Faculty of Economics and Management
Czech University of Life Sciences Prague

27[TH] International Conference

Mathematical Methods in Economics 2009

September 9 - 11. 2009

2009
MME

WHAT DOES THE WAVELET ANALYSIS TELL US ABOUT THE CENTRAL EUROPEAN STOCK MARKETS BEHAVIOR DURING THE CRISIS?[1]

Lukáš Vácha, Jozef Baruník

Institute of Economic Studies, FSV, Charles University in Prague.
Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague
(barunik@utia.cas.cz)

Abstract. In the proposed paper we would like to test for the different reactions of the stock markets to current financial crisis. We will focus on the Central European stock markets, namely Czech, Polish, Hungarian and compare them to German and U.S. benchmark stock markets.

Main method used for the analysis is the wavelet variance decomposition. Wavelet variance (or energy) decomposes a variance of stochastic process on scale basis and hence is important tool for analysing financial time series. The wavelet variance is a suitable alternative to the power spectrum analysis based on the Fourier transform. Such scale decomposition help us to track the different energies on the tested stock markets and their evolution in time.

The wavelet analysis of the tested stock markets shows different energies on scales during current financial crisis. Results indicate that each of the tested stock markets reacted differently to the current financial crisis. More important, Central European stock markets seem to have strongly different behaviour during the crisis. This may be in contradiction to common regional and liquidity similarities, which would indicate more common behaviour.

Keywords. Wavelet analysis, multiresolution analysis, Central European stock markets, financial crisis.

1. Introduction

Current stock market crisis offers applied researchers new possibilities for testing stock market behaviour from different perspectives. In this short paper we take advantage of this occasion and use wavelet analysis so we can see if it is able to uncover more information about the stock markets.

Wavelet analysis is a powerful mathematical tool for signal processing. Although wavelet analysis has recently shown diverse applications in many fields of research, it has received little attention in econometric analysis of financial data. The few authors dealing with this area of research are Vuorenmaa, Tommi A. (2005), Vacha and Vosvrda (2007) and Gallegati M., Gallegati M. (2007). In particular, the discrete wavelet transform is very powerful in decomposing time series into an orthogonal set of components associated with both time and scales (frequencies). Examining the relationship between high frequency and low-frequency fluctuations in stock returns of different countries, we can investigate the different behaviour that cannot be extracted using common econometric analysis. Moreover, an application of wavelet multiresolution analysis allows us to see even more deeply into the market behaviour structure. Main purpose of this paper is to use this analysis for comparison of Central European stock markets represented by Prague, Budapest, Warsaw and German indices, and U.S. market. We will compare decomposed signals of these markets during the current financial crisis, moreover we will compare evolution of their energies in time, which will allow us to see the possible differences in the behaviour.

The paper is organized as follows. We begin with the methodology description where we present wavelet analysis. Right after, we use the described methodology on the Central European Stock markets data and conclude the results.

[1] Support from the Czech Science Foundation under grant 402/09/H045, 402/08/P207, support from the Grant Agency of Charles University (GAUK) under project 46108, and Institutional Support by the Department of Education MSMT 0021620841 is gratefully acknowledged.

2. Wavelet Analysis

This section briefly introduces wavelets and wavelet variance. Wavelets are small waves that begin at a finite point in time and die out at a later finite point in time. This feature makes wavelike functions ideal for local approximation and time scale decomposition of time series under investigation. The time scale decomposition helps to recognize relationships between economic variables on the disaggregate (scale) level rather than at an aggregate level. Unlike the Fourier analysis, wavelets are suitable for detecting regime shifts, discontinuities and frequency changes. These features makes wavelets powerful tool for investigating the financial markets during current financial crisis. For a more comprehensive analysis of the topic see Gencay et al. (2002), Abramovich et al. (1999), Percival, Walden (2000).

There are two wavelets which form a pair in a wavelet family: father wavelets $\varphi(.)$ and mother wavelets $\psi(.)$. The father wavelet (scaling function) integrates to unity and is used for the trend components; on the other hand, the mother wavelet integrates to zero and is suitable for detection of all deviations from the trend. The mother wavelet is compressed or dilated in time domain, to generate cycles to fit the actual time series. The formal definition of wavelets is

$$\varphi_{j,k}(t) = 2^{-\frac{j}{2}}\varphi\left(\frac{t-2^j k}{2^j}\right), \quad \psi_{j,k}(t) = 2^{-\frac{j}{2}}\psi\left(\frac{t-2^j k}{2^j}\right) \tag{1}$$

where j is the scale (or dilatation) and k is the translation (or shift). Commonly, many types of wavelets can be possibly used, including Haar wavelet, Mexican hat, Morlet wavelet, Daubechies wavelet etc.

Any time series $x(t)$ can be built up as a sequence of projections onto father and mother wavelets indexed by both j, the scale, and k, the number of translations of the wavelet for any given scale, which is assumed to be dyadic. The wavelet coefficients are approximated by integrals

$$s_{J,k} \approx \int_{-\infty}^{\infty} x(t)\varphi_{J,k}(t)dt, \quad d_{j,k} \approx \int_{-\infty}^{\infty} x(t)\psi_{j,k}(t)dt, \tag{2}$$

$j = 1,2,...,J$, where J is the maximum scale. An important feature of a wavelet analysis is the possibility to decompose a time series into its constituent multiresolution components. The multiresolution analysis (MRA) of a time series $x(t)$ in $L^2(R)$ is given by the following formula:

$$x(t) = S_J + D_J + D_{J-1} + D_{J-2} + \ldots + D_1, \tag{3}$$

where

$$\begin{aligned} S_J &= \sum_k s_{J,k}\varphi_{J,k}(t), \\ D_j &= \sum_k d_{j,k}\psi_{j,k}(t), \quad j = 1,2,...,J, \end{aligned} \tag{4}$$

where the basis functions $\varphi_{J,k}(t)$ and $\psi_{j,k}(t)$ are assumed to be orthogonal. The sequence of terms S_J, called smooth and wavelet details $D_J, D_{J-1},...,D_1$ represents a set of signals components that provide representations of the signal at the different resolution levels.

2.1. Wavelet Variance

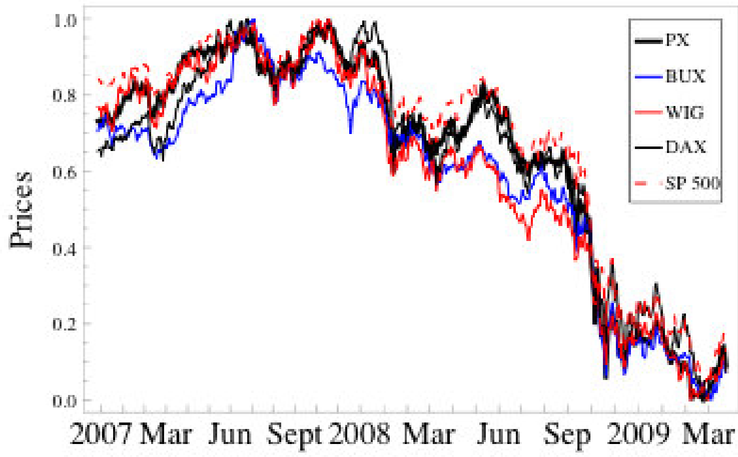
The wavelet variance is a concise alternative to the power spectrum based on the Fourier transform and is often easier to interpret than the frequency-based spectrum Percival, Walden (2000). For computation of the wavelet variance we use multiresolution components from time series decomposition. Such analysis helps us to track an evolution of the energy contribution at various scales, which is related to traders' investment horizons.

$$\|x\|^2 = \sum_{j=1}^{J}\|D_j\|^2 + \|S_J\|^2, \tag{5}$$

where the term "energy" means the sum of squared coefficients of a vector, i.e., $\sum_{t=0}^{N-1} x_t^2 = \|x\|^2$. In our analysis we are mainly focused on wavelet details vectors D_1, D_2, and D_3 which represents highest frequencies and energies of the examined time series, because these three scales (highest frequencies) have major energy contribution. For the MRA we use six scales decomposition ($J=6$) with the Daubechies *"db8"* wavelet filter.

3. Results

In this section we will apply the methodology of wavelet analysis described in previous text to a real world data set consisting of Central European, German and U.S. stock market indices. We use the sample of 512 daily prices from 20.12.2006 until 31.4.2009 of value-weighted indices PX, BUX, WIG, DAX and SP500, representing an approximation of Prague, Budapest, Warsaw, German and U.S. stock markets. *Graph 1* shows normalized prices of all indices from the sample. The prices are normalized to [0,1] interval. Reader can notice that the sample includes the current financial crisis of 2008. As these markets has different holidays and trading schedules, we use dummy variable for these days so we can exclude them from the analysis in order to be sure that each observation corresponds to the same day in whole sample of all five tested indices. After matching the daily observations we transform the prices into continuously compounded index returns.

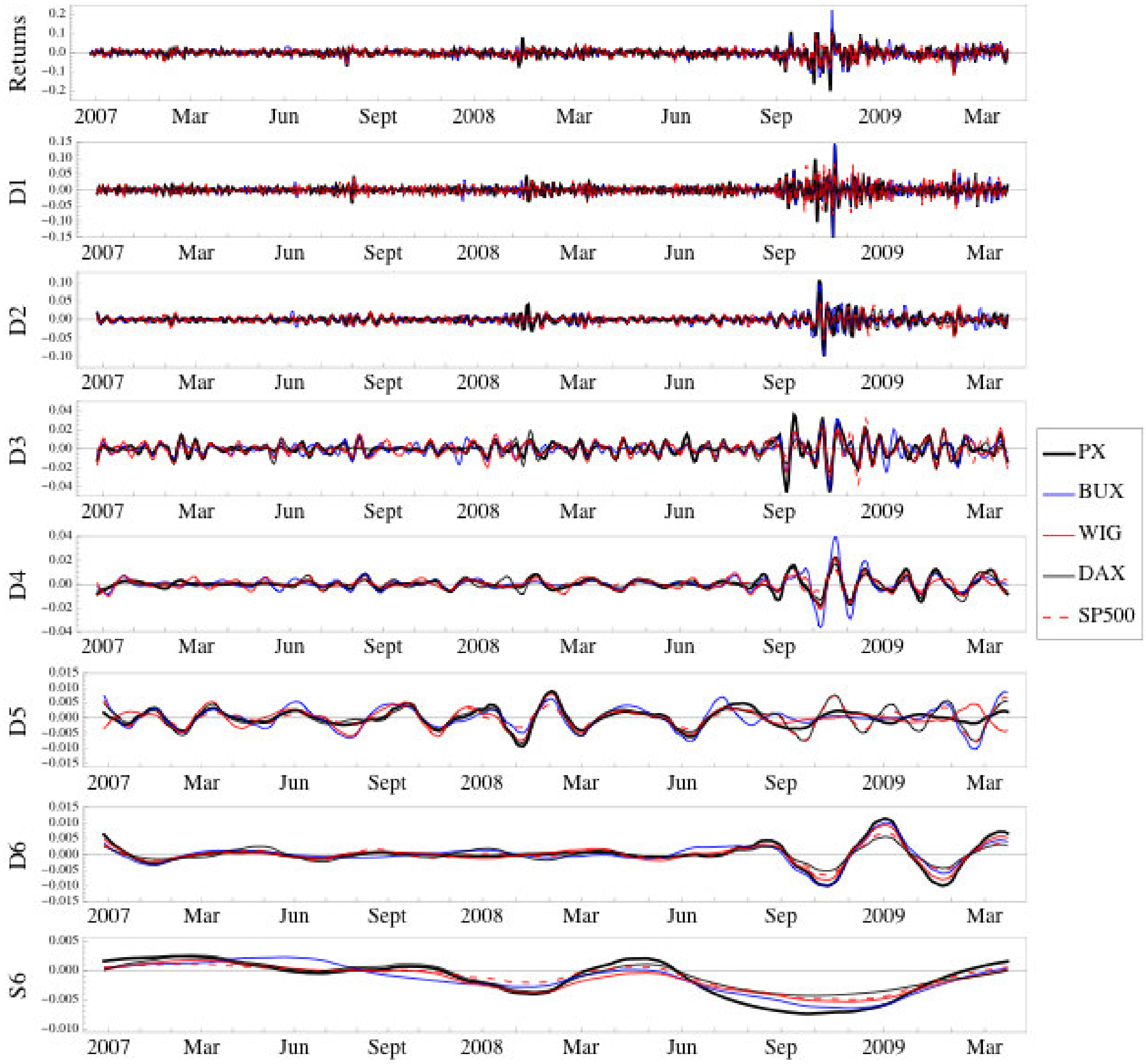


Graph 1: Prices of all series normalized to [0,1] interval

We begin the analysis with the discrete wavelet decomposition of all time series and the multiresolution analysis (MRA) with use of the Wavelet package for Mathematica written by Ian McLeod. *Graph 2* depicts the returns and the wavelet MRA of all five market indices. We use six-scale decomposition ($J=6$) from *Equation 3* with Daubechies *"db8"* wavelet filter. As we would like to compare all five indices, we use one illustrative plot. Reader should note different range of *y-axis* for all decomposed signals, as it is crucial to understand that higher frequencies, D_1 and D_2 represents much more energy than lower frequencies. It is interesting that the volatility of all signals substantially increased during the end of 2008 during the biggest drops of all stock markets. Thus we pick the greatest 2-week loss of whole tested period in October 2008, more precisely 6.10.2008 - 21.10.2008, and compare the contribution of each wavelet energy scale to the signal with the whole period of the sample. Increased volatility of all frequencies during this period indicates very uncommon situation for all stock markets, as all energies during this period seem to contribute to the signal much more than in other periods. Thus we will follow with the analysis of wavelet variance, which we use for analyzing the energies at different levels.

For the analysis of energies, we only compare D_1, D_2, and D_3 as they contribute to most of the variance of the signal. Lower frequencies are moreover long-term frequencies and have deniable contribution to the signal. *Graph 3* shows the sum of energies for wavelet detail vectors for whole tested period. During the whole period the highest frequency D_1 clearly dominates for all countries. There are also significant differences between markets. The highest percentage of D_1 has the US benchmark index SP500, followed by WIG and PX.

Graph 4 depicts the sum of energies for wavelet detail vectors D_1, D_2, and D_3 for all five examined market indices for the specific two-week large drop of October 2008 we have chosen because of the largest variance across all frequencies. Main feature is higher percentage of energy on scale represented by D_2. For BUX, DAX and PX the energy on D_2 was dominating in the two-week crash period. WIG and SP500 have D_1 dominant even in the short two-week period, but D_2 contributes with high percentage of variance also. It is surprising result mainly in comparison to the whole period.

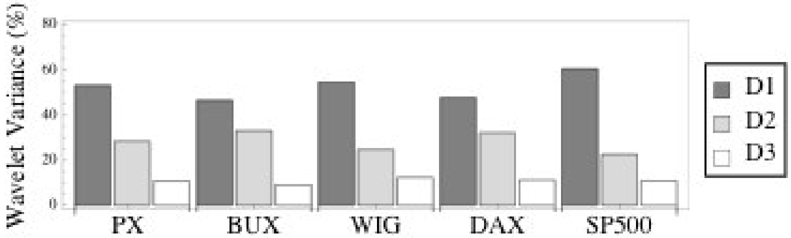


Graph 2: MRA of all series for whole tested period.

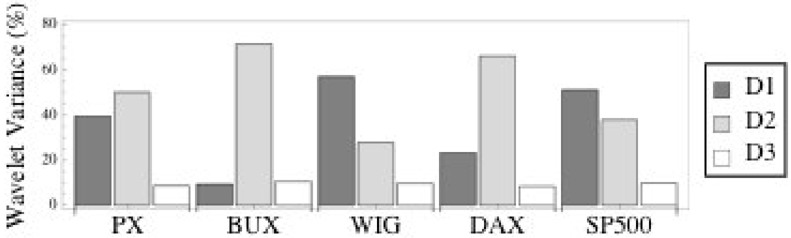
Hence *Graphs 3 and 4* give a clear comparison of market behaviour during the whole examined period and the short period of major market collapse. The highest frequency D_1 representing short-term variations has most important contribution to the overall variance of the series during the whole period (53%, 46%, 54%, 47%, 60% for PX, BUX, WIG, DAX and SP500 resp.), while D_2 which account for variations at a time scale of 4 days (2^2) has lower explanatory power (28%, 33%, 24%, 32%, 22% for PX, BUX, WIG, DAX and SP500 resp.), and D_3 (10%, 9%, 12%, 11%, 10% for PX, BUX, WIG, DAX and SP500 resp.), D_4, D_5, D_6 and S_6 accounts for the rest of the energy. This indicates that stock market movement is driven mainly by short-term fluctuations during the crisis. In contrast, the major two-week drop during October 2008 shows very different behaviour as D_1 represents the highest variance only for WIG and SP500. More precisely, it is 39%, 9%, 57%, 23%, 51% for PX, BUX, WIG, DAX and SP500 resp, while D_2 energy is much stronger for PX, and mainly BUX and DAX (50%, 71%, 27%, 66%, 37% for PX, BUX, WIG, DAX and SP500 resp.).

Thus during the period of the largest two week drop, stock markets behaviour changed significantly. Lower frequency component D_2 plays more important role for all CEE countries except of WIG including DAX, thus market is mostly driven by 4 days fluctuations. All markets reactions are also very different during this period. We can conclude that the SP500 seems to be most efficient as the highest frequency component explains most of the variance also during this period of large drops. On the other hand, component D_2 also plays important role in comparison with whole period. In this manner WIG seems to behave similarly to SP500 and holds its efficiency also during the two-week period. PX and BUX markets does not seem to hold efficiency during this short period, as the highest energy contribution comes from 4 day frequency. Surprise is DAX as it also behaves strongly inefficiently on the contrary to its counterpart SP500.

We have to remind the reader that these differences between markets can also be caused to some extend by the different structure of the tested indices. SP500 has very broad base of 500 stocks, while other indices contain different industry stocks, thus this debate about efficiency should be addressed with caution, as more rigorous analysis needs to be carried for stronger conclusion. This is albeit not the purpose of this short paper. Overall, the higher D_2 contribution is probably caused by the very pessimistic mood at all world markets, when short-term anticipations were negative.



Graph 3: Wavelet Variance of D1, D2, D3 for all tested countries for whole tested period



Graph 4: Wavelet Variance of D1, D2, D3 for all tested countries for the period of 6.10.2008 - 21.10.2008

The last part of our analysis is devoted to testing the differences of energies contribution across the markets. From the *Graph 3* we concluded that the short-term variations represented by D_1 frequency play most important role for all stock markets. Still, the question whether the stock markets are moving together during the crisis at all frequency levels remains to be answered.

For this we use variance equality tests, which evaluate the null hypothesis that the variances in all subgroups are equal against the alternative that at least one subgroup has a different variance. For general discussion on variance testing see Conover et al. (1981). More precisely we use Levene test, which is based on an analysis of variance of the absolute difference from the mean. The F-statistic for the Levene test has an approximate F-distribution with numerator degrees of freedom and denominator degrees of freedom under the null hypothesis of equal variances in each subgroup (Levene, 1960).

Levene test strongly rejects ($p < 0.01$) the null hypotheses of equal variances of D_1 frequencies of all testes markets, as well as it strongly rejects the null hypotheses of equal variances for all other frequencies. This result tells us that various frequencies across different countries are not the same. In other words, if we decompose the tested stock markets using MRA and compare the markets in the means of the frequencies, we arrive to the result that they have significantly different variances; hence they have significantly different energy contributions. This rigorously proves also our expectation from the *Graph 3*, where we could see the differences between frequencies across the countries.

4. Conclusion

In this short paper we applied a wavelet analysis for comparison of Prague, Budapest, Warsaw, German and U.S. stock markets. We use multiresolution analysis to decompose tested stock market indices into different frequency components. Moreover we compute variances of all frequencies, which represent energies contributions, and we use it for comparing the markets.

Looking at wavelet multiresolution analysis of all five stock market returns we can immediately see increased variance during the end of the year 2008 across all frequency components. This corresponds to the larges two-week drop from the tested period. Thus we compute variances of all frequencies for the whole period as well as for this short two-week period of large consecutive losses and we find out significant differences. For the whole period the highest frequency D_1 accounts for the highest energy in all five stock markets, while energy at lower four-day D_2 frequency is much lower. On the contrary, the same analysis on the short two-week period of large consecutive losses during October 2008 shows much different results. For PX, BUX and DAX markets, D_2 frequency has the highest energy, which indicates that the stock markets became highly inefficient during this period. For WIG and SP500 returns, D_1 still accounts for the highest energy, but also D_2 frequency remains at high levels. Moreover, we found differences of various energies between the compared stock markets.

Hence we showed that behaviour of the five tested stock markets differ across various decomposition levels during the current financial crisis. The view of a wavelet analysis allowed us to test the decomposed series and see exactly the contributions of each scales to the energies of the markets. This short paper is a pilot study of wavelet energies of Central European Stock Markets.

References

Abramovich F., Bailey T.C., Sapatinas T. (1999), *Wavelet Analysis and its Statistical Applications*, J. R. Statist. Soc. D, 48.

Conover, W. J., M. E. Johnson and M. M. Johnson (1981). A Comparative Study of Tests for Homogeneity of Variance with Applications to the Outer Continental Shelf Bidding Data, *Technometrics*, 23, 351–361.

Gallegati M., Gallegati M. (2007): *Wavelet Variance Analysis of Output in G-7 Countries*, Studies in Nonlinear Dynamics & Econometrics, Berkeley Electronic Press, vol. 11(3), pages 1435-1435.

Gencay, R., Selcuk, F., Whitcher, B. (2002), An Introduction to Wavelets and Other Filtering Methods in Finance and Economics. Academic Press.

Levene, H. (1960). “Robust Tests for the Equality of Variances,” in I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow, and H. B. Mann (eds.), *Contribution to Probability and Statistics*, Palo Alto, CA: Stanford University Press.

Percival, D. B., Walden, A. T. (2000), Wavelet Methods for Time series Analysis. Cambridge University Press.

Vuorenmaa, Tommi A. (2005) , A Wavelet Analysis of Scaling Laws and Long-Memory in Stock Market Volatility. Bank of Finland Research Discussion Paper No. 27/2005

Vacha. L., Vosvrda, M. (2007), *Wavelet Decomposition of the Financial Market*, Prague Economic Papers, no. 1,vol. 16,, pp. 38-54, ISSN 1210-0455

COOPERATIVE BEHAVIOR AND ECONOMIC GROWTH

Jaromir Baxa

Charles University in Prague, Faculty of Social Sciences, Institute of Economic Studies
Institute of Information Theory and Automation of the AS CR, Department of Econometrics,
jaromir.baxa@centrum.cz

Abstract. We investigate the importance of cooperative behavior for economic growth in simulated economies. Along with the other literature, the cooperative behavior and the ability to enforce cooperation are the key factors for long-term sustainable economic growth in our simulations. Interestingly, the effect of enforcement and punishment of piracy was not always positive: Introducing such mechanism caused elimination of the most successful agents without the positive effects on cooperation and productive economic activity. Hence, the income was lower for low enforcement rate than for the economies without any mechanism supporting cooperation. Similar effects occurred in the simulations of institutional change. In case of a discontinuous change, a radical enforcement mechanism was implemented in one point of time and it caused a sharp fall of wealth. Nevertheless, after some time the positive effects of cooperation dominated and economic growth emerged.
Keywords. Cooperation, iterated prisoner's dilemma, economic growth, institutional change, agent-based modeling

1. Introduction

Persisting cross-country differences in economic performance still pose a challenge to theories of economic development. Traditionally, there was an implicit assumption that all countries have the same growth trajectories and that the only difference between the developed and the underdeveloped countries is in their current stage of economic development.

The income gap, however, has not been closing but has widened even more during recent decades. Consequently, many researchers have started to ask whether there are any fundamental differences between rich and poor countries that can account for such persistence in the income gap. Following the tradition of institutional economics, many researchers believe that the key might be in different institutional structures that shape the direction and form of economic activity in these countries. For example, Douglass North, [4] argues that the inability of societies to develop effective low-cost enforcement of contracts is the cause of long-lasting stagnation and the current underdevelopment in many countries (on the role of trust and the culture of cooperation see [5], for example).

The underlying hypothesis behind the institutionalist point of view is as follows:

> "Without enforcement of property rights, participation in some market activity is discouraged by the prospect that anyone engaging in such activities is unlikely to receive its full benefits. Any expropriation of the proceeds of market activity by dishonest parties to a contract, bandits, or corrupt government officials is therefore likely to reduce incentives and opportunities for production, investment, and innovation" ([2]).

Gradually, as the share of the population involved in productive activities decreases, a different set of abilities and knowledge linked with predatory activities and piracy emerges and spreads through society. Redistribution of wealth starts to dominate its creation and steady stagnation begins.

