa_i\right) - \frac{\beta_i^2}{24}\kappa_i^2 + \frac{\beta_i}{20}\left(-5\sigma_w + \kappa_i^4\right) + \frac{\sigma_w}{4}\kappa_i^2 - \frac{1}{56}\kappa_i^6\right) + \\ &+\frac{\Delta_\sigma^2}{\sigma_w^2}\left(\frac{\sigma_w}{8} - \frac{\beta_i^2}{48} + \frac{\beta_i}{40}\kappa_i^2(3,4) + \frac{\alpha_i}{16}\kappa_i - \frac{1}{112}\kappa_i^4(5,8,9)\right) - \frac{1}{2}\log(2\pi\Delta\sigma_w^2)\Bigg] \end{aligned} \tag{8}$$

where $\Delta_\sigma = \Delta/\sigma^2$, $\alpha_i = \alpha_0 + \alpha_1\xi_i$, $\beta_i = \beta_0 + \beta_1\xi_i$, $z_i = (x_i/\sigma_w - \lambda)/\sigma$ and $\kappa_i^j(c) = \sum_{k=0}^{j} c_k\ z_{i+1}^k z_i^{j-k}$.

By differentiation of log-likelihood (8) with respect to α_0 and β_0 and laying down to zero, one can express these parameters wrt. other parameters

$$\hat{\alpha}_0 = \frac{2c_{13}^0c_{21}^0 - c_{12}^0c_{22}^0 + 2\beta_1(c_{22}^1c_{13}^0 - c_{22}^0c_{13}^1) + \alpha_1(4c_{13}^0c_{31}^1 - c_{22}^0c_{22}^1)}{(c_{22}^0)^2 - 4c_{13}^0c_{31}^0}, \tag{9}$$

$$\hat{\beta}_0 = \frac{2c_{12}^0c_{31}^0 - c_{21}^0c_{22}^0 + 2\alpha_1(c_{22}^1c_{31}^0 - c_{22}^0c_{31}^1) + \beta_1(4c_{31}^0c_{13}^1 - c_{22}^0c_{22}^1)}{(c_{22}^0)^2 - 4c_{13}^0c_{31}^0}, \tag{10}$$

where coefficients c_{ij} are as follows:

$$\begin{aligned}
c_{12}^k &= \sum_{i=1}^{n-1} \xi_i^k \left(\tfrac{1}{4}(z_{i+1}^2 - z_i^2) + \Delta_\sigma\left(-\tfrac{1}{4} + \tfrac{1}{20}\kappa_i^4\right) + \Delta_\sigma^2 \tfrac{\kappa_i^2(3,4)}{40}\right),\\
c_{13}^k &= \sum_{i=1}^{n-1} \xi_i^k \left(-\Delta_\sigma \tfrac{\kappa_i^2}{24} - \Delta_\sigma^2 \tfrac{1}{48}\right),\\
c_{21}^k &= \sum_{i=1}^{n-1} \xi_i^k \left(\tfrac{1}{2}(z_{i+1} - z_i) + \Delta_\sigma \tfrac{\kappa_i^3}{16} + \Delta_\sigma^2 \tfrac{\kappa_i}{16}\right),\\
c_{22}^k &= -\sum_{i=1}^{n-1} \xi_i^k \Delta_\sigma \tfrac{\kappa_i}{8},\\
c_{31}^k &= -\sum_{i=1}^{n-1} \xi_i^k \Delta_\sigma \tfrac{1}{8}.
\end{aligned}$$

Besides, one can express parameters α_1 and β_1 in a similar way to further reduce parameter space, which means that only parameters λ, σ and σ_w^2 needs to be estimated numerically. Employing of the transition density estimation allows to estimate variance of dW_t in contrast with stationary density estimation.

3.2 Probability of crash

Denote probability of extreme event by

$$\pi(x|x_0, \Delta; \theta, c) = \int_{-\infty}^{c} \exp[\ l(x|x_0, \Delta; \theta)]dx, \tag{11}$$

which for c corresponding to the antimode of the stationary density in bimodality case represents probability of switching to lower stable equilibria from starting point $x_0 > c$. If Cardan's discriminant (6) is lower then zero, than antimode c_a can be calculated as:

$$c_a = -2\ \mathrm{sign}(\alpha)\sqrt{\frac{\beta}{3}}\ \cos\left(\frac{1}{3}\arctan\left(\frac{2}{\mathrm{abs}(\alpha)}\sqrt{-\frac{\alpha^2}{4} + \frac{\beta^3}{27}}\right) + \arctan\left(\sqrt{3}\right)\right). \tag{12}$$

In general, $\pi(x|x_0, \Delta; \theta, c)$ can be used with an arbitrary c, when it represents standard risk measure approach. It measures large changes in unimodality case as well, however they occur exclusively due to the stochasticity of the process.

4 Empirical example

Following [10] and [7], the example for illustration of proposed methodology involves daily observed exchange rate between Icelandic Krona and British Pound in the period from July 1, 1999 to December 31, 2004, where as a market fundamental is used differential of one month interbank interest rate of the corresponding country[1]. The basic statistical characteristic is presented in Table 1. The interest rate differential is a commonly used macroeconomic indicator of the soundness of the banking sector as high interest rate differentials often precede increases in nonperforming loans (see [9]).

[1]Data from the Central Bank of Iceland and Bank of England.

	ISK/GBP	r^I/r^B
mean	128.97	1.0397
median	127.55	1.0421
std. dev.	9.94	0.0222
skewness	0.66	0.2694
kurtosis	2.71	1.8909
minimum	112.75	1.0076
maximum	156.3	1.0886

Table 1 Descriptive statistics of exchange rate and interest rate differential

Keeping the modeled quantities as usual (e.g. [6], [10]), we will use the logarithm of the exchange rate x_t and logarithm of the interest rate differential ξ_t, where variable parameters of drift (7) are given by $\alpha_t = \alpha_0 + \alpha_1 \xi_{t-1}$ and $\beta_t = \beta_0 + \beta_1 \xi_{t-1}$. Table 2 discloses the results of numerical maximum likelihood estimation of the diffusion process. It reveals that the stationary and transition estimates of parameters are similar, but the errors of estimates are remarkably higher for transition density estimation. There are two possible explanation for that: firstly the non-normality of estimates, but this holds for stationary density estimation as well, second explanation holds just for the transition model, where the last observation of exchange rate explains substantial part of distribution of next observation. However, Figure 1 shows, that the switching probability (11) indicated possibility of depreciation of the Krona in the first quarter of 2001, before the real depreciation came. Except of that Figure 1 displays evolution of Cardan's discriminant calculated according to both stationary and transition estimates, together with scaled exchange rate, interest rate differential and probability of adverse equilibria in a one-step ahead forecast.

parameter	stationary est.	stat. std. error	transition est.	trans. std. error
α_0	-0.529	0.054	-0.525	1.273
α_1	0.615	0.042	1.047	0.933
β_0	0.021	0.173	-0.054	1.562
β_1	2.312	0.099	2.018	1.328
λ	0.046	0.025	0.031	0.268
σ	0.906	0.018	0.755	0.105
σ_w			1.207	0.023

Table 2 Estimation results for stationary and transition density

5 Conclusion

This paper extends approximate maximum likelihood approach to the transition density estimation by time-varying parameters of the stochastic cusp model. It shows how to simplify the estimation and the measure for indicating endogenous crash of the system is introduced.

The example concerning Iceland Krona and British Pound exchange rate is used to illustrate the performance of proposed methodology, where interest rate differential serves as a market fundamental. Comparison with the stationary density approach reveals the behavior of the exchange rate close to the random walk, but the measure of crash was able to indicate the Krona depreciation in advance.

Acknowledgements

Supported by the grant No. P402/12/G097 of the Czech Science Foundations.

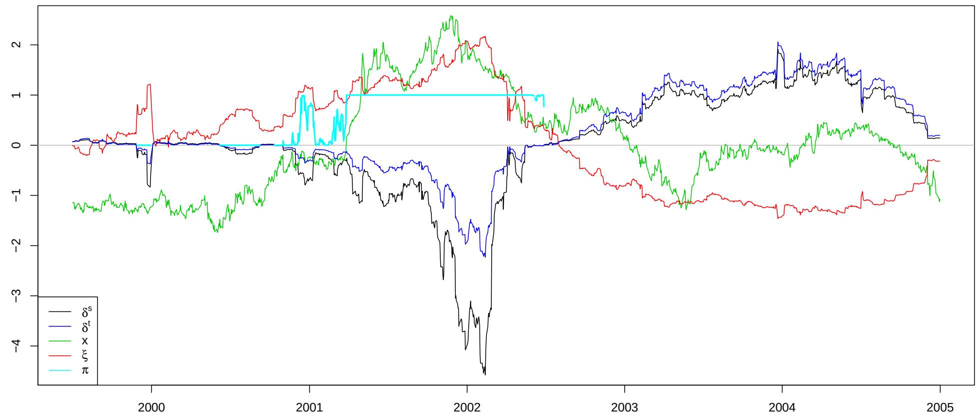


Figure 1 Stationary and transition evolution of Cardan's discriminant, exchange rate, interest rate differential and probability of adverse equilibria

References

[1] Aït-Sahalia, Y.: Testing continuous-time models of the spot interest rate. *Review of Financial Studies* **9** (1996), 385-426.

[2] Aït-Sahalia, Y.: Maximum Likelihood Estimation of Discretely Sampled Diffusion: A Closed-form Approximation Approach. *Econometrica* **70**, 1 (2002), 223–262.

[3] Aït-Sahalia, Y.: Closed-form Likelihood Expansions for Multivariate Diffusions. *The Annals of Statistics* **36**, 2 (2008), 906–937.

[4] Cobb, L., Koppstein, P., and Chen, N. H. : Estimation and Moment Recursion Relations for Multimodal Distributions of the Exponential Family. *Journal of the American Statistical Association* **78**, 381 (1983), 124–130.

[5] Creedy, J., and Martin, V. : Multiple Equilibria and Hysteresis in Simple Exchange Models. *Economic Modelling* **10**, 4 (1993), 339–347.

[6] Creedy, J., and Martin, V. : A Non-Linear Model of the Real US/UK Exchange Rate. *Journal of Applied Econometrics* **11**, 6 (1996), 669–686.

[7] Fernandes, M.: Financial crashes as endogenous jumps: estimation, testing and forecasting. *Journal of Economic Dynamics and Control* **30**, 1 (2006), 111–141.

[8] Jensen, B. and Poulsen, R. : Transition densities of diffusion processes: Numerical comparison of approximation techniques. *Journal of Derivatives* **9** (2002), 18-32.

[9] Kaminsky, G. L. and Reinhart, C : The twin crises: causes of banking and balance-of-payments crises. *American Economic Review* **89** (1996), 473–500.

[10] Koh, S. K., Fong, W. M., Chan, F. : A Cardans discriminant approach to predicting currency crashes. *Journal of International Money and Finance* **26** (2007), 131–148.

The Impact of Labor Taxation on the Economic Growth

Iva Vybíralová[1]

Abstract. Problems of taxation are not interesting just for economists, but also for common people, because taxation has influence on their lives in many ways. The aim of this article is to evaluate the impact of direct taxation on economic growth. Especially the impact of the tax burden on labour, represented by personal income tax and social security contributions. In this article, both the indicator of tax quota and the alternative indicator of tax burden (WTI) are used. In this work, a dynamic panel data analysis is employed. In this paper, the negative effect of labour taxation is confirmed. The negative influence of total personal income tax and social security contributions on economic growth is significant. In the case of model with tax quota, the positive effect of corporate taxation is verified. It could be caused by using the tax quota as an indicator of tax burden. In model with alternative WTI indicator, the negative effect of corporate taxation on economic growth is proved. The EViews 7 software is used.

Keywords: taxation, tax quota, WTI, dynamic panel data analysis, economic growth.

JEL Classification: H20, O40, C50
AMS Classification: 91G70

1 Taxation and the Economic Growth

This article is focused on the problem of labor taxation. The aim of this paper is to evaluate the effects of taxation on economic growth and provide recommentation for the tax policy makers who is dealing with taxation. This paper is worked with hypothesis of the negative impact of labor taxation to the economic growth. Two different types of tax burden are used. It means the tax quota, WTI and their subcomponents. It creates possibility to distinguish the burden of different types of taxes.

The issue of taxation of labor is an actual topic about every current period. Policy makers should be able to estimate level of taxation which is not harmful for state budget and which is not harmful for productivity of individuals due to excessive taxation. Authors Kessler and Norton [15] based their laboratory experiment on the idea that economic agents perceived taxation very negatively. In the experiment, they reduced wages of subject by two different ways. In the firs way, authors have reduced wages by a certain amount. In the second case, authors have reduced wages by taxation. On both cases the reduction of the same size. This experiment was confirmed the negative influence of taxation on productivity of individuals. Taxation had negative influence on disposable income of individuals. This influence determines their economic behavior. With the decline in disposable income can be expected restrictions on investment in education. This investments is in endogenous growth theory considered the engine of economic growth.

This work is based on endogenous growth theory. Romer [19] and Lucas [17] can be considered like main representatives authors of endogenous growth theory. They expanded the original neoclassical models by human capital and knowledge in their work. In endogenous growth models is permitted the opportunity on influence the growth of the economy due to taxes. Taxes had influence on growth variables. In endogenous growth models is permitted the opportunity on influence the growth of the economy in the form of tax effects on growth variables also after reaching a steady state. As reported Johansson, Head, Arnold, Brys and Vatria [10], taxes has influence on household decisions about the amount of savings, their labor supply, but also taxes is determining the amount of investment in human capital. In the paper of Leibfritz, Thornton and Bibbee [16] was stated, that the negative impact of taxation on growth lies in reducing savings and investments because of high taxation.

[1] Technical University of Ostrava, Faculty of Economics, Sokolská tř. 33, 721 00 Ostrava, Czech Republic, iva.vybiralova@vsb.cz.

1.1 Literature review

Many authors is devoted to issues of taxation and economic growth in their work. The effect of taxation on economic growth examined complexly for instance King and Rebelo [13]. According to their results, effect of taxation on the economic growth depend on the degree of openness of economies. Taxes has the greatest impact on economic growth in small open economies. The results of authors Easterly and Rebelo [7] did not confirmed a significant relationship between taxes and economic growth. They found negative effect of taxation on economic growth only in the case of personal taxes. Also Myles [18] does not consider the relationship between taxes and economic growth so evident. Authors who also confirmed the negative effect of taxation on economic growth included Karras and Furceri [11], Kotlán, Machová and Janíčková [15], Izák [9] or Romer and Romer [20]. The structure of taxation is also important as indicated works of Easterly and Rebelo [10], Widmalm [24] or Myles [21].According to the results of the authors Gemmell, Kneller and Sanz [8], taxes have influence on the gross domestic product (GDP), mainly through productivity factor, not through factor of accumulation.

Arnold [3], in his study of 21 OECD countries over the years 1971-2004, realized the negative impact of personal income tax on the amount of product. It also depends on the level of progressivity this tax. Lower output of the economy could be assumed with the higher progressivity. According to the result of the authors Acosta-Ormaechea and Yoo [1], the most damaging taxes for economic growth is personal income tax and social security contributions. The negative impact of personal income tax confirmed in their work also Johansson, Head, Arnold, Brys and Vatria [10] or Kotlán and Machová [14]. Dackehag and Hansson [5] admitted positive impact of personal income taxes and corporate taxes. But only on the condition that both these taxes are low. In another case, confirm the negative impact of taxation on economic growth. Daveri and Tabellini [6] odserved the relationship between unemployment and economic growth. According to their results, labor costs of companies were increased due to a high tax burden on labor. The consequence was not only higher unemployment but also reduce investment capital. Also Alesina, Ardagna, and Schiantarelli Perotti [2] in their work investigate to increase of income tax or social security. They examined the impact of this growth on labor costs and hence on the amount of future investments.

1.2 The data used and methodology

The models were estimated using generalized method of moments (GMM). Arellando and Bond estimator was used. The models were estimated in the software Eviews 7. As indicated Baltagi [4], the essence of dynamic panel is delayed of dependent variable and included it in the independent variable. The following models were dependent variable delayed of one season. Robust estimator “White Period” was used. The results of standard deviations did not include autocorrelations and heteroscedasticity due to this estimator [22].

It is important to observe the homogeneity of data. Results of the panel regression as accurate as possible due to homogeneity of data. For the study was chosen group of OECD countries. The OECD associating the most advanced countries of the world. The global databases were used in this work where can be assumed that the data collection and subsequent processing will support sample homogeneity. In this paper data from databases OECD, WB and from the pages of alternative indicators WTI were used. Variable GDP from the database WB Databank[2] as a variable of economic growth was set. Variable GDP is counted as GDP per capita expressed in constant 2005 prices in US dollars. The OECD.Stat[3] database was used for approximates the physical capital. Physical capital is represented in the following models as indicator of gross fixed capital expressed as a percentage of GDP. Human capital indicator is used from the OECDiLibrary[4] database. Indicator consists of several component parts, which are divided into three main components that relate to investment in human capital, its quality and learning outcomes. Taxes variables were taken from two sources. OECD.Stat[5] database was used in the case of model with tax quota. And in case of construction of a model with The alternative indicator of the tax burden was obtained from the WTI[6] database.

Indicators of tax rate as a tax variables were chosen. These indicators were chosen due to their simplicity and easy accessibility of data. It is the proportion of tax revenue of taxes to GDP. Tax quota as an inappropriate indicator for the tax burden was presented by Kotlán and Machová [14] and Kotlán, Machová and Janíčková [15]. Therefore, the alternative indicator WTI was used in this paper. WTI is a multi-criteria indicator of the tax burden. This indicator is combined hard and soft data. Tax revenue of taxes on income, profits and capital gains as % of GDP (TQ1200) as indicator of corporate taxation was chosen and index of corporate income tax (CIT) as an alternative. Excise taxes are represented in model with tax quota as tax revenue of general taxes on goods and

[2] http://data.worldbank.org/indicator/NY.GDP.PCAP.KD.
[3] https://stats.oecd.org/Index.aspx?DataSetCode=NAAG#.
[4] http://stats.oecd.org/BrandedView.aspx?oecd_bv_id=eo-data-en&doi=data-00717-en.
[5] https://stats.oecd.org/Index.aspx?DataSetCode=REV#.
[6] http://www.worldtaxindex.com/wti-dataset/.

services as % of GDP (TQ5110). Indicator of value added tax (VAT) was used in the case of the model with the WTI index. Also is this type of taxes represented by tax revenues of taxes on specific goods and services as % of GDP (TQ5120) and an index of other taxes on consumption (OTC) as alternative in case of a model with WTI index were chosen. Tax burden on labor in the model with tax quota was represented by the sum of tax revenue of taxes on income, profits and capital revenue and social security contributions as % of GDP (TQ1100_2000). Model with alternative indicator was used index of personal income tax (PIT).

There was made basic descriptive statistics. As revealed by the results, the highest tax burden measured by tax quota was found at taxation of labor. In the period under review (1995 – 2013), maximum of tax revenue as % of GDP was in 1997 in Sweden. Tax burden was gradually decreased over time. Highest average taxation exceeding 25% of tax revenues from taxes on income, profits and capital revenue and social security contributions (TQ1100_2000) was in Belgium, Sweden, Denmark and Finland. The lowest values (below 10%) were reached by Korea and Turkey. When the time series was shorted for five years (2000 – 2013), the results were very similar. Even in this period, the country with the lowest tax burden on labor were Korea and Turkey. There were changed the country rankings in the case of the highest taxation.

The average tax burden was the highest in labor taxes (PIT) in data sample of WTI. From the perspective of alternative indicator WTI was the highest value detected in the tax burden of value added tax (VAT), which is in models with tax quota represented by variable TQ5110.

With regard to the availability of data, models was constructed with data from 1995 until 2013. In the case of the model with the WTI index, data are available only since 2000. Model with tax quota was also constructed for the shorter period to allow comparison with model with WTI.

Econometric program Eviews 7 was used. Levin, Lin & Chu, Im, Pesaran and Shin and ADF - Fisher Chi-square test for testing stationarity of time series were used. [4]. First differences of time series (D1) was used in the case of nonstationary data. For fiscal variables OTC had to be used second differences (D2) of time series, since even after the first differences were not stationary time series. In order to interpret the results, all time series were logarithm (L). Empiric analysis was performed for 30 OECD countries.

2 Panel regression

Aim of this study is to evaluate the effect of the tax burden on economic growth. Estimated models was based on the endogenous growth theory. Which is expected effect of tax variables on individual growth variables.

Empirical specification constructed model may be in the form of equation (1):

$$HDP_{it} = \alpha HDP_{it-1} + \sum_{t=1}^{T} \beta_j X_{it} + \sum_{t=1}^{T} \gamma_k TAX_{it} + \varepsilon_{it} \tag{1}$$

where $i=1,\dots,N$ is the number of countries, $t=1,\dots,N$ is the number of years, X_{it} is set of control growth variable, TAX_{it} is set of taxes variable, ε_{it} is a residual component.

In equation (2) was extended equation (1) for the model using the tax quota. It could be written as follows:

$$HDP_{it} = \alpha HDP_{it-1} + \beta_1 K_{it} + \beta_2 H_{it} + \gamma_1 TQ1200_{it} + \gamma_2 TQ5110_{it} + \gamma_3 TQ5120_{it} + \gamma_4 TQ1100\_2000_{it} + \varepsilon_{it} \tag{2}$$

where K is an indicator of physical capital, H is indicator of human capital, *TQ1200* is ax revenue of taxes on income, profits and capital gains as % of GDP, *TQ5110* is tax revenue of general taxes on goods and services as % of GDP, *TQ5120* is tax revenues of taxes on specific goods and services as % of GDP, *TQ1100_2000* is sum of tax revenue of taxes on income, profits and capital revenue and social security contributions as % of GDP

Tax variables using partial indicators WTI can be written as equation (3):

$$HDP_{it} = \alpha HDP_{it-1} + \beta_1 K_{it} + \beta_2 H_{it} + \gamma_1 CIT_{it} + \gamma_2 VAT_{it} + \gamma_3 OTC_{it} + \gamma_4 PIT_{it} + \varepsilon_{it} \tag{3}$$

where *CIT* is index of corporate income tax, *VAT* is index of value added tax, *OTC* is index of other taxes on consumption, *PIT* is index personal income tax.

As is apparent from the performed literature search, there was assumed positive effects of growth variables - physical and human capital to the economic. In contrast, there were expected negative impact of fiscal variables on economic growth. The empirical analysis was based on a dynamic panel model. The model was estimated by generalized method of moments (GMM). There were used variables in the first difference (D1). All series were tested for the stationarity by the above three tests. The estimated model was statistically significant and all explanatory variables were statistically significant. Only delayed the response variables (GDP) was statistically significant at the 10% significance level. Cross-correlation was used. If was variable representing the tax burden on labor (TQ1100_2000) delayed of one period, the statistical significance of individual variables was remained statistics important. Also delayed explained variable became statistically significant at the 1% level of significance. The results of this model are presented in table 1. There was confirmed the negative impact of the tax burden from labor to economic growth. Personal income tax and social security contributions had quantified the highest negative impact on economic growth. Consumption taxes and corporate taxes have less impact on economic growth than on taxation on labor. For taxes on income, profits and capital gains (TQ1200) representing corporate tax has not been confirmed negative impact on growth, as foreseen in theory.

Statistical Verification		**Economic Verification**	
Dependent Variable D1_HDP_L			
Variable	**Coefficient**	**Theory**	**Empirical evidence**
D1_HDP_L(-1)	-0,07(-3,74)***	+	-
D1_K_L	0,28(19,18) ***	+	+
D1_H_L	5,60(9,63) ***	+	+
D1_TQ1200_L	0,05(9,57) ***	-	+
D1_TQ5110_L	-0,03(-3,01) ***	-	-
D1_TQ5120_L	-0,06(-8,23) ***	-	-
D1_TQ1100_2000_L(-1)	-0,11(-9,07) ***	-	-
Number of observations	451		
Number of instruments	30		
J-statistic	27,21		

Source: own calculations.

Note: the level of significance is created by number of stars 1% (***), 5% (**) a 10% (*).

Table 1 Model with Tax Quota (1995-2013)

The results of the model with the tax quota with shortened period are shown in table 2. There was estimated model with first differences of individual variables (D1). This model was estimated for the shortened period 2000 – 2013 in order to allow mutual comparisons with model WTI. Model was estimated again with the first differences (D1). The results are very similar when was compared this model with the model of tax quota with longer period of time. Model as a whole was statistically significant if was variables representing the tax burden on labor (TQ1100_2000) delayed of one period. All explanatory variables and delayed explained variable was statistically significant at a probability level of 1%. Quantitatively highest impact on economic growth had again a variable representing the tax burden on labor and there was again found positive impact of corporate taxation on economic growth. The same results were also in work of authors Kotlán, Machová and Janíčková [15]. Unlike their results, impact of this taxation on economic growth was not so strong.

Time series in model with WTI index were not stationary. There were used first differences (D1). Variable OTC was not stationary in this case again. There was used second difference (D2) of this variable. The remaining variables were in the first difference (D1). Variables CIT, OTC and PIT were statistically insignificant, when model was estimated without delay.

If the variable PIT was delayed of one period (-1), this variable became statistically significant but impact on economic growth has been positive. Variable OTC stayed longer statistically insignificant. Therefore, this model was rejected. According to the results of the cross correlation being variable OTC delayed by one period (-1) and variable CIT for two periods (-2). Aside variable representing labor taxation (PIT), remaining variables were statistically significant. This model was rejected. It was designed several models with different length of delay ensued from the cross correlation. Model where the taxes variables were delayed by two periods (-2) was accepted. All tested variables were statistically significant at the 1% level of probability. For all the tax variables were found negative effect on economic growth. Quantitatively highest impact on economic growth, according to the results had corporate taxation. According to the results, taxation of labor represented by variable PIT had second highest negative impact on economic growth. Also in this model was confirmed the positive effect of physical and human capital for economic growth.

Dependent Variable	D1_HDP_L		
Variable	**Coefficient**	**Variable**	**Coefficient**
D1_HDP_L(-1)	-0,19(-8,31)***	D1_HDP_L(-1)	-0,10(-3,45)***
D1_K_L	0,32(45,88)***	D1_K_L	0,39(25,49) ***
D1_H_L	10,09(6,95)***	D1_H_L	17,22(8,19) ***
D1_TQ1200_L	0,05(6,58)***	D1_CIT_L(-2)	-0,06(-11,13) ***
D1_TQ5110_L	-0,10(-8,51)***	D1_VAT_L(-2)	-0,03(-3,01) ***
D1_TQ5120_L	-0,16(-11,26)***	D2_OTC_L(-2)	-0,03(-4,44) ***
D1_TQ1100_2000_L(-1)	-0,24(-25,78)***	D1_PIT_L(-2)	-0,05(-2,62) ***
Number of observations	330	Number of observations	270
Number of instruments	30	Number of instruments	30
J-statistic	27,36	J-statistic	27,21

Source: own calculations.

Note: the level of significance is created by number of stars 1% (***), 5% (**) a 10% (*).

Table 2 Model with tax quota and model with WTI

3 Conclusion

The effect of taxation on economic growth is not in the spotlight only for professionals, but also among the general. Every economic entity must pay taxes. There are two basic options - either the tax is deducted directly from economic entities income (which carry very negatively), or as part of the product price. This hidden tax treatment should be less harmful for the economic growth. This conclusion was supported in work of Arnold [3], for example. Also Kessler and Norton [12] confirmed the negative view of economic entities on direct taxation based on the experiment.

The aim of this study was to evaluate the impact of labor taxation on economic growth. From the results of the analysis could be confirmed the negative impact of labor taxation on economic growth. In constructed models have been confirmed negative correlation between the variables representing this taxation. There was also detected quantitatively strongest relationship between taxes on labor and economic growth.

The model results with the use of indicators of the tax quota was confirmed theory in the case of labor taxes and consumption taxes (in models represented by TQ5110 and TQ5120). In the case of corporate tax has not been confirmed negative impact on economic growth. According to the results of the model, the corporate taxation should support economic growth. The negative impact of the tax burden of labor to GDP should be reflected with one year delay by the results of the model.

In the case of model with shorted period (2000 – 2013), the results were very similar. Again in the case of corporate tax has not been confirmed negative impact on economic growth. Quantitative impact of this taxation on the economic growth was lowest. There was confirmed negative impact of consumption taxes on economic growth. Again, the strongest negative influence was had indicator TQ1100_2000. This was again delayed by one period. Although there has been a change in the data base, the results were the same. Model can be considered robust.

The last constructed model was constructed by using the alternative indicator WTI as tax variable. Unlike previous models, this model fully confirms the economic theory that taxation has a negative impact on economic growth. In this model was found negative influence of corporate taxation (CIT) on economic growth. Quantitatively was this effect highest. Individual tax variables were delayed for two seasons (-2). During this delay, all the explanatory variables were statistically significant. During the construction of this model had to be variable OTC applied in the second difference (D2).

The basic recommendation for the makers of tax policy is to not raise tax revenues through taxation of labor. Results of the analysis confirmed the negative impact of this taxation. It offers the possibility of transferring the tax revenues from direct taxes to indirect taxes, which do not have such a negative impact on growth.

In the case of corporate taxation remains the question of appropriateness tax rate as an indicator of the tax burden. According to the results of the models would be recommendation for the tax policy makers increase this tax because of the positive effect on growth. However, as seen from the above literature, this taxation should have the highest negative effect. Therefore, it offers a suitable use of alternative indicator WTI, which confirmed the theoretical results

References

[1] Acosta-Ormaecha, S., and Yoo, J.: Tax Composition and Growth: A Broad Cross-Country Perspective. *IMF Working Paper No. 257*. International Monetary Fund, 2012.

[2] Alesina, A., Ardagna, S., Perotti, R., and Schiantarelli, F.: Fiscal Policy, Profits, and Investment. *NBER Working Papers 7207*. National Bureau of Economic Research, 1999.

[3] Arnold, J.: Do Tax Structures Affect Aggregate Economic Growth?: Empirical Evidence from a Panel of OECD Countries. *Working Papers No. 643*. Organization for Economic Cooperation and Development, 2008.

[4] Baltagi, B. H.: *Econometric analysis of panel data*. John Wiley & Sons Ltd., Chichester, 2005.

[5] Dackehag, M., and Hansson, Å.: Taxation of Income and Economic Growth: An Empirical Analysis of 25 Rich OECD Countries. *Working Paper 2012:6*. Department of Economics, 2012.

[6] Daveri, F., and Tabellini, G.: Unemployment, Growth and Taxation in Industrial Countries. *Working Paper No. 122*. InnocenzoGasparini Institute for Economic Research, 1997.

[7] Easterly, W., and Rebelo, S.: Fiscal policy and economic growth: An empirical investigation. *Journal of Monetary Economics* **32** (1993), 417–458.

[8] Gemmell, N., Kneller, R., and Sanz, I.: The growth effects of tax rates in the OECD. *Canadian Journal of Economics* **47** (2013), 1217–1255.

[9] Izák, V.: Vliv vládních výdajů a daní na ekonomický růst (empirická analýza). *Politická ekonomie* **59** (2011), 147–163.

[10] Johanson, Å., Heady, Ch., Arnold, J., Brys, B., and Vartia, L.: Tax and Economic Growth. *Working Paper No. 620*. Organization for Economic Cooperation and Development, 2008.

[11] Karras, G., and Furceri, D.: Taxes and growth in Europe. *South-Eastern Europe Journal of Economics* **7** (2009), 181–204.

[12] Kessler, J. B., and Norton, M. I.: Tax Aversion in Labor Supply. *Journal of Economic Behavior and Organization*, 2015.

[13] King, R. G., and Rebelo, S.: Public policy and economic growth: developing neoclassical implications. *NBER Working Papers 3338*. National bureau of Economic Research, 1990.

[14] Kotlán, I., and Machová, Z.: Horizont daňové politiky v zemích OECD. *Politická ekonomie* **62** (2012), 161–173.

[15] Kotlán, I., Machová, Z., and Janíčková, L.: Vliv zdanění na dlouhodobý ekonomický růst. *Politická ekonomie* **59** (2011), 638–658.

[16] Leibfritz, W., Thornton, J., and Bibbee, A.: Taxation and Economic Performance. *OECD Economics Department Working Papers No. 176*. Organization for Economic Cooperation and Development, 1997.

[17] Lucas, R. E.: On the Mechanics of Economic Development. *Journal of Monetary Economics* **22** (1988), 3–42.

[18] Myles, G. D.: Taxation and Economic Growth. *Fiscal Studies* **21** (2000), 141–168.

[19] Romer, P.: Increasing Returns and Long-Run Growth. *The Journal of Political Economy* **94** (1986), 1002–1037.

[20] Romer, Ch. D., and Romer, D. H.: The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks. *American Economic Revie* **100** (2010), 763–780.

[21] Widmalm, F.: Tax structure and growth: Are some taxes better than others? *Public Choice* **107** (2001), 199–219.

[22] Wooldridge, J. M.: *Introductory Econometrics: A Modern Approach*. South–Western CENGAGE Learning, Mason, 2009.

The similarity of the European Union countries as regards the production of energy from renewable sources

Katarzyna Warzecha[1]

Abstract. Sudden development of civilization causes the increase of demand for electricity. The particular situation applies to the European Union countries where the European Parliament and the Council defined with the use of the appropriate Directive the general objectives concerning the shares of energy from renewable sources in gross final energy consumption for every country regarding their individual national targets by the year 2020. One of the aims of the present study is to demonstrate how the European Union countries comply with the Directive.

The main objective of the research is the evaluation of changes occurring in production of the energy from renewable sources and the division of the countries into groups with similar structure of production of energy from renewable sources (with the use of Czekanowski's Diagram). To visualize multidimensional data Chernoff faces were used. The sample period was the years 2005 and 2014. The statistical data came from the Eurostat website and the calculations were made with Statictica and MaCzek computer programs.

Keywords: renewable energy, Czekanowski's Diagram, Chernoff faces

JEL Classification: Q42
AMS Classification: 62H86

1 Introduction

The technical and technological progress influencing the dynamic development of civilization is connected with the growing demand for energy. The use of traditional non-renewable energy sources, that is fossil fuels such as: coal, oil and natural gas, causes the increase of the environmental pollution, but first and foremost it leads to exhaustion of natural resources. Therefore, the European Union emphasizes to a large extent the use of more beneficial sources of energy which are renewable. This type of energy became the key element of the European Union energy policy and in 2009 The Climate and Energy Package was adopted [4] and it contained the main goals of energy policy by 2020. One of the objectives is to achieve in the UE at least 20% of gross final energy consumption which comes from the renewable energy sources, as well as, to anticipate the 20% reduction of primary energy consumption in comparison to projected levels.

The main objective of the research is the evaluation of the changes in production of energy which comes from renewable sources and dividing the EU countries into groups which have similar structure of the energy production from renewable sources. The aim of this study is also the presentation how the EU countries comply with the Directive of the European Parliament and of the Council of 2009 in the field of the share of renewables in the final energy consumption which was established by European Parliament for every country regarding their individual national targets by the year 2020.

The Czekanowski's diagram was used to establish the similarities between the EU countries in regard to the production of energy which comes from the renewable sources. The division of the examined EU countries into groups of similar structure of energy production from renewable sources allows the answer to the question whether the geographical position and the climate have the significant influence on the production of energy from the specific sources. To visualize the multidimensional data the Chernoff faces were adopted. The sample period is the year 2005 and 2014, the statistical data come from Eurostat webpage [10]. Statistica and MaCzek are the computer programs which were used in calculation.

2 The production of energy from renewable sources in the final energy consumption in EU countries in 2005 and 2014

In the Climate and Energy Package it was stated that at least 20% of gross final energy consumption should be fulfilled by the energy which comes from the renewable sources. Each EU country's share is established on individual level. In 2014, there was a total of 16% share of energy from renewable sources in gross final energy consumption for the European Union (28 countries) and it stated the raise equal to 7

[1] University of Economics in Katowice, Department of Econometrics, warzecha@ue.katowice.pl.

percentage points in comparison to 2005. To achieve the objective intended by the European Parliament 4 percentage points more are required. This means that within the space of 4 years the share has to increase by 25% of the value of 2014.

The target of the Polish energy policy is the increase of shares of energy from renewable sources in gross final energy consumption to the level of 15% in 2020. In 2014, there was a total amount of 11.45% share of energy from renewable sources in gross final energy consumption for Poland and it stated the raise equal to 4.55 percentage points in comparison to 2005. To achieve the intended objective 3.55 percentage points more are required. This means that within the space of 4 years the share has to increase by over 30% of the value of 2014.

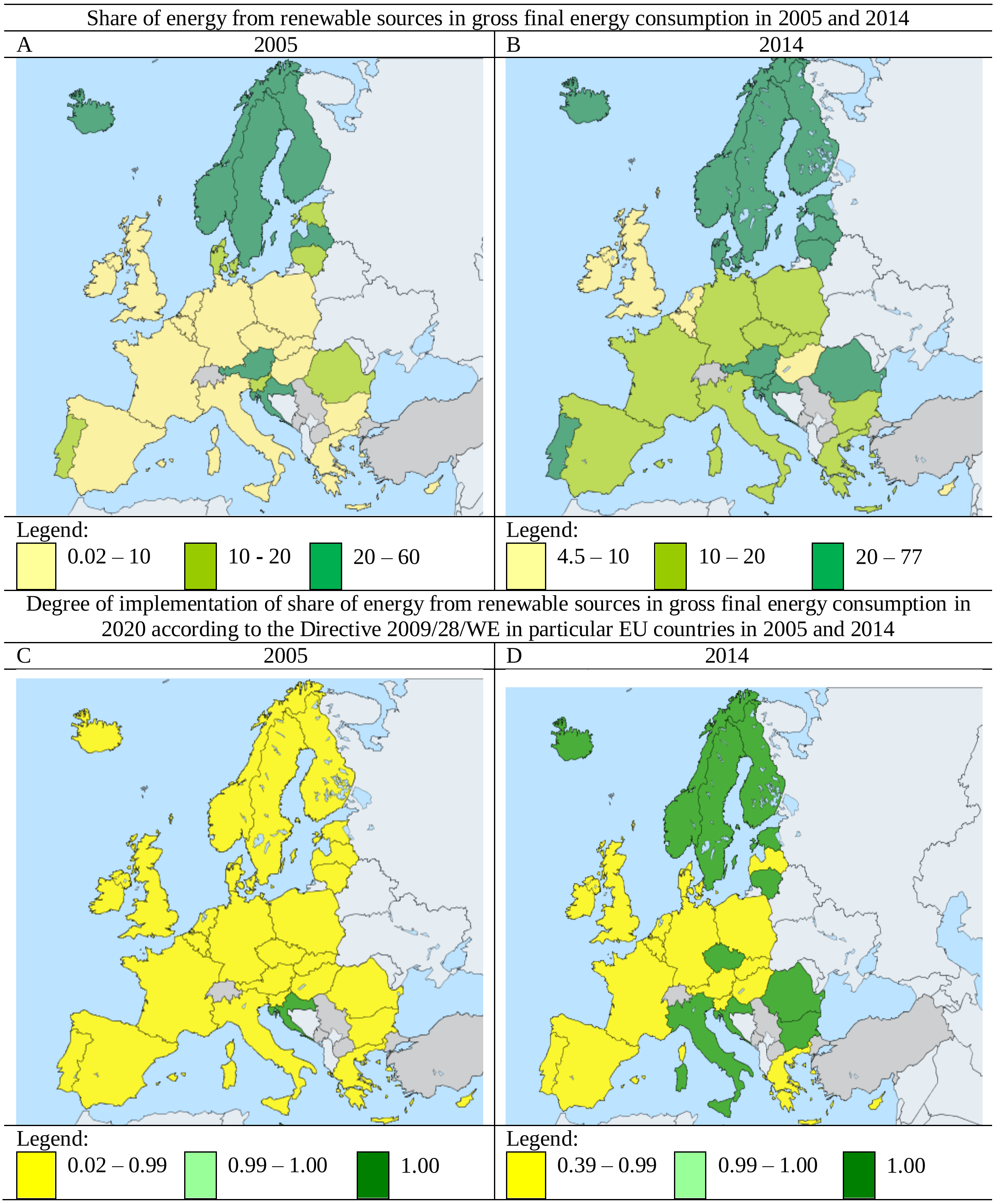


Figure 1. Share of energy from renewable sources in gross final energy consumption and the degree of implementation of this share in the EU countries in 2005 and 2014

The share of renewable energy in gross final energy consumption in 2005 and 2014 is presented on the maps (Figure 1 - A and B). The EU countries which are characterized by the lowest share of energy from renewable sources in gross final energy consumption (below the mean obtained in 28 EU countries which was equal to 9% in 2005 and 16% in 2014) are as follows: Ireland, the Netherlands, Luxembourg, Malta, Great Britain. From the data on maps A and B it is clearly visible in most countries that over the span of 10 years the share of renewable energy in gross final energy consumption increased which indicates that examined countries strive to obtain the objectives stated by the European Parliament for 2020.

The data demonstrated on maps C and D (Figure 1) present the results as follows:

- in 2005 Croatia was the only country, which achieved the aim stated in the EU Directive for the year 2020;
- in 2014 the aim stated in the EU Directive was achieved by more countries, i.e. Italy, Finland, Sweden, Estonia, the Czech Republic, Romania, Bulgaria and Lithuania.