In this paper, we investigate the importance of the ability to enforce cooperation and for economic growth and the effects of institutional change within a simulated economy using the framework of the agent based computational economics. The paper is organized as follows. Section 2 contains a description of the simulations, section 3 presents our results with constant ability to enforce cooperation and the last section discusses the effects of institutional change.

2. Simulation Design

The model is constructed as follows. We assume an initial population of agents living in an environment provided with an initial level of "natural" resources. These resources are a source of energy for the agents. They are assumed to be partially renewable and the speed of renewal influences how easy life is for the agents in their environment. They might be linked to grain or to any other potential resources for redistribution, for example.

To survive, agents need to acquire energy continuously. It can be acquired either from the environment directly (utilizing pieces of resources) or through interactions with other agents in order to produce or trade their goods. In each period, agents are allowed to move one step around. If they fail to find any resources or any other agent, one unit of their energy is lost. Hence, all the agents' effort is directed at getting enough energy to survive and not to starve. If the conditions are good they reproduce, whereas if their energy falls below zero, they die.

All agents are allowed to live infinitively. The only condition they have to satisfy is that their energy must always be strictly positive. Three types of agents were generated in our model: cooperating producers, predators and, finally, a number of random agents that mix the two basic strategies randomly (in 50% of the iterations they behave like producers and in the remaining 50% they behave like predators). All agents insist on their strategy for their whole lives. The population is growing at a rate that determines the number of new agents that invade the environment every 100 periods. These new agents choose their strategy randomly, with the same probability of choosing any of the set of strategies.

The interactions follow the simple Prisoner's Dilemma (PD) scheme. If producer meets another producer, they cooperate and share their joint profit equally. If producer meets predator, predator expropriates the whole production and finally if two predators meet, the result correspond to the Nash equilibrium of the PD game. The success of each strategy is reflected in the number and scores of agents following that strategy. Those who are unsuccessful lose their energy continuously and die out, whereas successful agents are able to acquire enough energy in environments with almost all resources consumed. This results in population dynamics in which the number of agents pursuing successful strategies steadily increases as both old and new agents survive. Therefore, the population dynamics replace the learning at the individual level.

The simulations were run in the NetLogo environment ([6]). The code or a complete NetLogo file can be sent by the author upon request. All simulations have taken about 1 minunte and 30 seconds.

The simulations differed in various aspects. First, the initial population might differ in size and in the shares of agents with their strategies. Furthermore, the environment might be either rich or poor in natural resources. Finally, agents might be able to detect those who defect rather than cooperate (the predators), and punish them. This was implemented as an exogenously given probability of detection here. In the case of detection, the predator is punished by a penalty of 50 units of energy. This size was chosen arbitrarily; usually it was high enough to cause the death of the punished agent.

3. The Results

In line with the intuition, growth occurred in simulations where cooperative behavior prevailed. Moreover, cooperation prevailed only when enforcement of cooperative behavior and punishment of predators was present. The relationship between welfare achieved and detection probability is summarized in Figures 1 and 2, which show the average welfare after 10,000 iterations for different detection probability values. Box-plot representation was chosen as it allows us to illustrate the distribution of the resulting values for all simulations with identical settings (we ran 30 simulations each time; statistical significance of increasing or decreasing changes of welfare with varying detection rate were tested using the Wilcoxon rank-sum test).

The main finding is that, in general, the effect of enforcement on welfare is positive and statistically significant when energy from the environment is similar to the potential benefits of economic activity with other agents. Interestingly, the dynamics from low income states to high income states are ambiguous for low detection probabilities. First, introducing an enforcement mechanism represented by the detection rate causes a decrease in welfare, as the wealth of predators is lost and the share of producers is not much affected by the change. On the other hand, a more radical increase of the detection rate has a clear positive effect on wealth.

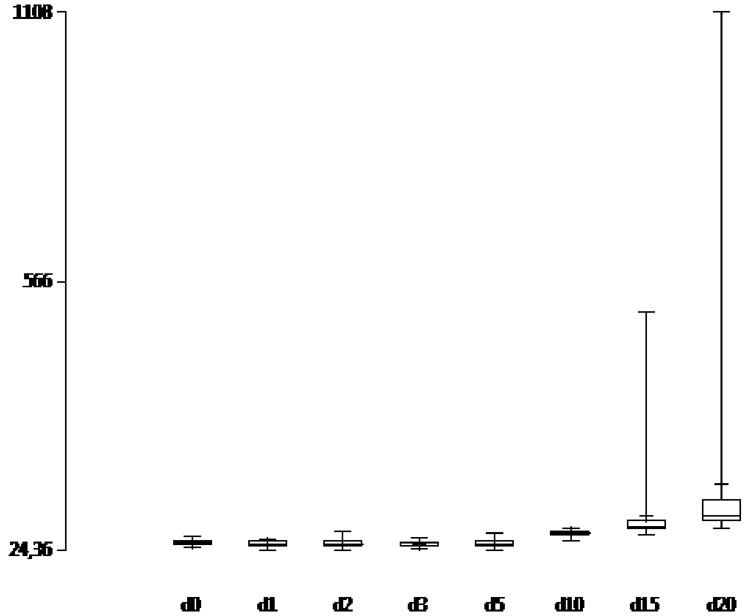


Figure 1: Average Welfare and Detection Rates – Poor Environment

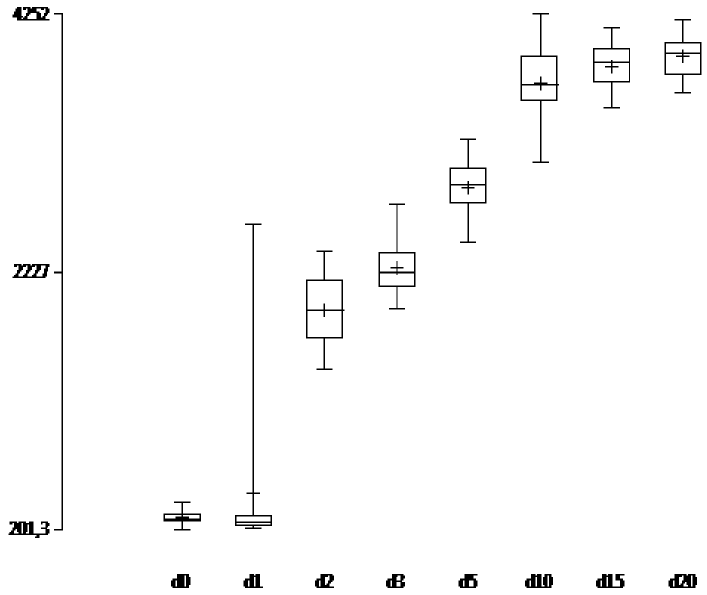


Figure 2: Average Welfare and Detection Rates – Rich Environment

As a matter of fact these simulations generate a J-curve that changes into an S-curve with increasing detection probability. The shape of the relationship between average wealth and the detection rate implies a dilemma of punishing the most successful members of society, and thus at the very beginning the newly established enforcement mechanism has a negative effect on welfare. This can complicate the shift from a "closed-eyes" policy, regardless of the uncertainty about the future effects of such a shift and the endangered status of the elites, which were mentioned before.

The relative importance of these two effects depends on the energy that can be acquired from the environment. The evidence for a decreasing effect is stronger in poor environments than in good ones. In the poor ones, the average wealth is rather small and the differences between old and new agents are not significant. If the detection rate is small, the proportion of producers generated is highly volatile, because it is influenced by new agents with predatory behavior. This volatility causes differences in the timing of growth and thus the observed average scores after 10,000 iterations exhibit the fat tails that occur in Figure 1. The volatility decreases when the detection probability exceeds 25%. This illustrates that communities living in an unpropitious environment have to be stricter in enforcing cooperative behavior in order to succeed and prosper.

For more favorable conditions, the observed trajectories followed the pattern presented in Figure 2. At low detection probabilities (2% for rich environment) cooperation starts to dominate piracy, although rent-seeking behavior is not eliminated. With a detection probability of 10% piracy dies out: after about 1,000 iterations the proportion of cooperative behavior exceeds 90%, and from that point on the population of pirates consists almost exclusively of new agents.

Our results were fairly robust over various parameterizations and even various initial distributions of agents. For the purpose of this paper all the simulation started with the population of 30 agents with 10 of each strategy. Details about the parameterization as well as discussion about the sensitivity analysis can be found in [1].

4. Institutional Change

The final simulations explore the process of institutional change. Suppose that the simulation starts without any enforcement mechanism. After 1,000 periods, a new enforcement mechanism is introduced and from that point on the external conditions is set in order to favor production against piracy. At this time, the average score of all the strategies is still growing and the predators' average payoff is about twice as high as the payoff of producers and random agents. However, its growth rates are gradually decreasing. Continuation of the same institutional setup will lead to stagnation and productive activities will be eliminated. Here, we abstract from cognitive aspects such as "How do they know that if they don't adopt any mechanism protecting producers they will face long-lasting stagnation?" (extensive discussion about this cognitive aspects of intended change of behavior see in [3]) For simplicity, it is assumed that similarly to the external nature of the enforcement mechanism, its implementation is given. Then the resulting dynamics are explored.

In line with the existing literature on institutional change, two types of institutional change are considered. The first case, presented in Figure 3, describes a radical discontinuous change. In this case, at time 1,000 the detection probability jumps from 0% to 10%. In response to this change, non-cooperative predators face important losses. First they lose their wealth, and then their number decreases, too. Following this change, the share of agents playing "Cooperate" increases from 55% to 80% within 500 iterations after the change. Nevertheless, the effects on output are devastating. In the particular simulation corresponding to Figures 3 and 4, the average score falls from 231 to 179 in 100 periods because of the falling average score of agents playing "Defect" and "Random." The growth of producers' income that could compensate for this fall starts after the next 200 periods, around 300 iterations after the change. From that point on, both the average payoff of producers and their number gradually increase. The total average score slowly recovers. After 800 iterations following the institutional change, the average score exceeds its previous level (at the time of the change).

Figure 3: Discontinuous institutional change

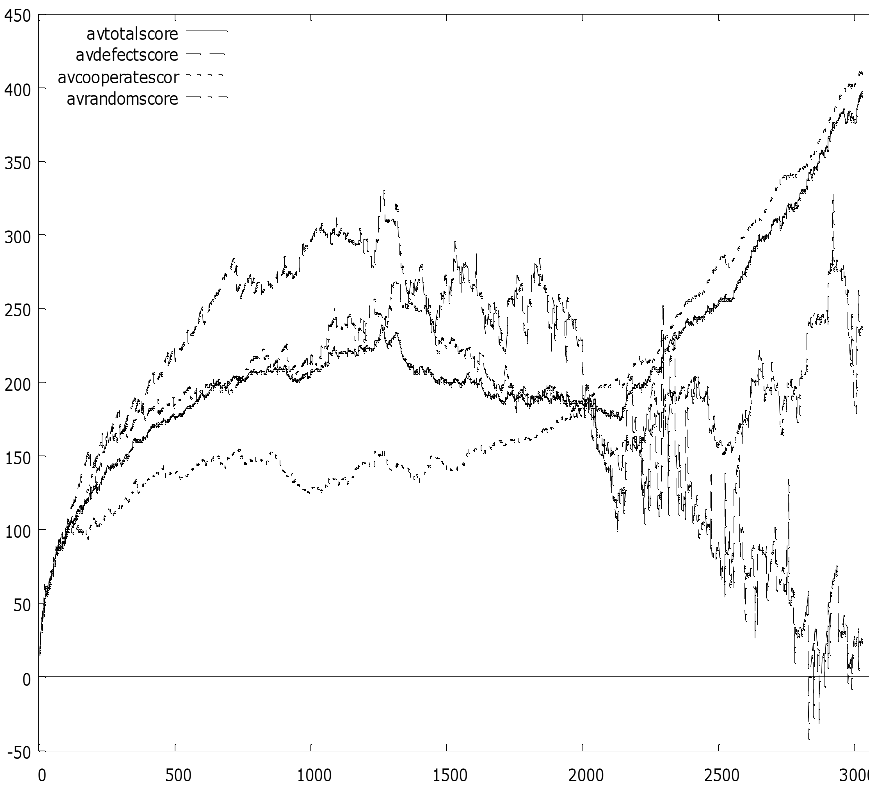


Figure 4: Continuous institutional change

The second institutional change was continuous. At time 1,000, a new rule was adopted and the detection probability was increased to 1%. After 100 iterations, the situation was repeated and the detection probability was increased again by 1% and so on up to the point when it reached 10%. The situation is shown in Figure 4. During hundreds of iterations nothing happens and the simulated economy seems to follow its original path. Also, the share of cooperation doesn't change significantly. Shortly after the detection rate increases to 3%, the average output falls. Again, most of this fall is related to agents who follow predatory and random strategies. Then a period of stagnation follows and the average output gradually decreases between periods 1,500 and 2,100. Later on, the trend reverses and the economy switches to a growing trajectory. At this time, it is 200 periods after the detection rate achieved the final rate of 10%, but at the same time the probability of cooperation exceeds 80%.

Clearly, the shift to an economy based on production is quite costly. The costs are distributed unequally and most of them are levied on predators, the former elites. One might object that this fall is not very realistic because it is a consequence of the rather simplistic assumptions of the model, namely, the inability of individuals to learn. On the other hand, learning and acquiring new knowledge takes some time, and at the organization level it is often a difficult and costly process connected with various risks. Hence, it is hard to expect quick adjustment to the new conditions. Moreover, the population dynamics together with the high number of iterations compensate for the lack of learning at the individual level.

5. Conclusions

This paper addresses these aspects of production and predation, economic growth, and institutional change explicitly within a framework of an agent-based economy. The main finding from the simulations is that no matter what the external conditions (opportunities for redistribution) are, productive activities based on cooperation among agents are the key source of growth. In some specific cases, production need not dominate predation, but without producers the income of the other agents stagnates. On the other hand, the worse the environment, the higher the need for cooperation for sustainable economic growth. However, the payoffs of the interactions were supposed to follow the Prisoner's Dilemma game, hence some enforcement mechanism that punishes predation was necessary to make production attractive and persist over time.

The simulations also show that the effects of adopting such an enforcement mechanism are mixed. For low enforcement rates, the effect on income was even negative: the most successful agents were the predators and they were punished. Yet the number of cooperative opportunities, expressed as the number of producers, was still very low to generate income high enough to compensate for the predators' loss. These results occurred in the simulations where the ability to enforce cooperation was constant or time varying (simulations of institutional change caused by an external change in enforcement ability). Thus, the fears of change that might have been perceived by the most successful agents, the elites, came true. The recovery came after a community of producers emerged; the delay was influenced mostly by the speed of inflow of new agents. Nevertheless, the effect of the change on producers' income was positive right from the very beginning.

Acknowledgements

I would like to thank T. Cahlík and M. Vošvrda for helpful comments. However, the usual disclaimer applies. Also the support of the Czech Science Foundation, grant no. 402/07/0890, the Institutional Support by the Department of Education MSMT 0021620841 and the Grant Agency of the Charles University in Prague (GAUK grant no. 139107) is gratefully acknowledged.

References

[1] BAXA, J. (2009): *Cooperative Behavior and Economic Growth.* Research Report, Institute of the Information Theory and Automation, Czech Academy of Science, June 2009.

[2] DABLA-NORRIS, E., Freeman, S. (2004). Enforcement of Property Rights and Underdevelopment. *Review of Development Economics*, 8:3, pp. 391-405

[3] MANTZAVINOS, Ch. (2001). *Individuals, Institutions and Markets.* Cambridge: Cambridge University Press.

[4] NORTH, D. C. (1990). *Institutions, Institutional Change and Economic Performance.* Cambridge: Cambridge University Press.

[5] PUTNAM, R. (1995). Bowling Alone: America's Declining Social Capital. *Journal of Democracy*, 6:1, pp. 65-78.

[6] WILENSKY, U. (1999). *NetLogo*. http://ccl.northwestern.edu/netlogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, IL.

Financial Development and Allocation of External Finance*

Jan Bena[1] and Peter Ondko[2]

[1] Financial Markets Group, London School of Economics and Political Science, London
J.Bena@lse.ac.uk
[2] Center for Economic Research and Graduate Education - Economic Institute (CERGE-EI), Praha
Peter.Ondko@cerge-ei.cz

Abstract. We investigate the role of financial system to facilitate efficient allocation of resources into perspective projects. Employing micro-level data on European firms in 1996-2005 period, we provide evidence that firms in industries with the best growth opportunities use more external finance in financially more developed countries. Our results are robust to the choice of proxy for growth opportunities and to controlling for technology determinants of external finance. Furthermore, we show that the importance of technological factors decreases once growth opportunities are controlled for, which confirm hypothesis that used measures of technological determinants of external finance are partly driven by growth opportunities.
Keywords. Financial development, external finance, allocation.

1 Introduction

The key role of a financial system is to acquire information about investment opportunities and facilitate the efficient allocation of resources into viable projects.[3] To perform this fundamental function, a financial system ought to identify firms with positive growth opportunities and channel external finance (EF) towards these firms.

Relying on industry-country data, recent empirical work presents an indirect evidence of such positive capital allocation function provided by more developed financial markets. [6] estimates the effect of financial development (FD) on the elasticity of aggregate investment with respect to growth opportunities. [3] measure the effect of FD on growth of industries with positive opportunities.

In this paper, we employ micro-level data to directly measure the effect of improved capital allocation function of more developed financial markets on firms. We show that the firms in industries with the best growth opportunities use more EF in more financially developed countries. Specifically, our regressions ask whether firms operating in industries with positive productivity shocks were able to respond to these new opportunities by increasing their external financing in countries with high levels of financial markets development. The focus of our analysis is on manufacturing sectors of a homogenous set of European countries with highly synchronized product markets where the key underlying assumption of global shocks to industry growth is arguably most likely to hold.

Thanks to the access to the detailed balance sheet data, we approximate the actual amount of external financing used by firms which makes us able to directly test the ability of financial markets to allocate external resources to the projects with growth prospects. We thus improve upon [6], who uses gross fixed capital formation on industry-level as the dependent variable and, thus, he investigates overall impact of the financial system on the allocation of capital, irrespective of the source of financing being external or internal.

Further, the proxy for industry growth opportunities [6] uses is actual growth of value added in a given industry-country pair [4]. In a country with perfectly developed financial market, up to unanticipated shocks, the realized growth fully reflects growth opportunities. If latent industry opportunities are positively autocorrelated, then using current growth to approximate for future growth opportunities is justified. However, it is less clear if this mechanism hold in countries where opportunities anticipated in the past aren't reflected in current growth due to financial or labor market frictions, which cast doubts on the use actual growth in value-added as a proxy for a future growth. Therefore we take a different approach and use two proxies for global component of industry-level growth opportunities that are common across countries: (i) industry value-added growth in the US and (ii) the change in global industry PE ratios. The former is based on the assumption that firms in financially the most developed countries, such as the US, are able to materialize growth opportunities caused by demand and productivity shifts, which are partly global. The later is justified if financial markets are sufficiently integrated to the extent that the common component of growth opportunities is priced in global industry portfolios[5].

* Partial financial support from GACR grant No. 402/09/1595 is gratefully acknowledged)
[3] See survey by [4].
[4] Using data available for US, he justifies the suitability of this proxy by its significant positive correlation with more traditional measures of growth opportunities such as average Tobin's Q, price-to-earnings ratio and sales growth.
[5] See [2].

We find that FDimproves the allocation of capital by channeling external finance to firms that operate in industries with high growth prospects. Both proxies yield estimates of similar economic significance and the economic effect is more than 4 times larger if we use the IV technique to correct for a measurement error in growth counterfactuals.

We also contribute to the literature on finance-growth nexus. Its main argument is that technology employed by firms is constant across industries and determines industry external finance dependence ([5]). We show that the ability to provide EF to firms in industries with the best opportunities still holds when we control for technology determinants of EF interacted with financial development. Furthermore, we find that the importance of technological determinants of EF use decreases by about 10 to 50 % once growth opportunities are controlled for. This suggests that the widely used measures of technological determinants of EF dependence are partly driven by growth opportunities that were financed and hence realized in countries with high FD(such as the US).

2 Methodology

Our aim is to test the hypothesis that financial development improves allocation of capital by channeling external finance towards firms in industries with best investment opportunities. Thus, our main regression specification is

$$EFU_{fic} = \alpha + \beta \ FD_c * GO_i + \sum_i \lambda_i D_i + \sum_c \lambda_c D_c + \gamma X_{fic} + \varepsilon_{fic}, \tag{1}$$

where EFU_{fic} stands for the average external finance use of firm f operating in industry i in country c for period 1996-2005, FD_c denotes country-level measure of financial development observed before our firm-level sample period, GO_i denotes proxy for industry-specific growth opportunities common that are common across countries, D_i is a fixed effect for industry i, D_c is a fixed effect for country c and X_{fic} is a vector of firm-level control variables.

We measure external finance use by yearly change in shareholder's capital (equity) plus yearly change in non-current liabilities (long term debt and other non-current liabilities) scaled by the total assets as of beginning of the year. We represent firm's external finance use the average of the annual measure over sample period.

We use period average US growth and global PE growth as proxies for GO_i. Due to unbalanced nature of our firm-level panel, we adjust the period used to compute growth opportunities counterfactual for every firm to match its period in the sample. Furthermore, we always control for a set of firm-level variables, measured as of the first year a firm enters the sample [6]. By doing so, we aim to eliminate initial differences in the within-industry distributions of firms along characteristics that have potential impact on the use of external finance. We always include full set of industry and country dummies into our specification.

Further, using our measure of EFU we are ready to directly test if firms operating in industries that are pre-destined for high dependence on external finance due to technological reasons are actually using more external finance in more developed financial markets. We estimate specification similar to 1, where GO_i is replaced by $Technology_i$, the industry-specific technological determinant of external finance use. We consider three candidates. First is external finance dependence (EFD), measured as in [5]. This is an all-encompassing measure of dependence on external finance and it's based on assumption that in highly developed financial market such as US, the industry differences in observed proportion of capital expenditures financed from external sources reflect underlying technological differences among industries. Second is R&D Intensity[7], because R&D Investments are often relatively large at the startup of firms or projects and they may be associated with longer gestation periods as it is likely that profits from R&D materialize later. Lastly, Investment Lumpiness is a proxy for the degree of mismatch between cash inflows and cash outflows[8].

The proxies for technological determinants of finance are calculated on US data over the period under investigation and, thus, they may as well be capturing underlying growth opportunity shocks specific to that period. To verify this, we estimate specification where we interact financial development with growth opportunities as well as with technological proxies,

$$\begin{aligned} EFU_{fic} = \alpha &+ \beta_1 \ FD_c * GO_i + \beta_2 \ FD_c * Technology_i \\ &+ \sum_i \lambda_i D_i + \sum_c \lambda_c D_c + \gamma X_{fic} + \varepsilon_{fic}. \end{aligned} \tag{2}$$

[6] We control for age, dummy for being quoted, stock of cash and measures of firm's leverage, asset tangibility and size

[7] The R&D Intensity is approximated by the average share of R&D expenditures on capital expenditures of a median firm in the US industry

[8] Investment Lumpiness as the average number of investment "spikes" in US industry over given period. "Spike" is defined as a year in which capital expenditures exceed 30 % of fixed assets.

If technology proxies are significantly contaminated by the growth opportunity shocks, we would expect β_2 to be smaller than their counterparts in specification without $FD_c * GO_i$. The magnitude of this decrease should be larger when GO_i is approximated by US Growth, because it controls for US specific shocks which are absent in global PE growth proxy.

3 Data

Firm-level panel data are obtained from 'TOP 200 thousand' version of BvD Amadeus, which contains balance sheet and income statement information for a large population of largest firms spanning whole Europe. The firm coverage is incomplete before 1996 and after 2005 and therefore we use only observations from 1996-2005. We exclude Romania from the sample due to large inconsistencies in the accounting data of its firms. Denmark and Norway have only few firms in the final sample and have been dropped too.

The value-added data for US used to compute our first proxy for growth opportunities are taken from OECD STAN database downloaded in 2009. We use index of volume of value added (VALK) for industries on 2-digit level defined by ISIC rev 3.1. The data for monthly series of global PE ratios are obtained from Datastream. We use Compustat to compute industry-level technological determinants of the need and ability to raise EF.

The country-level measures of FD we use are private credit by deposit money banks and other financial institutions to GDP, stock market capitalization to GDP and stock market total value traded to GDP, which are taken from World Bank Database described in [1]. Total capitalization is sum of credit to GDP and stock market capitalization to GDP. We complement measures of financial depth by a proxy for an institutional quality of financial markets as measured by an Accounting Standards index, which rates annual reports of companies in 1990 according to inclusion of 90 items in their balance sheets[9].

4 Results

We present basic estimates of our main specification of equation 1 in Table 1. Our results suggest that financial development improves allocation of external finance by channeling it to industries with high growth prospects. For example, if Bulgaria's banking sector were as developed as the one in Netherlands, the otherwise comparable firms that operate in an industry ranked at 75th percentile by US growth would use 0.6 percentage points (on average per annum) more external finance in proportion to their total assets relative to the firms that operate in an industry ranked at 25th percentile. When we approximate growth opportunities by global PE growth, we obtain the analogous estimate of 0.4 percentage points. Comparable economic effects obtained for other indicators of financial development range between 0.13 to 0.56 percentage points.[10] Further, If we use 2SLS[11] to alleviate measurement error, the economic effects increase more than 4 times.

In Panels C and D, we confirm that industries technologically dependent on external finance are actually using more of it in financially more developed countries. The results in Panel C confirm the hypothesis of the existence of a common factor in US growth opportunities and technological determinants of EF. The estimated coefficients on interactions of financial development and R&D intensity and EFD drops to almost half once interactions with US Growth are included. However, we actually observe drop in the estimated coefficient on interaction of financial development with US Growth, once corresponding interaction with Investment Lumpiness is included in the specification.

When we consider Global PE Growth as a proxy for growth opportunities, as in Panel D, the estimated coefficients on the interaction terms are statistically significant and very similar in magnitudes to their counterparts in Panel A of Table 1. The interactions of financial development with technological determinants in the joint specification remain statistically significant and the magnitude of the drop in the coefficients is less severe. This evidence suggests that the role of financial development with respect to allocation of external financing is two-fold. On the one hand it helps to channel external finance to industries which are presumably more dependent on it due to technological reasons. On the other hand, more developed financial markets are better in providing finance to the industries with greater global growth opportunities.

5 Conclusion

In this paper, we study allocation function of financial system. Using two alternative proxies for global industry-specific component of growth opportunities, we show that comparable firms with positive growth opportunities

[9] The index is produced Center for International Financial Analysis and Research, Inc.
[10] For the comparison, the sample mean and standard deviation of External Finance Use are 0.4 % and 3.8 %, respectively.
[11] In this method, one proxy is used as the instrument for the other.

Table 1. Financial Development and External Finance Use: Growth Opportunities

	Total Capitalization	Private Credit	Market Capitalization	Market Value Traded	Accounting Standards
			Panel A: Growth Opportunities		
FD * US Growth	0.027***	0.053***	0.038*	0.115***	0.084
	(0.01)	(0.016)	(0.02)	(0.04)	(0.054)
R^2	0.078	0.077	0.077	0.078	0.076
N	25,419	25,544	25,419	25,419	22,239
FD * Global PE Growth	0.026***	0.034***	0.070***	0.088**	0.043**
	(0.008)	(0.013)	(0.018)	(0.039)	(0.019)
R^2	0.078	0.077	0.078	0.077	0.076
N	26,526	26,665	26,526	26,526	23,196
			Panel B: Technology determinants		
FD * R&D Intensity	0.080***	0.147**	0.098*	0.269**	0.111
	(0.028)	(0.06)	(0.058)	(0.12)	(0.269)
R^2	0.077	0.075	0.075	0.075	0.073
N	26,259	24,629	24,517	24,517	21,495
FD * Investment Lumpiness	0.247***	0.473***	0.309***	0.760***	0.43
	(0.058)	(0.134)	(0.11)	(0.236)	(0.504)
R^2	0.078	0.077	0.077	0.077	0.075
N	26,514	26,653	26,514	26,514	23,187
FD * EFD	0.091**	0.158**	0.136	0.305*	0.502
	(0.04)	(0.074)	(0.083)	(0.183)	(0.437)
R^2	0.077	0.076	0.076	0.076	0.075
N	26,259	26,392	26,259	26,259	22,975
			Panel C: US Growth vs. Technology		
FD * US Growth	0.026**	0.050***	0.035	0.111**	0.087
	(0.011)	(0.018)	(0.023)	(0.045)	(0.06)
FD * R&D Intensity	0.048*	0.095*	0.05	0.127	-0.007
	(0.025)	(0.049)	(0.051)	(0.124)	(0.274)
FD * US Growth	0.013	0.028	0.019	0.074*	0.078
	(0.011)	(0.018)	(0.021)	(0.043)	(0.055)
FD * Investment Lumpiness	0.218***	0.410***	0.272**	0.599**	0.296
	(0.061)	(0.139)	(0.112)	(0.251)	(0.509)
FD * US Growth	0.024**	0.047***	0.033	0.104**	0.076
	(0.011)	(0.017)	(0.021)	(0.042)	(0.055)
FD * EFD	0.056	0.094	0.092	0.158	0.388
	(0.038)	(0.064)	(0.08)	(0.192)	(0.436)
			Panel C: Global PE Growth vs. Technology		
FD * Global PE Growth	0.027***	0.035**	0.076***	0.098**	0.047**
	(0.008)	(0.014)	(0.019)	(0.041)	(0.021)
FD * R&D Intensity	0.063**	0.127**	0.048	0.205*	0.079
	(0.028)	(0.059)	(0.055)	(0.122)	(0.27)
FD * Global PE Growth	0.022***	0.027**	0.063***	0.071*	0.043**
	(0.008)	(0.013)	(0.019)	(0.039)	(0.019)
FD * Investment Lumpiness	0.218***	0.440***	0.220**	0.667***	0.384
	(0.06)	(0.137)	(0.11)	(0.24)	(0.503)
FD * Global PE Growth	0.026***	0.034**	0.071***	0.092**	0.046**
	(0.008)	(0.013)	(0.019)	(0.041)	(0.02)
FD * EFD	0.070*	0.131*	0.086	0.226	0.447
	(0.039)	(0.073)	(0.081)	(0.184)	(0.435)

Note: The dependent variable is the time average of annual firm-level External Finance Use (EFU). All specifications are linear regressions with outliers removed (observations outside the 1-to-99 percentile range of the dependent variable), include a constant, firm-level controls and 3-digit ISIC industry and country dummies. In panels C and D, the R^2 and N are similar to their equivalents in panels A and B and they're not reported. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively, based on robust standard errors clustered at the industry-country level, which are reported in brackets.

are more likely to obtain external finance in countries with more developed financial markets. Given that our sample consists of relatively large and well-established firms which are shown to be less affected by financial development, it is likely that economic significance of our result is even larger in overall population.

Our findings complement existing literature which focused on the implications of financial development on capital expenditures and growth. However, the most important role of financial system in the link from growth opportunities through investment to growth, is to provide external finance, test of which is a primary question of our paper. In light of the outlined mechanism, we plan to complement our analysis by a second step in which we check whether the improved capital allocation function of more developed financial markets leads to higher corporate investment and growth precisely through more extensive use of external finance.

References

1. T. Beck, A. Demirgüç-Kunt, and R. Levine, *A New Database on Financial Development and Structure*, World Bank Economic Review, 14 (2000), pp. 597–605.
2. G. Bekaert, C. Harvey, C. Lundblad, and S. Siegel, *Global growth opportunities and market integration*, The Journal of Finance, 62 (2007), pp. 1081–1137.
3. R. Fisman and I. Love, *Financial development and growth in the short and long run*, NBER Working Papers 10236, National Bureau of Economic Research, Inc, 2004.
4. R. Levine, *Finance and growth: Theory and evidence*, in Handbook of Economic Growth, P. Aghion and S. Durlauf, eds., vol. 1, 2005, ch. 12, pp. 865–934.
5. R. Rajan and L. Zingales, *Financial Dependence and Growth*, The American Economic Review, 88 (1998), pp. 559–586.
6. J. Wurgler, *Financial markets and the allocation of capital*, Journal of Financial Economics, 58 (2000), pp. 187–214.

QUALITATIVE STABILITY OF STOCHASTIC PROGRAMS WITH THIRD-DEGREE STOCHASTIC DOMINANCE CONSTRAINT INDUCED BY MIXED-INTEGER LINEAR RECOURSE

Martin Branda*

Charles University in Prague
Faculty of Mathematics and Physics
Department of Probability and Mathematical Statistics
Sokolovska 83, Prague 8, 186 75, Czech Republic
branda@karlin.mff.cuni.cz

Abstract. We deal with stochastic programs with third-degree stochastic dominance constraint induced by mixed-integer linear recourse function. We study qualitative stability of the set of feasible solutions and the optimal value function with respect to weak convergence of underlying probability measures.
Keywords. Stochastic optimization models, qualitative stability, mixed-integer value function, third-degree stochastic dominance.

1 Introduction

Stochastic dominance represents a very useful tool in decision making. It is related to the expected utility theory, where stochasticaly dominating random variables are optimal with respect to a set of utility functions with certain properties, [11]. Third degree stochastic dominance was introduced in [12] as a natural extension of stochastic dominances of lower orders.

In this paper we extend some results valid for stochastic programs with first and second degree stochastic dominance constraints induced by mixed-integer linear recourse. These models were recently introduced, studied and applied to power market in [4, 5]. We study qualitative stability of the set of feasible solutions defined by third-degree stochastic dominance constraint induced by mixed-integer linear recourse and the optimal value function with respect to weak convergence of underlying probability measures. The results are based on theory of parametric programming [1] and convergence of probability measures [2].

The paper is organized as follows. In Section 2 third-degree stochastic dominance is defined. Basic properties of mixed-integer value function are reviewed and extended in Section 3. In Section 4, optimization model with third degree-stochastic dominance constraint induced by mixed-integer value function is introduced and its stability is studied. Section 5 contains some concluding remarks and topics for future research.

2 Third-degree stochastic dominance

To define stochastic dominance of kth order, we need to consider the following auxiliary functions which are derived from distribution function:

$$F_X^{(k)}(\eta) = \int_{-\infty}^{\eta} F_X^{(k-1)}(\xi) d\xi, \ \forall \eta \in \mathbb{R}, \ k = 2, 3, \tag{1}$$

where $F_X^{(1)} = F_X$ is equal to the distribution function of the random variable X. Formal definition of third degree stochastic dominance follows.

Definition 1.
Let X, Y be two random variables on $(\Omega, \mathcal{F}, P)$ with distribution function F_X, F_Y. We say that X dominates Y in the sense of Third Degree Stochastic Dominance (TSD), denoted by $X \succeq_{TSD} Y$, if and only if

$$F_X^{(3)}(\eta) \leq F_Y^{(3)}(\eta), \ \forall \eta \in \mathbb{R}. \tag{2}$$

* This work was supported by the grants GACR 402/09/H045 and GAUK ?/?/?.

We say that X strictly dominates Y in the sense of third degree stochastic dominance, denoted by $X \succ_{TSD} Y$, if and only if $X \succeq_{TSD} Y$ and $Y \succeq_{TSD} X$ does not holds.

Let X be a random variable with finite second moment, i.e. $\mathbb{E}X^2 < \infty$. Then the auxiliary function can be expressed as (see [6, Theorem 3.1])

$$F_X^{(3)}(\eta) = \frac{1}{2}\int_{-\infty}^{\eta} [\eta - \xi]_+^2 dF_X(\xi),$$

where we denote $[\eta - \cdot]_+^2 = (\max\{0, \eta - \cdot\})^2$.

Let $\mathcal{U}_3$ be a class of real-valued differentiable functions with $u' > 0$, $u'' < 0$, and $u''' > 0$. Such functions are usually refered as decreasing absolute risk aversion (DARA) utility functions. The relation $X \succeq_{TSD} Y$ is equivalent to the condition that $\mathbb{E}u(X) \geq \mathbb{E}u(Y)$ holds for all utility functions $u \in \mathcal{U}_3$ for which both expectations are finite, see [7, 8].

3 Mixed-integer value function

We consider the following *mixed-integer value function* with random right-hand side vector h (random demand)

$$f(x, h(\omega)) = c^T x + \Phi(h(\omega) - Tx),\ x \in C \tag{3}$$

with the second stage problem which is a parametric mixed-integer linear problem

$$\Phi(z) = \min\left\{q^T y + q'^T y' : Wy + W'y' = z,\ y \in \mathbb{Z}_+^m,\ y' \in \mathbb{R}_+^{m'}\right\}, \forall z \in \mathbb{R}^s, \tag{4}$$

where $c \in \mathbb{R}^n$, $q \in \mathbb{R}_+^m$, $q' \in \mathbb{R}_+^{m'}$ are vectors, $T(s\times n)$, $W(s\times m)$, $W'(s\times m')$ are matrices with mentioned dimensions, and $\mathbb{Z}_+, \mathbb{R}_+$ denote nonnegative integers and real numbers. The function f appears as objective function in expectation in two-stage stochastic models with mixed-integer linear recourse, c.f. [9, 10].

We assume that matrices W, W' have only rational entries, $C \subseteq \mathbb{R}^n$ is a nonempty closed set (possibly partly restricted to integers), the right-hand side vector $h \in \mathbb{R}^s$ is a random vector on a probability space $(\Omega, \mathcal{F}, P)$ with $\int_\Omega \|h(\omega)\|^2 P(d\omega) < \infty$. We denote $\mu = h^{-1} \circ P$ the image measure on $\mathbb{R}^s$ which belongs to the general class of Borel measures $\mathcal{P}(\mathbb{R}^s)$. The function $\Phi(z)$ is real-valued on $\mathbb{R}^s$ if we further assume (see [9]):
(A1) complete recourse: $W(\mathbb{Z}_+^m) + W'(\mathbb{R}_+^{m'}) = \mathbb{R}^s$,
(A2) dual feasibility: $\{u \in \mathbb{R}^s : W^T u \leq q,\ W'^T u \leq q'\} \neq \emptyset$.
Previous assumptions ensure also that $f(x, h)$ is lower semicontinuous jointly in (x, h) on $C \times \mathbb{R}^s$, i.e. its epigraph is closed. To handle the mixed-integer value function in optimization model with TSD we need to show that its second moment is finite.

Proposition 1. *Let the assumptions (A1), (A2) be fulfilled and $\int_{\mathbb{R}^s} \|h\|^2 \mu(dh) < \infty$. Then $\mathbb{E}\left\{|c^T x + \Phi(h(\omega) - Tx)|^2\right\} < \infty,\ \forall x \in C$.*

Proof. Our assumptions ensure that $\Phi(0) = 0$. Using this fact we obtain

$$\begin{aligned}
\int_{\mathbb{R}^s} |\Phi(h - Tx)|^2 \mu(dh) &= \int_{\mathbb{R}^s} |\Phi(h - Tx) - \Phi(0)|^2 \mu(dh) \\
&\leq \int_{\mathbb{R}^s} \left(c_1\|h - Tx\| + c_2\right)^2 \mu(dh) \\
&= \int_{\mathbb{R}^s} c_1^2\|h - Tx\|^2 + 2c_1c_2\|h - Tx\| + c_2^2\ \mu(dh) \\
&= c_1^2 \int_{\mathbb{R}^s} \|h - Tx\|^2 \mu(dh) + 2c_1c_2 \int_{\mathbb{R}^s} \|h - Tx\| \mu(dh) + c_2^2
\end{aligned}$$

where the inequality follows from pseudo-Lipschitz continuity of the mixed-integer value function (4), see [3], for some positive constants c_1, c_2. Due to our assumptions the last expression is finite for any $x \in C$.

4 Optimization model with third-degree stochastic dominance constraint

We consider the following optimization model

$$\varphi(\mu,\nu) = \min\{g(x): \ x \in C(\mu,\nu)\},$$

where g is a real valued objective function on $\mathbb{R}^n$ and the set of feasible solutions is defined by third degree stochastic dominance constraint and depends on image probability measures

$$C(\mu,\nu) = \{x \in C : -f(x,h(\omega)) \succeq_{TSD} t(\omega)\}.$$

The mixed-integer value function $f(x,h(\omega))$ stands for the loss variable, and $t(\omega)$ is a random treshold with $\int_\Omega |t(\omega)|^2 P(d\omega) < \infty$, in our settings a minimal acceptable income. Let $\nu = t^{-1} \circ P$ denote the image measure. According to previous section, the constraint $-f(x,h(\omega)) \succeq_{TSD} t(\omega)$ is equivalent to

$$\int_{\mathbb{R}^s} [\eta - (-f(x,h))]_+^2 \mu(dh) \le \int_{\mathbb{R}} [\eta - t]_+^2 \nu(dt), \ \forall \eta \in \mathbb{R}. \tag{5}$$

One of the main results is stated in the following proposition. Reasonable behaviour of the set of feasible solutions with respect to weak convergence of underlying image measures will lead to stability of the optimal value function of our optimization problem.

Proposition 2. *Let the assumptions (A1), (A2) be fulfilled. Then the set of feasible solutions $C(\cdot,\cdot)$ is outer semicontinuous in any (μ,ν), where $\mu \in \mathcal{P}(\mathbb{R}^s)$, and $\nu \in \mathcal{P}_{\rho,R}(\mathbb{R}) \equiv \{\nu \in \mathcal{P}(\mathbb{R}) : \int_{\mathbb{R}} |t|^{2+\rho}\nu(dt) < R\}$ for some $\rho > 0$, $R > 0$, i.e. for arbitrary admissible $\mu_n \in \mathcal{P}(\mathbb{R}^s)$, $\nu_n \in \mathcal{P}_{\rho,R}(\mathbb{R})$ such that $\mu_n \xrightarrow{w} \mu$, $\nu_n \xrightarrow{w} \nu$, and $x_n \to x$ with $x_n \in C(\mu_n,\nu_n)$ it follows that $x \in C(\mu,\nu)$.*

Proof. Let $\eta \in \mathbb{R}$ be arbitrary and fixed, and denote $\tilde{f}(x,h) = [\eta + f(x,h)]_+^2$. We see that the function $\tilde{f}$ is nonnegative and lower semicontinuous jointly in both variables, thus we have

$$\begin{aligned}\int_{\mathbb{R}^s} \tilde{f}(x,h)\mu_n(dh) &\le \int_{\mathbb{R}^s} \liminf_{k\to\infty} \tilde{f}(x_k,h)\mu_n(dh) \\ &\le \liminf_{k\to\infty} \int_{\mathbb{R}^s} \tilde{f}(x_k,h)\mu_n(dh),\end{aligned} \tag{6}$$

where the first inequality follows from the definition of lower semicontinuity of $\tilde{f}(\cdot,h)$ and the second inequality is a consequence of Fatou's lemma. Using this result, we obtain

$$\begin{aligned}\int_{\mathbb{R}^s} \tilde{f}(x,h)dP(h) &\le \liminf_{n\to\infty} \int_{\mathbb{R}^s} \tilde{f}(x,h)\mu_n(dh) \\ &\le \liminf_{n\to\infty} \liminf_{k\to\infty} \int_{\mathbb{R}^s} \tilde{f}(x_k,h)\mu_n(dh) \\ &\le \liminf_{n\to\infty} \int_{\mathbb{R}^s} \tilde{f}(x_n,h)\mu_n(dh) \\ &\le \lim_{n\to\infty} \int_{\mathbb{R}} [\eta - t]_+^2 \nu_n(dt),\end{aligned}$$

where the first inequality is a consequence of Lemma 2.1 in [5], which states lower semicontinuity of integral functional in itegrating measure, the second one follows from (6), the third one can be obtained setting $n = k$, the fourth one follows from the denition of the set $C(\mu_n,\nu_n)$. The following convergence is shown bellow and finishes the proof

$$\lim_{n\to\infty} \int_{\mathbb{R}} [\eta - t]_+^2 \nu_n(dt) = \int_{\mathbb{R}} [\eta - t]_+^2 \nu(dt). \tag{7}$$

Without loss on generality, we choose an $\alpha > 0$ such that $\nu(\{t : [\eta - t]_+^2 = \alpha\}) = 0$. We define the "cut-off" function $d_\alpha(x) = x$ if $|x| < \alpha$ and $d_\alpha(x) = 0$ otherwise. Since $\nu_n \xrightarrow{w} \nu$, and the function $d_\alpha([\eta - \cdot]_+^2)$ is continuous and bounded, it holds

$$\lim_{n\to\infty} \int_{\mathbb{R}} d_\alpha([\eta - t]_+^2)\nu_n(dt) = \int_{\mathbb{R}} d_\alpha([\eta - t]_+^2)\nu(dt).$$

For the difference of integrals holds

$$\int_{\mathbb{R}} [\eta - t]_+^2 \nu(dt) - \int_{\mathbb{R}} d_\alpha([\eta - t]_+^2)\nu(dt) = \int_{\{t:[\eta-t]_+^2 \geq \alpha\}} [\eta - t]_+^2 \nu(dt).$$

The same formula is valid if we use ν_n instead of ν. Hence

$$\begin{aligned}
&\limsup_{n\to\infty} \Big| \int_{\mathbb{R}} [\eta - t]_+^2 \nu_n(dt) - \int_{\mathbb{R}} [\eta - t]_+^2 \nu(dt) \Big| \\
&\leq \sup_n \int_{\{t:[\eta-t]_+^2 \geq \alpha\}} [\eta - t]_+^2 \nu_n(dt) + \int_{\{t:[\eta-t]_+^2 \geq \alpha\}} [\eta - t]_+^2 \nu(dt) \\
&\leq \sup_n \int_{\{t:(\eta-t)^2 \geq \alpha\}} (\eta - t)^2 \nu_n(dt) + \int_{\{t:(\eta-t)^2 \geq \alpha\}} (\eta - t)^2 \nu(dt) \\
&= \sup_n \int_{\{t:t^2 \geq \alpha\}} t^2 \nu_n(dt) + \int_{\{t:t^2 \geq \alpha\}} t^2 \nu(dt) \\
&\leq \alpha^{-\rho} \sup_n \int_{\mathbb{R}} t^{2+\rho} \nu_n(dt) + \alpha^{-\rho} \int_{\mathbb{R}} t^{2+\rho} \nu(dt).
\end{aligned}$$

Then, if we pass $\alpha \to \infty$, we get the equality (7). This finishes the proof.

Similar results were obtained in [4, 5] for the set of feasible solutions with first and second degree stochastic dominance constraint.

Remark 1. A natural question arise: how to ensure that $P_n \xrightarrow{w} P$ implies $\mu_n \xrightarrow{w} \mu$ and $\mu_n \xrightarrow{w} \mu$, where the image measures are defined by $\mu_n = h^{-1} \circ P_n$, $\nu_n = t^{-1} \circ P_n$. The answer gives [2, Theorem 5.5]. Let D_h and D_t denote the sets of discontinuity points of $h(\cdot)$ and $t(\cdot)$. If $P(D_h) = 0$ and $P(D_t) = 0$ then $P_n \xrightarrow{w} P$ implies the weak convergence of image measures $\mu_n \xrightarrow{w} \mu$ and $\mu_n \xrightarrow{w} \mu$.

As an immediate consequence of Proposition 2 we obtain closedness of the set of feasible solutions. Moreover, if we relax the integer conditions, the set of feasible solutions is convex. The later fact follows again from the alternative description (5) of the set of feasible solutions and convexity of continuous value function.

Finally, qualitative stability of the optimal value function can be get.

Proposition 3. *Let the function g be lower semicontinuous, and let $C \subset \mathbb{R}^n$ be a compact set. Then the optimal value function $\varphi(\mu, \nu)$ is lower semicontinuous in each $(\mu, \nu) \in \mathcal{P}(\mathbb{R}^s) \times \mathcal{P}_{\rho,R}(\mathbb{R})$ for which the set of feasible solutions $C(\mu, \nu)$ is nonempty, i.e.*

$$\varphi(\mu, \nu) \leq \liminf_{\mu_n \xrightarrow{w} \mu, \nu_n \xrightarrow{w} \nu} \varphi(\mu_n, \nu_n).$$

Proof. If φ were not lower semicontinuous, there would exists sequences $\mu_n \xrightarrow{w} \mu$, $\nu_n \xrightarrow{w} \nu$, $x_n \in C$, and an $\varepsilon > 0$ such that $x_n \in C(\mu_n, \nu_n)$, and $g(x_n) \leq \varphi(\mu, \nu) - \varepsilon$. Obviously this implies $x_n \in C(\mu_n, \nu_n) \setminus C(\mu, \nu)$ for all n. Since C is compact there exists an accumulation point $\overline{x}$ of the sequence x_n. By the outer semicontinuity of $C(\cdot, \cdot)$, see Proposition 2, it follows that $\overline{x} \in C(\mu, \nu)$. The lower semicontinuity of g however leads to a contradiction

$$\varphi(\mu, \nu) \leq g(\overline{x}) \leq \liminf_{n\to\infty} g(x_n) \leq \varphi(\mu, \nu) - \varepsilon.$$

5 Conclusion

We considered stochastic optimization model with third degree stochastic dominance constraint induced by mixed-integer value function. We studied qualitative stability of this model with respect to weak convergence of underlying probability measures. We obtained outer semicontinuity of the set of feasible solutions and lower semicontinuity of the optimal value function. Further research will be adressed into finding an algorithm for solving this problems and to possible extensions to dynamic settings.