3 The use of Czekanowski's diagram to establish the similarities between EU countries in regard to structure of production of energy which comes from the renewable sources

The structure of production of energy from renewable sources in the examined countries is diversified, which is a result of geographical conditions, and consequently the climatic conditions of a particular country and the capability of this country to use its resources. In EU-28 as a whole, in 2014 in comparison to 2005 (Figure 2) there was a substantial decrease in share of solid biofuels in the total production of energy from renewable sources. The decrease was from 55.5% to 44.56% (about 20% decrease). There was also the decrease of hydropower in the total production of energy from renewable sources. Here, the reduction was from 21.18% to 16.02% (about 25% decrease). Over the span of 10 years, there was a significant increase of wind energy share in the total production of energy from renewable sources from 4.77% to 10.81% (about 130% increase), the rise of solar energy in the total production of energy from renewable sources from 0.55% to 2.02% (about 270% increase). The rise can be noticed as well in other energy carriers, which are as follows: municipal waste, bio gasoline, biodiesel and the tidal energy.

2005	2014

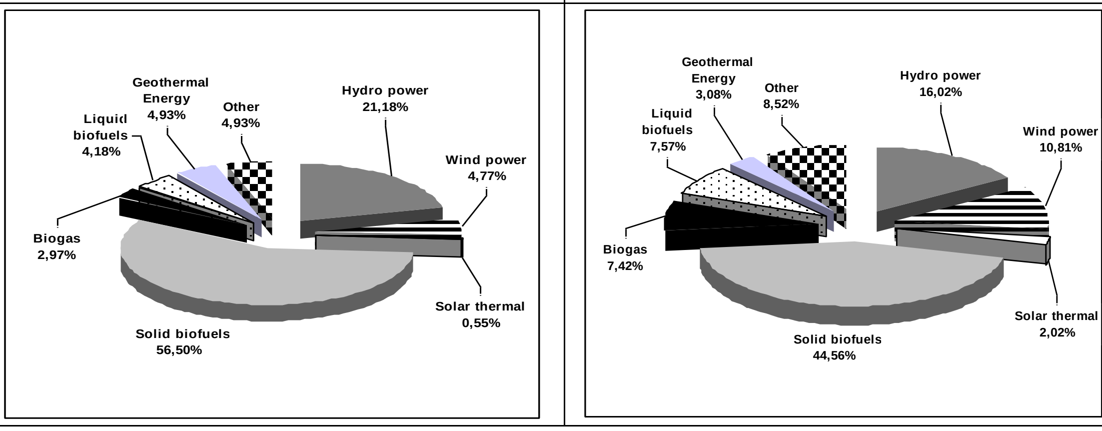


Figure 2. Share of the various energy carriers in production of energy from renewable sources in EU-28 in 2005 and 2014

To discover the similarities and differences between the EU countries as regards the structure of production of energy from renewable sources the Czekanowski's method was applied and it was a method primarily proposed by the Author for the anthropology demand [2, 3]. The description of Czekanowski's method can be found in: [7; 6; 8; 9]. The calculations were conducted to establish the similarities of the EU countries and they applied to shares of electricity produced from the renewable energy sources which come from various points of supplies (such as water, wind, solar energy, solid biofuels, biogas, geothermal energy and others) in the production of electricity from renewable energy carriers. According to the rules of Czekanowski's method all of the examined variables were transformed, all of the variables were stimulants (i.e. variable (X) is a stimulant if, the higher is the value of this variable the better is the situation of a particular object as regards the examined point of view). Using the zero unitarization method the variables were normalized according to the formula (1):

$$z_{ij} = \frac{x_{ij} - \min x_{ij}}{\max x_{ij} - \min x_{ij}} \qquad (i=1,2,\ldots,n; \;\; j=1,2,\ldots,m) \qquad (1)$$

where: $\min x_{ij}$ - min of variable x_j ; $\max x_{ij}$ – max of variable x_j;
z_{ij} –standardized value of variable x_j for i- object.

The countries were considered to be similar if they were situated close to each other in the multidimensional space of explanatory variables. The distance between the examined countries in multidimensional space was defined with the use of Euclidean distance formula (2):

$$d_{il} = \sqrt{\sum_{j=1}^{m} \left(z_{ij} - z_{lj}\right)^2} \qquad (i=1,2,\ldots,n;\ \ j=1,2,\ldots,m) \qquad (2)$$

where: d_{il} – the distance between i-object and l-object;
z_{ij} – standardized value of variable for i-object;
z_{lj} – standardized value of variable for l-object.

The graphic image of the discussed method is Czekanowski's diagram which allows the division of examined EU countries for the groups which are characterized by a high level of similarity. The diagrams were constructed with the use of computer program MaCzek [11].

In the case of the partially corresponding countries the division was made according to the shortest mean Euclidean distance of inconclusively defined country and other countries from possible groups of its affiliation.

Particular columns and rows of the Czekanowski's diagram correspond to a particular EU country. The bigger the symbol on the crossing of the column and row the more similarity is attributed to countries as regards the shares of production of electricity from the renewable energy sources which come from various points of supplies in the general production of energy from the renewable energy carriers. The most similar objects are situated as close as possible to the main diagonal.

In regard to the limitation of the number of article's pages, the Czekanowski's diagram presented below corresponds to the most current data- the data from 2014.

Czekanowski's diagram and the EU countries in the division for the groups of similar countries as regards the shares of production of electricity from the renewable energy sources in 2014.

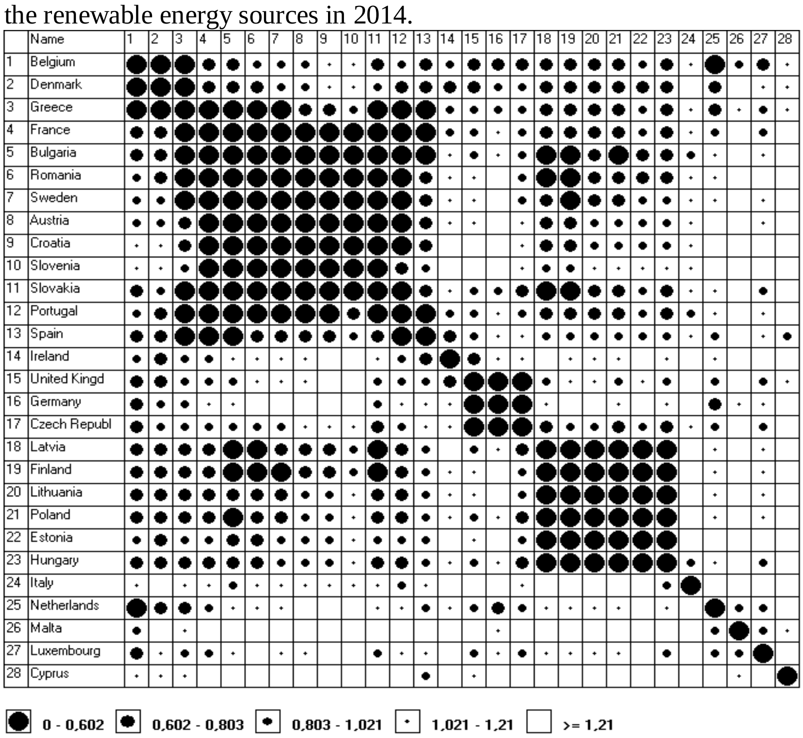


Figure 3. The similarity of the EU countries as regards the structure of production of energy from renewable sources

From the data in Figure 3 it is visible that:

- group I consists of the countries (Belgium, Denmark, Greece) which are characterized by the great share of the energy production from solid biofuels (on average about 48% of the renewable energy production- which is below the mean in EU) and the great share of the energy production from wind power (significantly above the EU mean- on average about 17% of energy production from renewable energy sources);
- group II is most numerous when it comes to the countries (France, Bulgaria, Romania, Sweden, Austria, Croatia, Slovenia, Slovakia) which are characterized by the great share of the energy production from water (significantly above the EU mean- on average about 31% of energy production from renewable energy

sources) and the share of the energy production from solid biofuels (on average about 51% of renewable energy production - slightly above the EU mean);

- group III belongs to the countries (Portugal, Spain) which are characterized by the great use of solar energy (significantly above the EU mean- on average about 7% of energy production from renewable sources), wind and water energy (both kinds of energy are significantly above the EU mean- on average about 22% of renewable energy production) and the geothermal energy as well (significantly above the EU mean- on average about 2% of energy production from renewable energy sources). The energy obtained from solid biofuels equals on average to 36% which is significantly below the EU mean.
- group IV consists of countries (Latvia, Finland, Lithuania, Poland, Estonia, Hungary) which are characterized by the great share of energy production from biogases (significantly above the EU mean- on average about 18% of energy production from renewable energy sources) and wind energy (above the EU mean- on average about 13% of renewable energy production). The energy from solid biofuels equals on average to 44% of energy production from renewable energy sources which is below the EU mean);
- group V consists of countries (Iceland, Italy, the Netherlands, Malta, Luxembourg and Cyprus do not create any groups) which use almost exclusively the solid biofuels (significantly above the EU mean- on average about 82% of energy production from renewable energy sources).

4 Chernoff faces as a way of visualization of countries as regards the structure of energy production from renewable energy sources

Figure 4 presents a very interesting way of visualization of multidimensional data which describe the structure of energy production from renewable energy sources, i.e. the Chernoff faces [1]. The similarity or dissimilarity of the multidimensional data is reflected here in picture facial expressions. The caption in Figure 4 indicates which face elements correspond to particular attributes of data set.

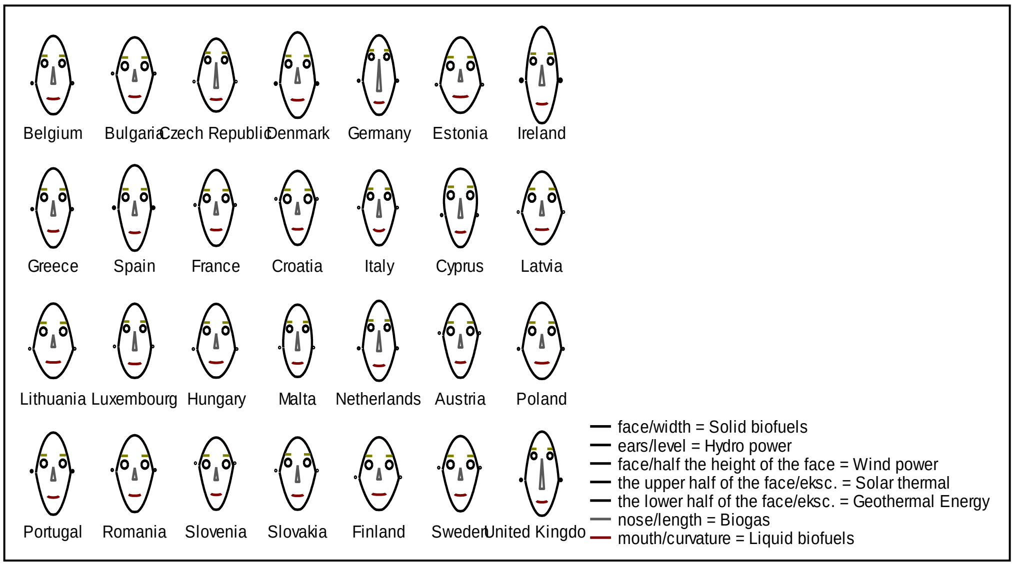


Figure 4. The picture chart in the form of so called Chernoff faces enabling the analysis of the multidimensional data which describe the structure of energy production from renewable energy sources in the EU countries in 2014

It results from the picture that there are two groups of countries with similar structure of energy production from the renewable sources as it was in Czekanowski's diagram. The definite similarity is observed in such countries as: Belgium, Denmark, Greece, Portugal and Spain; Great Britain and Germany; whereas the Czech Republic is not too similar in facial pictures to this group of countries which were indicated before in Czekanowski's diagram. The countries' faces which belong to other groups, set apart in Figure 3, are similar to each other as well. It is possible to find effortlessly the countries in the chart that are dissimilar to other countries on the basis of the presented face- the face in completely different size and different face expression, i.e. Malta and Ireland in this case. It can be stated then that the charts of this kind are the appropriate statistical tool which presents in easy, interesting and visual way the diversity of examined objects as regards the particular phenomenon.

5 Conclusions

The increase of energy production from the renewable sources of energy becomes a necessity taking into consideration the significant decrease in natural resources and the increase of environmental pollution. This particular situation applies to the European Union countries where the European Parliament and the Council defined with the use of the appropriate Directive the general objectives in the shares of energy from renewable sources in gross final energy consumption in 2020. It results from the conducted analyses that some of the countries (Croatia, Italy, Finland, Sweden, Estonia, the Czech Republic, Romania, Bulgaria and Lithuania) already achieved the aims stated in the EU Directive. However, such countries as: Ireland (in 2014 the share of energy from renewable sources in gross final energy consumption equaled to 13.8% where the aim for 2020 is 18%), the Netherlands (in 2014 the share of energy from renewable sources in gross final energy consumption equaled to 5.5% where the aim for 2020 is 14%), Luxembourg (in 2014 the share of energy from renewable sources in gross final energy consumption equaled to 4,5% where the aim for 2020 is 11%) or Malta (in 2014 the share of energy from renewable sources in gross final energy consumption equaled to 4.7% where the aim for 2020 is 10%) can have problems with fulfilling the requirements because their share of energy from renewable sources in gross final energy consumption is very low when compared to the requirements stated for 2020. In 2014 the share of the energy from renewable sources in gross final energy consumption in Poland equaled to 11.45% and it increased by 4.55 percentage points in comparison to 2005, thus to achieve the required aim it is necessary to obtain 3.55 percentage points more. About 3% below the required European Parliament aim in the scope of the share of energy from renewable sources in gross final energy consumption in 2020 can also be found in Denmark, Latvia and Austria.

The studies show that the shares of energy from renewable energy carriers which come from different points of supplies (solid biofuels, water, wind, biogas, solar and geothermal energy and others) differ from each other and they change in various periods.

The research demonstrated that the research tools used in the process (the Czekanowski's diagram and the picture chart in the form of so called Chernoff faces) can be used successfully to define the groups of countries similar to each other as regards the structure of energy production from renewable sources.

References

[1]. Chernoff H., The Use of faces to represent points in k-dimensional space graphically, *Journal of the American Statistical Association*, vol. 68, no. 324 (Jun., 1973), pp. 361-368.

[2]. Czekanowski J., *Zur Differenzionaldiagnose der Neandertallgruppe*, Korespondemz-Blatt der Deutsche Geselschaft fur Antropologie und Urgeschichte, Braunschweig, (1910), pp. 9-12.

[3]. Czekanowski J., *Zarys metod statystycznych w zastosowaniu do antropologii (Outline of statistical methods in application to anthropology)*, Prace Naukowego Towarzystwa Warszawskiego (The works of Warsaw Learned Society), no. 5 (1913).

[4]. Dyrektywa Parlamentu Europejskiego i rady Europy nr 2009/28/WE w sprawie promowania stosowania energii ze źródeł odnawialnych zmieniająca i w następstwie uchylająca dyrektywy 2001/77/WE oraz 2003/30/WE (Directive 2009/28/WE of the European Parliament and of the Council on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/WE and 2003/30/WE), (Act of 5 June 2009), Official Journal L 140/16.

[5]. Dziechciarz J., *Ekonometria. Metody, przykłady, zadania*. (*Econometrics. Methods, examples and tasks)*, Wydwanictwo Uniwersytetu Ekonomicznego we Wrocławiu (University of Economics in Wroclaw), Wrocław, (2012).

[6]. Heffner K., Gibas P., *Analiza ekonomiczno-przestrzenna* (*Economics and Spatial Analysis*), Wydawnictwo Akademii Ekonomicznej w Katowicach (University of Economics in Katowice), Katowice, (2007).

[7]. Pociecha J., Podolec B., Sokołowski A., Zając K., *Metody taksonomiczne w badaniach społeczno-ekonomicznych (Taxonomic methods in social and economic research)*, PWN, Warszawa, (1988).

[8]. Mielecka-Kubień Z., Warzecha K., *Sytuacja demograficzno-społeczna wybranych miast województwa śląskiego w latach 2002 i 2009* (The *socio-demographic situation of chosen cities in Śląskie Province in the years 2002 and 2009*), Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie (Scientific Notebooks of University of Economics in Cracow), 2013, no. 923, (2013), pp. 5-21

[9]. Sołtysiak A., Jaskulski P., *Czekanowski's diagram. A method of multidimensional clustering* , [w:] New Techniques for Old Times. CAA 98. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 26th Conference, Barcelona, March 1998, Chapter:, Publisher: BAR International Series 757, Editors: J.A. Barceló, I. Briz, A. Vila, (1998), pp.175-184

[10]. European Commission. *Eurostat - Share of energy from renewable sources in gross final energy consumption* [Online]. Available at: http://ec.europa.eu/eurostat/tgm/mapToolClosed.do?tab=map&init=1&plugin=1&language=en&pcode=t2020_31&toolbox=legend [accessed 28.04.2016]

[11.] Description of the program MaCzek [Online]. Available at: http://softadvice.informer.com/ext/www.antropologia.uw.edu.pl/MaCzek%2Fmaczek.html [accessed 28.04.2016]

The evaluation of factors influencing flight delay at popular destinations of Czech tourists

Martina Zámková[1], Martin Prokop[2]

Abstract. The main goal of this paper was assessment of factors influencing flights delay at selected international airports including Antalya, Burgas, Charles de Gaulle, Heraklion, Hurghada, Palma de Mallorca and Rhodes. Data of an Airline company, whose flights are mainly oriented on favourite destinations of Czech tourists, were used for needs of this article. Results from analysis showed that delays are the most frequent in Palma de Mallorca and the least often in Burgas airport. It was proved that delays caused by delay of preceding flight are the most common cause of delay at all. Only at Charles de Gaulle airport there was also found out that delay is also quite often caused by problems at the airport and in the flight traffic, by unusual events and also by airlines and suppliers. Delayed preceding flights cause mainly longer delay of connecting flight.

Contingency tables present an easy way to display the relationship between categorical data. We used the Pearson's chi-square test; the null hypothesis of the test assumes independence of variables. Correspondence analysis aims to reduce the multidimensional space of row and column profiles and to save the original data information.

Keywords: flights delay, correspondence analysis, contingency table, Pearson chi-squared test

JEL Classification: C30, R40
AMS Classification: 62H17, 62H25

1 Introduction

The main goal of this paper was assessment of factors influencing flights delay at selected international airports including Antalya, Burgas, Charles de Gaulle, Heraklion, Hurghada, Palma de Mallorca and Rhodes. We selected these airports because the airline that we monitor is concentrated mainly on these destinations. Based on the carried out dependency test in contingency tables and depicted correspondence maps, dependence of delay on various factors were examined. There were subsequently designed recommendations how to eliminate the causes of delay and how to mitigate not only the financial implications.

The article [2] deals with a similar problematics, in particular it analyses dependence of flights delay on destination airport. More details are offered in the article [4] by strategy for reduction of delays, for example setting scheduled times for completion of a process, increasing the number of service counters, and priority service for emergent flights. The economics including possible optimizing is analysed for example in an article [5]. An article [6] analyses a model describing optimization of subsequent flights in order to prevent spreading and chaining of delays. Similar analysis of factors influencing delays at Czech international airports was carried out in the article [7]. The article [8] is focused also on financial implications on airlines.

2 Methods and Materials

Contingency tables present an easy way to display the relationship between categorical data. We used the Pearson's chi-square test; the null hypothesis of the test assumes independence of variables. The condition that a maximum of 20% of the expected frequencies are less than five must be met; see [1]. Correspondence analysis that was used for this study is a multivariate statistical technique, which allows the display and summary of a set of data in two-dimensional graphic form. This analysis aims to reduce the multidimensional space of row and column profiles and to save the original data information as much as possible; see [3]. Unistat and Statistica software was used for primary data processing.

The primary data were examined in the busiest season for the selected airline company (1. 6. 2013 – 30. 9. 2013). Data contained information about length of delay and cause of delay according to IATA[3] codes, which

[1] College of Polytechnics Jihlava, Department of Mathematics, Tolstého 16, 58601 Jihlava, martina.zamkova@vspj.cz.
[2] College of Polytechnics Jihlava, Department of Mathematics, Tolstého 16, 58601 Jihlava, Martin.Prokop@vspj.cz.
[3] IATA – International Air Transport Association

were adjusted by the airline company according to its needs. Substantial part of data is categorical or suitable to categorization. Data was gathered from an internal database of the monitored airline.