References

1. Bank, B., Guddat, J., Klatte, D., Kummer, B., Tammer, K.: Non-linear parametric optimization. Akademie-Verlag, Berlin (1982)
2. Billingsley, P.: Convergence of probability measures. John Wiley & Sons, New York (1968)
3. Blair, C.E., Jeroslow, R.G.: The value function of a mixed integer program: I. Discrete Mathematics 19, 121-138 (1977)
4. Gollmer, R., Neise, F., Schultz, R.: Stochastic programs with first-order stochastic dominance constraints induced by mixed-integer linear recourse. SIAM Journal on Optimization, Vol. 19, No. 3, 552-571 (2008)
5. Gollmer, R., Gotzes, U., Schultz, R.: A note on second-order stochastic dominance constraints induced by mixed-integer linear recourse. Mathematical Programming, Ser. A, to appear (2009)
6. Gotoh, J.-Y., Konno, H.: Third degree stochastic dominance and mean-risk analysis. Management Science, Vol. 46, No. 2, 289-301 (2000)
7. Levy, H.: Stochastic dominance: Investment decision making under uncertainty. Second edition, Springer, New York (2006)
8. Muller, A., Stoyan, D.: Comparison Methods for Stochastic Models and Risks. Wiley, Chichester (2002)
9. Schultz, R.: Rates of convergence in stochastic programs with complete integer recourse. SIAM Journal on Optimization, Vol. 6, 1138-1152 (1996)
10. Schultz, R.: Stochastic programming with integer variables. Mathematical Programming, Ser. B 97, 285-309 (2003)
11. Tesfatsion, L.: Stochastic dominance and the maximization of expected utility. The Review of Economic Studies, Vol. 43, No. 2, 301-315 (1976)
12. Whitmore, G. A.: Third-Degree Stochastic Dominance. The American Economic Review, Vol. 60, No. 3, 457-459 (1970)

MODEL OF THE COLLECTION CENTRES IN COMMUNES OF SLOVAK REPUBLIC[1]

Ivan Brezina, Juraj Pekár, Zuzana Čičková

University of Economics in Bratislava
Faculty of Economic Informatics
Department of Operations Research and Econometrics
brezina@euba.sk, pekar@euba.sk, cickova@euba.sk

Abstract: Presently the current legislation that is based on environmental factors has motivated the development of model approaches, which support the collection process. Presented model is aimed at accessibility of collection centres in communes of Slovak Republic. The goal is to motivate the population to collect the used products, components, packaging and redundant stocks for their dismantling and reuse. The model is aimed at projecting of collecting system that can be characterized by the same availability for the whole population (based on distance or time) so that the number of collection centres is as minimal as possible.

Keywords: Collection Centres, Location, Set Covering Problems

Introduction

Presently the social and environmental processes respect the legislation of European Union (represented by European Commission, European Parliament, Council of the European Union), which comprises many directives that are the base for legislation of EU members. Related legislation on waste collection and elimination led to emplacement of waste collection systems for households and also on public places.

The importance of collection is emphasized by the directive of European Union. EU member states have to follow guidelines for implementation of the environmental directive. Special directives are e.g. Directive on Waste Electrical and Electronics Equipment – WEEE, (submitted 2001), Directive on the Restriction of the Use of Certain Hazardous Substances in Electrical and Electronic Equipment - RoHS, (2002/95/EG) etc. (e.g. WEEE - EU Directive 2002/96/EG on Old Electrical and Electronic Devices - Waste of Electrical and Electronic Equipment and the associated Amendment 2003/108/EG). The targets for member states are:

1. to minimize the use of dangerous substances and the number of plastics, some materials must be phased out,

2. to promote the design for recycling,

3. producers have to take responsibility for end of life of the products and provide information to the processors about appropriate recycling,

4. the systems for separate collection should be put in public place and to household for free,

5. producers have to set up and finance appropriate systems to ensure processing and recycling of products and they have to be responsible for that products in the place of collection,

6. collection services must be offered by sellers and also by municipalities with an aim to collect 4 kg of waste per head of population yearly,

7. recovery quotas are set from 70% to 90% of collected waste, depending on the products category etc.

On the national level, the models are usually focused on modelling of collection of recyclable products. The models that deal with before mentioned problem are usually oriented on positioning of collection centres. The goal is to achieve the efficiency in collection of recyclable products from the final users. After separation in collection centres, the products are distributed to processing centres, which location is also the part of optimization in process of minimizing of distribution costs. In processing centres the products are machined on the basis their usableness (or unusableness) and they are distributed to remanufacturing centers, where they are

[1] This research was supported by the grant project of Slovak Ministry of Education VEGA 1/4588/07 „Reverse Logistics Modeling – Optimization of Recycling and Disposal Processes“.

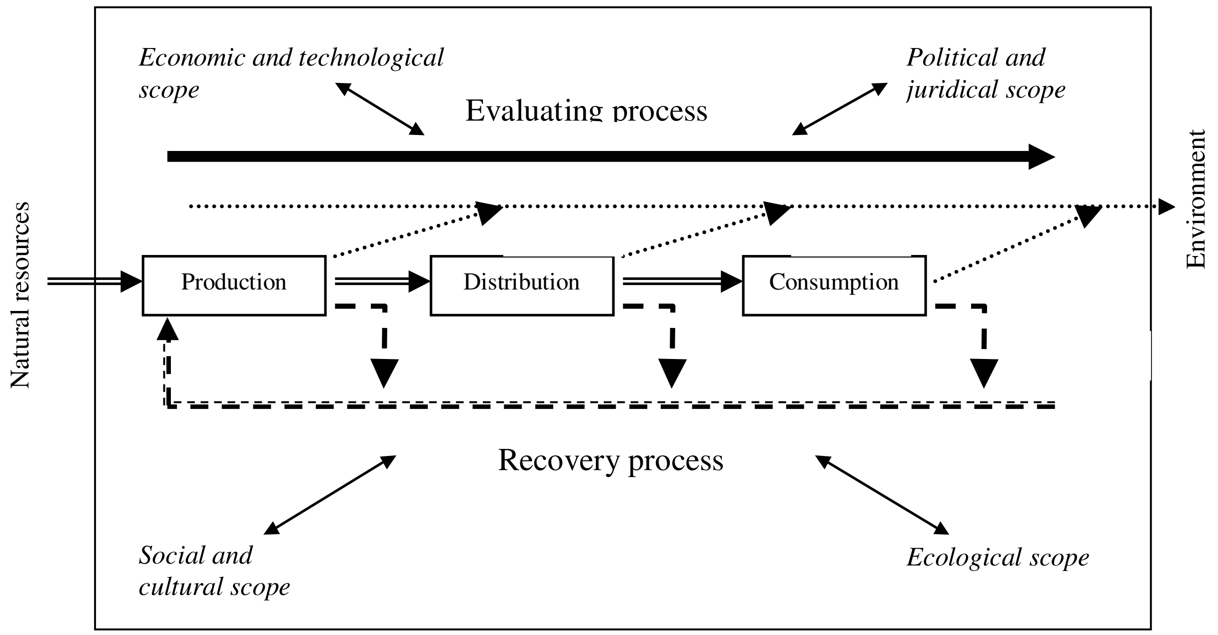
usually remanufactured to raw material that will be use in production process or they are brought to specialized waste disposal.

1. Model of collection centres in communes of Slovak Republic

On the national level, the models are usually focused on modelling of collection of recyclable products. Their development is motivated mainly by relatively extensive environmental legislation that forces government to realize the recycling process within the frame of waste management. In the every state a lot of producers are from foreign countries that come under the legislation of both the European Union and the Slovak republic and the government has to take responsibility for such products (so called OEM problem – Original Equipment Manufactures Problem).

Gradually, the recycling process evolved in such part of logistics that monitors backwards flows of goods, material, covers and reused material from customers to distributor, respectively to producer, with aim to ensure the reclaim, repair, reuse, recycling or liquidation in accord with statutory instruments and directives in an ecological attractive way (Figure 1).

Figure1: Material flows of reverse logistics



The three main levels are known in optimization of reverse logistics. The first is oriented on collection of discarded vehicles (ELV – End-of-Life Vehicles), the second is pointed on collection of waste of electrical and electronic equipment (Waste Electrical and Electronic Equipment). Both are determined by legislation and there exist directives on relevant waste collection. The third level is focused on paper collection that is in general most frequently recyclable product. Related optimization models are based on linear programming (Linear Programming – LP), mixed integer programming (Mixed Integer Programming – MILP), special heuristic algorithms and also another approaches.

The goal of modelling of collection centres is aimed at accessibility of collection (sorting) centres for whole population in Slovak Republic for the purpose to motivate the population to collect the used products, components, packaging and redundant stocks for their dismantling and reuse. In general, there exist two basic conceptions – the first one is aimed on covering the whole population with the minimal count of service channels (Location Set Covering Problem – LSCP), the goal of the second one is to reach maximal covering by limited number of service channels (Maximal Covering Location Problem – MCLP).

Presently, there exist 2 916 communes in Slovak republic (according to Statistical Office of Slovak Republic). The goal of the model is to locate the collection centres in communes so that the accessibility of collection centres is preserved and the count of collection centres is as minimal as possible.

The problem can be formulated as binary programming task with the variables $x_j \in \{0,1\}, j = 1, 2, ..., n$, where n is the number of communes in Slovak Republic (2 916). If the variable is set to 0, the collection centre is not located in corresponding commune, if the variable is set to 1, the collection centre will be opened in corresponding commune. Considered above mentioned criteria, the objective function could be formulated as follows:

$$f(x) = \sum_{j=1}^{n} c_j x_j \rightarrow \min$$

The coefficients in objective function assume the value 1.

The constraints need to ensure that the distance between every commune and the closest collection centre is lesser than or equal to K. Parameters d_{ij} represent distances between communes i a j, so matrix **D** (size n x n) is the matrix of minimal distances between all pairs of communes. Based on matrix **D**, the matrix **A** can be defined (also size n x n), where the elements a_{ij} are set to 0, if the distance between i-th and j-th communes is greater than K, or 1 otherwise.

$$a_{ij} = \begin{cases} 0, d_{ij} > K \\ 1, d_{ij} \le K \end{cases} \quad i, j = 1, 2, ..., n$$

Further on, the accessibility of collection centre for every commune need to be met within the distance K that can be formulate as follows:

$$\sum_{j=1}^{n} a_{ij} x_j \ge 1, \qquad i = 1,2,...,n$$

The model needs to deal with binary variables:

$$x_j \in \{0,1\}, \quad j = 1,2,...,n \; .$$

The formulation of the problem:

$$f(x) = \sum_{j=1}^{n} c_j x_j \rightarrow \min$$

$$\sum_{j=1}^{n} a_{ij} x_j \ge 1, \qquad i = 1,2,...,n$$

$$x_j \in \{0,1\}, \quad j = 1,2,...,n$$

2. The solution of the model

Presented model was applied on data in Slovak Republic. We used data from the year 2001, when the whole number of population was 5 378 511 inhabitants. The goal was to minimize the number of residents that need to travel to collection centres (Pekár – Brezina (2008)).

Input data:

n – number of communes in Slovak Republic (2 916 communes),

D (n x n) – matrix of minimal distances between all communes,

K – maximal distance from closest collection centre 30 km,

c_j – potential location of collection centres in all communes, coefficients in objective function are equal to 1 (c_j = 1, j = 1, 2, ..., n),

Output data:

The above mentioned model was solved by software product GAMS.

For distance K = 30 the collection centres are:

Malacky, Malinovo, Drahovce, Trnava, Gabčíkovo, Očkov, Senica, Komárno, Pribeta, Koniarovce, Úľany nad Žitavou, Šaľa, Trenčianske Teplice, Dolné Kočkovce, Mužla, Lehôtka pod Brehmi, Demandice, Tekovské Nemce, Nitrica, Veľký Krtíš, Domaníky, Vígľaš, Banská Bystrica, Stará Halič, Bravăcovo, Kokava nad Rimavicou, Šurice, Tornaľa, Vernár, Ratková, Spišské Podhradie, Helcmanovce, Gemerská Poloma, Slovenská Ves, Chmeľnica, Košické Oľšany, Moldava nad Bodvou, Trnkov, Bardejov, Luhyňa, Soľnička, Nacina Ves, Žalobín, Mestisko, Svetlice, Sobrance, Príslop, Svinia, Raková, Breza, Nižná, Podhorie, Kraľovany, Ivachnová, Východná, Malý Čepčín.

It is evident that efficiency of collection centre depends on character of collected products (it is different in case of cars, fridges, freezers or in case of car batteries) so we considered with different variants of distances *K* that were set to 50, 40, 30, 20, 15 and 10 km. Another modification follows from the reasoning that one of the requirements is to locate the collection centre to region cities (administrative centres of regions). Further on, that strict requirement was modified by the weight that was based on proportion between number of population in the region city and the whole population in all cities. The goal was to minimize the number of residents that need to travel to collection centres.

If the models consider about potential location of collection centres in all communes, coefficients in objective function are equal to 1 ($c_j = 1$, j = 1, 2, ..., *n*). If the model considers about location of collection centres in all communes simultaneously with requirement that the collection centre must to be located in every region city, where the set of region cities is N_k, then the coefficients in objective function are equal to 1 ($c_j = 1, j = 1, 2, ..., n$) and decision variables for region cities $x_j = 1, j \in N_k$. The last approach was to consider about potential location of collection centres in all communes with respect to number of population in region cities, so that number of population in 2 916 communes was P = 5 378 511 and P_j was number of population of *j*-th city, coefficients of objective function was calculated as $c_j = 1 - \frac{P_j}{P}, j = 1,2,\ldots,n$, where c_j represents potential number of population that need to travel to collection centre.

Conclusion

Presented approach can be used as a base for efficient decision making about location of collection centres considered the special criteria as well as the given parameters of accessibility and the importance for population. The model can be used in general for different input data. Results of this paper show that presented model is relatively flexible to adjust the changed conditions. The accessibility is usually only the one from the set of parameters that determines the location of collection centre and it is possible to come out of other restrictions.

Literature

[1] Arakaki, R. G. I. – Lorena, L. A. N. *A Constructive Genetic Algorithm for the Maximal Covering Location Problem.* In: MIC 2001 – 4th Metaheuristics International Conference, Porto, Portugal 2001, pp. 13–17

[2] Chuang, C. – Lin, R. *A Maximum Expected Covering Model for an Ambulance Location Problem.* In: Journal of the Chinese Institute of Industrial Engineers, Vol. 24, No. 6/2007, p. 468-474

[3] Ivaničová, Z. – Chocholatá, M. – Stolze, A. *Operations Research für Manager.* Vydavateľstvo EKONÓM, Bratislava 2007, ISBN 978-80-225-2347-9

[4] Pekár, J. – Brezina, I. *Problém umiestnenia zberných miest.* In: Medzinárodný seminár mladých vedeckých pracovníkov, Praha, 2008. ISBN 978-80-245-1405-5

[5] Plastria, F. – Vanhaverbeke, L. *Maximal Covering Location problem with Price Decision for Revenue Maximization in a Competitive Environment.* In: www.vub.ac.be/MOSI/papers/PlastriaVanhaverbeke2006_MaxcoverRevenue.pdf

[6] Turner, D. *Implementing the Flow-Covering Location-Allocation Model with Geographic Informations Systems.* Master´s Project. In: epps.utdallas.edu/mgis/prj_mstrs/2005/Fall/Turner/DT_master_FINAL.ppt

[7] Vidović, M. – Dimitrijević, B. – Ratković, B. *Locating Collection Centres for ELV.* In: XXXV Simpozijum o operacionim istraživanjima, Zbornik radova, Soko Banja, 2008, ISBN 978-86-7395-248-2

REAL ESTATE PRICES USING ARTIFICIAL NEURAL NETWORKS OVER TIME

Caridad y Ocerin, José Mª; Núñez Tabales, Julia ;Ceular Villamandos, Nuria; Millán Vázquez de la Torre, G.

Dpt. of Stastistics and Econometrics University of Córdoba. Spain.
ccjm@uco.es

Abstract. Econometric models in the estimation of real estate prices is a useful and realistic approach for buyers and for local and fiscal authorities. From the classical hedonic models to more data driven, based on Artificial Neural Networks, many papers have appeared in economic literature trying to compare the results with both methods. We insist in the use of ANN, when there is enough statistical information, with some comparisons in a medium size city in the South of Spain, with an extensive set of data spanning over several years. Exogenous variables include external and internal data from the dwelling, both numerical and qualitative, as of the building in which it is located and its surroundings. Alternative models are estimated for several time intervals, being able to compare the effects of the rising prices during the bull market over the last decade, and could be used to asses the downturn of last couple of years.
Keywords: real estate prices, artificial neural Networks, econometric modelling

1. Introduction

Last decade Spain presented a seller's real estate market. There is a strong tradition of home ownership in the country, coupled to fiscal stimuli and low interest rates associated to the euro zone. Prices have continuously gone up, and a growing part of the family's income was devoted to attend mortgages and housing related costs. This trend has come abruptly to and end in 2008, causing turmoil in the country economy. In this context, an objective way to assess the prices of properties coming to the market is needed. Its demand come from different actors: buyers and sellers of family flats and second homes, agents in the real state sector, investors, value agency and financial institutions, and fiscal authorities at national and municipal levels. Hedonic and ANN models could fill the gap, although they must be updated and maintained, with the additional benefit of their usefulness in detecting and measuring price changes.

Real estate prices were studied since the fifties, and two decades later, Rosen proposed the use of regression models, called 'hedonic models', and applied them in urban areas. In Spain there are macroeconomic and sectorial statistics, that try to explain investments in building and home buying. The use of ANN has been used in previous studies in the South of Spain, and we return to this approach to used in a dynamic framework, applying it to a case study. Sample data from more than 10000 transactions have been collected from 2002 to 2006 enabling us to establish temporal comparisons.

2. Artificial Neural Networks

Artificial Intelligence methods started to be applied to real estate prices last decade, with a growing interest. Several studies are related in table 1. In Spain several studies have been done by Caridad y Ceular (2001), García Rubio (2004), Gallego (2004) and Lara (2005), using data from different urban areas.

ANN were introduced by McCulloch and Pitts (1943) as an alternative to algorithmic programming with some previous work by Kart Lashley in the twenties. In Rumelhart and McClelland (1986) an ANN is presented as a ser of operators or neurons, with a small amount of storage capacity, one-way numerically connected by some links so called axons. Nodes operate with local data supplied trough the axons, weighted with some parameters, w_{ij}, linking neurons i and j. A propagation rule establish how are valued and processed the inputs to a neuron. The basic model has an input layer, one or several hidden layers formed by non observable variables, and an output layer containing the dependent variables. Like any model, the specification is the core task when using ANN: it is necessary to define the network topology (number of hidden layers, and the neurons in each of them), the propagation rule, the transformation of input explanatory variables, and so on. The activation function and the learning rule should be specified, also. An excessive number of neurons will be a cause of lack of forecasting power, due to overparametrisation. Computer time to estimate the ANN parameters, learning process, is becoming less relevant. There are several learning procedures to estimate the parameters being a widely used the back-propagation method.

TABLE 1. USE OF AI METHODS IN REAL ESTATE PRICE ESTIMATION

AUTHORS	DATE	GEOGRAPHICAL AREA
Borst	1991	New England
Tay y Ho	1992	Singapur
Do and Grudnitski	1992	California
Collins and Evans	1994	U.K.
Worzala, Lenk and Silva	1995	Colorado
Mc Cluskey	1996	
Rossini	1997	Australia
Haynes y Tan	1998	Australia
Bonissone	1998	
Cechin	2000	Porto Alegre (Brasil)
Karakozova	2000	Helsinki (Finland)
Nguyen and Cripps	2001	Tennessee
Kauko	2003	
Limsombunchai	2004	New Zeland
Liu, Zhang and Wu	2006	
Selim	2009	Turkye

An ANN is like a non linear regression or multivariate regression model, with non observable variables, that, nevertheless, once the topology and the parameters of the network are specified, it amounts to an ordinary statistical or econometric model. Her we use a multilayer perception network (MLP) with one hidden layer.

2. Model specification

The estimation of the selling price of a flat or apartment, *y*, is the objective of the ANN model. A sample is obtained in a medium size city of the South of Spain, on the main urban area. Exogenous variables include internal data about the dwelling and of the building and its location. The population is composed by 130563 properties in the area, and over 75% are the main residence, about 14% non used, and less than 10% secondary residences. It has augmented, in 2006 to 135920 flats. Price data are recorded by the national statistics institute (INE), and by the municipality, for fiscal purposes, but they are not focussed on precise valuation of individual properties. The market intermediaries are the main source of precise data; the main company, Grupo Barin, has supplied 10124 cases of transactions distributed as follows: in 2002, 772 cases, in 2003, 1685, in 2004, 1399, in 2005, 3380, and in 2006, 2888. This company is active in the whole area, with 18 offices scattered trough the town. A complementary sample was used from several similar smaller companies.

In table 2 are found the exogenous variables used. The price of the transactions are the real market price (not the declared price, nor the offer price), avoiding, thus duplicities in the data really used.

TABLE 2. EXOGENOUS VARIABLES AND ATTRIBUTES

<table>
<tr><th colspan="3">INTERNAL</th><th colspan="2">EXTERNAL</th></tr>
<tr><td>BASIC</td><td colspan="2">Area
Bedrooms
Bath
Complimentary baths
Terrace (*)
Communications (*)
Wardrobes (*)
Garage (*)
Storage room(*)
Climatization</td><td>GENERAL</td><td>Building year
Lift(*)
Laundry(*)</td></tr>
<tr><td rowspan="2">GENERAL</td><td>QUALITY</td><td>Floors
Window's type
Interior wooden furniture(*)
Kitchen furniture(*)</td><td rowspan="2">EXTRAS</td><td rowspan="2">Pool (*)
Tennis(*)
Garden(*)</td></tr>
<tr><td>REFORMS</td><td>Reformed(*)</td></tr>
<tr><td>ORIENTATIÓN</td><td colspan="2">Orientationr(*)</td><td rowspan="2">LOCALIZATION</td><td rowspan="2">Zone</td></tr>
<tr><td>ECONOMICS</td><td colspan="2">Common expenses
Market price</td></tr>
</table>

The (*) variables are binary, and are used to elaborate numerical indexes in the range (0; 1) that could be easily interpreted (table 3). These have been validated with a posterior sample of real estate market agents.

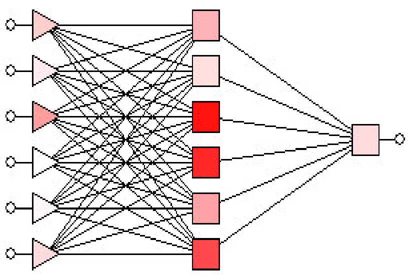
TABLE 3.INDEXES ASSOTIATED TO EACH PROPERTY

	VARIABLES USED
QUALITY INDEX	Floors, Windows type, Wooden, Kitchen and reforms
FACILITIES INDEX	Pool, tennis, garden
BUILDING INDEX	Age, lift, laudry
EXTERNAD DATA INDEX	Orientation, terrace
EXTRAS INDEX	Garage , Storage
LOCATION INDEX	Geographical position within the city

3. Results

A*MLP, Multi Layer Perceptron*[1] 6:6-6-1:1 is used with the following input variables: square meters of the house. age of the building, localisation index, extras index, community expenses and quality index. They were selected after an identification process with several alternatives. The output layer includes only the transaction price and there are six neurons in the hidden layer (as showed in figure 1). Trajan Neural Networks software was used, with a linear activation function in the input layer and logistic activation function in the output layer. The error function used is the residual sum of squares.

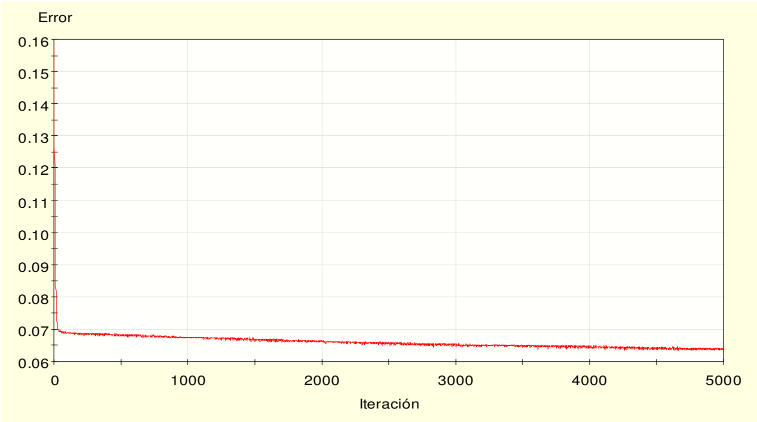
FIGURE 1.ANN 6:6-6-1:1*TOPOLOGY*



[1] Freeman and Skapura (1993), Haykin (1999), García Rubio (2004).

The training set is composed by 80% of the sample[2] and the back propagation method was used with a 0.1 learning rate. The error evolution is showed in graphic 1.

GRAPHIC 1 EVOLUTION OF THE RESIDUAL ERROR DURING THE TRAINING PROCESS



In tables 5 and 6 are presented the weights and thresholds of the input and hidden layer neurons.