3 Research results

It is evident from the frequency table (see Tab. 1) that delays occur most often at Palma de Mallorca airport where more than one half of departures are delayed. It is a tourist destination where the intensity of traffic is the biggest in the monitored time period and the airport is very busy. Over 30% of departures are delayed from Charles de Gaulle, Heraklion, Hurghada and Rhodes airports. Charles de Gaulle airport is a big international transit airport with an intensive year-round traffic. Heraklion, Hurghada and Rhodes airports are smaller airports with only one take-off runway and one terminal, therefore delays as consequence of insufficient capacity occur here. Burgas airport achieved the best result (approximately 23% delayed departures). The result is probably caused by the fact that the high season is only from July to August in this destination due to the geographical location. BOJ airport is minimally utilized in June and September and delays almost don't occur here.

Destination	Country	CODE	Delayed flights	All flights	Percentage
Antalya	Turkey	AYT	387	1324	29.23%
Burgas	Bulgaria	BOJ	177	785	22.55%
Charles de Gaulle	France	CDG	214	516	41.47%
Heraklion	Greece	HER	223	666	33.48%
Hurghada	Egypt	HRG	155	358	43.30%
Palma de Mallorca	Spain	PMI	228	412	55.34%
Rhodes	Greece	RHO	192	520	36.92%

Table 1 Relative frequencies – Selected airports and delayed flights

It is clear from the row relative frequencies table (see Tab. 2), that the biggest problems with delay at night are in Hurghada where more than half of flights are delayed. Worse situation is also at Charles de Gaulle airport where almost 30% of flights are delayed. On the contrary, at Heraklion, Palma de Mallorca and Rhodes airports less than 5% of departures are delayed at night. During the morning the differences are less significant, worse situation is at Burgas, Heraklion and Antalya airports (around 40%). On the contrary, at Rhodes, Hurghada a Charles de Gaulle airports it is only slightly over 10%. In the afternoon the differences are even smaller. The worst situation is at Burgas, Heraklion and Charles de Gaulle airports (always over 30%), at the rest of airports is the ratio of delayed departures 20-30%. In the evening the differences are significant again. At Rhodes and Palma de Mallorca airports is the rate of delayed flights more than 50%, in case of Heraklion and Antalya airports approximately one quarter of departures are delayed and at Burgas, Hurghada and Charles de Gaulle airports the ratio of delayed departures is only around 10%. Dependency is statistically significant; p-value is less than 0.0001.

Row relative frequencies	00:01-06:00	06:01-12:00	12:01-18:00	18:01-24:00
AYT	11.11%	37.73%	24.29%	26.87%
BOJ	10.17%	46.89%	33.90%	9.04%
CDG	28.97%	10.75%	48.60%	11.68%
HER	2.69%	37.67%	32.74%	26.91%
HRG	54.84%	12.26%	20.00%	12.90%
RHO	0.52%	11.98%	26.04%	61.46%
PMI	4.39%	20.18%	23.25%	52.19%

Table 2 Contingency table **–** Daytime and airport

From correspondence map (see Fig. 1) it is evident that the delayed departures at night are very frequent at Hurghada airport. The most frequent delays in the evening occur at Rhodes and Palma de Mallorca airports. Delays at Burgas airport occur most often in the morning.

It is furthermore evident from the row relative frequencies table (Tab. 3) that delays caused by problems at the airport and in the air traffic as well as caused by unusual events (transport of handicapped passengers and necessary healthcare for passengers with sudden healthy problems) are the most frequent at Charles de Gaulle airport (approximately 21% of delayed flights). It is the busiest airport, that's why more unusual events can occur there. On the contrary these causes of delay occur the least frequently at Heraklion, Burgas, Hurghada, Antalya and Rhodes airports (up to 6% of delayed flights). The best situation at all is at Burgas airport (only 2% of delayed flights). It is the least busy airport with minimum problems in the air traffic and minimum unusual events. Delays on airline's side (technical maintenance, defects, crew standards) and suppliers' side (handling, supply of fuel, catering) are the most frequent again at Charles do Gaulle airport (27% of delayed departures). These causes occur very rarely at other airports (up to 6% of delayed departures). It is even less than 1% of delayed departures at Heraklion, Hurghada and Rhodes airports. Aircraft have very short downtime at these airports, technical problems are solved only in necessary cases and therefore they mostly don't cause delays. Similarly catering is complemented mostly at home base airports. Delay caused by delay of preceding flight is the most common cause of delays at all. They occur most often at Antalya, Heraklion, Palma de Mallorca, Burgas, Hurghada and Rhodes airports, approximately 90% of all delays. Charter flights, which have higher probability of delays in total, fly mostly to these holidays' destinations. That is why chaining of delays originate here. This cause is less common only at Charles de Gaulle airport and includes 53% of all delayed flights. Mostly regular flights which are less often delayed fly to this airport and the chaining doesn't originate. The dependence is statistically significant; p-value is less than 0.0001. Similar result is visible also in the correspondence map; see Fig. 2, where all airports except are concentrated around delay caused by preceding delay. Charles de Gaulle and Palma de Mallorca airports are closer to other causes of delay.

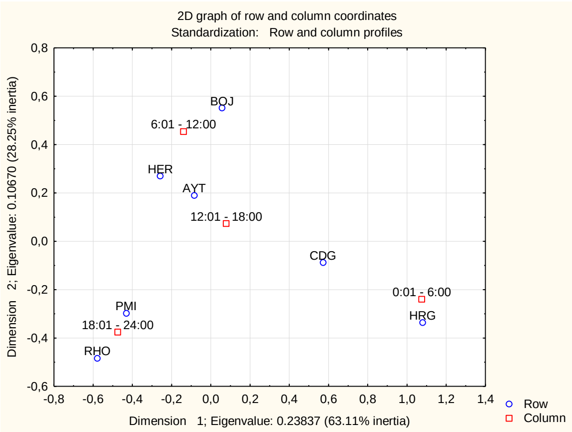


Figure 1 Correspondence map - Daytime and airport

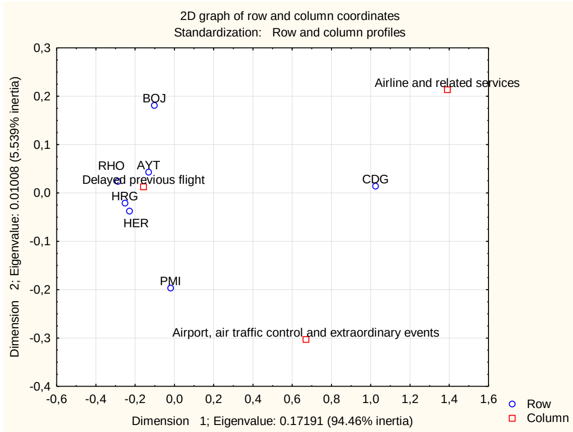


Figure 2 Correspondence map – Causes of delay and airport

Row relative frequencies	Airport, air traffic control and extraordinary events	Airline and related services	Delayed previous flight
AYT	5.17%	3.88%	90.96%
HER	5.83%	0.90%	93.27%
CDG	20.56%	26.64%	52.80%
PMI	12.28%	3.07%	84.65%
BOJ	2.26%	6.21%	91.53%
HRG	5.16%	0.65%	94.19%
RHO	3.65%	0.52%	95.83%

Table 3 Contingency table **–** Causes of delay and airport

In case of dependence of length of delay on particular airport the differences are not so big like in previous cases, see Tab. 4. Regarding the length of delay, the best situation is at Heraklion a Charles de Gaulle airports, where long delays over 1.5 hour occur rarely (up to 7% of all delayed flights), on the contrary short delays up to 30 minutes prevails at these airports (around 50% of all delays). With regard to the capacity and number of departure runways of CDG airport only short delays originate here although the intensity of traffic is high. There are bigger manipulation opportunities. Dependence is statistically significant, p-value is 0.0043.

It is evident from the row relative frequencies table (see Tab. 5) that delays caused by problems at the airport and in the air traffic as well as by unusual events (transport of handicapped passengers and necessary healthcare for passengers with sudden healthy problems) are the most frequent in case of short delays up to 30 minutes (11% of delayed flights). With increasing length of delay these causes occur less often (with an exception of the longest delays over 2 hours). Delays on airline's side (technical maintenance, defects, crew standards) and suppliers' side (handling, supply of fuel, catering) are the most frequent again in case of short delays. And again the ratio of these causes decrease with increasing length of delay with an exception of the longest delays (over 2 hours). On the contrary, the ratio of delays caused by delay of preceding flight mostly slightly increase with increasing length of delay. Delayed preceding flights cause mostly longer delays of connecting flights, whereas operational and technical reasons cause rather shorter delays. It is probably due to possible sanctions for suppliers and also duty to pay out compensations to passengers. Dependence is statistically significant; p-value is less than 0.0001.

Row relative frequencies	00:15-0:30	00:31-1:00	01:01-1:30	01:31-2:00	02:01 and more
AYT	46.51%	24.29%	11.89%	6.72%	10.59%
BOJ	41.24%	24.29%	12.43%	11.86%	10.17%
CDG	52.80%	25.70%	9.35%	7.01%	5.14%
HER	49.33%	30.04%	11.66%	4.48%	4.48%
HRG	39.35%	26.45%	14.84%	9.03%	10.32%
PMI	34.21%	33.33%	14.91%	7.02%	10.53%
RHO	41.67%	31.77%	8.85%	10.42%	7.29%

Table 4 Contingency table – Length of delay and airport

Row relative frequencies	Airport, air traffic control and extraordinary events	Airline and related services	Delayed previous flight
00:15-0:30	11.08%	8.63%	80.29%
00:31-1:00	6.41%	3.89%	89.70%
01:01-1:30	3.72%	3.72%	92.55%
01:31-2:00	0.82%	2.46%	96.72%
02:01 and more	2.24%	5.22%	92.54%

Table 5 Contingency table – Cause of delay and length of delay

During the time period we monitor, from July to September there are only minimal differences concerning causes of delay, which is evident in the row relative frequencies table (see Tab. 6). Statistical dependence was also not proven (p-value is $p = 0.96$).

Row relative frequencies	Airport, air traffic control and extraordinary events	Airline and related services	Delayed previous flight
June	7.36%	5.94%	86.70%
July	7.46%	5.04%	87.50%
August	6.83%	6.83%	86.34%
September	7.61%	6.23%	86.16%

Table 6 Contingency table – Cause of delay and selected time period

It is then apparent from the row relative frequencies table (see Tab. 7), that differences between particular airports are not so significant during the monitored time period. The best situation in June is at Burgas airport (only 17% of all delayed flights). The highest number of delayed flights in June is at Antalya, Heraklion and Hurghada airports (over 30% of all delayed flights). In July the worst situation is in Burgas and Hurghada airports (around 40% of all delayed flights). In August the worst situation is in Burgas and Charles de Gaulle airports (over 30% of all delayed flights). The situation is the best in September overall. The worst situation is at Antalya, Charles de Gaulle, Palma de Mallorca and Rhodes airports (over 20% of delayed flights). The best situation is in Hurghada in September (only 7% of all delayed departures). It can be stated overall that at Burgas airport delays originate the most often during the high season (July, August). The high season lasts only from July to August at this airport unlike other airports, delays almost don't occur in other months due to small traffic intensity. On contrary at Hurghada airport the bigger problems with delays occur in June and July. At Heraklion airport the situation improves significantly in September, the intensity of tourism decreases. These significant differences don't occur at the rest of airports. Dependence is statistically significant; p-value is less than 0.0001. From correspondence analysis (see Fig. 3) it is also evident, that delays are often at Antalya and Heraklion airport in June. In August delays are frequent at Charles de Gaulle airport. Burgas airport is the closest to July and August, and in these months delays are really the most frequent according to frequencies table.

Row relative frequencies	June	July	August	September
AYT	32.30%	23.51%	20.93%	23.26%
BOJ	16.95%	40.68%	31.64%	10.73%
CDG	23.36%	24.30%	32.71%	19.63%
HER	30.04%	27.35%	27.35%	15.25%
HRG	30.32%	40.00%	22.58%	7.10%
PMI	22.37%	29.39%	26.32%	21.93%
RHO	26.56%	26.56%	24.48%	22.40%

Table 7 Contingency table – Airports and monitored time period

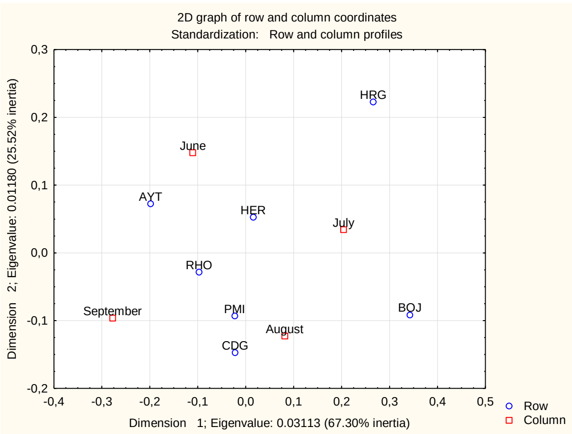


Figure 3 Correspondence map – Monitored time period and airport

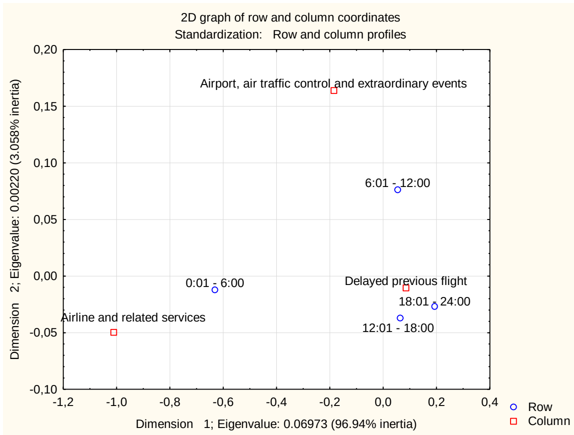


Figure 4 Correspondence map – Causes of delay and day time

It is evident from the relative frequencies table (see Tab. 8), that delays caused by problems at the airport and in the air traffic as well as caused by unusual events (transport of handicapped passengers and necessary healthcare for passengers with sudden healthy problems) are the most frequent are the most frequent at night and in the morning (around 10% of all delayed flights). Number of delays on airline's side (technical maintenance, defects, crew standards) and suppliers' side (handling, supply of fuel, catering) is the highest at night (20% of all delayed flights). Otherwise these causes are very rare (up to 5% of all delays). Delays caused by delay of preceding flight are the least frequent at nights. There are only few flights at night in general and the air space is unoccupied, that's why chaining of delays doesn't originate. In the morning there is generally busy traffic, therefore delays often occur by reason of problems at airport and difficulties in air traffic management. Delays caused by suppliers and service companies occur often at night, probably due to lack of night shift employees. Dependence is statistically significant; p-value is less than 0.0001. It is proved by correspondence analysis (see Fig. 4), that delays on airline's' and suppliers' side are the most frequent at night. Delays caused by delay of preceding flights are the closest to afternoon and evening time.

Row relative frequencies	Airport, air traffic control and extraordinary events	Airline and related services	Delayed previous flight
00:01-6:00	10.22%	20.44%	69.33%
06:01-12:00	8.96%	4.25%	86.79%
12:01-18:00	6.02%	4.73%	89.25%
18:01-24:00	5.63%	1.73%	92.64%

Table 8 Contingency table – Day time and causes of delay

4 Conclusions and discussion

Factors influencing delay of flights at selected international airports were examined in this article. These factors were identified by contingency tables analysis, which was created from the primary data, and also with use of airline's database which includes various causes of delays. It was proved from the carried out analysis that delays occur most frequently at Palma de Mallorca airport. It is a tourist destination where the intensity of traffic is the highest in the monitored season and the airport is very busy. Delay caused by delay of the preceding flight is the most frequent cause of delay at all. This cause occurs at Antalya, Heraklion, Palma de Mallorca, Burgas, Hurghada and Rhodes airports at most, it presents approximately 90% of all delays. Delays caused by other reasons occur minimally at all airports. Charter flights mostly fly to holidays destinations. Mostly regular flights fly to Charles de Gaulle airport, where delays are less common and chaining of delays doesn't originate. Regular flights always take precedence over the charter flight. Concerning the length of delay, the best situation is at Heraklion and Charles de Gaulle airports, where long delays over 1.5 hours occur rarely. Regarding the capacity and the number of departure runways at CDG airport, although the intensity of traffic is high, only short delays originate here thanks to bigger manipulation possibilities. Delayed preceding flights cause usually longer delays of connecting flights, whereas operational and technical reasons cause rather short delays. It is probably because of penalties for suppliers or necessity of compensations paying out to passengers. Delays caused by delay of preceding flights are the least frequent at night. At night there is low number of departures overall and unoccupied air space, that's why chaining of delays doesn't originate. In the morning there is generally busy traffic, therefore delays often occur by reason of problems at airport and difficulties in air traffic management.

In the article [2] there was proven dependence of delay on destination aerodrome and in our analysis the significant dependence on destination airport was also proven. The article [4] similarly offers different variants of strategy for reduction of delays, but again from another angle, the author recommends for example increasing the number of service counters during checking-in. However delays caused by passengers and their baggage occur very rarely at airports we monitored, except the international transit airport CDG. Similar to the article [5] also the airline we monitor tries to optimize the operating costs. In the article [6] the authors describe optimization of connecting flights in order to prevent chaining of delays. Our research showed that the most frequent cause of delays was particularly delay of preceding flight, approximately 90% of all delayed flights. In the article [8] it was also found out that delays caused by delay of preceding flight occur very frequently in general, they are much more frequent in Brno and Ostrava, around 80%, in Prague around 41%. Whereas in the article [7] the dependence of length of delay on particular airport in Czech Republic was not proven. Dependence of length of delay on selected destination airports is statistically significant.

It is necessary for airline to eliminate longer delays, because paying out of compensations to passengers is very expensive. There are nowadays specialised companies searching for passengers from delayed flights and recover financial compensations by legal way. In order to eliminate delay we would recommend borrowing an alternate aircraft in summer months to have the opportunity to replace an aircraft delayed on arrival to home base airport by an alternate aircraft. We would furthermore suggest suppliers and service companies hiring a certain number of seasonal workers for these months. It is beneficial also for these companies, because they can be also penalised for delay caused by their employees.

References

[1] Agresti, A.: *Categorical Data Analysis*. John Wiley a Sons, New York. 1990.

[2] Fleurquin, P., Ramasco, J.J., and Eguiluz, V.M.: Characterization of Delay Propagation in the US *Air-Transportation Network. Transportation journal* **53** (2014), 330–344.

[3] Hebák, P. et al.: *Vícerozměrné statistické metody 3.* Informatorium, Praha, 2007.

[4] Hsu, CI., Chao, CC., and Hsu, NW.: Control strategies for departure process delays at airport passenger terminals. *Transportation planning and technology* **38** (2015), 214–237.

[5] Pan, W., Wang, Z., and Zhou, G.: Flight Delay Cost Optimization based on Ground Power Supply. In: *Proceedings of the 2015 international informatics and computer engineering conference*. Atlantis Press, China, 2015, 336–339.

[6] Xing, Z., Yu, J., and Lu, H.: Optimal Control of Flight Delays Allocation at Airport. In: *Proceedings of the International Conference on Advances in Materials Science and Information Technologies in Industry.* Trans Tech Publications, China, 2014, 513–517.

[7] Zámková, M., and Prokop, M.: The Evaluation of Factors Influencing Flights Delay at Czech International Airports. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis* **63** (2015), 2187–2196.

[8] Zámková, M., and Prokop, M.: Příčiny a finanční důsledky zpoždění letů v České republice. *Acta academica karviniensia* **3** (2015), 110–120.

Decision of a Steel Company Trading with Emissions

František Zapletal[1], Martin Šmíd[2]

Abstract. We formulate a Mean-CVaR decision problem of a production company obliged to cover its CO2 emissions by allowances. Certain amount of the allowances is given to the company for free, the missing/redundant ones have to be bought/sold on a market. To manage their risk, the company can use derivatives on emissions allowances (in particular futures and options), in addition to spot values of allowances. We solve the decision problem for the case of an real-life Czech steel company for different levels of risk aversion and different scenarios of the demand. We show that the necessity of emissions trading generally, and the risk caused by the trading in particular, can influence the production significantly even when the risk is decreased by means of derivatives. The results of the study show that even for low levels of the risk aversion, futures on allowances are optimal to use in order to reduce the risk caused by the emissions trading.

Keywords: CVaR, emissions trading, optimization, allowances, EU ETS

JEL classification: C61
AMS classification: 90C15

1 Introduction and the current state of art

Since its launching, the European emissions trading system (EU ETS) became the object of interest of both researchers and managers of affected companies. The system began to work in 2005 with the purpose to decrease the amount of CO_2 emissions by forcing EU industrial companies to cover their emissions by emissions allowances (1 ton of CO_2 must be covered by 1 emissions allowance).