TABLE 4. PARAMETERS: INPUT AND HIDDEN LAYERS

	2.1	**2.2**	**2.3**	**2.4**	**2.5**	**2.6**
threshold	-0,8637	-5.3193	0,8587	1.3877	-0,8773	-2.9494
Area	1,4581	-2.4698	-3,1902	0.8425	-0,3820	-0.6004
Common expenses	3,4714	-1.6389	1,7720	-0.5584	0,6896	-1.0893
Age	-0,0090	-0.1852	0,8528	0.7126	-1,2623	2.4097
Anex index	2,1949	-1.4684	-0,0846	-0.2235	-2,8127	1.4203
Localization	-1,1403	-0.7805	-5,0060	-1.5345	0,3433	-2.8736
Quality index	0,3414	-0.7721	0,2906	1.4205	-1,0161	-0.0384

TABLE 5.PARAMETERS: HIDDEN AND OUTPUT LAYER

	3.1
threshold l	-0.9370
2.1	3.9678
2.2	-3.9225
2.3	-3.7264
2.4	1.8743
2.5	0.9303
2.6	-2.5079

Sensitivity analysis allows to evaluate the influence of the exogenous variables, using its errors ratio (table 6), obtained as the error in the model without the variable over the model including it. As expected, it is the size variable the most important (1.2953), followed by the localisation index (1.2009), and by the common expenses (1.1577) as show on table 6.

TABLE 6. SENSITIVITY ANALYSIS

[2] *All cases had no missing data*

INPUT	RATIO	ORDER
Area	1.2953	1
Common E.	1.1577	3
Age	1.0395	6
Anex index	1.0414	5
Localisation	1.2009	2
Quality index	1.0644	4

The other three variables are less important, but if taken out of the model, the final results are less likable. The R^2 = 86.05% and the *RMSE* = 39540.36 € . The mean absolute error is 28551.34 €, and in relative terms 13.69%. On graphic 2 are shown the real vs the estimated price, with a good fit all over the spectrum of properties (this can seldom be achieved with classical hedonic models).

Using the same model with 2002-5 data, again with 80% of the cases as training set, and similar tuning parameters, the results are similar.

GRAPHIC 2. REAL AND ESTIMATED PRICES

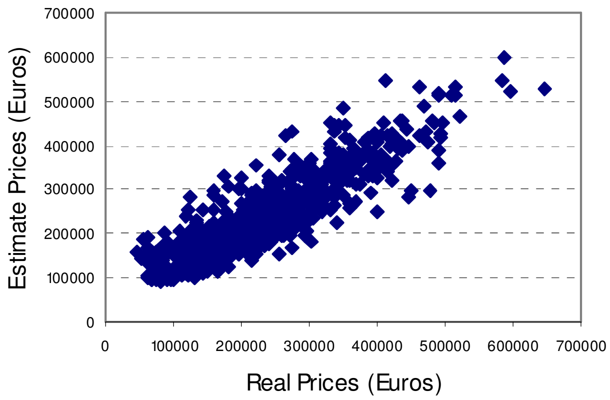
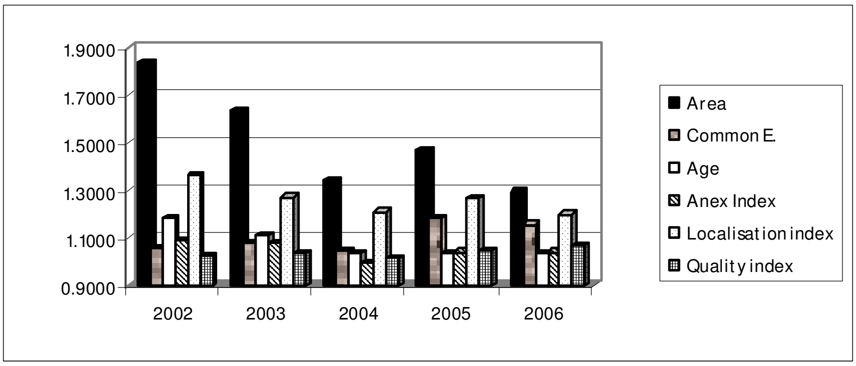


TABLE 7. SAMPLE SIZE FOR THE YEARLY MODELS

	CASES	TRAINING	TEST
2002	470	376	94
2003	914	731	183
2004	791	633	158
2005	1686	1349	337

Sensitivity analysis allows to evaluate the influence of the input variables upon the dwelling's market price, In graphic 3, the variables are ordered with their relative influence, with some temporal variability. In each model the area variable appears in the first place, followed by the localisation of the property. Then comes a third variable that can be either the age of the building, the maintenance cost or the quality index. The rest of the variables show less influence.

GRAPHIC 3. COMPARATIVE SENSIBILITY ANALYSIS OF THE INPUTS



Different goodness of fit statistics are presented on table 8.

TABLE 8. FIT MEASURES 2002-2006

	R^2	*RMSE*	*RESIDUAL SE*	*MAE*	*RME*
2002	90,21%	21.652,25	21.634,04	14.900,48	13,78%
2003	84,75%	25.884,77	25.738,99	18.999,93	14,75%
2004	90,23%	30.983,78	28.908,32	24.157,36	16,26%
2005	81,12%	32.825,44	32.755,04	23.562,59	13,82%
2006	86,05%	39.540,36	39.102,13	28.551,34	13,69%

4. Conclusions

The use of ANN is more flexible than classical econometric models when enough data are available. Some common problems, in hedonic models, such as non linearities in the extreme range of prices in real estate market, are well estimated using ANN, even adapting to properties that could be labelled as outliers, although, some authors[3] criticize this 'black box' approach. In our case study with a MLP, estimates fit well with the real market situation. The 'dark side' of this methodology is the sampling effort needed, that can be realistically overcome, only by firms with a broad presence in the markets, so that they can use their internal databases, where the uncertainty about the exact identification of each dwelling is eliminated. Most papers include as explanatory variables, the house area[4], its location, age and availability of garage. Sensibility analysis confirms these asserts. The distance to the city center is useful in monocentric towns. Variability is present even in the same street/building. In Spanish clogged cities, the availability of garage enhances dramatically the final price. Common expenses are related to the facilities provided by the building, so this variable is highly collinear with other used, and less limitative in the transactions.

In the validation process, ANN present more uncertainties, although in hedonic models, multicollinearities introduce a significant difficulty in the interpretation of hedonic prices. In the models proposed, these hedonic prices are obtained as marginal non linear functions derived once the topology of the network is specified; it is somewhat cumbersome process, but the nonlinearities of the marginal implicit prices conforms better to the reality (for example, a decreasing implicit price with the property's size).

The same Network topology is used in different years, with R^2 between 0.81 and 0.91. Once applied the 2006 model to new data from 2008 and 2009, the estimated price with these networks are around 12% higher than the real transaction data, showing the clear downturn of the real estate market in Spain.

These models could be useful not only to potential buyers/sellers and their agents, but for foreclosure valuations, not uncommon in the actual bear market, and for fiscal purposes, but also in sectorial studies, to assess the importance of an structural change due to turn points in the cyclical evolution of the building industry

.

[3] *Allen and Zumwalt (1994) o Worzala, Lenk y Silva (1995).*
[4] *García Rubio (2004)*

References

[1] **Borst, R**. (1991) *Artificial Neural Networks: The Next Modelling/ Calibration Technology for the Assessment Community?* Property Tax Journal, IAAO, 10 (1), pp. 69-94.

[2] **Caridad, J. M. y Ceular, N.** (2001) *Un análisis del mercado de la vivienda a través de Sistemas de Redes Neuronales*. Revista de Estudios de Economía Aplicada 18, pp. 67-81.

[3] **Collins, A. y Evans, A.** (1994) *Artificial Neural Networks: an application to residential valuation in the U.K.* Journal of Property Valuation and Investment, vol. 11, nº2, pp. 195-204.

[4] **Do, A. y Grudnitski, G.** (1992) *A Neural Network Approach to Residential Property Appraisal* The Real Estate Appraiser, Vol.58 (3), pp. 38-45.

[5] **Gallego Mora-Esperanza, J.** (2004) *La inteligencia artificial aplicada a la valoración de inmuebles. Un ejemplo para valorar Madrid* Revista *CT/Catastro* nº 50, pp.51-67

[6] **Lara Cabeza, J.** (2005) *Aplicación de las redes neuronales artificiales al campo de la valoración inmobiliaria* Revista Mapping, nº 104, pp. 64-71.

[7] **Nguyen, N. y Cripps, A.** (2001) *Predicting housing value: a comparison of multiple regression analysis and artificial neural networks* Journal of Real Estate Research, 22 (3), pp. 314-336.

[9] **Selim, H.** (2009) *Determinants of house prices in Turkey: Hedonic regresión versus artificial neural network* Expert Systems with Applications, nº36, pp.2843-2852.

A SMALL NEW KEYNESIAN MODEL OF SLOVAKIA

Veronika Cigáňová, Osvald Vašíček

Masaryk University, Faculty of Economics and Administration
Department of Economics, Lipová 41a, 602 00 Brno
ciganova.veronika@gmail.com, osvald@econ.muni.cz

Abstract. This paper deals with a small open economy New Keynesian Dynamic Stochastic General Equilibrium (NK DSGE) model and its ability to describe behaviour of the Slovak economy. The model framework consists of three home subjects: households, firms and a monetary authority. The representative household maximizes its utility function which contains a real rigidity in consumption in a form of a habit formation. The monopolistic competitive firms maximize their net present value of dividends (their profits) with respect to Calvo-type price setting. The sticky price setting of home producers and importers represents nominal rigidities in the model. Monetary policy is conducted according to a Taylor-type rule. The model includes exogenous foreign sector which is represented by Euroarea-12. The framework assumes a perfectly competitive labour market and a complete financial market. The log-linear model is estimated using Bayesian method (we use the Dynare toolbox for Matlab). Evaluation of the model for Slovakia is based on an analysis of impulse response functions and an assessment of estimated parameters. We compare consistency of our results with similar studies.

Keywords. NK DSGE model, real rigidity, nominal rigidity, small open economy, Slovakia.

1 Introduction

Many authors are interested in estimating a small open economy NK DSGE models[1] but only [8] uses data of Slovakia for estimation. In this paper only one third of parameters are estimated, remaining parameters are calibrated. The aim of this paper is to describe a basic SOE NK DSGE and estimate it for the Slovak economy. Estimation exploits the model proposed in [3] and [4].[2]

2 A Simple Small Open Economy Model

As we mentioned before, the model assumes the structure developed mainly in [3]. Main features of the model are optimizing agents (a representative household, domestic producers and importers), monetary policy conducting according to a modified Taylor rule and an exogenous foreign sector. The model also incorporates nominal rigidities and a real rigidity due to structural parameters like sticky-price parameters of domestic producers and importers and habit formation. The structural equations of the model are:[3]

$$c_t - hc_{t-1} = E_t\left(y_{t+1}^* - hy_t^* + \sigma^{-1}(1-h)q_{t+1}\right) + \sigma^{-1}(1-h)\left(r_t - E_t\pi_{t+1} + \upsilon_{g,t}\right) \tag{1}$$

$$(1-\alpha)c_t = y_t - \alpha\eta(2-\alpha)s_t - \alpha\eta\psi_t - \alpha y_t^* \tag{2}$$

$$s_t - s_{t-1} = \pi_{F,t} - \pi_{H,t} \tag{3}$$

$$q_t = \psi_t + (1-\alpha)s_t \tag{4}$$

$$\pi_{H,t} - \delta\pi_{H,t-1} = \theta_H^{-1}(1-\theta_H)(1-\beta\theta_H)mc_t + \beta E_t(\pi_{H,t+1} - \delta\pi_{H,t}) \tag{5}$$

[1] For example in [2], [7] and [8].
[2] The structure of the model is different from [8].
[3] Detailed description of parameters is in Appendix (Table 2).

where

$$mc_t = \varphi y_t - (1+\varphi)\upsilon_{a,t} + \alpha s_t + \sigma(1-h)^{-1}(c_t - hc_{t-1})$$

$$\pi_{F,t} - \delta\pi_{F,t-1} = \theta_F^{-1}(1-\theta_F)(1-\beta\theta_F)\psi_t + \beta E_t(\pi_{F,t+1} - \delta\pi_{F,t}) \quad (6)$$

$$(i_t - E_t\pi_{t+1}) - (i_t^* - E_t\pi_{t+1}^*) = E_t q_{t+1} - q_t + \upsilon_{s,t} \quad (7)$$

$$i_t = \rho\ i_{t-1} + (1-\rho)(\psi_\pi \pi_t + \psi_y y_t) + \varepsilon_{M,t} \quad (8)$$

$$\pi_t = \pi_{H,t} + \alpha(s_t - s_{t-1}) \quad (9)$$

$$\upsilon_{a,t} = \rho_a \upsilon_{a,t-1} + \varepsilon_{a,t} \quad (10)$$

$$y_t^* = \rho_y y_{t-1}^* + \varepsilon_y \quad (11)$$

$$\pi_t^* = \rho_\pi \pi_{t-1}^* + \varepsilon_\pi \quad (12)$$

$$i_t^* = \rho_i i_{t-1}^* + \varepsilon_i \quad (13)$$

$$\upsilon_{g,t} = \rho_g \upsilon_{g,t-1} + \varepsilon_g \quad (14)$$

$$\upsilon_{s,t} = \rho_s \upsilon_{s,t-1} + \varepsilon_s \quad (15)$$

All variables are expressed as log deviations from their respective steady state values. Equation (1) is the combination of the consumption Euler equation and the international risk sharing condition. This equation makes connection between present and future consumption and also between foreign and domestic consumption level. Second equation represents the goods market clearing condition. Equation (3) expresses changes in the terms of trade and a positive gap of terms of trade could be interpreted as an increase in competitiveness of home producers. The relation among real exchange rate gap, law of one price gap and terms of trade is formulated by equation (4). Equations (5) and (6) are the New Keynesian Philips curves of home producers and importers, respectively. The uncovered interest rate parity is given by equation (7). Equation (8) is a Taylor rule with monetary policy shock. Domestic inflation and inflation of home producers are related according to equation (9). The exogenous foreign sector is expressed by equations (11) - (13) as AR (1) process of foreign output gap, inflation and nominal interest rate. Disturbances in the model, which are expressed by equations (10), (14) and (15), are also AR (1) processes.

3 Data and Results of Estimation

Quarterly data from 1Q1999 to 4Q2008 of Slovakia and Euro-Area 12[4] on output, inflation, interest rate and real exchange rate are used to estimate the parameters of the model. All data enter the estimation as deviations from the long run trend. Trends are computed by HP filter. Data are not annualized. The source of the data is Eurostat database.

- y_t and y_t^* are macroeconomic productivity gaps of seasonally adjusted real GDP of Slovakia and Euroarea-12 from their long-term trend;
- π_t and π_t^* are seasonally adjusted q-o-q HCPI inflation gaps of Slovakia and Euroarea-12 from their long-term trend;
- i_t and i_t^* are 3-month nominal interest rate gaps of BRIBOR and EURIBOR from their long-term trends;
- q_t is a gap of real exchange rate SKK/EUR from its long-term trend.

[4] EA12- BE, DE, IE, GR, ES, FR, IT, LU, NL, AT, PT, FI

The model is estimated using Bayesian method which is incorporated in Dynare toolbox for Matlab. Table 1 represents results of the estimation of parameters as well as their prior settings (the distribution with their mean and their standard error). Table 1 also consists of the specification and the estimation of standard deviation of shocks. The issue of prior settings for Slovakia exists because of a lack of microeconomic studies so the prior specification is taken from [3] and a similar study of a small open economy in [4]. Only a discount factor is calibrated to value 0.99.

At the beginning, we could notice that almost all posterior marginal densities of parameters are more concentrated compared with their prior counterparts. Only the inverse elasticity of labour supply and the elasticity of substitution between home and foreign goods are exceptions because they highly depend on prior specification.[5] The posterior mean of the degree of habit formation is 0.11 which means that the representative household changes its structure of consumption significantly. It has to consume 11% more compared to the previous consumption to obtain the same utility excluding an effect of a long-run growth in consumption. In comparison with the studies for the Czech Republic in [6] and [7], this parameter is low.[6] The mean of the inverse elasticity of the intertemporal substitution in consumption is estimated at 0.47 so the elasticity of intertemporal substitution is approximately 2.13. The value of this parameter is close to the estimated value in [7].[7] The degree of openness is estimated at 0.38 which means that the percentage of imports on household's consumption is 38%.

The part of the non-optimizing firms which index their prices according to previous inflation is 0.7. This value is close to the value estimated by [7]. It is in contrast with the values estimated in [3] where they were less than 0.4. The estimation of the Calvo parameters for home producers and import retailers implies that prices of home producers last approximately 3 quarters and prices of imports last 4 quarters. More flexible prices were found out in [8].[8]

Monetary policy is conducted by a modified Taylor rule. The smoothing parameter of interest rate is 0.85 which is higher compared to estimated value for the Czech economy in [6].[9] The relative weights of monetary policy on inflation and output gap are estimated at 1.23 and 0.28 respectively. The weight on inflation is more than one so the Taylor principle is hold. The parameters of the monetary policy reaction function confirm the inflation-targeting regime conducted by the central bank because the weight on inflation is higher than the weight on the output gap.

The foreign sector autoregressive parameters of output gap, inflation and nominal interest rate are estimated at 0.63, 0.62 and 0.43 respectively. The autoregressive parameter of foreign output is relatively less than in the similar studies of [6], [7] and [8].[10] Table 1 also includes parameters of inertias of technology, cost-push shock and risk premium shock. The most significant influence of shocks on the Slovak economy are technology shock, cost-push shock and risk premium shocks because of the estimated posterior mean of standard deviations of these shock.

[5] We tried several estimations with different prior specifications.
[6] They estimated value about 0.9.
[7] The estimated value of inverse elasticity of intertemporal substitution in consumption in [7] is 0.53.
[8] The prices of home producers and importers last 2.17 and 1.85 quarters respectively.
[9] The value of the smoothing parameter was estimated at 0.64.
[10] The estimations of this parameter in [6], [7] and [8] are 0.80, 0.91 and 0.91 respectively.

Table 1: Estimation of the parameters

	prior mean	posterior mean	90% confidential interval		prior distribution	standard deviation
h	0.50	0.11	0.02	0.20	beta	0.20
σ	1.00	0.47	0.19	0.78	gamm	0.50
φ	1.00	0.85	0.20	1.48	gamm	0.50
α	0.70	0.38	0.19	0.56	beta	0.20
δ	0.50	0.70	0.49	0.93	beta	0.20
θ_h	0.50	0.68	0.57	0.79	beta	0.20
θ_f	0.50	0.75	0.59	0.92	beta	0.20
η	1.00	1.02	0.27	1.67	gamm	0.50
ρ	0.50	0.85	0.80	0.91	beta	0.20
γ_π	1.50	1.23	0.92	1.52	gamm	0.20
γ_y	0.20	0.28	0.18	0.38	gamm	0.05
ρ_a	0.60	0.59	0.39	0.78	beta	0.15
ρ_y	0.40	0.63	0.46	0.80	beta	0.15
ρ_i	0.40	0.62	0.46	0.79	beta	0.15
ρ_π	0.40	0.47	0.28	0.66	beta	0.15
ρ_g	0.70	0.54	0.32	0.75	beta	0.15
ρ_s	0.70	0.65	0.51	0.80	beta	0.15
σ_a	1.00	3.06	1.43	4.86	invg	Inf
σ_M	1.00	0.28	0.21	0.34	invg	Inf
σ_g	1.00	1.22	0.75	1.72	invg	Inf
σ_s	1.00	0.92	0.55	1.27	invg	Inf
σ_y	1.00	0.42	0.34	0.51	invg	Inf
σ_i	1.00	0.15	0.12	0.18	invg	Inf
σ_π	1.00	0.20	0.16	0.23	invg	Inf

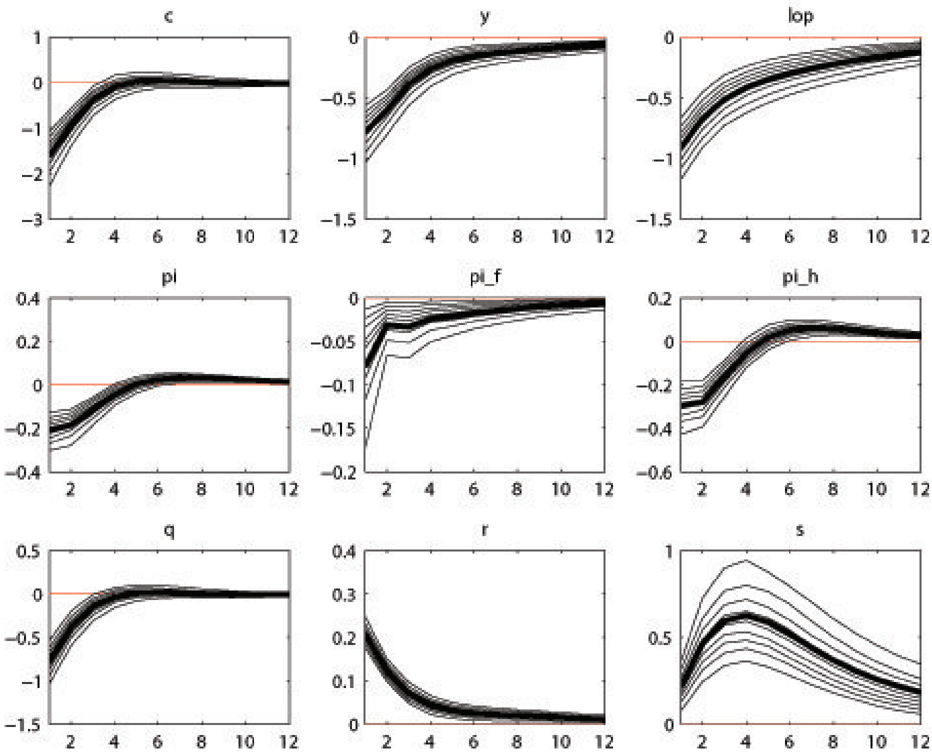
4 Analysis of behaviour

This section examines behaviour of the estimated model by impulse response function analysis. Model structure incorporates seven shocks – technology shock (Eq. (10)), cost-push shock (Eq. (14)), risk premium shock (Eq. (15)), monetary policy shock (Eq. (8)) and three foreign shocks of output gap, inflation and nominal interest rate (Eq. (11), (12) and (13)), so that seven impacts of these shocks could be analysed. In this paper we discuss effects of monetary policy shock.

The response of the variables to a positive monetary policy shock is depicted in Fig.1. Monetary authority immediately reacts to the shock by increasing nominal interest rate by about 0.2%. The positive interest rate differential attracts foreign investors, which is expressed by appreciation of the domestic currency. The appreciation of real exchange rate causes a fall of l.o.p[11] under the long run trend. The consumption is also reduced by the influence of appreciation and by postponing its current consumption because of its relatively high cost to the future consumption. The output gap also drops under the long run trend. Apart from consumption, the negative l.o.p. gap has also a negative effect on output gap which is decreased by a positive gap of terms of trade. The increase of the domestic goods' competitiveness at international market is produced by the fact that inflation gap of domestic producers decrease more than inflation gap of the importers (the rate of q-o-q imported inflation fall by 0.075 percentage points and q-o-q inflation rate of domestic goods decrease by 0.3 percentage points). It is worth noticing that the impulse response functions do not have a hump-shape, especially consumption (like in [6], [7] and [8]), which is caused by the low value of parameter h.

[11] l.o.p means the law of one price.

Figure 1: Impulse response functions of positive monetary shock



5 Conclusion

This paper estimates New Keynesian model of a small open economy for Slovakia. The model incorporates real and nominal rigidities according to [3]. The existence of a real rigidity – the habit formation – has a small effect in Slovak economy. On the other hand, the nominal rigidities cause that the price contracts of domestic producers and importers are fixed for 3 and 4 quarters, on average. The most significant shock is the technology shock. The ability of the model to describe the dynamic development of the economy is assessed by the impulse response functions, specifically on the monetary policy shock. This shock has the largest impact on consumption and home inflation reaches its steady state level after ten quarters. The model needs to be adjusted in order to include the acceptation of Euro for the future research.