Our paper investigates the influence of the emissions trading on profit and production portfolio plan of a steel company participating in the EU ETS. The aim of the paper is to assess the impact of the EU ETS system on the company and, also, to determine the optimal emissions management for a single one-year-long period. Because two types of emissions allowances exist (European Union Allowaces EUA, and Certified Emissions Reduction CER) and, moreover, EUA's derivatives exist (options and futures), our goal is to determine which allowance type and which financial instruments are optimal to use. The model is applied using data of one Czech steel company.

Several models optimizing the output of a company with respect to a duty of emissions trading have been already published. In [4], the deterministic models has been presented. In [9], the stochastic model has been built, but without involving any risk measure. This paper follows [8] where the single-stage single-period mean-variance model has been presented. This present contribution uses a different risk measure - Conditional Value at Risk (CVaR). In particular, a mean-CVaR model is formulated. Unlike the variance, the CVaR measure belongs to the group of coherent risk measures; moreover, it takes into account possible fat tails and an assymetry of a distribution, see e.g. [1]. CVaR has been used in both static ([6]) and dynamic (multi-stage) models, see e.g. [5]. However, the onlyapplication of the CVaR in relation to the EU ETS system has been presented by [3]; however, that paper is aimed to determine the dynamic relationship between electricity prices and prices of emissions allowances.

The core of this paper consists in the formulation of the optimization model and its verification and sensitivity analysis at the end. In order to satisfy the requirements concerning paper's length, some details about the conditions of the EU ETS have to be skipped; however, they can be found e.g. in [7]

[1]VSB - Technical University of Ostrava, Faculty of Economics, Sokolská 33, 702 00 Ostrava, Czech Republic, frantisek.zapletal@vsb.cz
[2]The Institute of Information Theory and Automation of the CAS, Pod Vodárenskou věží 4, 182 08 Prague, Czech Republic, smid@utia.cas.cz

or [2].

The paper is organized as follows. After the introduction, the model is designed in Section 2. Section 3 presents the results of the model verification and the sensitivity analysis is done there. The last part of the paper is the conclusion where the main findings are pointed out and some proposals for further research are given.

2 The optimization model

In this section, the model assumptions are given and the Mean-CVaR optimization model is designed.

Assumptions

Production

- There is single (unit) decision period.
- There are n products produced.
- Demand for the products $d \in \mathbb{R}^n_+$, selling prices $p \in \mathbb{R}^n_+$ and the unit production costs (of final products) $c \in \mathbb{R}^n_+$ are known at the beginning of the period.
- Raw production x which is necessary for final production y is given by $x = Ty$ where $T \in \mathbb{R}^{n \times n}$ is an inverted technological matrix.
- The final production is non-zero, the limits of the raw production are given by vector $w \geq 0$, i.e.,

$$y \geq 0 \qquad Ty \leq w. \qquad (P)$$

Emissions

- The vector of CO_2 emissions resulting from raw aproduction x is given by $h^T x$ where h is a vector.

Finance

- Selling prices are collected at $t = 1$.
- The production costs are funded by a credit with (low) interest rate ι.
- Other costs are funded by loans payable at $t = 1$ with (high) interest rate $\rho > \iota$.
- If there is excess cash at $t = 0$, it may be deposited up to $t = 1$ with interest ι.
- Insufficiency of the unit of cash at $t = 1$ is penalized by constant σ (perhaps a prohibitive interest rate).

Emissions trading

- r EUA permits are obtained for free.
- At $t = 0$, the company may buy s^E_0 EUA spots and s^C_0 CER spots, f futures, $\phi_1, \phi_2, \dots \phi_k$ call EUA options with strike prices $K_1 < K_2 < \cdots < K_k$ respectively, and/or $\psi_1, \psi_2, \dots \psi_l$ put EUA options with strike prices $L_1 > L_2 > \cdots > L_l$, respectively.
- Short sales are not allowed, i.e.

$$s^E_0 \geq -r, s^C_0 \geq 0, f \geq 0, \phi \geq 0, \psi \geq 0 \qquad (F)$$

- Relative margin ζ is required when holding a future, for simplicity we assume that margin calls can occur only once a month and the margin cannot be decreased (for more on futures margins, see e.g. http://www.investopedia.com/university/futures/futures4.asp).

- At $t = 1$, the company may buy s_1^E EUA spots and s_1^C CER spots, short sales are not allowed:
$$s_0^C + s_1^C \geq 0, \quad r + s_0^E + s_1^E \geq 0. \qquad (G)$$
- Banking of permits is not allowed, i.e.
$$r + s_0^E + s_0^C + s_1^E + s_1^C + f + \sum_{i=1}^{k} \phi_1 - \sum_{i=1}^{l} \psi_i = hTy. \qquad (E)$$
- Only limited number of CERs may be applied, in particular,
$$s_0^C + s_1^C \leq \eta hTy \qquad (C)$$
where $0 \leq \eta < 1$.
- The company does not speculate: in particular, they would not buy more spots than needed neither they would buy more put options than the initial number of spots, i.e.,
$$s_0^E \leq hTy, \qquad s_0^C \leq \eta hTy, \qquad \sum \psi_i \leq r, \qquad (S)$$
(an alternative to (S) would be an assumption of no arbitrage).

Summarized, the vector of the decision variables is (y, ξ) where
$$\xi = (s_0^E, s_0^C, s_1^E, s_1^C, f, \phi_1, \ldots, \phi_k, \psi_1, \ldots, \psi_l)$$
(note, however, that the components of ξ are dependent due to (E)).

The gross balance from financial operations at $t = 0$, excluding margins of futures, thus equals to
$$E_0 = -s_0^E p_0^E - s_0^C p_0^C - \sum_{i=1}^{k} \phi_i p_0^{c,K_i} - \sum_{i=1}^{l} \psi_i p_0^{p,L_i}$$
where p_0^E and p_0^C are EUA, CER, respectively, prices at $t = 0$, p_0^{c,K_i} and p_0^{p,L_i} are the prices of call options with strike price K_i, put option with strike price L_i, respectively, at $t = 0$.

The costs of futures maintenance amounts to fM where
$$M = \sum_{j=0}^{11} \rho_{12-j} \left[M_j - M_{j-1}\right]^+, \qquad M_j = \max\left(M_{j-1}, \left[p_0^f - (1-\zeta) P_{j/12}^f\right]\right),$$
$$j = 0, 1 \ldots, 11.$$
Here, $\rho_j = (1+\rho)^{j/12} - 1$ is the $j/12$ time units interest rate, p_0^f is the future price at $t = 0$ and $M_{-1} = 0$ by definition.

The cash balance at $t = 1$ resulting from production is
$$B = p^T \min(d, y) - (1 + \iota) c^T y,$$
while the balance resulting from financial operations at $t = 1$ is
$$E_1 = -s_1^E P_1^E - s_1^C P_1^C - f p_0^f - \sum_{i=1}^{k} \phi_i \min(P_1^E, K_i) + \sum_{i=1}^{l} \psi_i \max(P_1^E, L_i).$$
Summarized, the value function is given by
$$V(y, \xi) = V(y, \xi, P_1^E, P_1^C, M_\rho) = g_{1+\sigma,1}\left(g_{1+\rho,1+\iota}(E_0) - Mf + E_1 + B\right)$$
where
$$g_{\alpha,\beta}(x) = \begin{cases} \beta x & x \geq 0 \\ \alpha x & x \leq 0 \end{cases}.$$
Consequently, the decision problem is formulated as
$$\begin{array}{ll} \max_{y,\xi} & (1-\lambda)\mathbb{E}V(y,\xi) - \lambda \mathrm{CVaR}_\alpha\left(-V(y,\xi)\right) \\ & \text{s.t. } (P), (F), (G), (E), (C), (S) \end{array} \qquad (1)$$
where $0 \leq \lambda \leq 1$ is a level of risk aversion.

3 Application of the model

3.1 Data

The following data sources were used for verification of the model:

- database of ICE stock exchange for data on emissions prices and related financial indeces (see theice.com);
- business data of certain Czech steel company for values related with production (i.e. margins, costs, production coefficients); these data have been slightly modified;
- database Carbon Market Data for amount of emissions and allocated allowances (see carbonmarketdata.com);
- database of Czech Steel Federation for historical data on steel market.

Due to the limited length of the contribution, it is not possible to present the complete set of input data. The same input values as in [8] have been used for corresponding parameters (e.g. data on demand or production) with the difference that a longer period of permit prices was used. Prices of EUA's and CER's spots at the beginning of the period are $p_0^e = 5$ EUR and $p_0^c = 0.38$ EUR. For the sake of simplicity, only 5 call and put options have been involved with different strike prices ($K = L = \{3, 4, 5, 6, 7\}$ EUR). Prices of those options are shown in Tab. 1.

Type of option/strike price	3	4	5	6	7
Call option	2.31	1.95	1.12	0.77	0.53
Put option	0.15	0.56	1.06	1.71	2.48

Table 1 Prices of EUA options with different strike prices [EUR] (source: theice.com)

The distribution of the emissions prices at the end of the period was assumed to be log-normal; in the actual computation more than 1000 scenarios were used with the following mean values: $p_1^e \doteq 5.29$ EUR, $p_1^c \doteq 0.43$ EUR and $M \doteq 0.78$ EUR.

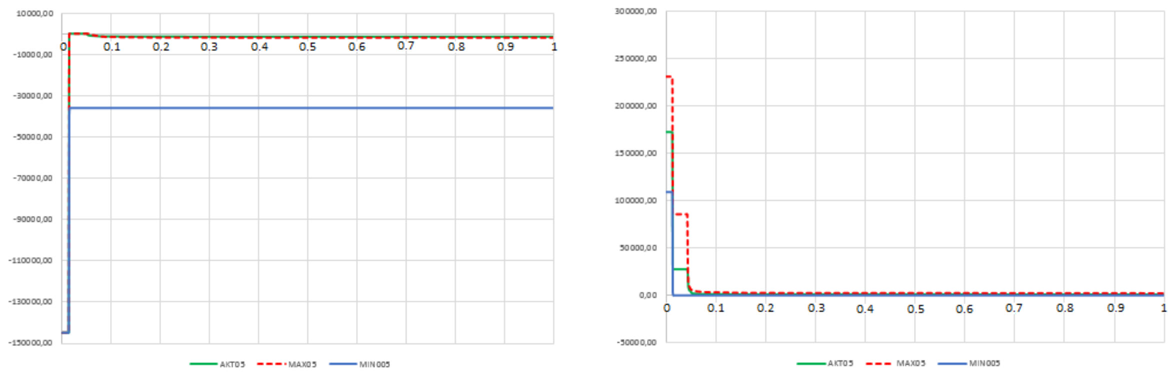
Unfortunately, no precise data on interests rates ι, ρ, σ were available. Therefore, their values are just estimated to be meaningful and to satisfy the assumptions listed in Section 2. In particular, $\iota = 0.01$, $\rho = 0.1$, $\sigma = 1$ were set (essentially, σ rate should be prohibitive enough). For the future research, those values will be surveyed and set more accurately.

It is reasonable to expect that the portfolio of financial instruments on allowances will depend on the fact if a company needs some additional allowances to purchase or if the amount of EUA's received for free is sufficient to cover all CO_2 emissions. That is why three different scenarios of the demand for production are considered. The first one $S1$ corresponds to the current level (i.e. level of 2014), when the company has a "slight" shortage of allowances. The second scenario $S2$ counts with the values of demand corresponding to production capacities of the company (i.e. the state when the company needs the maximum number of additional allowances). And the last scenario ($S3$) supposes the demand at the level of 50% of production capacities (in this case, the company has a surplus of permits).

The last parameter whose value left to be set is the risk aversion parameter λ. In order to avoid setting a particular range of values, the model will be run with all possible values, i.e. for $\lambda \in [0, 1]$. CVaR will be computed at 5% level of α.

3.2 Results

The model is solved using GAMS software. Similarly to other studies mentioned in the introduction (e.g. [8]), emissions trading does not influence the amount of the production. Moreover, the amount of production of all products is equal to the levels of demand considered, regardless a degree of risk-aversion. Thus, the only matter to analyze is the optimal portfolio of financial instruments for emissions allowances. In none of the three scenarios considered under any value of λ, EUA options are traded. All the remaining variables of allowances type differ with scenarios and a value of risk-aversion coefficient.



(a) p_0^e (b) p_1^e

Figure 1 EUA's traded at the beginning and end of the period

Most of the allowances are traded in a form of EUA spots. Optimal numbers of EUA spots traded at the beginning and the end of the period are shown in Fig. 1a and Fig. 1b, respectively. It can be seen that the company would sell all the allowances obtained for free (i.e. value of r) in all three scenarios for very low values of risk-aversion (for λ from 0 to values of around 0.02). The reason is, that the mean value of EUA price at the end of the period is slightly higher than the current (certain) value. With values of λ greater than 0.02, the company prefers higher certainty given by the current price of EUA and it sells the EUA spots only when no additional EUA's have to be bought (i.e. $S3$). At the end of the period, for small values of λ, all EUA's needed for the whole period are bought, when a risk-aversion increases, optimal values of p_1^e drops to zero (let us remind the expected increase in EUA price till the end of the period again).

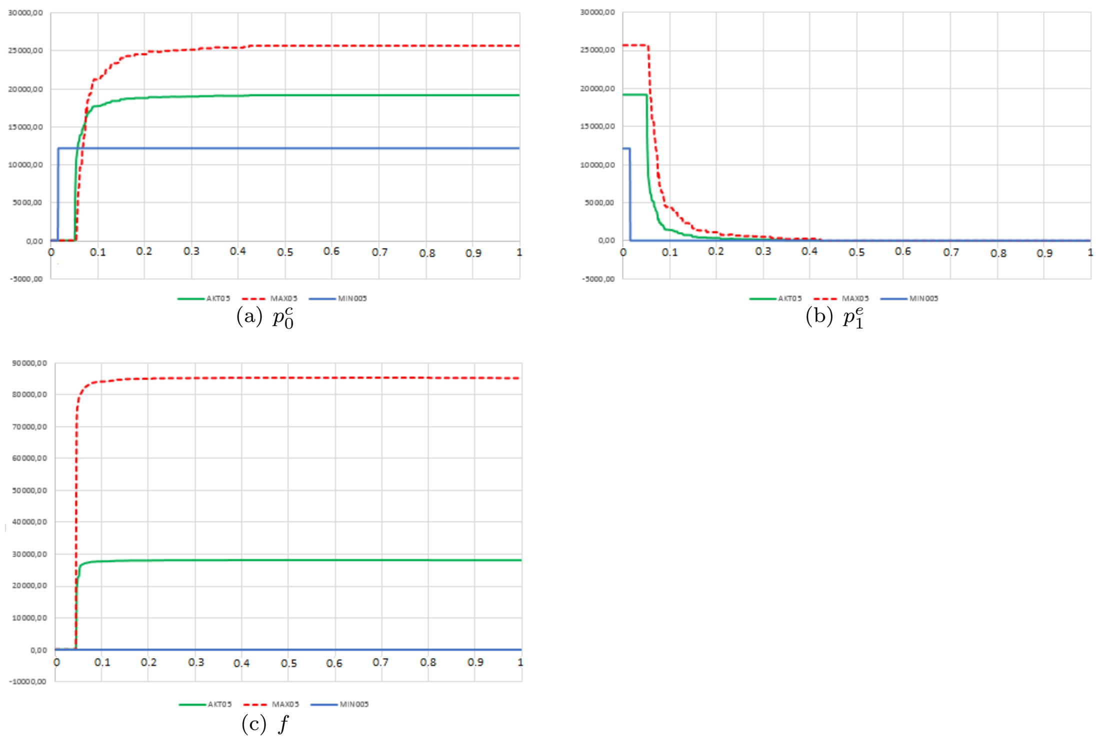
Due to very low prices of CER's in comparison with EUA's, the highest possible amount of CER's (given by $\eta = 10\%$) is always used by the company. The higher value of λ, the more CER's are bought at the beginning of the period and the less at its end, see Fig. 2a and Fig. 2b.

Out of all the EUA derivatives, only the futures are possibly used by the company. The optimal values of f when changing λ parameter are depicted in Fig. 2c. Under the scenario $S3$, no futures should be purchased at all. There is no reason for this because no additional EUA's are needed due to the low level of demand. The results differ when the number of allowances allocated to the company for free (r) is not sufficient to cover all CO_2. When assuming a lack of allowances $S1$ and $S2$, no futures should still be bought for $\lambda < 0.046$. But, for stronger risk-aversion, futures are used to decrease the risk related with emissions trading. In scenario $S1$, the amount of around 15% of all allowances needed should be purchased in the form of futures, meanwhile, under the scenario $S2$, the share of the futures goes even up to one third.

4 Conclusions

This paper was devoted to the effects of emissions trading within the EU ETS system on steel companies. The single-stage mean-CVaR optimization model has been formulated. The model has been applied using the data of the real steel company. For the sake of better understanding the modeled system, different scenarios of demands have been considered and all possible values of the decision-maker's risk aversion have been investigated. In the line with past studies, emissions trading does not affect the amount of production. The emphasis has been put on the analysis of optimal portfolio of financial instruments on emissions allowances. EUA options (both, call and put) have been considered to be inefficient to use. The optimal strategy regarding EUA spots differ with the value of risk-aversion coefficient. The company would always use 10% limit for CER allowances. When the number of allowances allocated to company for free is not sufficient, EUA futures should be used to decrease the risk.

Two ways how to extend the model to the future are expected. The first lies in improvement of the presented model (e.g. data on interest rates could be further investigated and other sensitivity analyses could be performed) and the second one consists in the extension to the multi-stage (dynamic) version.



(a) p_0^c (b) p_1^e (c) f

Figure 2 CER's traded at the beginning and end of the period and amount of futures purchased

Acknowledgements

This paper was supported by grant No. GA 16-01298S of the Czech Science Foundation, SGS project No.SP2016/116 and the Operational Programme Education for Competitiveness Project CZ.1.07/2.3.00/20.0296. The support is greatly acknowledged.

References

[1] Artzner, P. et al.: Coherent measures of risk, *Mathematical Finance* **9**, (1999), pp. 203228.

[2] European Commission: Directive 2003/87/ec of the european parliament and of the council of 13 october 2003. Union, European **46**, (2003), 32–46.

[3] Fell, H.: EU-ETS and Nordic Electricity: A CVAR Analysis, *The Energy Journal* **31**(2), (2010), 219–239.

[4] Letmathe, P., Balakrishnan, N.: Environmental considerations on the optimal product mix, *European Journal of Operational Research* **167**(2), (2005), 398–412.

[5] Kozmík, V., Morton, D. P.: Risk-Averse Stochastic Dual Dynamic Programming *Mathematical Programming* **152**(1), (2015), 275–300.

[6] Rockafellar, R. T., Uryasev, S.: Conditional value at risk for general loss distributions, *Journal of Banking & Finance* **26**, (2002), 1443–1471.

[7] Zapletal, F., Němec, R.: The Usage of Linear Programming to Construction of Ecologic-economical Model for Industrial Company Profit Optimization, *Proceedings of 30th International Conference Mathematical Methods in Economics*, Karviná, (2012), 1010–1015.

[8] Zapletal, F., Šmíd, M.: Mean-Risk Optimal Decision of a Steel Company under Emission Control, *Central European Journal of Operations Research* **24**(2), (2016), 435-454.

[9] Zhang, B., Xu, L.: Multi-item production planning with carbon cap and trade mechanism. *International Journal of Production Economics* **144**(1), (2013), 118–127.

Using Semi-dynamic Input-Output Model for the Analysis of Large Investment Projects

Jaroslav Zbranek[1], Jakub Fischer[2], Jaroslav Sixta[3]

Abstract. The use of Input-Output analysis for various macroeconomic studies is being more important in recent years. Due to more efficient computing technologies, these approaches provide extensive possibilities for various modifications of Leontiefs IO model. Static IO analyses based on constant technical coefficients are not sufficient for studying impacts of huge external impulses like investment shock. For these purposes, deeper analysis is needed. More sophisticated models such as CGE and DSGE offer solution, which is relatively complicated and their results are very much aggregated. The impact is even more significant in cases when the duration of investment shock is split into more years. The aim of the paper is to introduce a modified semi-dynamic IO model that we developed and successfully used for de-tailed analysis. Our model takes into account development in current and previous period including changes caused by investments realized in particular industry. The model permanently updates input coefficients without occurrence of any imbalances in the system. This allows reaction on gradual structural changes in the economy induced by the implementation of investment project.

Keywords: Input-Output, technical coefficients, models, macroeconomic analysis.

JEL classification: C67, O11
AMS classification: 65C20

1 Introduction

Input-Output analysis (IOA) is a continually expanding tool used especially for macroeconomic analyses. Due to continuous improvement of computer technology[4], it is possible to make IOA based on quite large Input-Output models (IOM). Such IOMs can be based on more detailed structures of commodities or industries on the one hand and they can also reflect various factors significantly influencing the results of the analyses on the other hand.