References

[1] CALVO, G.: *Staggered Contracts in a Utility-Maximising Framework*, Journal of Monetary Economics,12, 1983, pages 383-398

[2] CARAIANI, P.: *The impact of monetary policy shocks in a small open economy*, Ekonomický časopis/Journal of Economics, Volume 56, 2008, No. 6: 551-564

[3] JUSTINIANO, A., PRESTON, B.: *Small open economy DSGE models: Specification, estimation and model fit*, Department of Economic, Columbia University, Working Paper, 2004

[4] LIU, P.: *A small New Keynesian model of the New Zealand economy*, Reserve Bank of New Zealand, Reserve Bank of New Zealand Discussion Paper Series DP2006/03, 2006

[5] MONACELLI, T.: *Monetary policy in a low pass-through environment, Journal of Money, Credit, and Banking*, Vol. 37, No 6, The Ohio State University, 2005

[6] MUSIL, K., VAŠÍČEK, O.: *Behavior of the Czech Economy: New Open Economy Macroeconomics DSGE Model*, CVKS MU, working paper No 23/2006, Brno 2006

[7] REMO, A.: *Monetary policy and stability of Czech Economy: Optimal Commitment Policy in NOEM DSGE Framework*, CVKS MU, working paper No 28/2008, Brno 2008

[8] ZEMAN J., SENAJ, M.: *Modeling the development of the Slovak economy using a basic DSGE model*, Biatec volume 16 No. 3/2008

Appendix

Table 2: Description of parameters

Parameters	Description of parameters
$h \in (0;1)$	degree of habit formation
$\beta \in (0;1)$	discount factor; calibrated
$\sigma > 0$	inverse elasticity of intertemporal substitution in consumption
$\varphi > 0$	inverse elasticity of labour supply
$\alpha \in (0;1)$	degree of openness
$\delta \in (0;1)$	degree of price indexation
$\theta_H \in (0;1)$	fraction of non-optimizing home producers
$\theta_F \in (0;1)$	fraction of non-optimizing importers
$\eta > 0$	elasticity of substitution between home and foreign goods
$\rho \in (0;1)$	smoothing of interest rate
$\gamma_\pi > 0$	weight of monetary policy on inflation
$\gamma_y > 0$	weight of monetary policy on output gap
$\rho_a, \rho_y, \rho_i, \rho_\pi, \rho_g, \rho_s \in (0;1)$	technology shock, foreign output, foreign interest rate, foreign inflation, cost-push shock and risk premium shock AR1 parameters

GLOBAL SENSITIVITY ANALYSIS OF A DSGE MODEL OF THE CZECH ECONOMY

Jan Čapek and Osvald Vašíček*

ESF MU, VCKCE ESF MU
capek@econ.muni.cz

Abstract. The paper shows some interesting results of Global Sensitivity Analysis on a particular dynamic stochastic general equilibrium model. The key behavior of the Czech economy is approximated by Lubik and Schorfheide model, which is a small-scale structural general equilibrium model of a small open economy. The sensitivity analysis class of methods presented in the paper consists of various individual analyses. Stability mapping analysis detects the parameters which are responsible for potential instability or indeterminacy in the model. Mapping the fit is a useful tool to learn about the linkages that drive the fit of trajectories of particular variables to data. Information provided by the results of mapping the fit can be used to unveil possible trade-offs between the fit of individual observables and maybe also to amend model structure or calibrate parameters properly in order to increase the fit of variables of researcher's interest. Other individual analyses are just mentioned.
Keywords. Global sensitivity analysis, DSGE model, stability mapping, mapping the fit, high dimensional model representation, Morris sampling .

Introduction

This contribution introduces the reader to Global Sensitivity Analysis due to Saltelli et al. (2004 and 2008). The prototypical model for the exercise is the Lubik and Schorfheide (2003) model introduced in subsection 1.1. The same model is used by Ratto (2008a) with Canadian data set. This paper uses Czech data set and offers comparisons to Ratto's (2008a) paper.

The term "Global Sensitivity Analysis" is used in this contribution in the same sense that it is established and used in Saltelli et al. (2004 and 2008) or Ratto (2008a). The word "Global" means that the analysis is not "local", that is, it doesn't approximate solutions around one given point in space (like Taylor approximation does). Global methods use more points judiciously drawn from space and therefore overcome problems when the model is not linear.

A possible definition of sensitivity analysis is the following: *"The study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input"* (Saltelli et al. 2004). Sensitivity analysis is therefore not only the well-know ad-hoc exercise, when researcher arbitrarily changes inputs of the model and observes changes in output. It also encompasses methods/procedures that can somehow describe relations between inputs and output.

Global Sensitivity Analysis is therefore an open group of individual analyses or methods, some of which are introduced in this paper (in subsections 2–4).

1 Preliminaries

1.1 The model

This paper presents some results of an analysis of a single model, which is a model of Lubik and Schorfheide (2003)[1]. It is a small-scale structural general equilibrium model of a small economy. This paper uses Czech data set so that model equations (1)–(8) describe elementary behavior of the Czech economy. Generally, Δ denotes first difference so that e. g. $\Delta\xi_t = \xi_t - \xi_{t-1}$, star superscript (*) relates to a foreign economy,

* We gratefully acknowledge support by MSMT project Research Centers 1M0524 and funding of specific research at ESF MU.

[1] Cited from Ratto (2008a, p. 123); (hereafter LS model). The citation of the model equations is not literal, because sixth equation on page 123 of Ratto (2008a) is "$y_s = \rho_{y^*} y_{t-1} + e_{y^*,t}$", which is obviously an error. Also, in order to prevent confusion, an explanation to variables that have letter e in its notations follows here: e_t in equations (3) and (4) means nominal exchange rate, whereas $e_{\cdot,t}$ in equations (4)–(8) stands for exogenous shocks. Although the notation varies in the number of subscripts, the difference might not be obvious at first glance.

subscript t denotes (relative) time and E_t denotes rational expectations made in time t. Also note that $rr = -400 \cdot \log(\beta)$.

$$\begin{aligned} y_t = {} & E_t y_{t+1} - [\tau + \alpha(2-\alpha)(1-\tau)](R_t - E_t \pi_{t+1}) \\ & - \alpha[\tau + \alpha(2-\alpha)(1-\tau)] E_t \Delta q_{t+1} - \alpha(2-\alpha)\frac{1-\tau}{\tau}\Delta y^*_{t+1} - E_t z_{t+1} \end{aligned} \tag{1}$$

$$\pi_t = \beta E_t \pi_{t+1} + \alpha\beta E_t \Delta y_{t+1} - \alpha \Delta q_t + \frac{k}{\tau + \alpha(2-\alpha)(1-\tau)}(y_t - \bar{y}_t) \tag{2}$$

$$\pi_t = \Delta e_t + (1-\alpha)\Delta q_t + \pi^*_t \tag{3}$$

$$R_t = \rho_R R_{t-1} + (1-\rho_R)(\psi_1 \pi_t + \psi_2(y_t - \bar{y}_t) + \psi_3 \Delta e_t) + e_{R,t} \tag{4}$$

$$\Delta q_t = \rho_q \Delta q_{t-1} + e_{q,t} \tag{5}$$

$$y^*_t = \rho_{y^*} y^*_{t-1} + e_{y^*,t} \tag{6}$$

$$\pi^*_t = \rho_{\pi^*} \pi^*_{t-1} + e_{\pi^*,t} \tag{7}$$

$$z_t = \rho_z z_{t-1} + e_{z,t} \tag{8}$$

Equation (1) is an open economy IS curve. If $\alpha = 0$, the equation becomes closed economy variant of IS equation. If $\tau = 1$, the world output shocks Δy^*_{t+1} drops out of IS equation and since it is not present in any other equation but AR1 process (6), it drops out of the system completely.

The open economy Phillips curve (2) also collapses to closed economy version if $\alpha = 0$. Consumer price index CPI is introduced in (3) with an assumption of a relative version of purchasing power parity.

Equation (4) is a monetary rule, or in another words, a nominal interest rate equation. It describes, how the monetary authority sets its instrument, when inflation or output depart from their targets or when the currency appreciates or depreciates.

Remaining model equations are just AR1 processes that describe the course of terms of trade, foreign output and inflation, and technological progress.

1.2 The data

The data span from the second quarter of 1996 to the fourth quarter of 2008. The source of all data is Czech Statistical Office and are per cent. There are five time series used, these are: output growth, inflation, interest rate, change in nominal exchange rate, and the change in terms of trade.

1.3 Used software

The main tool used in the analysis is the software package Dynare version 4.0.3. The analysis also requires Global Sensitivity Analysis (hereafter GSA) toolbox by Marco Ratto, which is – according to Dynare site – beginning to be added to Dynare version 4. This toolbox is downloadable from Euro-area Economy Modelling Centre web pages `http://eemc.jrc.ec.europa.eu/`. Documentation for these software packages is semifinished and is in Griffoli (2007) for Dynare version 4 and in Ratto (2008b) for GSA.

2 Stability mapping

2.1 Theory

Stability mapping helps to detect parameters X_i that are responsible for possible "bad behavior" of the model. First step of the computation is to define two subsets of a full domain: subset B produces behavior (= good behavior of the model), subset $\overline{B}$ produces non-behavior (= bad behavior of the model).

N Monte Carlo simulations are then run over the domain, which results in two subsets, $(X_i|B)$ of size n and $(X_i|\overline{B})$ of size $\overline{n}$, where $n + \overline{n} = N$. The two sub-samples may come from different probability density functions (PDFs) $f_n(X_i|B)$ and $f_{\overline{n}}(X_i|\overline{B})$. Corresponding cumulative distribution functions (CDFs) are $F_n(X_i|B)$ and $F_{\overline{n}}(X_i|\overline{B})$.

If $F_n(X_i|B)$ and $F_{\overline{n}}(X_i|\overline{B})$ differ for a given parameter X_i, the parameter may drive bad behavior of the model if its value falls within $\overline{B}$ subset. The shape of $F_{\overline{n}}(X_i|\overline{B})$ indicates, whether rather smaller or higher values of X_i drive the non-behavior. If the non-behavior CDF is to the left from behavior CDF, it indicates that rather smaller values of X_i are more likely to drive non-behavior. On the other hand, if the non-behavior CDF is to the right from the behavior CDF, it suggests that rather bigger values of X_i drive non-behavior.

In order to obtain also numerical results, a statistic that computes the greatest distance between behavior and non-behavior CDFs is computed. More formally, the (so-called) Smirnov d statistic is defined as

$$d_{n,\overline{n}}(X_i) = \sup ||F_n(X_i|B) - F_{\overline{n}}(X_i|\overline{B})||$$

The Smirnov d statistic has a domain $[0, 1]$, where 0 means that the two (behavior and non-behavior) CDFs perfectly overlap and 1 means that the two underlying subsets B and $\overline{B}$ have no common elements. In other words, $d = 1$ means that one of the CDFs reaches unity before the other increases from zero.

This analysis doesn't use data, so the results are just a matter of model relations (equations) and parameter calibration, not the data itself. The results are therefore the same as in Ratto (2008a).

3 Mapping the fit

3.1 Theory

Since DSGE models consist of a number of observed variables, which should fit the data as well as possible, mapping the fit may be a useful tool to learn about the linkages that drive the fit of trajectories of particular variables to data. Information provided by the results of mapping the fit can be used to unveil possible trade-offs and maybe also amend model structure or calibrate parameters properly in order to increase the fit of variables of interest.

The procedure is carried out as follows:

1. Structural parameters are sampled from posterior distribution,
2. RMSE[2] of 1-step-ahead prediction is computed for each of observed series,
3. 10 % of lowest RMSE is defined as behavioral and B is defined as a subset of parameter values producing these behavioral results and
4. the calculations results in a number of distributions $f_j(X_i|B)$ that represent the contribution of parameter X_i to best possible fit of j-th observed series.

Plotting the distributions (or better the CDFs) is one step further to trace possible trade-offs. A trade-off is present, when at least two distributions differ from posterior distribution (denoted in Figures as `base`) and differ from each other.

3.2 Results for LS model

Ratto (2008a, p. 126) lists these parameters as the ones bearing biggest trade-offs: $\psi_1, \psi_3, \rho_R, \alpha, k, \rho_q, \rho_{y^*}$. This subsection compares and contrasts results obtained by Ratto and by this paper.

In Ratto (2008a), parameters ψ_1 and ψ_3 both represent similar trade-offs, albeit a bit smaller in volume in case of parameter ψ_3. Both parameters should be rather smaller in order to fit inflation π and rather larger in order to fit the change in nominal exchange rate Δe. Realization of the LS model on Czech data looks similarly – see figure 1, panel one and three. ψ_1 and ψ_3 should be smaller in order to fit inflation optimally and larger in order to fit the change in nominal exchange rate, as in Ratto's realization on Canadian data. The magnitude of these trade-offs is however visibly smaller. Also, larger value of ψ_1 than its posterior distribution supports a better fit of output and interest rate.

Indications of trade-offs associated with parameter ρ_R are similar in Canadian and Czech realization of the LS model. Both realizations suggest that ρ_R should be lower in order to fit inflation better and higher in order to fit interest rate and the change in nominal exchange rate better. However, the magnitude is different. A deviation of the parameter from its posterior distribution is higher in the Czech model in case of interest rate and the change in nominal exchange rate: see Fig. 1, panel 4.

Parameter α fits all variables rather well in both country realizations. Deviations are small and different in the two countries.

Parameter k is much more interesting. Ratto (2008a) states that all observed series have a preference for a larger value of k. The Czech model displays no preference for bigger k from all variables. Interest rate and the change in nominal exchange rate prefer lower k then posterior distribution. For details see Fig. 2, panel 2.

[2] root mean square error

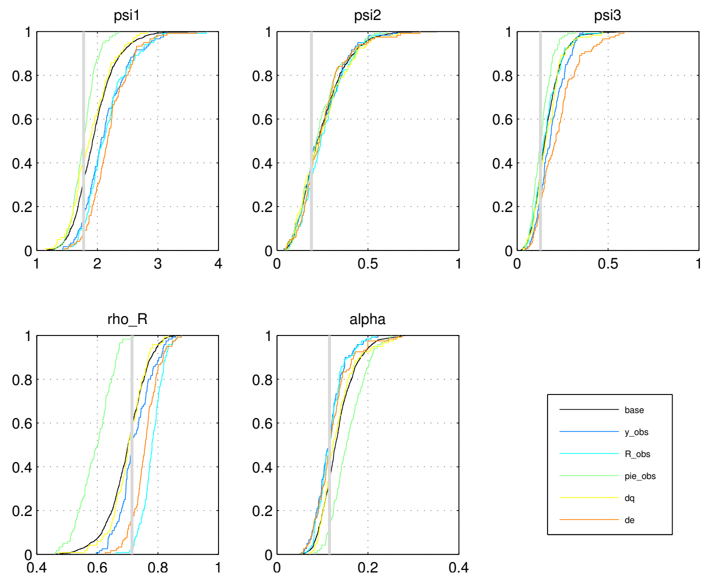


Fig. 1. Cumulative posterior distributions (base) and the distributions of the filtered samples corresponding to the best fit for each singular observed series. Grey vertical lines denote posterior mode. (1 of 2)

Parameter ρ_q has also interesting trade-offs. Δq alone would imply a given value of the parameter approximately 0.5, which is almost at posterior mode. Other variables fit data rather well with ρ_q at posterior distribution in both country realizations.

No other parameter causes major conflicts between fit of the variables and that holds for both Canadian and Czech data realization.

4 High dimensional model representation / Reduced form mapping

4.1 Theory

Due to the lack of space, the reader is referred e.g. to Ratto (2008a) for theoretical background.[3] We shall only define the most important measures here.

The first order sensitivity index can be defined as $S_i = V_i/V$, which is a scalar measure that shows the relative importance of structural parameter X_i on the variance of Y (V being the symbol for variance). A reduced form of a DSGE model can be written as $y_t = Ty_{t-1} + Bu_t$ and generic output Y can be the entries of matrices T or B. These generic outputs Y are also called reduced form coefficients.

4.2 Results for LS model

According to the left panel of Fig. 3, the most influential parameters are ψ_1, ρ_R, k and τ. The least influential parameters seem to be ψ_2, rr and ρ_{y^*}. Parameter ψ_1 has highest median, but upper quartile and upper whisker isn't much higher. Such characteristics could mean that ψ_1 is important for many reduced form coefficients, but is rarely very important. Parameter ρ_R has highest upper quartile and upper whisker, but it has lower median than ψ_1. This backward-looking parameter of the monetary rule is therefore quite important for many reduced form coefficients and very important for some, too. On the other hand, lower median would suggest that the number of reduced form coefficients for which is ρ_R important is lower than in the case of ψ_1. Parameters k and τ are even less important than the two just discussed. Both have lower values of median and upper quartile. Upper whisker is somewhat lower, too.

Both parameters ψ_2 and rr have all sensitivity indices virtually zero, which should suggest that these parameters are unimportant for all possible reduced form coefficients. Boxplots of $\alpha, \rho_q, \rho_A, \rho_{y^*}$ and ρ_{π^*} represent rather peculiar results. All of these parameters have median and upper quartile virtually zero, but have some high outliers. Such results mean that these parameters are unimportant for most reduced

[3] For primary literature, see e.g. Li et al. (2002 and 2006) and Sobol' (1993).

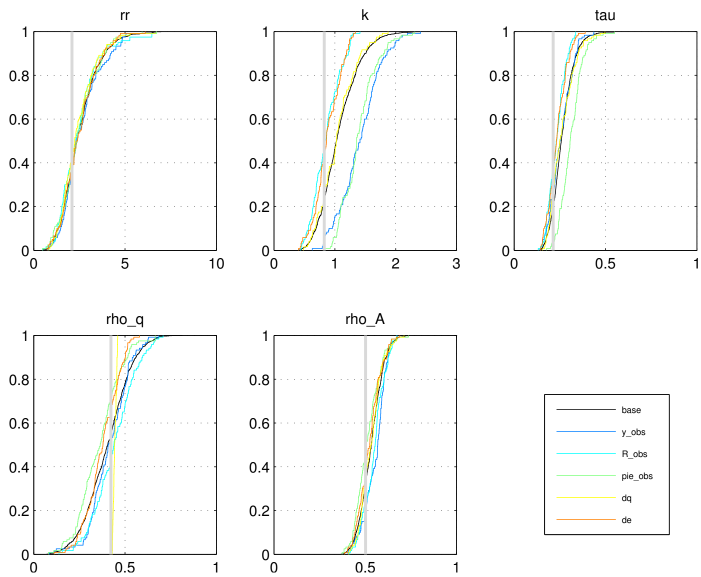


Fig. 2. Cumulative posterior distributions (base) and the distributions of the filtered samples corresponding to the best fit for each singular observed series. Grey vertical lines denote posterior mode. (2 of 2)

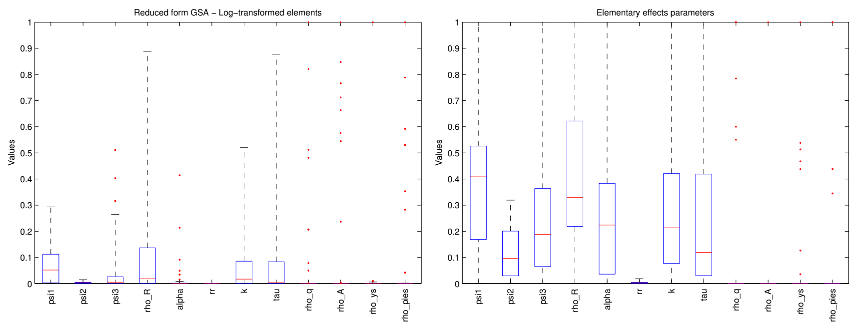


Fig. 3. Left panel: Boxplots of sensitivity indices. (All endogenous variables vs. all exogenous and all lagged endogenous with $\log(-Y)$ transformation). Right panel: Elementary effects computed as a screening procedure with Morris sampling.

form coefficients, but have important (in case of α) or very important (in case of the remaining four parameters) influence on some reduced form parameters.

Right panel of Fig. 3 shows similar results, but computed as a screening procedure with the so-called Morris sampling.[4] This screening can be computed as a preliminary check of importance of parameters. Its main advantage is that it takes about 200 times less time to compute in comparison with the full results (left panel of Fig. 3). Its understandable drawback is that the results are only approximate. The patterns (most influential parameters, noninfluential parameters) are similar in both panels, but the actual values of sensitivity indices/elementary effects differ. These details are not discussed here due to the lack of space.

Conclusion

Global sensitivity analysis helps to better understand linkages that drive the behavior of a DSGE model. Various individual tools of GSA are used to illuminate dependencies in separate parts of DSGE models and together form a unified picture. This paper presented mainly Mapping the fit, which helps to detect trade-offs among parameter values in order to attain the best possible fit of an observable variable. Less space is devoted to the distribution of sensitivity indices, which can be used as a measure of the importance of individual parameters.

[4] See Morris (1991).

References

ČAPEK, J., VAŠÍČEK, O. *Sensitivity Analysis of a DSGE model*, CVKS working paper No. 1/2009

JUILLARD, M. et al. *Dynare software package.* Homepage: `http://www.cepremap.cnrs.fr/dynare/`

GRIFFOLI, T. M. *Dynare v4 – User guide*, 2007, cited draft: March 2007.

LI, G. et al. *Practical approaches to construct RS-HDMR component functions.* Journal of Physical Chemistry 106, 8721-8733.

LI, G. et al. *Random Sampling-High Dimensional Model Representation (RS-HDMR) and orthogonality of its different order component functions.* Journal of Physical Chemistry A 110, 24742485.

LUBIK, T., SCHORFHEIDE, F. *Do Central Banks Respond to Exchange Rate Movements? A Structural Investigation*, Economics Working Paper Archive, n. 505, The Johns Hopkins University, Department of Economics, 2003

MORRIS, M. D. *Factorial sampling plans for preliminary computational experiments.* Technometrics, 33, 161174.

RATTO, M. *Analysing DSGE Models with Global Sensitivity Analysis*, Computational Economics, Volume 31, Number 2 / March, 2008, pp 115-139, ISSN 0927-7099 (Print) 1572-9974 (Online)

RATTO, M. *Sensitivity Analysis Toolbox for DYNARE.* version: April 18, 2008.

SALTELLI, A. et al. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*, John Wiley & Sons, Ltd., 2004, ISBN 0-470-87093-1

SALTELLI, A. et al. *Global sensitivity analysis. The Primer.* John Wiley & Sons, Ltd., 2008, ISBN 978-0-470-05997-5

SOBOL', I. *Sensitivity analysis for non-linear mathematical models.* Mathematical Modeling & Computational Experiment (Engl. Transl.), 1993, 1, 407414.

EVALUATION OF THE INFLATION TARGETING EFFECTIVENESS

Ondřej Čížek

University of Economics, Prague
Faculty of Informatics and Statistics, Econometrics and Operations Research
cizeko@vse.cz

Abstract. The goal of this paper is to analyze the impact of inflation targeting on disinflation process from the empirical point of view. The panel data methods are applied to estimate the modified Phillips curve, which describes the evolution of the inflation. The case of the transition economies is treated separately from developed market economies.
In transition economies, there was often problem with high inflation, so the main purpose of the inflation targeting in these countries was to get inflation under control. However, in developed market economies the inflation has been stable in the last 10 years, so the inflation in these economies is modelled slightly differently from the case of transition economies.
It is stressed the importance of correct economic interpretation of the results from the regression. In the case of transition economies, it may look like from the first point of view that the inflation targeting strategy is enormously effective way how to get control over the inflation, but more detailed examination shows that it is not true due to the above mentioned differences between the two groups of countries. In the case of developed market economies, the results from the regression are biased to the other side. From the first point of view, it looks like the inflation targeting strategy gives the same results as other monetary regimes. But in the analogy to the first case of the transition economies, it is then shown that the results are more complicated to interpret.
Keywords. Inflation targeting, panel data

1. Introduction

The goal of the paper is to evaluate the effectiveness of the inflation targeting policy in transition economies. To do this, panel data approach will be used. The data set comprises of eight transitive countries. For each unit (country) ten observations are available from 1998 to 2007, so the data are balanced. Because the selection of the units to the data set is not random, the fixed effect model is used.

The structure of the paper is as follows. The basic model based on the Phillips curve is formulated in the 2. chapter. In the next chapter the method of the parameter estimation is briefly described. In the 4. chapter the results are interpreted. The final 5. chapter concludes.

2. Formulation of the Model

The model is based on the Phillips curve, which relates unemployment (or product) with inflation. This basic Phillips curve is modified here to encompass the following additional variables:

time,
dummy variable for assessing inflation targeting effectiveness,
dummy variable for individual effects.

The econometric model is formulated as follows

$$\pi_{it} = \alpha_i + \beta \cdot t + \gamma \cdot CI_{it} + \delta \cdot U_{it} + \eta_{it}, \qquad (0.1)$$

where π_{it} is inflation in the country i at time t,

α_i is individual effect in the country i,

t is time,

CI_{it} is dummy binary variable with value 1 meaning that in the country i at time t the inflation targeting policy is applied,

U_{it} is the rate of unemployment in the country i at time t,

η_{it} is random error in the country i at time t.

All the transition economies in the central and eastern Europe had after transition problems with high inflation, which was caused by the price deformations in central planned economies. The goal of all these countries was therefore this high inflation gradually decrease so the expected sign for the parameter β is negative.

According to economic theory analysing relation between unemployment and inflation we will assume that the parameter δ will be negative. Bud primarily, we will be interested in the sign of the coefficient γ, because this coefficient measures the effectiveness of the inflation targeting policy. The negative sign of this coefficient means that the inflation targeting strategy works better to lower inflation.

It is important to stress at this point that the panel data structure has the advantage that it can describe the dynamics. In other words, the cross-section data can say only if inflation targeting countries has at a certain point in time lower inflation than other countries. But with panel data, we also have the information whether or not the lower inflation at these countries is correlated with the time of introducing the inflation targeting regime.

The variable time describe the effort of the inflation targeting countries to achieve convergence criteria. The variable unemployment model the transmission channel between economic activity and inflation and α_i comprises all the other things that cannot be modelled explicitly. The variable CI model the influence of the implementation of the inflation targeting regime.

3. Data

The data set comprises of eight countries, which are the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia. For each country ten observations are available from 1998 to 2007, so the data are balanced.

The source of the data on inflation and unemployment was Eurostat. Which countries target inflation and the year of implementation was found in internet. The inflation targeting countries are the Czech Republic (from 1997), Hungary (from 2001), Poland (from 1998) and Slovakia (from 2005).

The main reason for the data to begin in 1998 is the fact that most transitive economies had very high inflation in the first half of the nineties caused by the transition and so the development of these countries in these years was determined by other factors than those used in this paper. The other reason is that data on the unemployment was not available before this year.

The data on inflation and unemployment are for completeness summarized in the following tables.

Inflation	***Czech Republic***	***Estonia***	***Hungary***	***Latvia***	***Lithuania***	***Poland***	***Slovakia***	***Slovenia***
1998	9,7	8,8	14,2	4,3	5,4	11,8	6,7	7,9
1999	1,8	3,1	10	2,1	1,5	7,2	10,4	6,1
2000	3,9	3,9	10	2,6	1,1	10,1	12,2	8,9
2001	4,5	5,6	9,1	2,5	1,6	5,3	7,2	8,6
2002	1,4	3,6	5,2	2	0,3	1,9	3,5	7,5
2003	-0,1	1,4	4,7	2,9	-1,1	0,7	8,4	5,7
2004	2,6	3	6,8	6,2	1,2	3,6	7,5	3,7
2005	1,6	4,1	3,5	6,9	2,7	2,2	2,8	2,5
2006	2,1	4,4	4	6,6	3,8	1,3	4,3	2,5
2007	3	6,7	7,9	10,1	5,8	2,6	1,9	3,8

Table 1: Inflation

Unemployment	Czech Republic	Estonia	Hungary	Latvia	Lithuania	Poland	Slovakia	Slovenia
1998	6,4	9,2	8,4	14,3	13,2	10,2	12,6	7,4
1999	8,6	11,3	6,9	14	13,7	13,4	16,4	7,3
2000	8,7	12,8	6,4	13,7	16,4	16,2	18,8	6,7
2001	8	12,4	5,7	12,9	16,5	18,3	19,3	6,2
2002	7,3	10,3	5,8	12,2	13,5	20	18,7	6,3
2003	7,8	10	5,9	10,5	12,5	19,7	17,6	6,7
2004	8,3	9,7	6,1	10,4	11,4	19	18,2	6,3
2005	7,9	7,9	7,2	8,9	8,3	17,8	16,3	6,5
2006	7,2	5,9	7,5	6,8	5,6	13,9	13,4	6
2007	5,3	4,7	7,4	6	4,3	9,6	11,1	4,9

Table 2: Unemployment

4. Parameter estimation

The parameters of this model will be estimated by the Least Squares Dummy Variable method (LSDV), which is in fact application of the least squares to the regression

$$\pi = D\vec{\alpha} + X\vec{\beta} + \eta, \tag{0.2}$$

where

$$D = \begin{bmatrix} 1 & 0 & \dots & 0 \\ & \vdots & & \\ 1 & 0 & \dots & 0 \\ \hline 0 & 1 & \dots & 0 \\ & \vdots & & \\ 0 & 1 & \dots & 0 \\ \hline & \vdots & & \\ \hline 0 & 0 & \dots & 1 \\ & \vdots & & \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad \vec{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_8 \end{bmatrix}, \qquad X = \begin{bmatrix} 1 & CI_{1,1} & U_{1,1} \\ & \vdots & \\ 10 & CI_{1,10} & U_{1,10} \\ \hline 1 & CI_{2,1} & U_{2,1} \\ & \vdots & \\ 10 & CI_{2,10} & U_{2,10} \\ \hline & \vdots & \\ \hline 1 & CI_{8,1} & U_{8,1} \\ & \vdots & \\ 10 & CI_{8,10} & U_{8,10} \end{bmatrix}, \quad \vec{\beta} = \begin{bmatrix} \beta \\ \gamma \\ \delta \end{bmatrix}.$$

However, practically all econometric software has procedures for parameter estimation of the panel data models. For example, with software PcGive it can be done as follows. In *Package* the item *Panel Data Models* will firstly be activated. Then in the *Model* the *Static Panel Methods* will activate the window in which the model can be formulated. In the next steps, the method called LSDV will be chosen.

5. Results

The software PcGive was used to obtain the results which are as follows[1]

DPD(1) Modelling pi by LSDV

	Coefficient	*Std.Error*	*t-value*	*t-prob*
t	-0.496213	0.1325	-3.75	0.000
CI	-4.07800	0.8082	-5.05	0.000
U	-0.583041	0.1428	-4.08	0.000
I0	14.259 1	1.318	10.8	0.000
I1	12.6814	1.671	7.59	0.000
I2	17.0476	1.190	14.3	0.000
I3	13.7451	1.874	7.33	0.000
I4	11.6875	1.950	5.99	0.000
I5	20.6950	2.467	8.39	0.000
I6	19.9112	2.542	7.83	0.000
I7	12.1981	1.297	9.40	0.000

Sigma	2.138193	sigma^2	4.571871
R^2	0.6135782		
RSS	315.45906972	TSS	816.3595
no. of observations	80	no. of parameters	11

Before we interpret the results, the test whether or not it was correct to use the fixed effect model rather than pooled OLS regression will be performed. The application of the OLS regression to the model $\pi_{it} = \alpha + \beta \cdot t + \gamma \cdot CI_{it} + \delta \cdot U_{it} + \eta_{it}$ gives coefficient of determination $R^2_{OLS} = 0.178067$ and so

$$F = \frac{\dfrac{R^2_{MMP} - R^2_{MN\check{C}}}{n-1}}{\dfrac{1-R^2_{MMP}}{nT-n-k}} = \frac{\dfrac{0.613678 - 0.178067}{8-1}}{\dfrac{1-0.613578}{80-8-3}} = 11.1.$$

The 95% quantile of the Fisher distribution[2] $F(7,77)$ is 2.1 and therefore the null hypotheses saying that there are no individual effects is refused.

The statistical interpretation of the above results is straightforward. The t-tests say that all the parameters in the regression are statistically significant even in the 1% level of significance.

The economic interpretation is also straightforward. The estimated coefficient before the variable time (t) and unemployment (U) has the expected signs. The important however is that the estimated parameter before the variable CI is negative which means that the inflation targeting regime is more effective than the other ones. On the other hand, the absolute value of this coefficient is too high to believe that the result are not somehow distorted, because it says that the implementation of the inflation targeting regime lowered inflation more than 4% per year.

Let's look at the above formulated model in more details. It was already said that the panel data model can describe dynamics. In other words, with panel data structure it is possible to say whether decline of the inflation in the inflation targeting countries came at the same time as the implementation of the inflation targeting regime. If this is true than it seems that the inflation targeting policy is more effective than other policies. But this

[1] A shortcut DPD indicates that panel data package was used to estimate the parameter.
[2]The Fisher distribution is a probability distribution well known from statistics.

argument is not correct as it is highly possible that countries which finally decided to implement inflation targeting had bigger problems with inflation after transition than other countries. However, in these countries it is possible to expect more dramatic decline in inflation after realization of big problems with it. And so this dramatic decline of inflation would probably happen even if the inflation targeting regime was not implemented.

The outlined hypothesis is supported by the data. The Hungary (targeting inflation from 2001) and Slovakia (targeting inflation from 2005) had from 1998 to 2005 almost highest inflation from the selected countries. The same applies for Poland (targeting inflation from 1998) from 1998 to 2001. The Czech Republic in 1998 (which is also the year of the implementation of the inflation targeting) had after Hungary and Poland the highest inflation.

Let's therefore look at the individual effects of the inflation targeting countries

Czech Republic (from 1997)	14.2,
Hungary (from 2001)	17.0,
Poland (from 1998)	20.7,
Slovakia (from 2005)	19.9,

and at the individual effects of the other countries

Estonia	12.7,
Latvia	13.0,
Lithuania	11.7,
Slovenia	12.2.

The main difference is straightforward. Individual effects in the inflation targeting countries are much higher than in the other countries. The question is how to interpret this result correctly. Assume that the estimated individual effects are good estimates of the true parameters in the Czech Republic, Hungary, Poland and Slovakia and let's ask a question what the inflation in these countries would be if they didn't implement the inflation targeting regime and had approximately same level of the unemployment. The answer is straightforward – they would have higher inflation. This is in line with the already mentioned argument (the inflation targeting countries implemented inflation targeting regime because they had bigger problems with inflation) explaining high (in the absolute value) estimated coefficient before variable CI .

So we can conclude that higher inflation in the Czech Republic, Hungary, Poland and Slovakia caused high estimated individual effects as well as high (in the absolute value) estimated coefficient before variable CI . The panel data regression therefore correctly recognized the fact that the inflation targeting countries had bigger problems with inflation, but was unable to detect the fact that these countries would have lowered their inflation even without implementation of the inflation targeting regime. The consequence is that the estimated coefficients before variable CI are overestimated and therefore the same applies even for the estimated individual effects.

6. Conclusion

The panel data regression was distorted by the fact that the inflation targeting regime was implemented primarily in countries with bigger problem with inflation. So it is necessary to interpret the results from the regression with caution and it would be nonsense to say that the implementation of the inflation targeting regime will lower inflation more than 4%, which was the very first interpretation of the obtained results.

With regard to the evaluation of the effectiveness of the inflation targeting regime in the transition economies, it can be said only that this regime was successful in lowering high inflation, which was the consequence of the transition. However, it is always necessary to bear in mind that this result can never be interpreted as a causality, but only as a correlation.

References

[1] Hušek, R.: Aplikovaná ekonometrie. VŠE, Praha, 2001

[2] Johnston, J., DiNardo, J.: Econometric Methods, McGraw-Hill, New York, 1997.

A NOTE ON METHODS OF EFFICIENCY MEASUREMENT USED IN HEALTH CARE SERVICES

Martin Dlouhý

University of Economics Prague
School of Informatics and Statistics, Department of Econometrics,
dlouhy@vse.cz

Abstract. The study reviews major groups of methods of quantitative economic analysis that can be applied to efficiency measurement in health care services. Four major groups of methods that are discussed in this study are ratio analysis, data envelopment analysis, econometric modelling, and survivor analysis. Each method has its advantages and disadvantages. It is important for an analyst being aware of these advantages and disadvantages and combine methods by taking into account their limits. In my view, from four groups of method, the usefulness of survivor analysis may be questioned. The method is based on strong economic assumptions about the power of market and is not universal regarding type of efficiency analysis. It should be noted that efficiency or inefficiency in the health care sector does not depend only on managerial abilities of health care managers, but also on regulation environment and payment systems.
Keywords. Efficiency, Ratio Analysis, Data Envelopment Analysis, Econometrics, Survivor Analysis, Health Care.

1. Introduction

Efficiency measurement in health care is both theoretical and practical problem. Health care is an application field with specific context and characteristics, for example final output (health improvement) is hard to measure; the causality between input and output is not always certain; apart from the efficiency of the system, the equity and moral principles have to be taken into account; and there is a high variation in demand for services [8].

The objective of this study is to explore how the different methods of quantitative economic analysis can be successfully applied to efficiency measurement and evaluation in health care services. I will distinguish four major groups of methods: ratio analysis, data envelopment analysis, econometric modelling, and survivor analysis. The brief description of all four methods follows in section two. Advantages and disadvantages of different methods are discussed in section tree.

2. Methods

2.1. Ratio analysis

Ratio analysis is based on calculation of various ratios between one input and one output variable. Examples from health care are ratios such as the number of patients per physician, cost per day, cost per case, number of beds per nurse, etc. Naturally, ratios cannot easily accommodate situations with multiple inputs and outputs [20]. In business practice, individual companies almost universally rely on simple output–input ratios in efficiency and productivity evaluation. Instead of single ratio, a manager or analyst can work with the set of ratios in order to obtain more comprehensive information on efficiency.

2.2. Data Envelopment Analysis

The data envelopment analysis (DEA) was developed by Charnes, Cooper and Rhodes [4] in 1978. DEA defines the relative technical efficiency of a production unit as the ratio of its total weighted physical output to its total weighted physical input. The method allows each production unit to determine its own weights of inputs and outputs in order to maximize its efficiency score. For each production unit, DEA calculates the efficiency score; determines the relative weights of inputs and outputs; and identifies peers for each unit that is not technically efficient. The peers of an inefficient production unit are efficient units with similar combinations of inputs and outputs. The peers can serve as benchmarks, which show potential improvements that the inefficient unit can attain. The formulation of the dual input-oriented DEA model in the matrix form is:

$$\begin{aligned}
\text{minimize} \quad & \theta_q - \varepsilon(e^T s^+ + e^T s^-), \\
\text{subject to} \quad & X\lambda + s^- = \theta_q x_q, \\
& Y\lambda - s^+ = y_q, \\
& \lambda \geq 0,\ s^+ \geq 0,\ s^- \geq 0,
\end{aligned}$$

where θ_q is the efficiency score, λ is the vector of variables, s^+, s^- are vectors of slack variables, x_q and y_q are vectors of inputs and outputs of the evaluated unit, X is a matrix of m inputs and Y is a matrix of r outputs, $e^T = (1, 1, \ldots, 1)$, and ε is infinitesimal constant. The production unit q is technically efficient if the optimal value of variable θ^*_q is one and the values of all slack variables s_i^{+*} and s_i^{-*} equal zero. According to the returns to scale, the theory distinguishes four alternative DEA models: the constant-returns-to-scale model, the variable-returns-to-scale model, the non-increasing-returns-to-scale-model, and the non-decreasing-returns-to-scale model. These models differ in restrictions for the vector λ.

The literature reviews [14, 19] show that DEA and related methods are frequently used for efficiency evaluation in health care. The studies [7, 9, 18] are examples of hospital efficiency evaluation from the Czech Republic.

2.3. Econometric Modelling

An econometric model is formulated in the form of one or more stochastic equations. Econometric modelling (regression analysis) is a parametric method that requires to specify an input-output relationship. Regression analysis allows existence of random noise, offers to estimate confidence intervals and offers more stable accuracy of efficiency estimates because estimates are not dependent on a small subset of units [22]. The shortcoming of regression approach is that it concentrates on the estimation of average behaviour, not on the best (frontier) practice. In economic theory, a production function is defined as a maximum output obtainable from given resources, not as an average or usual level of output. To deal with this problem, the classical regression analysis of production function was extended to the frontier analysis [1, 16, 17]. Let us assume that production model can be written as

$$y_i = f(x_i;\beta) \cdot TE_i,$$

where y_i is the scalar output of *i*-th production unit, x_i is the vector of *n* inputs used by *i*-th unit, $f(x_i;\beta)$ is the production frontier, β is the vector of unknown technology parameters to be estimated, and TE_i is the output-oriented technical efficiency of production unit *i*. The output-oriented technical efficiency TE_i is defined as the ratio of observed output y_i to the maximum feasible output, which is determined by the production frontier $f(x_i;\beta)$. We have

$$TE_i = \frac{y_i}{f(x_i;\beta)}.$$

If the production unit is technically efficient, then technical efficiency $TE_i=1$; if the unit is not technically efficient, then $TE_i<1$. This type of the frontier model is called the deterministic frontier analysis (DFA) because the difference between the observed output y_i and maximum feasible output $f(x_i;\beta)$ is explained by the technical inefficiency. There is no stochastic term, which could represent random shocks in production. In comparison, the stochastic frontier analysis (SFA) assumes the existence of random shocks that affect efficiency of the production process. The deterministic model is modified to a stochastic model as

$$y_i = f(x_i;\beta) \cdot \exp\{v_i\} \cdot TE_i,$$

where $f(x_i;\beta)\exp\{v_i\}$ is the stochastic production frontier, which consists of two parts: the deterministic part $f(x_i;\beta)$ and the stochastic part $\exp\{v_i\}$, where v_i is the two-sided random component with the normal distribution $N(0,\sigma_v^2)$. The output-oriented technical efficiency is given by the ratio

$$TE_i = \frac{y_i}{f(x_i;\beta)\exp\{v_i\}}.$$

Similarly as in the deterministic frontier model, the production unit is technically efficient if TE =1 and inefficient if $TE_i<1$. Advocates of the stochastic frontier analysis argue that it is not possible to ignore random shocks as the deterministic frontier analysis proposes, and therefore, the stochastic frontier analysis should be preferred.

One of the possible estimation methods of deterministic frontier analysis is the Corrected Ordinary Least Squares. This method estimates production frontier in two steps. In the first step, the ordinary least squares method is applied to obtain consistent and unbiased estimates of the slope parameters β_j and consistent but biased estimate of intercept β_0. In the second step, the intercept is “corrected” by adding a residual with maximal value

$$\hat{\beta}_0^* = \hat{\beta}_0 + \max_i\{\hat{u}_i\}.$$

The production unit with a maximum value of residual $\hat{u}_i$ is from the definition efficient and lies on the production frontier. Residuals of OLS regression are corrected by

$$-\hat{u}_i^* = \hat{u}_i - \max_i\{\hat{u}_i\}.$$

The technical efficiency of the production unit is calculated by $TE_i = \exp\{-\hat{u}_i^*\}$. The main advantage of the method is its simplicity; the disadvantages are ignoring the stochastic term and assuming that the efficient production frontier is parallel to the OLS regression. SFA assumes that a difference between the observed and the feasible output can be partitioned into two components, one representing technical inefficiency and one representing effects of random shocks. If we assume the case of the Cobb-Douglas production function then we can write the stochastic frontier analysis model as

$$\ln y_i = \ln\beta_0 + \sum_{j=1}^{k}\beta_j \ln x_{ij} + v_i - u_i ,$$

where v_i is the two-sided random component with the normal distribution $N(0,\sigma_v^2)$ and u_i is the non-negative technical efficiency component of error term with the non-negative half normal distribution $N(0,\sigma_u^2)$. The error components u_i and v_i are independent of each other, and of the regressors. The composed error term ε_i is defined as $v_i - u_i$. The composed error term is asymmetric since $u_i \geq 0$. The distribution of u_i is usually assumed to be non-negative half normal, truncated normal, exponential or gamma [16]. Such restrictive assumption about the distribution of u_i is a weakness of SFA.

If a set of observations of the same production units over several time periods is available, econometric methods of panel data analysis can be applied. The panel data contain more information than the cross-sectional data because the panel data include a time dimension. The analysis of panel data is a special part of econometric modelling that contains a variety of advanced econometric models [2]. Moreover, the panel data analysis can be combined with methods of frontier analysis.

2.4. Survivor analysis

A survivor analysis is recommended by health economists as a simple alternative to estimating cost functions [12, 13]. The idea of the method is straightforward: those categories that grow relative to the rest of the industry are assumed to have some advantage over the other ones. In the long-run, the distribution of providers should tend toward an optimum, which is, by the analysis, identified as the category with the fastest growth. Categories may be defined by the size of hospital or of group practice (when estimating economies of scale), by the specialty, by the type of ownership, by location, and so forth.

An advantage of classical, univariate survivor analysis is that the method includes both the factor to be investigated and all other factors. The analysis thus includes factors that are hard to measure in econometric studies of cost function. On the other hand, a limitation of the survivor analysis is that it is not able to isolate the effects of those factors [10]. This limitation can be moderated by taking an explicit account of such factors in an expanded, multivariate survivor analysis. The linear version of the multivariate survivor analysis takes the form

$$s_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + u_i,$$

where s_i is the change in market share of group i, and $x_{1i}, x_{2i}, \ldots, x_{ki}$ are the explanatory variables (factors). An alternative is a binary growth model, but in this type of model, the information is lost in converting a continuous variable into a binary one.

The univariate survivor analysis has one advantage over other methods: simplicity. But both the univariate and multivariate survivor analyses, like the other methods of cost analysis, are not able to overcome the fact that the governments, not the market forces play a significant role in the health sector. Changes in the structure of the health care market may rather demonstrate planned governmental interventions than a real economic struggle of providers for their survival [10]. However, the survivor analysis of governmental policy seems to be also an interesting application.

3. Discussion

The economic theory distinguishes four types of efficiency: technical efficiency that is derived from production function, cost efficiency, revenue efficiency, and profit efficiency (Table 1). The type of efficiency is determined by the measurement of inputs and outputs: non-monetary (physical) or monetary units (Table 1). Ratio analysis and econometric modelling may be applied to evaluation of all types of efficiency. Data envelopment analysis was originally developed as a method of technical efficiency evaluation, but it is now used by researchers to measure also other types of efficiency. Survivor analysis was developed as a method of cost efficiency evaluation. From this view, survivor analysis is not a universal method.

Table 1: Methods of Efficiency Analysis and its Typical Application by Type of Efficiency.

Method	Type of Efficiency			
	Technical Efficiency	Cost Efficiency	Revenue Efficiency	Profit Efficiency
Ratio Analysis	++	++	++	++
Econometric modelling	++	++	++	++
DEA	++	+	+	+
Survivor Analysis		++		

The main advantages of ratio analysis are: it relatively is easy, no special mathematical skills are needed; and it is frequently used in practice, so benchmarks are at disposal. The critique of ratio analysis is based on the argument that ratio analysis, unlike DEA, is not found to be suitable for setting targets due to the fact that DEA takes simultaneous account of all inputs and outputs in assessing efficiency while ratio analysis relates only one input to one output at a time. However, the two methods can support each other if used jointly [21]. Sherman [20] argues that ratio analysis is very useful in identifying which aspect of a hospital operation is out of line with norm.

The main strengths of DEA are: the ability to deal with multiple inputs and outputs and the ability to identify the real-world peers (benchmarks) for inefficient production units. The weakness of DEA is a deterministic nature of method because it does not include an error term in the model, so there is a possibility of a measurement error. Hypothesis testing is much less developed in comparison to econometric analysis, although the statistical properties of DEA estimators can be established. The main advantage of econometric modelling over DEA and related methods is a highly developed methodology of hypothesis testing. The selection of explanatory variables and of functional form can be tested and serve as a feedback to a researcher. On the other hand, it is more complicated to deal with multiple inputs and outputs in econometric analysis than in non-parametric methods. It is also necessary to cope with many problems related to the estimation method and the nature of data such as the multicollinearity, heteroscedasticity, etc. An analyst using DEA can avoid dealing with such issues, but then a choice of model and input and output variables is not based on sound arguments [8, 11]. A priori assumptions about returns to scale and convexity in DEA models can be avoided by using the Free Disposable Hull model.

Stochastic frontier analysis is an econometric method that focuses on the best practice as DEA and moreover it is able to incorporate random shocks in efficiency evaluation. In my view, the main disadvantage of the method is that a priori assumptions about the distribution of efficiency have to be made (half-normal, exponential distributions). The use of more general distributional forms has partially alleviated the problem (truncated normal, two-parameter gamma). This weakness can be also solved if panel data are available because no assumptions about the distribution of efficiency have to be made in the panel-data model.

Bryce, Engberg and Wholey [3] empirically tested three methods of efficiency analysis: data envelopment analysis, stochastic frontier analysis, and the fixed-effects panel-data model. For this purpose, they analysed a data set containing 585 health maintenance organizations operating from 1985 through 1994. All three methods identified similar trends for the industry as a whole; however, the methods differed in assessing relative technical efficiency of individual production unit. Thus, these techniques are limited for either benchmarking or setting rates because the production units identified as technically efficient may be a consequence of model selection rather than actual efficiency. Chilingerian and Sherman [5] suggest that methods of frontier analysis may be helpful to understand the behaviour of the entire population of health care providers, and DEA might be used when research focuses more on individual health care providers.

An advantage of univariate survivor analysis is that it includes both the factor to be investigated and all other factors that are hard to measure in econometric studies of cost function. But Koutsoyiannis [15] argues that the survivor analysis suffers from serious limitations because it assumes that the firms pursue the same objectives (but hospitals are both private and public); the firms operate in similar environments so that they do not have locational or other advantages (e.g., public hospitals receive subsidies); prices of factors and technology are not changing (e.g., prices of health technology grow rapidly); the firms operate in a very competitive market structure, that is, there are no barriers to entry or collusive agreements, since under such conditions inefficient (high-cost) firms would probably survive for long periods of time (hospitals are local monopolies, the entry is regulated). Another shortcoming of the survival analysis is that it is not able to explain cases where the size distribution of firms remains constant over time. If the share of the various plant sizes does not change over time, this does not imply that all scales of plant are equally efficient. The survivor analysis indicates only the broad shape of the long-run cost curve, but it does not show the actual magnitude of economies of scale.