The aim of this paper is to introduce developed IOM which is adapted to analyses of impact of large multi-annual investment projects. This IOM called Semi-dynamic Input-Output model takes into account primary investment shock and includes also calculation of impact of induced effects in subsequent years. These effects come from increased wages and salaries and increased profits. The Semi-dynamic IOM uses automatically updated technical coefficients after the first and each of following years. This approach allows to adapt the following calculation to the changed situation in the economy due to interim results after finished stages of analyzed investment shock. Moreover, Semi-dynamic IOM retains detailed structure of commodities all the time of the process of analysis.

[1]University of Economics in Prague, Dep. of Economic Statistics, Nam W. Churchilla 4, Prague 3, Czech Republic, jaroslav.zbranek@vse.cz
[2]University of Economics in Prague, Dep. of Economic Statistics, Nam W. Churchilla 4, Prague 3, Czech Republic, fischerj@vse.cz
[3]University of Economics in Prague, Dep. of Economic Statistics, Nam W. Churchilla 4, Prague 3, Czech Republic, sixta@vse.cz
[4]Improvements of ICT products confirm also results of the study in [3].

2 Semi-dynamic Input-Output Model

The reason for the development of Semi-dynamic IOM was the need for analyzing the impacts of the huge investment shock that are spread over several years in the economy. As Symmetric Input-Output tables as input data are broken down by 82 commodities, using DSGE approach would need to aggregate input data in order to be able to reach results, [1]. Therefore Semi-dynamic IOM seems to be good approach concerning loss of information due to aggregation of input data which is not desirable for application of Semi-dynamic IOM. For this purpose, static Leontief's IOM, as describes Hronová, at al. [4] or Miller and Blair [5], seems to be insufficient because it is needed to take into account also changes in the economy due to investment shock made and induced effects as well. The base for the developed Semi-dynamic Input-Output model is a simple static IOM based on Leontiefs inverse matrix, as following formula 1:

$$x = (I - A)^{-1} y, \tag{1}$$

where

x	output vector,
y	vector of final use,
$(I - A)^{-1}$	Leontief inverse.

Beside primary investment shock, Semi-dynamic IOM takes into account two types of induced effects in each year of the investment project. The first induced effect is an increase of *final household consumption expenditures (FHCE)*, defined in [2], due to increased *wages and salaries (W)*, also defined in [2], of employees after primary investment shock. In case of induced effect based on W we assume that investment in particular year has impact only 50% to increase of *FHCE* in the same year. The remaining 50% has impact to the *FHCE* in the next year, according our assumptions. The estimate of induced increase of *FHCE* is based on the assumption of *average propensity to consume (APC)*, explained in [7]. The coefficient of APC is estimated according to formula 2:

$$APC = \frac{FHCE_{2010}}{W_{2010}}, \tag{2}$$

where

$FHCE_{2010}$	final household consumption expenditures in 2010 according to national accounts,
W_{2010}	wages and salaries in 2010 according to national accounts.

Induced increase of *FHCE* is calculated according to formula 3:

$$\Delta FHCE_t = 0.5 \cdot APC \cdot \Sigma\big(\Delta w_{j,t} + \Delta^* w_{j,t-1}\big), \tag{3}$$

where

$\Delta w_{j,t}$	increase of *wages and salaries* due to primary investment shock in the year *t*,
$\Delta^* w_{j,t-1}$	total increase of *wages and salaries* in the year *t-1*.

The commodity structure of induced increase of *FHCE* was calculated using *FHCE* matrix broken down by *CPA codes* and by *COICOP codes*, described in [2]. It is necessary to note that induced increase of *FHCE* is assumed especially in connection with products which can be signed as luxury products, which describes [7]. Products and services according to *COICOP* classification, described in [2], which can be considered as luxury products were selected in *FHCE* matrix for this purpose. Information and results of the study [6] were taken into account for the compilation of structural matrix of *FHCE.*

The second induced effect taken into account is further increase of investment, at macroeconomic level called *gross fixed capital formation (GFCF)*, defined in [2], due to increased *operating surplus* and *mixed income (OSMI)*, also defined in [2], which indicate operating profit of enterprises at macroeconomic level. The induced increase of *GFCF* takes into account also the increase of *OSMI* and based on the first mentioned induced effect, i.e. induced increase of *FHCE*. The estimate of induced increase of *GFCF* is based on the coefficient of *investment rate (i)* for this purpose defined according formula 4:

$$i_j = \frac{GFCF_{j,2010}}{OSMI_{j,2010}}, \tag{4}$$

where

$GFCF_{j,2010}$	*gross fixed capital formation* of industry *j*, in 2010 according to national accounts,
$OSMI_{j,2010}$	*operating surplus* and *mixed income* in industry *j*, in 2010 according to national accounts.

Induced increase of *GFCF* is calculated according to formula 5:

$$\Delta GFCF_{j,t} = i_j \cdot \left(\Delta OSMI_{I,j,t} + \Delta OSMI_{II,j,t}\right), \tag{5}$$

where

$\Delta OSMI_{I,j,t}$	increase of *operating surplus* and *mixed income* in industry *j*, in the year *t* due to primary investment shock,
$\Delta OSMI_{II,j,t}$	increase of *operating surplus* and *mixed income* in industry *j*, in the year *t* due to induced *FHCE*.

The commodity structure of induced increase of *GFCF* was done using structural matrix of *GFCF* for the year 2010 which is compiled and described as *AN approach* in [8]. This *GFCF* matrix is broken down by commodities in rows and by industries in columns.

The total impact is done by all effects together; primary investment shock, induced increase of *FHCE* and induced increase of *GFCF*. The principle of operation of Semi-dynamic IOM in each analyzed year shows Figure 1. More detailed description of Semi-dynamic IOM can be found in [9].

Before the analysis of impacts in the next year, there are always updated technical coefficients. The principle of their update is the same as their calculation for the simple static IOM. There are only changed input data, i.e. after including all impacts to the economy in the previous year. Input data for update of technical coefficients is still in form of Symmetric Input-Output Tables and thus this procedure does not cause any imbalances.

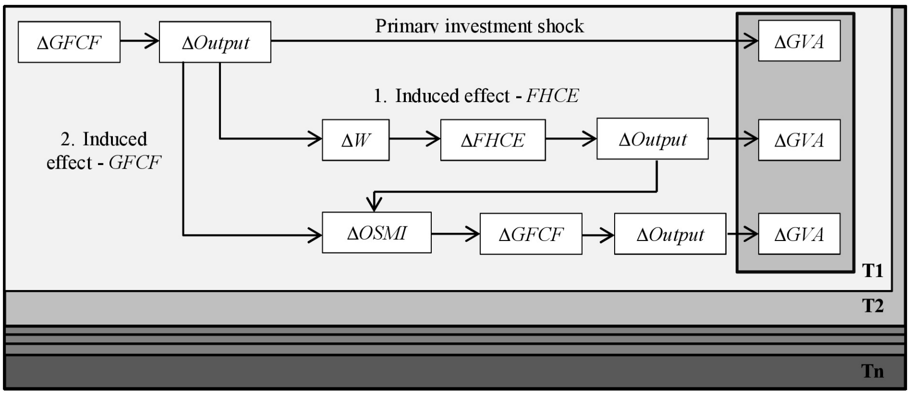


Figure 1 Overview of Semi-dynamic Input-Output Model

3 Input-Output Analysis Based on Semi-dynamic IOM

There are many possibilities of primary shock that can be used for application of Semi-dynamic IOM to make an Input-Output analysis. We decided to analyze a model situation when there is a large investment project, investment into the infrastructure in the Czech Republic. The duration of this project is four years and part of the large investment shock occurs in each year of the period. For illustration purposes, our calculations are proceed for the period of five years and the fifth year is devoted to induced effects only. The distribution of total investment shock more than 50 CZK billion according to product and year of investment shows Table 1.

CPA code	Product	T1	T2	T3	T4
42	Constructions and construction works for civil engineering	5 000	13 000	16 000	16 000
62	Computer programming, consultancy and related services	0	0	500	500
71	Architectural and engineering services; technical testing and analysis services	1 875	625	0	0

Table 1 Distribution of investment shock in time by commodity, CZK mil.

There are also planned capital expenditures connected with project documentation during the first two years (CPA 71), see Table 1. There is also gradual increase in capital expenditures directly to infrastructure. During last two years there are planned capital expenditures aimed at creation of necessary software connected with following use of constructed infrastructure.

3.1 Impact of the Investment Shock on the Gross Value Added

The impact of investment shock distributed during the period of four years can be measured by the change of gross value added (GVA). As it was mentioned with regard to induced effects results of application of Semi-dynamic IOM are presented on the period with duration five years, as it shows Figure 2.

The Figure 2 shows the development of change in GVA based on primary investment shock including both described induced effects. During the investment period, i.e. from T1 to T4 increased GVA varies between 87% in T1 and 91% in T4 in relation investment shock in particular year. Without inclusion of any induced effect it would be only between 74% in T1 and 72% in T4. The resulting figure is below 100% since the increase of investments stimulates also imports.

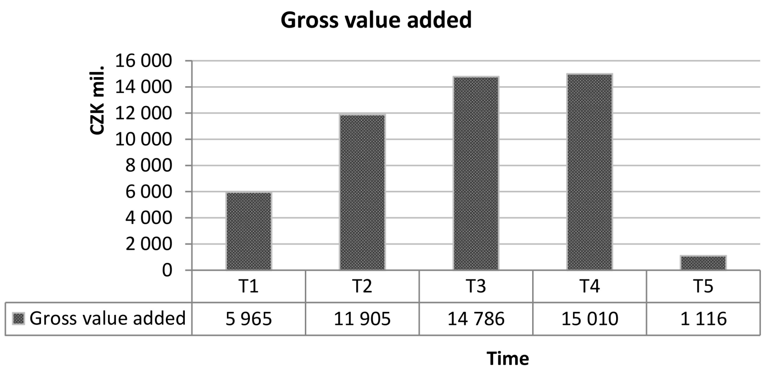


Figure 2 Distribution of impact to Gross value added during the analyzed period, CZK mil.

3.2 Impact of the Investment Shock on the output

Another approach how to evaluate the impact of investment shock distributed during the period of four years consists in using the change of output broken down by products (commodities). Similarly as in the case of GVA, the results of application of Semi-dynamic IOM are presented on the period with duration of five years, see Table 2. The table shows commodity structure aggregated on the section level of this classification. All calculations were done on the level of 2-digits CPA codes.

It is seen a significant increase of output connected with products related to construction industries, see Table 2. During the investment period from T1 to T4 there is total increase of output including both induced effects roughly three times higher than primary investment shock in particular year. Without including of any induced effect, it would be only around 2.5 times. Table 2 shows also changes in output structure in the fifth year when there is the highest output related to products of manufacturing industry.

The reason of this phenomenon lies in the occurrence of induced increase mainly of *FHCE*. The induced increase of *GFCF* is insignificant in this year.

CPA[5]/ Year	**T1**	**T2**	**T3**	**T4**	**T5**
Total	19 650	41 561	51 180	51 782	2 997
A	96	254	352	379	135
B	178	440	541	543	9
C	2 352	5 374	6 891	7 150	1 288
D	190	398	496	507	59
E	45	102	126	128	11
F	10 191	25 161	30 813	30 846	161
G	781	1 540	1 898	1 937	195
H	426	993	1 252	1 283	155
I	118	283	398	435	184
J	379	691	1 517	1 567	248
K	363	640	762	776	71
L	363	625	765	777	62
M	3 721	4 037	4 018	4 044	128
N	343	763	979	1 003	121
O	12	22	27	27	2
P	18	33	43	45	8
Q	1	2	3	3	1
R	30	93	141	159	86
S	43	110	158	173	73

Table 2 Distribution of impact on output by products during analyzed period, CZK mil.

3.3 Impact of the Investment Shock on the Employment

Important measure for analyses of impact of large investment projects is the change in employment. In connection with investment project, there is always expected positive impact on the employment. The impact on change of employment in assumed example is shown on Figure 3.

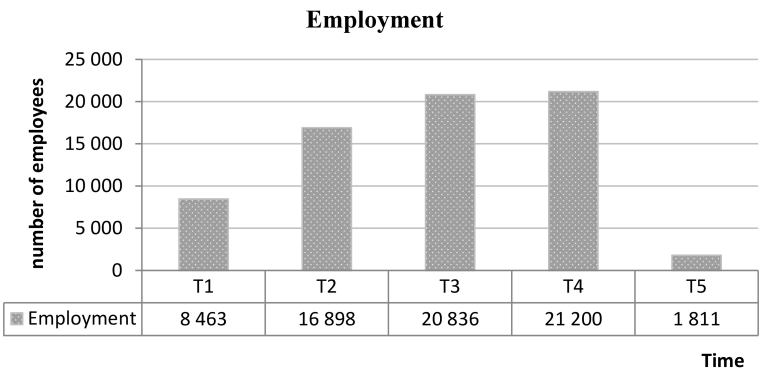


Figure 3 Distribution of impact to employment during the analyzed period

The Figure 3 clearly shows the highest impact to the employment in the year T4 with total increase 21.2 thousands of employees. This difference is caused by the inclusion of induced effects. From another point of view, one million of primary investment shock including both induced effects needs to hire around 1.25 employees to the production process. Without inclusion of any induced effect it would be maximally 1.03 employees.

[5] **A** Products of agriculture, forestry and fishing; **B** Mining and quarrying; **C** Manufactured products; **D** Electricity, gas, steam and air conditioning; **E** Water supply; sewerage, waste management and remediation services; **F** Constructions and construction works; **G** Wholesale and retail trade services; repair services of motor vehiclesand motorcycles; **H** Transportation and storage services; **I** Accommodation and food services; **J** Information and communication services; **K** Financial and insurance services; **L** Real estate services; **M** Professional, scientific and technical services; **N** Administrative and support services; **O** Public administration and defence services; compulsory social security services; **P** Education services; **Q** Human health and social work services; **R** Arts, entertainment and recreation services; **S** Other services.

4 Conclusion

In comparison with simple static IOM, semi-dynamic IOM provides more complex information in connection with external shock.. Inclusion of induced effects represents the most important benefit of Semi-dynamic IOM. The first effect is connected with induced change of final household consumption expenditures due to the change in wages and salaries after investment shock. Changes in wages and salaries imply changes in disposable income of households which is used for consumption expenditures or for savings. The second effect is connected with induced change in gross fixed capital formation due to change in operating profit of enterprises. This is caused by both initial investment shock and induced effect of final household consumption expenditures. Changes in operating profit of enterprises imply changes in investment behavior of subjects in the economy. Beside simple static Input-Output analysis and inclusion of induced effects, the Semi-dynamic IOM updates technical coefficients when moving to the next period of analysis. Updates of technical coefficients take into account structural change in the economy based on the investment shock. That is why that such model is useful for deep economic analysis especially for the large multiannual investment projects.

This paper brings an illustrative example of possible way of application of the Semi-dynamic IOM. This model was used for analysis of impacts of large investment shock in the Czech Republic. We have been developing Semi-dynamic IOM since spring 2013. It is based on published Symmetric Input-Output Tables, the last for 2010. All other necessary input data used for the development of the Semi-dynamic IOM come from national accounts of the Czech Republic for the year 2010. The currently used dimension of 82×82 products provides sufficient details for deep and relatively reliable analysis

Acknowledgements

Supported by the long term institutional support of research activities by Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] Bissová, S., Javorská, E., Vltavská, K., Zouhar, J.: Input-output interactions in a DSGE framework, In: *Mathematical Methods in Economics 2013*, Jihlava, 11.09.2013 - 13.09.2013. Jihlava : College of Polytechnics Jihlava, 2013, p. 55-59. ISBN 978-80-87035-76-4.

[2] European Union: *European System of Accounts ESA 2010*. Luxembourg, Publications Office of the European Union, [online], 2010.
URL: http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-02-13-269

[3] Fischer, J., Vltavská, K., Doucek, P., Hančlová, J.: Vliv informačních a komunikačních technologií na produktivitu práce a souhrnnou produktivitu faktorů v České republice. *Politická ekonomie* **5** (2013), p. 653–674. Text in Czech. ISSN 0032-3233.

[4] Hronová, S., Fischer, J., Hindls, R., Sixta, J.: *Národní účetnictví. Nástroj popisu globální ekonomiky.* C. H. Beck, Prague, 2009.

[5] Miller, R. E. and Blair, P. D.: *Input-Output Analysis: Foundations and Extensions.* Cambridge University Press, New York, 2009. ISBN-13 978-0-511-65103-8.

[6] Sixta, J., Vltavská, K., Hronová, S., Hindls, R.: Struktura spotřeby českých domácností 1970–2012. *Politická ekonomie* **6** (2014), p. 725–748. Text in Czech. ISSN 0032-3233.

[7] Soukup, J., Pošta, V., Neset, P., Pavelka, T., Dobrylovský, J.: *Makroekonomie: Moderní přístup.* Management Press, Prague, 2007. ISBN 978-80-7261-174-4.

[8] Safr, K.: Pilot Application of the Dynamic Input-Output Model. Case Study of the Czech Republic 2005-2013, *Statistika* **"forthcoming"**.

[9] Zbranek, J.: *Konstrukce a využití časových input-output tabulek v kontextu dynamizovaného input-output modelu.* Prague, 2014. Dissertation thesis. University of Economics in Prague. Faculty of Informatics and Statistics. Department of Economic Statistics. Text in Czech.

Tax burden and economic growth in OECD countries: Two indicators comparison

Petr Zimčík[1]

Abstract. The aim of this paper is to examine possible effects of tax burden on real economic growth in 32 OECD countries. The research is focused on individual components of two tax burden indicators. A tax quota is used in most of the similar studies and World Tax Index is an alternative. Instead of concentrating on the aggregate effects of tax burden, we rather focus on the effects of individual taxes in both indicators. These individual indicators are separately added into the augmented endogenous growth model with physical capital approximation as well as other control variables. The dynamic panel regression is used as a method to analyze any tax burden effects in a time period of 2000-2013. Main findings of this paper are that product taxes show the most significant adverse effects on the economic growth. Results also indicate that property taxes can positively stimulate economic growth rate. There is a certain contradiction between our results and majority of similar empirical studies. Any distortionary tendencies of corporate income tax quotas are identified as statistically insignificant in analysis and have no effects on economic growth.

Keywords: Economic growth, tax burden, tax quota, World Tax Index.

JEL Classification: E62, H20
AMS Classification: 62P20

1 Introduction

Taxes represent a crucial instrument in every economy. Governments need financial funds to operate and to conduct their policies. So they use taxes to collect these funds. More the governments collect, more public goods and services can be provided to citizens. This is a basic idea behind tax policy. However taxes have negative implications on real economy because they drain money from private sector. They are mandatory for citizens and companies and represent their contributions from wages, profits, properties and consumption. This could create distortions and slow down economic growth rate. All this can be interpreted as a tax burden, which is perceived by all economic subjects.

Governments stand before difficult challenge: to find balance between collecting necessary funds to cover their expenditures, and also not to exceed tax burden of domestic economic subjects. Tax burden is more apparent in developed (richer) economies because they usually have an extensive public sector and require higher volume of revenues. There are many ways how to measure tax burden in a specific country.

I use two indicators in this paper, tax quota with an alternative indicator, which is World Tax Index (WTI). More about WTI can be found in Kotlán and Machová [11]. Tax quota represents a share of total tax revenues in a relation to a nominal GDP. WTI is an indicator, which consists of tax quota hard data but also using soft data regarding expert opinions of economic institutions, researchers and others. It sets index of assumed tax burden in advanced countries.

The aim of this paper is to examine possible effects of tax burden on real economic growth in 32 OECD countries using two presented indicators. Overall tax burden is analyzed together with effects of individual taxes.

2 Literature review

Theoretical framework used to study any long-term impacts of taxation is the endogenous model of growth pioneered by Barro [4]. Fiscal policy in this model has impacts not only on the economic output but can also influence a steady-state growth rate. Barro [4] predicted either negative or neutral effects of taxation based on type of a tax classification.

Fölster and Henrekson [8] used econometric analysis to get estimates of taxation effects on growth rate. They discovered a negative correlation between those two variables. Additional testing of their estimates showed that

[1]Department of Economics, Faculty of Economics and Administration, Masaryk University, Lipová 41a, Brno, Czech Republic, e-mail: petr.zimcik@mail.muni.cz.

they were not robust. More recent studies showed robust and significant outcomes. Bergh and Karlsson [5] estimated effects of total taxation in OECD countries for periods 1970-1995. They result showed that an increase in total taxation of 1 percentage point (pp) induced decrease in growth rate by 0.11 percentage point. Bergh and Karlsson [5] also extended observed period to 1970-2005, which changed results only slightly. Taxation had 0.10 pp adverse effect on economic growth rate.

Romero-Avila and Strauch [12] pointed their interest on 15 EU countries in period 1960-2001. They found weaker negative effects than [5]. A one pp increase in taxation was associated with downturn in economic growth by 0.06-0.07 pp. Afonso and Furceri [1] used panel regression on OECD and EU member states to analyze both group of countries separately. They studied composition of taxation and its overall effect. In all regressions the effects were robust and negative. Results for both EU and OECD countries were similar and estimated to be about -0.12 pp, which is higher than two previous studies.

Not all empirical studies set down effects of rising taxation to be negative. Colombier [7] worked with a new regression estimator and fiscal variables. He has found effects of taxation on per capita growth rate to be either insignificant or small but positive. Bergh and Öhrn [6] presented some criticism of Colombier´s work in their paper. Criticism was not point to a new estimator but to the fact that certain control variables were omitted from regression. This could lead to bias in results. Bergh and Öhrn [6] re-estimated his results with proper control variables and obtained robust negative results from taxation increase.

Some authors were interested to make a scale, which would rank distortionary effects of individual type of taxes. Arnold [3] checked effects of different tax categories on GDP per capita. He used panel of 21 OECD countries in his analysis and found out that income taxes are in general accompanied with lower economic growth than taxes on property and production taxes. His analysis enabled him to create a ranking of taxes in view of economic growth. Property tax, especially recurrent tax on immovable property, has the most growth-friendly effects. Production taxes followed by personal income taxes indicated adverse effects. Corporate taxes have according to [3] the strongest negative effects on GDP per capita.

Xing [14] examined results in Arnold [3] and presented some doubts about his ranking system. In Xing´s analysis shifts in total tax revenue towards property taxes are indeed accompanied with higher income level per capita. However increase of tax revenues from production taxes, corporate and personal income taxes are associated with lower per capita income in the long run. Xing [14] haven´t found any proof to prioritize personal income taxes over corporate ones or favoring consumption taxes over any income taxes as in [3].

2.1 Tax burden

Tax policy as one form of government´s fiscal policy focuses on ensuring macroeconomic stability, which is essential for achieving and maintaining economic growth. According to IMF [9] a well-design tax policy can boost employment, productivity and investments through several channels. Excessive tax burden can have an opposite effect.

The tax and benefit systems, which include social transfers and labor income tax, influence decision making whether the person will be willing to work or not. They also determine how much (long) they will work. Capital income taxes influence private savings and investment. Decrease of savings returns on the individual level could also discourage small entrepreneurs from starting their own business. Governments can also use tax incentives to encourage private research spending. Introduction of this support design could provide high economic and social returns from research and development.

I follow model in Turnovsky [13] in my analysis of tax burden. The individual agent in a market economy purchases consumption out of the after-tax income generated by his labor and his holdings of capital. His objective is to maximize:

$$\int_0^\infty \frac{1}{\gamma}\left(cl^\theta G_c^\mu\right)^\gamma e^{-\rho t}dt, \tag{1}$$

where γ is a parameter related to inter-temporal elasticity of substitution, c is individual private consumption, l is leisure and G_c denotes the consumption services of a government provided consumption goods. Parameters θ and μ measure impacts of leisure and public consumption on welfare of the private agent. Economic agent is therefore a subject to the individual accumulation equation:

$$\dot{k} = (1-\tau_w)w(1-l) + (1-\tau_k)r_k k - (1+\tau_c)c - T/N \tag{2}$$

r_k is return to capital, w real wage rate, τ_w tax on wage income, τ_k tax on capital, τ_c consumption tax, T/N is a share of lump-sum taxes of an agent and k refers to agent´s holdings of capital. All types of taxes have negative repercussions on growth according to this model.

3 Methodology

We follow methodology in Bergh and Öhrn [6] and similar works. Tax variables from Turnovsky model (eq. 2) were added into endogenous model of growth. These fiscal variables represent approximation of tax burden in economy. Control variables used in Bergh and Öhrn [6] were also taken into consideration. We get a final econometric equation in a form of:

$$g_{i,t} = \alpha_i + \sum_{i=1}^{k} \beta_i Y_{i,t} + \sum_{i=1}^{m} \gamma_j X_{j,t} + u_{i,t}, \qquad (3)$$

where $g_{i,t}$ represents annual GDP growth, $Y_{i,t}$ are non-fiscal control variables and $X_{j,t}$ is approximation of tax burden. β_i and γ_j are respective coefficients of variables, $u_{i,t}$ is an error term. I use both overall tax quota and WTI to better determine possible effects of tax burden on economy and also to assess, which indicator is better. Second set of regressions are used to examine effects of individual categories of taxes.

3.1 Data

I used annual data from period 2000-2013 in analysis, which focuses on 32 OECD countries.[2] Data were collected from OECD statistics[3] and WTI, which can be found in Kotlán and Machová [11]. Dependent variable is the real GDP annual growth in percent. To sufficiently compare WTI and Tax quota, small adjustment had to be made. Tax quotas are in relation to GDP expressed in percent. WTI is constructed as an index and its nominal value is lower than one. WTI was multiplied by 100 to better compare these two variables. This changes interpretation of the results, which will be presented in next section.

Not only aggregate tax quota is observed. Individual tax categories were also examined for more detailed analysis. We used 4 main tax categories as seen as in table 1. These tax categories have their counterparts in WTI dataset, which were also multiplied by 100 to achieve comparable results.

Control variables are similar to those in Bergh and Öhrn [6]. Individual types of taxes were divided as follows: OECD have five main categories of taxes, which were transformed to respond to the WTI division. Eventually four groups of taxes were compared in analysis.[4] These groups with all explanatory variables and their description can be found in Table 1:

Control Variables	**Description**
Capital formation	Gross fixed capital formation growth (in %)
Employment	Annul change in employment (in %)
Inflation	Inflation rate measured by CPI
Unemployment	Unemployment rate (in %)
Trade	Annual growth rate of international trade (in %)
Fiscal Variables	**Description**
WTI	World Tax Index (multiple by 100)
Tax quota	Share of tax revenues to a nominal GDP
PIT	Personal income taxes
CIT	Corporate income taxes
Property	Individual property taxes
Product	Taxes on consumption (VAT included)

Table 1 Description of independent variables

Stationarity of all variables was tested using ADF test used in Levin [10]. Dependent variable same as control variables didn´t indicate any stationarity problems. In all fiscal variables, which are in form of indexes or quotas, was detected possible presence of a unit-root. Fiscal variables were change into first differences form and

[2] Chile and Mexico, as only two OECD member states, are not included in research due to data unavailability.
[3] Available [online] at http://stats.oecd.org/
[4] WTI has also five tax categories, value added tax and other taxes on consumption were merged to single category. All WTI categories were also multiplies by 100 to better correspond to their tax quota counterparts.

additional testing confirmed, that this procedure solved non-stationarity problem. In tables 2 and 3, you can see these differentiated variables with “d_” form.

3.2 Method

Two-step generalized method of moments (GMM) was used as an estimation technique. I applied Arellano-Bond estimator [2], which use dynamically add number of variable instruments for increasing t. Using this robust estimator in calculating the covariance matrices ensures that standard deviation of estimates and hypothesis tests are right to possible presence of heteroscedasticity and autocorrelation. Using appropriate instruments eliminates risk of endogeneity of dependent variable´s lagged values, independent variables with a random component. Starting from t-2, lagged values of the dependent variable were used as instruments.

Sargan–Hansen test was conducted to confirm use of appropriate instruments. J-statistic can be found in both sets of regressions in result section. Value of J-statistic does not reject a null hypothesis of instruments validity.

4 Empirical results

Results of two-step Arellano-Bond estimator can be found in Table 2. Dependent variable is real GDP growth rate. Lagged value is used as explanatory variable in dynamic panels. Coefficients of all control variables are consistent with the theory and the result in Bergh and Öhrn [6]. Capital formation, employment and openness are in direct relation with an economic activity and promote growth. From inflation and unemployment one expects opposite tendencies, which can be also seen in Table 2.

Only overall tax burden effects are estimated in this set of regressions. Coefficients of both indicators were negative, which correspond to an indirect relationship between economic growth and tax burden. Measuring burden with tax quota brings stronger negative tendencies than WTI, if you compare value of both coefficients. Ten pp increase of tax quota should be accompanied with 3.48 pp decrease in GDP growth rate ceteris paribus. Results for WTI show milder effects. Increase of WTI by 0.10 point is associated with 0.78 pp drop in growth rate.

Dependent variable	**Real GDP growth rate**			
Variable	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
GDP(-1)	0.036	1.126	0.019	0.612
const	-0.097	-4.838***	-0.128	-6.235***
Capital formation	0.095	6.831***	0.095	6.754***
Employment	0.356	5.336***	0.337	4.962***
Unemployment	-0.045	-0.958	-0.062	-1.297
Inflation	-0.077	-2.611***	-0.056	-1.867*
Trade	0.215	14.680***	0.207	13.900***
d_Tax quota	-0.348	-4.571***		
d_WTI			-0.078	-2.191**
Observations	384		384	
Number of instruments	85		85	
J-statistic	26.72		26.60	

Table 2 Effects of overall tax burden[5]

Overall effect of tax burden was decomposed to individual effects of personal income tax (PIT), corporate income tax (CIT), individual property tax (Property) and product taxes (Product) in Table 3. Control variables have similar coefficients as in Table 2. If you compare results for individual tax categories, main difference is the statistical significance. In WTI division only effects of product taxes are statistically significant at 5% level. Tax quotas have significant result for PIT, property and product taxes, which PIT and Product at 1% significance level. Based on this criterion, tax quotas give more relevant information in case of individual tax effects on economic growth.

[5] t-statistics are adjusted for heteroscedasticity and autocorrelation; *, **, *** indicate significance levels of 10 %, 5 %, and 1 %, respectively. Source: Author´s calculations. Please note: same is applied for Table 3.

Interpretation of coefficient is exactly same as in case of Table 2. Product taxes have the most distortionary tendencies. The 10 pp increase in product tax quota induces a 6.7 pp drop in GDP growth rate ceteris paribus. PIT has a milder effect estimated to be 4.74 pp. Individual property tax shows a direct relationship with economic growth. These results are consistent with previous study by Xing [14]. However corporate tax quota indicated no effect on growth even after being expected to have one of the strongest distortionary effects.

Dependent variable	**Real GDP growth rate**			
Indicator	**Tax quota**		**WTI**	
Variable	**Coefficient**	**t-statistic**	**Coefficient**	**t-statistic**
GDP(-1)	0.038	1.143	0.017	0.527
const	-0.096	-4.643***	-0.125	-6.074***
Capital formation	0.099	7.146***	0.097	6.851***
Employment	0.329	4.948***	0.336	4.954***
Unemployment	-0.048	-1.024	-0.056	-1.179
Inflation	-0.062	-2.135**	-0.056	-1.867*
Trade	0.209	13.870***	0.207	13.910***
d_PIT	-0.474	-2.868***	0.003	-0.053
d_CIT	0.025	0.163	-0.010	-0.135
d_Property	0.841	1.987**	-0.231	-1.435
d_Product	-0.670	-4.265***	-0.094	-2.206**
Observations	384		384	
Number of instruments	88		88	
J-statistic	24.09		27.17	

Table 3 Effects of individual tax burden

5 Conclusion

The aim of this paper was to examine possible effects of tax burden on real economic growth in 32 OECD countries in period 2000-2013. Two tax burden indicators were used. First is overall tax quota, second is World Tax Index. The latter is created not only with hard data but also consists of qualified opinions of institutions and domestic experts. Dynamic two-step GMM was used as technique to estimate both overall and individual effects of tax burden.

Tax quota has, according to our results, a significant negative correlation with GDP growth rate, while WTI has this negative correlation also but weaker In second set of regressions overall tax burden was decomposed to four tax categories – personal income tax, corporate income tax, individual property tax and product taxes, which include value added tax. The latter has the most distortionary effects followed by personal income tax. Property tax indicated a positive relation with economic growth. These results are similar to other empirical studies with one exception. In our case the corporate income tax did not show any significant effects.

These results are related to tax quota division only. In WTI division only product tax had significant negative effects. Compering these two indicators, tax quota presented more significant and relevant results. This can be explained by definition of WTI. WTI is constructed to contrast tax burden comparison among OECD countries. Value of index is in interval <0;1> and does not frequently change in time. Therefore it is not suitable for dynamic analysis. Tax quota is simpler but can be better use for studying impacts of tax burden on economic growth.

Acknowledgements

This work was supported by funding of specific research at Faculty of Economics and Administration - project MUNI/A/1055/2015. This support is gratefully acknowledged.

References

[1] Afonso, A., and Furceri, D.: Government Size, Composition, Volatility and Economic Growth. *Working paper No. 849*. ECB, 2008.

[2] Arellano, M., and Bond, S.: Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies* **58** (1991), 277-297.

[3] Arnold, J.: Do Tax Structures Affect Aggregate Economic Growth?: Empirical Evidence from a Panel of OECD Countries, *OECD Economics Department Working Papers*, No. 643, OECD Publishing.

[4] Barro, R.: Government Spending in a Simple Model of Endogenous Growth. J*ournal of Political Economy* **98** (1990), 103–125.

[5] Bergh, A., and Karlsson, M.: Government Size and Growth: Accounting for Economic Freedom and Globalization. *Public Choice* **142** (2010), 195-213.

[6] Bergh, A., and Öhrn, N.: Growth Effects of Fiscal Policies: A critical Appraisal of Colombier´s (2009) Study. *IFN Working Paper No. 865* (2011).

[7] Colombier, C.: Growth Effects of Fiscal Policies: An Application of Robust Modified M-Estimator. *Applied Economics* **41** (2009), 899-912.

[8] Fölster, S., and Henrekson, M.: Growth effects of government expenditure and taxation in rich countries. *European Economic Review* **45** (2001), 1501–1520.

[9] IMF. Fiscal Policy and Long-Term Growth. (2015). *IMF Policy Paper*. Washington D.C.

[10] Levin, A., Lin, C. F., and Chu, C.: Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *Journal of Econometrics* **108** (2002), 1–24.

[11] Machová, Z., and Kotlán, I.: World Tax Index: New Methodology for OECD Countries, 2000-2010. DANUBE: *Law and Economics Review* **4**, 165-179.

[12] Romero-Ávila, S., and Strauch, R.: Public finances and long-term growth in Europe: Evidence from a panel data analysis. *European Journal of Political Economy* **24** (2008), 172-191.

[13] Turnovsky, S.: Fiscal policy, elastic labor supply, and endogenous growth. *Journal of Monetary Economics*. **45** (2000), 185-210.

[14] Xing, J.: Does tax structure affect economic growth? Empirical evidence from OECD countries. *Centre for Business Taxation* WP 11/20.

Effectiveness and Efficiency of Fast Train Lines in the Czech Republic

Miroslav Žižka[1]

Abstract. The article deals with the evaluation of long-distance railway lines in the Czech Republic using Data Envelopment Analysis. The object of the research was 28 lines which are operated under a contract of public service obligation between the Ministry of Transport and Czech Railways. The article is divided into three parts. The first part focuses on a literature review of the use of DEA in evaluating transport services. The next part of the article describes the research methodology. The DEA model with uncontrollable inputs (fees for the use of infrastructure), variable returns to scale and input focused (NC-IV model) was used. The output of the first model was traffic performance. The result of DEA was cost efficiency. The outputs of the first model then enter the second model, in which service effectiveness was analyzed using the BCC-I model. Based on a comparison of yields with the initial inputs, overall efficiency was determined. In the third part of the article, the results of an analysis of each of the fast train lines are presented. For each line, there were determined the scores of technical efficiency, service effectiveness and overall effectiveness.

Keywords: fast train lines; DEA analysis; BCC model; non-controllable input; cost efficiency; service effectiveness; overall efficiency.

JEL Classification: R41, C61
AMS Classification: 90C90

1 Introduction

The article deals with evaluating the efficiency of long-distance rail transport in Czechia and the factors affecting it. Rail transport is operated either as a public service or on a commercial basis. On a commercial basis, there are EC, IC and SC trains that are currently operated by Czech Railways (CD) and open access connections operated by RegioJet and LEO Express operators. The article focuses on long-distance railway lines which are operated under a contract of public service [13]. The contract between the Ministry of Transport (MoT) and CD was signed for the period of 2010-19 and concerns 26 lines of long-distance transport in the categories of an express train and a fast train. The range of transport performance and the amount of compensation is annually specified by the amendments to the contract. For the year 2011, which is the subject of the analysis, the range of 35.2 million of train-kilometers was agreed. To cover demonstrable losses, the operator was paid CZK 4,005 million [12]. Two lines of fast train transport R14 and R16 have separate contracts as they were the subject of a public competition in 2005. Transport performance of these two lines accounted for 1.5 million of train-kilometers with the compensation of CZK 66.2 million. Thus, in total, there are operated 28 fast train lines in the public interest (see Table 1). In connection with the liberalization of the European passenger rail transport market, it is expected that instead of directly awarding the contract to one company, the contract to operate services after 2019 will be tendered. Currently, the tender schedule of 6 lines is already known [11]. Since a relatively high amount of public financial means attracting the interest of private operators are allocated for the reimbursement of provable losses, the debate was raised whether these funds are spent transparently and efficiently. One of the tools that can assess the efficiency of public services is the Data Envelopment Analysis (DEA). This article aims to evaluate the efficiency of fast train lines in terms of both an operator and in terms of passengers. From the viewpoint of the operator, the efficiency is connected with the transformation of inputs to transport performance. From the viewpoint of passengers, it is their willingness to pay the fare for the offered transport performance. The overall efficiency then measures revenues for the sold transport performance and costs paid at the input.

2 Literary overview

To measure the efficiency of rail transport, parametric and non-parametric approaches can be used. The parametric approach seeks to estimate the parameters of the production function. The non-parametric method is particularly DEA. The comparison of parametric and non-parametric approach was carried out on a sample of 89 commuter rail systems. The results revealed a relatively large consensus of both approaches [6]. Given that this arti-

[1] Technical University of Liberec, Faculty of Economics, Studentská 2, Liberec, miroslav.zizka@tul.cz.

cle is based on DEA approach, a research of the usage of this method was carried out when evaluating the efficiency of rail transport. The issue of evaluation of performance and efficiency in the transport sector can be divided into two areas - the evaluation of the performance of operators and the evaluation of the efficiency of lines.

The performance of railway companies can be broken down into three components: technical efficiency, service effectiveness and technical effectiveness [17]. It is related with the fact that the outputs can be divided into two categories: oriented inside the company (process of production services) and customer-oriented (business process). Technical efficiency measures the relation between inputs and product outputs. These outputs are transport performance (e.g. seat-kilometers, train-kilometers). Using these outputs then generates revenues which can be understood as the consumption of the service and also as the output of a business process. If there are no financial indicators at a disposal, passenger-kilometers can be used to measure consumption of the service. Service effectiveness evaluates the consumption of a service related to the offered transport performance. Technical efficiency puts into proportion transport performance and physical inputs. Relations between inputs and outputs often form a network. For measuring the performance of railway companies, it is therefore appropriate to use the network and multi-activity DEA approaches. The results were empirically tested on a sample of 20 railway companies in Europe [17], or in other research on 40 companies worldwide [16]. A research of efficiency of 8 high-speed rail systems was based on a similar principle. The system efficiency was divided in production efficiency and service effectiveness. The inputs of the production process were the capacity of the fleet and the length of the network and the output was the performance of the fleet. The output of the first process also served as an input to the second process. Carriage performance (in passenger-kilometers) and the number of passengers were the output of a consumer process [4]. In the survey of 24 US operators of suburban transportation, performance was divided into 4 sequential stages: management, operations, routing and scheduling, marketing and sales. The basic input at the beginning of the first process were operating costs, the final output were the revenues. The solution was the efficiency scores of individual operators in the time series 2001-10 and an assessment of the factors that influence this efficiency. It was found that systems with a high degree of subsidies are less organizational efficient [9]. Extensive research focused on the indicators that are used to evaluate the effectiveness and efficiency of public transport was carried out in the work [3]. The research showed that the inputs can be divided into three categories: physical, capital and operational. Physical inputs most often include fleet, employees and fuel consumption. Capital inputs mean the price of capital. Operating costs mostly include the cost of labor and fuel. Outputs can also be divided into 3 categories: providing service, service consumption and revenues. In addition, other external variables often occur in studies. These variables can be divided into 5 categories: quality (network length, average speed, average fleet age), socio-demographic and geographic (population density, locality), management (type of contract), subsidies and externalities (accidents, emissions).

The inspiration for the evaluation of individual lines rather than the performance of operators can be found in other modes of transport. Roháčová evaluated the efficiency of 28 lines of urban transport in Slovakia using a DEA hybrid model. Based on the correlation analysis, 4 inputs were identified (line length, no. of connections, the capacity, operating costs) and 2 outputs (total no. of passengers and sales). The input line length was considered uncontrollable as no modifications to existing lines are possible [15]. Another example is the evaluation of 15 airlines in Taiwan. The performance was evaluated from 3 perspectives: cost efficiency, service effectiveness and cost effectiveness. The inputs of the first production process were fuel cost, personnel cost and cost aircraft cost, outputs were the number of flights and seat-miles. These outputs were also inputs to the follow-up process of service consumption. The final outcome was passenger-mile and number of passengers [1].

3 Data and methodology

The basic source of data were statements [14] published by the MoT after a dispute with the RegioJet company in 2013. The statements include the expense and revenue items and data on transport performance in train-kilometers for each fast train lines. However, the MoT refused to disclose the figures for subsequent years for the reason of trade secret. The data from the MoT were further complemented by the length of each fast train lines in kilometers, the number of connections per week (taking into account the fact that some lines are not operated every day) and an average travel speed (km/h). These data were found on the timetable of CD in 2011. In addition, the numbers of inhabitants who may be potentially served by fast train lines were found. These are inhabitants from municipalities where the corresponding service stops. The data source was the statistics of municipalities [12]. The research methodology was divided into the following steps:

1. **Definition of inputs and outputs** – when providing railway services these inputs are important – fleet, personnel, energy and rail infrastructure. Fleet costs can be expressed by means of depreciation which expresses the amortization of vehicles as a result of their use to provide services to passengers. In Czechia, the ownership and maintenance of infrastructure are separated from railway transport operation. For using transport infrastructure rail operators pay fees to Railway Infrastructure Administration. In terms of DEA models, this fee has a nature of uncontrollable input because the operator can influence the amount in a very limited way.

The fee consists of two parts: the first part depends on the traveled distance and the second part depends on the weight of the train (in gross ton-kilometers). If a given line has a certain fixed length, the operator may affect only a component dependent on the weight of the train, using lighter train units. From the statements of the MoT, the detailed data on the cost structure were available (traction power and fuel, power consumption, personnel costs, depreciation, fees for use of railway infrastructure, etc.). From the viewpoint of the operator, sales are also compensations expressing the difference between the economically justified costs and revenues. There are two types of customers for operators. Passengers who pay the fare for the service provided and the MoT that will order a certain extent of transport performance to provide transport services in the territory. Sometimes these compensations are referred to as subsidies. The subsidies were not considered to be the output in DEA because they automatically compensate for the difference between costs and revenues. Empirically, it was verified that in this case the coefficient of overall efficiency in all lines would equal one or would be very close to this value.

Inputs and outputs for DEA should be independent and reflect the reality [5]. Relations between the input and output variables were assessed using Pearson product moment correlation coefficients. Strongly correlated variables were excluded from the analysis (r > 0.95, p-value < 0.0001): total cost, production consumption, operating and administrative overheads. On the output side, there is quite strong and direct relationship between revenues and transport performance in train-kilometers (r = 0.8777, p-value < 0.0001). Further, direct and strong correlation between the number of connections and realized transport performance (r = 0.8155, p-value < 0.0001) is logical. Only moderate correlation was found between revenues and number of connections (r = 0.5785, p-value = 0.0013) and average travel speed (r = 0.5119, p-value = 0.0054). A strong correlation with all inputs was found at the transport performance indicator in train-kilometers (r > 0.76, p-value < 0.0001). A similar result also applies to sales; the strongest correlation is between revenues and rail infrastructure fees (r = 0.9449, p-value < 0.0001). It can be interpreted in the way that the lines with high yields are operated on tracks where higher fees are paid for their use. Such a result is expected because the rail infrastructure fees depend on the category of the railway line (corridor, national, regional line). The most profitable lines are operated in corridor lines where the highest fee is paid for their use. Only a moderate correlation of travel speed on inputs is quite surprising. From this perspective, stronger correlation (r = 0.5066, p-value = 0.0059) was found only between travel speed and rail infrastructure charges. In fact, the fastest connections are operated along the corridor lines where the fees are higher. In the event of correlation of speed and depreciation, only weak and insignificant correlation was found (r = 0.3249, p-value = 0.0916). This can be explained by the fact that the travel speed depends mainly on the number of stops; not only on the type of power vehicle. The number of connections shows the strongest correlation with personnel costs (r = 0.8010, p-value <0.0001) – high frequency of connections requires a larger number of staff. Similarly, relationship of a number of connections with the rail infrastructure fees (r = 0.7408, p-value <0.0001) and energy consumption (r = 0.7248, p-value < 0.0001) is strong and logical.

2. **Construction of mathematical models** – for the first stage of cost efficiency, a model was developed with the assumption of variable returns to scale, with a focus on inputs and one uncontrollable variable (NC-IV model). In the mathematical model, see the equations (1) and (2), it is necessary to divide the inputs into controllable (I_K) and uncontrollable (I_N). **Controllable inputs** for calculating the *CE* **(model 1)** were consumption of fuel and energy, personnel expenses and depreciations. **Uncontrollable input** (model 1) was the rail infrastructure fees. Slack variables belonging to uncontrollable inputs are set to equal zero [7]. **Outputs** were transport performances in train-kilometers and number of connections. In the service effectiveness stage **(model 2),** classic BCC model with a focus on inputs was used, see relations (3) and (4). Two outputs from the first model were used as **inputs** (transport performance in train-kilometers; number of connections). Revenues were **output.** For the last stage **(model 3)** of overall efficiency, there was also used NC-IV model with one **output** (revenues) and the **inputs** from the first model of analysis (controllable inputs: fuel and energy, personnel expenses, depreciations; uncontrollable input: infrastructure fees). Models with a focus on inputs were used because rail transport is one of the sectors with a derived demand [8]. This means that the demand can be influenced by changing the range of transport service. The open source OSDE-GUI SW was used to solve all models.

$$z = \theta_q - \varepsilon\left(\sum_{i \in I_k} s_i^- + \sum_{i \in O} s_i^+\right) \ldots \min \tag{1}$$

$$\begin{aligned} \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- &= \theta_q x_{iq}\ ,\ i \in I_K \\ \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- &= x_{iq}\ ,\ i \in I_N \end{aligned} \tag{2}$$

$$\sum_{j=1}^{n} y_{ij}\lambda_j - s_i^+ = y_{iq}\ ,\ i \in O$$

$$\lambda_j \geq 0, s_i^+ \geq 0, s_i^- \geq 0,\ \sum_{j=1}^{n}\lambda_j = 1$$

$$s_i^- = 0,\ i \in I_N$$

$$z = \theta_q - \varepsilon\left(\sum_{i \in I} s_i^- + \sum_{i \in O} s_i^+\right) \ldots \min \tag{3}$$

$$\sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- = \theta_q x_{iq}\ ,\ i \in I$$

$$\sum_{j=1}^{n} y_{ij}\lambda_j - s_i^+ = y_{iq}\ ,\ i \in O \tag{4}$$

$$\lambda_j \geq 0, s_i^+ \geq 0, s_i^- \geq 0,\ \sum_{j=1}^{n}\lambda_j = 1$$

3. **Analysis of the effects of exogenous factors** – the regression analysis examined other factors that could affect the scores of efficiency and effectiveness. As exogenous factors were considered line length, average travel speed, the population living in municipalities which are served by individual lines and the amount of compensation to cover losses. Given that the rate of efficiency/effectiveness of DEA method takes values in the interval (0; 1), it is often recommended that instead of using "classical" regression based on the method of least squares, use the Tobit model. McDonald, however, shows that there are no practical reasons for that. Neither the scores of efficiency nor effectiveness are generated by censored processes [10].

4 Research results

Table 1 provides an overview of evaluated of express and fast train lines including initial and terminal destination. For each line, there are given coefficients of cost efficiency *CE* (the result of NC-IV model solution), service effectiveness *SE* (BCC-I model) and the overall efficiency *OE* (NC-IV model).

Table 1 shows that the cost efficiency is quite high. To achieve cost efficiency, inputs would need to decrease on average by only 4%. The lines R21, R22 and R15, where controllable inputs would need to be reduced by up to 47% while maintaining the existing volume of outputs, were the least cost efficient. For each inefficient line peer lines are given. Due to the lack of space, weights of the sample units are not included. For example, for the line R21, peer units are the lines R14 (0.064), R16 (0.686) and R23 (0.25), data in parentheses are weights. To achieve cost-efficiency, it would be necessary to reduce fuel costs by 53%, labor costs by 49% and depreciation by 47% in the line R21. Fees for using the infrastructure remain the same as uncontrollable input.

Service effectiveness in the second stage is relatively low; inputs in the second phase would have to be reduced by an average of 49%. Only four lines lie on the frontier - Ex1, Ex3, Ex4 and N28. The lowest level of service effectiveness was reached by R26, R14 and R11 lines where inputs would have to be reduced by up to 77%. The overall efficiency indicates the need to reduce inputs by 10% on average. On the efficient frontier are 19 lines. The least efficient are the lines R15, R21 and R27 where it would be necessary to reduce inputs by up to 56%. Table 2 shows the necessary level of reduction of individual cost items to achieve effective frontier. The results indicate that in terms of all levels of efficiency and effectiveness, the best rated are Ex3, Ex4 and N28 lines. These are lines operated on corridor lines.

In the last section, there were examined more exogenous factors which could affect the individual scores of efficiency and effectiveness (table 3). It was found that in the first model none of the considered factors have influence on the cost efficiency score. For service effectiveness score, there was found moderately strong dependence on the length of line (r = 0.4550) and weaker, but significant dependence on the average travel speed (r = 0.3759). It can be concluded that the longer the route of the train, the more customers it serves, which should have a positive impact on sales. Similarly, a higher speed should attract more passengers and increase revenues on the line. In the last model, there was found moderately strong dependence of overall efficiency score and a travel speed (r = 0.5544). Conversely, there was no evidence that subsidies are a factor that affects the efficiency and effectiveness scores. Apparently this is due to the fact that they are provided automatically to cover the difference between the economically justified costs and the recorded revenues.

5 Conclusion

Each model reflects efficiency and effectiveness from a different perspective. The first model shows that inputs were used quite efficiently to create service offerings from the perspective of the operator. The second model shows how passengers used this offer, or better say how successfully this offer was sold to passengers. This effectiveness was relatively low. It appears that on some lines, there is either a potential to increase the output (e.g. by improvement of provided services), or to reduce the scale of operations. The third model assesses the efficiency on an aggregated basis from the point of view of initial inputs and final output. It can also be assessed from the perspective of society as a whole because the loss from the lines operation is compensated to the operator by the order party. The overall efficiency was relatively high; however, there are fast train lines with a very low overall efficiency score. In terms of practical applications, it is interesting that when the MoT was looking for savings in 2012, it abolished the order of connections on the lines R17 and R25 which showed a relatively high score of efficiency and effectiveness in this analysis. The least efficient yet were the lines R27 and R15 where the scale of operation remained unchanged. The discussion in professional journals suggesting little willingness of the MoT to publish data on loss rates for individual lines shows that little attention is paid to the issue of the efficiency of individual lines. DEA is a fairly powerful tool that could help to objectively compare the effectiveness and the efficiency of fast train lines operated in the public interest.

References

[1] Chiou, Y. Ch., and Chen, Y. H.: Route-based performance evaluation of Taiwanese domestic airlines using data envelopment analysis. *Transportation Research* **42** (2006), 116–127.

[2] CZSO: *Městská a obecní statistika* [Municipal statistics][online]. Czech Statistical Office, Prague, 2016. Available at https://vdb.czso.cz/mos.

[3] Daraio, C., Diana, M., Di Costa, F., Leporelli, C., and Matteucci, G.: Efficiency and effectiveness in the urban public transport sector: A critical review with directions for future research. *European Journal of Operational Research* **248** (2016), 1–20.

[4] Doomernik, J. E.: Performance and Efficiency of High-Speed Rail Systems. *Transportation Research Procedia* **8** (2015), 136–144.

[5] Düzakin, E., and Düzakin, H.: Measuring the Performance of Manufacturing Firms with Super Slacks based Model of Data Envelopment Analysis: An Application of 500 Major Industrial Enterprises in Turkey. *European Journal of Operational Research* **182** (2007), 1412–1432.

[6] Graham, D. J.: Productivity and efficiency in urban railways: Parametric and non-parametric estimates. *Transportation Research* **44** (2008), 84–99.

[7] Jablonský, J., and Dlouhý, M.: *Modely hodnocení efektivnosti produkčních jednotek* [Models of efficiency assessment of decision-making units]. Professional Publishing, Prague, 2004.

[8] Jitsuzumi, T., and Nakamura, A.: Causes of inefficiency in Japanese railways: Application of DEA for managers and policymakers. *Socio-Economic Planning Sciences* **44** (2010) 161–173.

[9] Mallikarjun, S., Lewis, H. F., and Sexton, T. R.: Operational performance of U.S. public rail transit and implications for public policy. *Socio-Economic Planning Sciences* **48** (2014), 74–88.

[10] McDonald, J.: Using least squares and tobit in second stage DEA efficiency analyses. *European Journal of Operational Research* **197** (2009), 792–798.

[11] MoT: *Aktualizovaný harmonogram soutěží provozovatelů dálkové osobní železniční dopravy* [Updated tender schedule for operators of long-distance passenger rail transport][online]. Ministry of Transport, Prague, 2014. Available at http://www.mdcr.cz/cs/verejna-doprava/nabidkova-rizeni.

[12] MoT: *Dodatek č. 1 ke Smlouvě o závazku veřejné služby v drážní osobní dopravě ve veřejném zájmu* [Amendment Nr. 1 to the contract of public service obligation in rail passenger transport in the public interest] [online]. Ministry of Transport, Prague, 2009. Available at http://www.mdcr.cz/cs/verejna-doprava.

[13] MoT: *Smlouva o závazku veřejné služby v drážní osobní dopravě ve veřejném zájmu* [Contract of public service obligation in rail passenger transport in the public interest][online]. Ministry of Transport, Prague, 2009. Available at http://www.mdcr.cz/cs/verejna-doprava/Smlouvy+o+ZVS/ smlouva-o-zavazku-verejne-sluzby.htm.

[14] MoT: *Výkaz nákladů a výnosů z přepravní činnosti ve veřejné drážní osobní dopravě po jednotlivých linkách, období 1. 1. – 31. 12. 2011* [Statement of expenses and revenues from transport activities in the public passenger rail transport by individual lines, period 2011-01-01 – 2011-12-31][online]. Ministry of Transport, Prague, 2013. Available at http://www.mdcr.cz/NR/rdonlyres/EAA58F20-01D1-4951-BF16-C4D2D546B91D/0/245_odpovII.pdf.

[15] Roháčová, V.: A DEA based approach for optimization of urban public transport system. *Central European Journal of Operations Research* **23** (2015), 215–233.

[16] Yu, M. M.: Assessing the technical efficiency, service effectiveness, and technical effectiveness of the world's railways through NDEA analysis. *Transportation Research* **42** (2008), 1283–1294

[17] Yu, M. M., and Lin, E. T. J.: Efficiency and effectiveness in railway performance using a multi-activity network DEA model. *Omega* **36** (2008), 1005–1017.
[18] Zhao, Y., Triantis, K., Murray-Tuite, P., and Edara, P.: Performance measurement of a transportation network with a downtown space reservation system: A network-DEA approach. *Transportation Research* **47** (2011), 1140–1159.

Nr	FROM – TO	CE	Peer	SE	Peer	OE	Peer
Ex1	PHA–OSTRAVA– PL/SK	0.9739	3, 5, 19, 28	1.0000		1.0000	
Ex2	PHA–LUHACOVICE/ZLIN	1.0000		0.7557	1, 4, 28	1.0000	
Ex3	DE–USTI–BRECLAV–AT	1.0000		1.0000		1.0000	
Ex4	BRECLAV–OSTR–BOHUMIN	1.0000		1.0000		1.0000	
R5	PHA–USTI–CHEB	1.0000		0.4200	4, 28	1.0000	
R6	PHA–PLZEN-CHEB/KLATOVY	1.0000		0.3726	4, 28	1.0000	
R7	PRAHA–CB–AT	1.0000		0.3958	4, 28	1.0000	
R8	BRNO–BOHUMIN	0.9591	5, 7, 9, 19	0.4067	4, 28	0.9562	5, 9
R9	PHA–HB–BRNO	1.0000		0.4029	4, 28	1.0000	
R10	PHA–TRUTNOV	1.0000		0.3712	4, 28	0.7874	12, 18, 19, 28
R11	BRNO–PLZEN	1.0000		0.2819	1, 4, 28	1.0000	
R12	BRNO–SUMPERK	1.0000		0.3774	4, 28	1.0000	
R13	BRNO–BRECLAV–OLOMOUC	0.9868	7, 12, 19, 23	0.3640	4, 28	0.9850	3, 5, 19, 23
R14	PARDUBICE–LBC	1.0000		0.2774	4, 28	1.0000	
R15	USTI–LBC	0.8383	12,14,16, 24	0.3483	4, 28	0.4370	4, 14, 16, 28
R16	PLZEN–MOST	1.0000		0.4561	4, 28	1.0000	
R17	PARDUBICE–JIHLAVA	1.0000		0.9329	4	1.0000	
R18	PHA–VSETIN/LUHACOVICE	1.0000		0.3663	1, 4, 28	1.0000	
R19	PHA–CT–BRNO	1.0000		0.3441	4, 28	1.0000	
R20	PHA–USTI–DECIN	0.9748	7, 12, 23	0.4642	4, 28	0.9944	1, 5, 12, 23
R21	PHA–TANVALD	0.5291	14, 16, 23	0.3669	4, 28	0.4656	4, 14, 16, 28
R22	KOLIN–RUMBURK	0.6899	14, 16, 23	0.3214	4, 28	0.5492	4, 14, 16, 28
R23	KOLIN–USTI	1.0000		0.3278	4, 28	1.0000	
R24	PHA–RAKOVNIK	1.0000		0.4506	4, 28	0.5579	3, 4, 14
R25	RAKOVNIK–CHOMUTOV	1.0000		0.7766	4	1.0000	
R26	PHA–PRIBRAM–CB	1.0000		0.2343	4, 28	1.0000	
R27	OSTRAVA–KRNOV–JESENIK	1.0000		0.3453	4, 28	0.4916	3, 4, 23
N28	PHA–OLOMOUC–PL/SK	1.0000		1.0000		1.0000	
	AVERAGE	0.9626		0.5057		0.9009	
	SD	0.1031		0.2472		0.1886	

Table 1 Coefficients of cost efficiency, service effectiveness and overall efficiency

DMU Nr	Consumption of energy	Personnel expenses	Depreciations
R8	4.38	5.96	50.81
R10	21.26	21.26	21.26
R13	1.50	1.50	10.10
R15	56.30	56.30	56.30
R20	0.56	11.21	0.56
R21	53.44	53.44	53.44
R22	45.08	45.08	45.08
R24	44.21	44.21	44.74
R27	50.84	50.84	71.50

Table 2 Cost reduction (%) necessary for achieving overall efficiency

Dependent variable	Independent variable	Slope	t-statistic	p-value
CE	Length	1.3559E-4	1.1625	0.2556
	Speed	0.0025	1.7367	0.0943
	Population	2.0440E-8	0.7307	0.4715
	Subsidy	1.9258E-7	1.0679	0.2954
SE	**Length**	**6.65678E-4**	**2.6052**	**0.0150**
	Speed	**0.0069**	**2.0687**	**0.0487**
	Population	8.3751E-8	1.2736	0.2141
	Subsidy	1.0389E-7	0.2353	0.8158
OE	Length	3.65408E-4	1.7660	0.0891
	Speed	**0.0078**	**3.3964**	**0.0022**
	Population	6.1199E-8	1.2163	0.2348
	Subsidy	3.7413E-7	1.1365	0.2661

Table 3 Regression analysis of exogenous factors

Title	34th International Conference Mathematical Methods in Economics 2016 Conference Proceedings
Author	composite authors participants of the international conference
Intended	for participants of the international conference for researchers and professional public
Publisher	Technical University of Liberec, Studentská 1402/2, Liberec

Authorized by rectorate of the Technical University of Liberec, reference number RE 31/16, as on the 25th July 2016

Published	in September 2016
Pages	951
Edition	1.
Press	Vysokoškolský podnik Liberec, spol. s r.o., Studentská 1402/2, Liberec
Reference number of publication	55-031-16

This publication has not been a subject of language check.

All papers passed a double-blind review process.

ISBN 978-80-7494-296-9

TECHNICAL UNIVERSITY OF LIBEREC
www.tul.cz

ČSOV

34th International Conference

Mathematical Methods in Economics

September 6th – 9th, 2016, Liberec, Czech Republic

ISBN 978-80-7494-296-9