Although the application of methods of efficiency analysis in health care is not without its problems (theoretical assumptions, practical interpretation of results), I believe that methods I have discussed are able to help us with identifying best and worst practices. It is no surprise that each method has its advantages and disadvantages. It is important for an analyst being aware of these advantages and disadvantages and combine methods cleverly by taking into account their limits. In my view, from four groups of method, only the usefulness of survivor analysis may be questioned. The method is based on strong economic assumptions about the power of market and is not universal regarding the type of efficiency analysis.

Finally, the efficiency or inefficiency in the health care sector does not depend only on managerial abilities of health care managers that can be evaluated by quantitative methods, but also on regulation environment and payment systems [6] that influence overall efficiency of health system as a whole. Additional qualitative information and observation from more than one period will strengthen the analysis and its findings.

Acknowledgement

The work was supported by project no. 402/09/0231 "Models for evaluation of efficiency and performance of decision-making units and their applications", funded by the Czech Science Foundation.

4. References

[1] AIGNER, D. J., LOVELL, C. A. K., SCHMIDT, P. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 1977, 6:21–37. ISSN 0304-4076

[2] BALTAGI, B. H. *Econometric Analysis of Panel Data*. Chichester, John Wiley & Sons 2001.

[3] BRYCE, C. L., ENGBERG, J. B., WHOLEY, D. G. Comparing the agreement among alternative models in evaluating HMO efficiency. *Health Services Research* 2000, 35:509–528. ISSN 0017-9124

[4] CHARNES, A, COOPER, W. W., RHODES, E. Measuring the inefficiency of decision making units. *European Journal of Operational Research* 1978, 2:429–444. ISSN 0377-2217

[5] CHILINGERIAN, J. A., SHERMAN, H. D. Health Care Applications. In COOPER, W.W., SEIFORD, L.M., ZHU, J. (Eds). *Handbook on Data Envelopment Analysis*. Norwel, Kluwer Academic Publishers 2004.

[6] DLOUHÝ, M. Financování nemocnic: léta devadesátá [Financing hospitals in the nineties], *Politická ekonomie* 1997, 6:862–868. ISSN 0032-3233

[7] DLOUHÝ, M. Matematický model restrukturalizace odvětví a firmy [A mathematical model of sector/firm restructuralization], *Politická ekonomie* 2001, 4:540–546. ISSN 0032-3233

[8] DLOUHÝ, M. Production Functions and Efficiency Evaluation in Health Care. In BRAILSFORD, S., HARPER, P. (Eds): *Operational Research for Health Policy: Making Better Decisions*. Bern, Peter Lang 2007.
[9] DLOUHÝ, M., JABLONSKÝ, J., NOVOSÁDOVÁ I. Využití analýzy obalu dat pro hodnocení efektivnosti českých nemocnic [Use of data envelopment analysis for efficiency evaluation of Czech hospitals]. *Politická ekonomie* 2007, 1:60–71. ISSN 0032-3233
[10] DLOUHÝ, M., FLUSSEROVÁ, L. Three Approaches to the Analysis of Cost Function in Health Care. *Ekonomický časopis* 2007, 55(1):69–78. ISSN 0013-3035
[11] DLOUHÝ, M. Efficiency and Productivity Analysis in Health Services. *Prague Economic Papers* 2009, 18(2): 176–184. ISSN 1210-0455
[12] FELDSTEIN, P. J. *Health Care Economics*. 5th Edition, Delmar Publisher 1998.
[13] FOLLAND, S., GOODMAN, A. C., STANO, M. *The Economics of Health and Health Care*. 3rd Edition. Upper Saddle River: Prentice Hall 2001.
[14] HOLLINGSWORTH, B. Non-parametric and parametric applications measuring efficiency in health care. *Health Care Management Science* 2003, 6:203–218. ISSN 1386-9620
[15] KOUTSOYIANNIS, A. *Modern Microeconomics.* Second Edition, London and Basingstoke: The Macmillan Press 1979.
[16] KUMBHAKAR, S. C., LOVELL, C. A. K. *Stochastic Frontier Analysis*. Cambridge, Cambridge University Press 2000.
[17] MEEUSEN, W., VAN DEN BROECK, J. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* 1977, 18:435–444. ISSN 0020-6598
[18] NOVOSÁDOVÁ I., DLOUHÝ, M. Hodnotenie technickej efektívnosti nemocníc a odmeňovanie zdravotníkov [Evaluation of technical efficiency of acute hospitals and its relation to wages of health personnel]. *Ekonomický časopis* 2007, 55(8): 783–792.
[19] O'NEILL, L., RAUNER, M., HEIDENBERGER, K., KRAUS, M. A cross-national comparison and taxonomy of DEA-based hospital efficiency studies. *Socio-Economic Planning Sciences* 2008, 42(3):158–159. ISSN 0038-0121
[20] SHERMAN, H. D. Hospital Efficiency Measurement and Evaluation: Empirical Test of a New Technique. *Medical Care* 1984, 22(10):922–938. ISSN 0025-7079
[21] THANASSOULIS, E., BOUSSOFIANE, A. DYSON, R. G. A comparison of data envelopment analysis and ratio analysis as tools for performance assessment. *Omega* 1996, 24(3):229–244. ISSN 0305-0483
[22] THANASSOULIS, E. A comparison of Regression Analysis and Data Envelopment Analysis as Alternative Methods for Performance Assessments. *The Journal of Operational Research Society* 1993, 44(11):1129–1144. ISSN 0160-5682

STOCHASTIC GEOMETRIC PROGRAMMING: APPROACHES AND APPLICATIONS

Jitka Dupačová

Charles University in Prague
Faculty of Mathematics and Physics
Department of Probability and Mathematical Statistics
Sokolovská 83, Prague 8, 18675, Czech Republic
dupacova@karlin.mff.cuni.cz

Abstract. During the last years, an increasing interest in geometric programming (GP) can be observed. Advances in numerical methods allow to solve large GPs and new areas of successful applications have emerged: besides of technical applications, there are also GPs for optimal production planning, finance, etc. In real-life applications of GP, some of coefficients and/or exponents need not be precisely known. Stochastic geometric programming can be used to deal with such situations. In this paper, we shall indicate which of general stochastic programming techniques and under which structural and distributional assumptions do not destroy the favorable structure of GPs. Both the already recognized and new approaches will be presented in connection with formulation of the optimization problem. The short note below should serve as an introduction to basic concepts and references.
Keywords. Stochastic geometric programming, applications, statistical sensitivity analysis

1 Geometric programming

Geometric programs introduced by [4] are a special type of nonlinear programming problems in which the objective function and/or some of constraints are *posynomials*:

$$\text{minimize } g_0(\boldsymbol{t}) \text{ subject to } g_k(\boldsymbol{t}) \leq 1,\ k = 1, \dots, K,\ \boldsymbol{t} \in \mathbb{R}^M_{++} \tag{1}$$

with

$$g_k(\boldsymbol{t}) = \sum_{i \in I_k} c_i \prod_{j=1}^{M} t_j^{a_{ij}} = \sum_{i \in I_k} u_i(\boldsymbol{t}),\ k = 0, \dots, K. \tag{2}$$

We denote Q the total number of *monomials* $u_i(\boldsymbol{t}) = c_i \prod_{j=1}^{M} t_j^{a_{ij}}$ in the formulation of geometric program (1), (2) and $\{I_k, k = 0, \dots, K\}$ is a decomposition of $\{1, \dots, Q\}$ into $K + 1$ disjoint index sets. The exponents a_{ij} are arbitrary real numbers and the coefficients c_i are positive. Notice that simple box inequality constraints can be written as inequalities for monomials.
The special structure of geometric program (1)–(2) allows to derive a numerically tractable *dual problem*:

$$\max_{\delta,\lambda} v(\delta, \lambda) := \prod_{i=1}^{Q} (c_i/\delta_i)^{\delta_i} \prod_{k=1}^{K} \lambda_k^{\lambda_k} \tag{3}$$

subject to

$$\sum_{i \in I_0} \delta_i = 1,\ \delta_i \geq 0,\ i = 1, \dots, Q,$$

$$\sum_{i=1}^{Q} a_{ij}\delta_i = 0,\ j = 1, \dots, M,\ \sum_{i \in I_k} \delta_i = \lambda_k,\ k = 1, \dots, K.$$

The optimal solutions $\boldsymbol{t}^*$ of (1) and δ^*, λ^* of (3) are related as follows:

$$\delta_i^* = \frac{u_i(\boldsymbol{t}^*)}{g_0(\boldsymbol{t}^*)} = \frac{u_i(\boldsymbol{t}^*)}{v(\delta^*, \lambda^*)} \text{ for } i \in I_0$$

$$\delta_i^* = \lambda_k^* u_i(\boldsymbol{t}^*) \text{ for } i \in I_k,\ k = 1, \dots, K.$$

Hence $\frac{\delta_i^*}{\lambda_k^*}$, $i \in I_k$ is the proportional contribution of the i-th monomial to the value of posynomial g_k at the optimal solution $\boldsymbol{t}^*$. Numerical solution of small size geometric programs based on solution of their relatively simple duals exploits these duality relations.

The *degree of difficulty* of a geometric program is defined as $\Delta = Q - M^* - 1$ where M^* denotes the rank of the (Q, M) matrix $\boldsymbol{A} = (a_{ij})$. It refers to the dimensionality of the set of feasible solutions of the dual program. For $\Delta = 0$, i.e. for the *zero degree of difficulty geometric programs*, there is a unique solution of the system $\sum_{i \in I_0} \delta_i = 1$, $\sum_{i=1}^{Q} a_{ij}\delta_i = 0$, $j = 1, \dots, M$. If this solution is nonnegative, then it is the optimal solution of the dual problem and it is possible to get an explicit representation of the optimal value function of (3) in terms of coefficients c_i. Moreover, its logarithm is a linear function in coefficients c_i.

Geometric programs (GP) can be reformulated as convex programming problems: Using the exponential substitution $z_j = \log t_j \; \forall j$ the posynomials (2) are transformed to

$$h_k(\boldsymbol{z}) = \sum_{i \in I_k} c_i \exp\{\sum_{j=1}^{M} a_{ij} z_j\}, \; k = 0, \dots, K. \tag{4}$$

An additional log transform of functions h_k is frequently recommended.
The resulting transformed GP is then the *convex* program

$$\text{minimize } h_0(\boldsymbol{z}) \text{ subject to } h_k(\boldsymbol{z}) \leq 1, \; k = 1, \dots, K, \; \boldsymbol{z} \in \mathbb{R}^M. \tag{5}$$

See e.g. [2, 4, 16, 17] for these and related results.
The early applications of geometric programming were connected mainly with mechanical engineering but they include also economic and managerial problems, cf. [16], chemical equilibrium and nonlinear network flow problems. In these areas, more sophisticated applications have been further developed and extended to inventory control, production system optimization, computational finance etc. The presently prevailing field of applications seems to be in digital circuit design.
The recently observed growing interest in GP stems from the fact that various practical problems can be reformulated as geometric programs and there are solution methods which solve even very large-scale GPs efficiently and reliably. With a basic interior-point method which exploits sparsity of the generic geometric program (1)–(2) the reported efficiency is close to that of linear programming solvers. We refer to [3] for an up-to-date survey of various applications and an extensive list of references and to [18] for an interesting reformulation of an entropy optimization problem emanating from computational finance to a dual of a tractable GP.

2 Stochastic geometric programming

In applications, some of coefficients c_i and/or exponents a_{ij} need not be known precisely and their incomplete knowledge may be modeled as random. As in general stochastic programming problems one deals with the distribution problem or focuses on decision problems. The question is which of stochastic programming approaches and under which distributional assumptions do not destroy the favorable structure of the (generalized) geometric programs.

The origins of *stochastic* geometric programming (SGP) are connected with paper [1], where the exponents a_{ij} are deterministic and the coefficients c_j are positive random variables. The main result of the paper are numerically tractable bounds for the optimal value of (1); see also [13, 20] for their further elaboration and application.

Construction of confidence bounds for the optimal value of a geometric program, deriving its moments or probability distribution is a task belonging under *distribution problem* of stochastic geometric programming. It was developed first for zero degree of difficulty geometric programs in connection with lognormal distribution of coefficients c_i and fixed exponents a_{ij}. Then the logarithm of the optimal value function in (3) is an affine linear function in $\log c_i$, hence, for a lognormal distribution of c_i, one gets lognormal distribution of the optimal value. For extensions of these results to other probability distributions and to problems with degree of difficulty $\Delta > 0$ see e.g. [8, 19].

Individual probabilistic constraints have been applied under assumption of deterministic exponents and normally distributed, mostly uncorrelated coefficients c_i; see e.g. [11, 17]. It means that the constraints of (1)

$$\sum_{i \in I_k} c_i \prod_{j=1}^{M} t_j^{a_{ij}} \leq 1, \; k = 1, \dots, K$$

are replaced by

$$P\{\sum_{i\in I_k} c_i \prod_{j=1}^{M} t_j^{a_{ij}} \le 1\} \ge 1-\varepsilon_k,\ k=1,\dots,K,$$

with prescribed tolerances ε_k. For independent normally distributed coefficients $c_i \sim N(Ec_i,\sigma_i^2)\,\forall i$ these constraints are equivalent to

$$\sum_{i\in I_k} Ec_i \prod_{j=1}^{M} t_j^{a_{ij}} + \Phi^{-1}(1-\varepsilon_k)\sqrt{\sum_{i\in I_k}\sigma_i^2 \prod_{j=1}^{M} t_j^{2a_{ij}}} \le 1,\ k=1,\dots,K,$$

where $\Phi^{-1}(1-\varepsilon_k)$ is quantile of the standard normal distribution $N(0,1)$. Each constraint is then split into two constraints that involve posynomials in $t_j \forall j$ and a common additional slack variable.
Of course, the assumption of normally distributed costs c_i is not in agreement with the required positivity of coefficients in (1). For general probability distributions of coefficients c_i [10] suggests to approximate the probabilistic constraints by one-sided Chebyshev inequality. A similar approximation is used also in [14] for optimization of stochastic activity networks with random durations characterized by mean values and standard deviations of the posynomial form.

In various engineering and economic applications of GP random character of exponents can be observed as well. Consider for example production functions of the Cobb-Douglas type used to describe requirements or to formulate the objective function. In the simplest situation, the constraint on production is

$$Ct_1^{a_1}t_2^{a_2} \ge k \tag{6}$$

where t_1, t_2 are inputs. The common assumption that the coefficient C and exponents a_1, a_2 are given constants is not quite realistic. Hence, one gets interested in sensitivity of results on small changes of these "constants"; the classical sensitivity analysis, cf. [15, 16] is the first step. It is not enough, however, when the coefficients *and* exponents of posynomials are random, being e.g. differentiable functions of statistical estimates of true parameter values. In comparison with randomness present only in the coefficients c_i, a substantially higher level of difficulty arises. In general, one can design simulation experiments to get an idea about the probability distribution of the optimal value, to evaluate approximate confidence bounds and moments of the optimal value, etc. However, such experiments are computationally expensive and do not provide sufficient information about the optimal solutions or their logarithms. In the sequel, we shall review some other techniques.

A possibility which applies to SGP with random parameter, say β, only in the objective function and to a discrete distribution of these parameters is to use a *tracking model* related with the goal programming; cf. [6].

For random costs and exponents in the objective function and in constraints of (1), a *penalization or two-stage approach* was suggested in [12]. First of all, using an additional constraint and an additional variable t_0, geometric program (1) can be rewritten to have a nonrandom linear objective function:

$$\min\{t_0 \,:\, t_0^{-1}g_0(\boldsymbol{t},\beta) \le 1,\ g_k(\boldsymbol{t},\beta) \le 1,\ k=1,\dots,K,\ t_0>0,\ \boldsymbol{t}\in \mathbb{R}^M_{++}\}. \tag{7}$$

The constraints of (7) can be further split to

$$u_i(\boldsymbol{t},\beta)\theta_{ik}^{-1} \le 1,\ i=1,\dots,Q,\ k=0,\dots,K \tag{8}$$

with $\theta_{ik}>0$, $\sum_{i\in I_k}\theta_{ik}=1$ interpreted as the proportional contribution of i-th monomial to the value of k-th posynomial.
The first stage decisions are $t_0>0$, $\boldsymbol{t}\in\mathbb{R}^M_{++}$ and $\theta_{ik}>0$, $i\in I_k\,\forall k$, and $\sum_{i\in I_k}\theta_{ik}=1\forall k$ are the first stage constraints. After observing realizations of random coefficients and exponents, possible violation of constraints (8) can be corrected for an additional cost. Logarithmic penalty function is suggested and the case of multivariate discrete or normal distribution of parameters c_i, a_{ij} is discussed.

For GP of the form (7) one may consider a *robust reformulation*

$$\min\{t_0 \,:\, t_0^{-1}g_0(\boldsymbol{t},\boldsymbol{u}) \le 1,\ g_k(\boldsymbol{t},\boldsymbol{u}) \le 1,\ k=1,\dots,K,\ t_0>0,\ \boldsymbol{t}\in\mathcal{T},\ \boldsymbol{u}\in\mathcal{U}\}$$

where $\mathcal{U}$ denotes a prespecified uncertainty set. Various possibilities how to approach such semiinfinite problems are discussed e.g. in [9].

In [6, 7] a technique for construction of *confidence bounds* for optimal value and optimal solution of SGP has been proposed. It is based on sensitivity results for deterministic geometric programming [15] and on stochastic sensitivity analysis [5]. The motivation comes from metal cutting problems where the tool life affects substantially the total machining costs. Due to nonhomogeneity of the machined and cutting material variability of the tool life occurs even at fixed machining conditions. It can be influenced by a careful choice of cutting conditions in accordance with the postulated technical relation: The tool life is a monomial in cutting speed, feed and depth whose parameters can be obtained as statistical estimates of the true values. Derivatives of the minimal total cost with respect to the parameters, regression analysis and Delta theorem lead to an approximate confidence interval for the minimal machining costs, the tool life, etc.
This technique can be evidently applied to decision problems involving estimated production, demand or utility functions of the posynomial form such as the Cobb-Douglas production function in (6); see [16] for instances of deterministic versions of such problems. Among others, the lower and upper bounds on the system's cost obtained in this way are an important information for the purpose of economic decision making.

Acknowledgement. The research was partly supported by the project "Methods of modern mathematics and their applications" – MSM 002620839 and by the Grant Agency of the Czech Republic (grant 402/08/0107).

References

1. Avriel, M., Wilde, D.J.: Stochastic geometric programming. In: H. W. Kuhn (ed.), Proc. of the Princeton Symp. of Math. Progr., Princeton Univ. Press (1970)
2. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming (Theory and Algorithms), 2nd edition, Wiley, New York (1993)
3. Boyd, S. et al.: A tutorial on geometric programming. Optim Eng 8, 67–127 (2007)
4. Duffin, R.J., Petterson, E.L., Zener, C.: Geometric Programming, Wiley, New York (1967)
5. Dupačová, J.: Stability in stochastic programming with recourse - Estimated parameters. Math. Progr. 28, 72–83 (1984)
6. Dupačová, J.: Stochastic geometric programming with an application. Submitted (2009)
7. Dupačová, J., Charamza, P., Mádl, J.: On stochastic aspects of a metal cutting problem. In: Kall, P., Marti K.(eds.), Stochastic Programming: Numerical Methods and Engineering Applications, LNEMS 423, Springer, Berlin, pp.196–209 (1995)
8. Ellner, P.M., Stark, R.M.: On the distribution of the optimal value for a class of stochastic geometric programs. Naval Res. Log. Quart. 27, 549–571 (1980)
9. Hsiung, K-L., Kim, S-J., Boyd, S.: Tractable approximate robust geometric programming. Optim Eng 9, 95–118 (2008)
10. Hsiung, K-L., Kim, S-J., Boyd, S.: Power allocation with outage probability specifications in wireless shadowed fading channels via geometric programming. RR, Information Systems Laboratory, Stanford University (2008)
11. Iwata, K. et al.: A probabilistic approach to the determination of the optimal cutting conditions. J. of Engineering for Industry Trans. of ASME 94, 1099–1107 (1972)
12. Jagannathan, R.: A stochastic geometric programming problem with multiplicative recourse. Oper. Res. Letters 9, 99–104 (1990)
13. Jha, N.K.: Probabilistic cost estimation in advance of production in a computerized manufacturing system through stochastic geometric programming. Computers ind. Engng 30, 809–821 (1996)
14. Kim, S-J. et al.: A heuristic for optimization stochastic activity networks with applications to statistical digital circuit sizing. Optim Eng 8, 397–430 (2007)
15. Kyparisis, J.: Sensitivity analysis in geometric programming: Theory and Computations. Ann Oper Res 27, 39–64 (1990)
16. Luptáčik, M.: Geometric programming.Method and applications. OR Spektrum 2, 129–143 (1981)
17. Rao, S.S.: Engineering optimization: Theory and practice, 3rd edition. Wiley-Interscience, New York (1996)
18. Rajasekera, J.R., Yamada, M.: Estimating the firm value distribution function by entropy optimization and geometric programming. Ann Oper Res 105, 61–75 (2001)
19. Stark, R.M.: On zero-degree stochastic geometric programs. JOTA 23, 167–187 (1977)
20. Wiebking, R.D.: Optimal engineering design under uncertainty by geometric programming. Manag. Sci. 6, 644–651 (1977)

MODELING AND SOLVING OF DOUBLE AUCTIONS

Petr Fiala, Lenka Flusserová

University of Economics, Prague
Faculty of Informatics and Statistics, Department of Econometrics
pfiala@vse.cz, lflusser@vse.cz

Abstract. Auctions are important market mechanisms for the allocation of goods and services. An auction provides a mechanism for negotiation between buyers and sellers. In forward auctions a single seller sells resources to multiple buyers. In a reverse auctions, a single buyer attempts to source resources from multiple suppliers, as is common in procurement. Auctions with multiple buyers and sellers are called double auctions. Auctions with multiple buyers and sellers are becoming increasing popular in electronic commerce. It is well known that double auctions in which both sides submit demand or supply bids are much more efficient than several one-sided auctions combined. Attention is devoted to double combinatorial auctions. Combinatorial auctions are those auctions in which bidders can place bids on combinations of items. The paper presents models and solutions for double combinatorial auctions. Combinatorial double auctions can be transformed to combinatorial single-sided auctions and solved by methods for these auctions.
Keywords. Combinatorial auctions, forward and reverse auctions, double auctions, iterative approach

1. Introduction

Auctions are important market mechanisms for the allocation of goods and services. Auction theory has caught tremendous interest from both the economic side as well as the Internet industry. Auctions have emerged as a particularly interesting tool for negotiations. An auction provides a mechanism for negotiation between buyers and sellers. We consider so called combinatorial auctions where bidders can place bids on combinations of items.

There is possible to formulate single-sided combinatorial auctions, forward auctions and reverse auctions. In forward auctions a single seller sells resources to multiple buyers. In reverse auctions, a single buyer attempts to source resources from multiple suppliers, as is common in procurement. Auctions with multiple buyers and sellers, so called double auctions, are becoming increasing popular in electronic commerce. It is well known that double auctions in which both sides submit demand or supply bids are much more efficient than several one-sided auctions combined. The aim of the paper is to propose models and solving approaches for double auctions

The paper presents a model for double auctions. Special cases of the double auction model are single-sided auctions. The formulated combinatorial double auction can be transformed to a combinatorial single-sided auction. The model can be solved by methods for single-sided combinatorial auctions. We propose to use an iterative approach for combinatorial auctions based on primal-dual algorithm.

2. Combinatorial auctions

Combinatorial auctions (see [1], [2]) are those auctions in which bidders can place bids on combinations of items, so called bundles. The advantage of combinatorial auctions is that the bidder can more fully express his preferences. This is particular important when items are complements. The auction designer also derives value from combinatorial auctions. Allowing bidders more fully to express preferences often leads to improved economic efficiency and greater auction revenues.

Most of the auctions studied in the literature are one-sided: either multiple buyers compete for commodities sold by one seller, or, multiple sellers compete for the right to sell to one buyer. The problem, called the winner determination problem, has received considerable attention in the literature. The problem is formulated as: Given a set of bids in a combinatorial auction, find an allocation of items to bidders that maximizes the seller's revenue. It was introduced many important ideas, such as the mathematical programming formulation of the winner determination problem, the connection between the winner determination problem and the set packing problem as well as the issue of complexity.

Forward auction

Many types of combinatorial auctions can be formulated as mathematical programming problems. From different types of combinatorial auctions we present a forward auction of indivisible items with one seller and multiple buyers. Let us suppose that one seller S offers a set R of r items, $j = 1, 2, \ldots, r$, to n potential buyers B_1, B_2, ..., B_n (see Fig. 1).

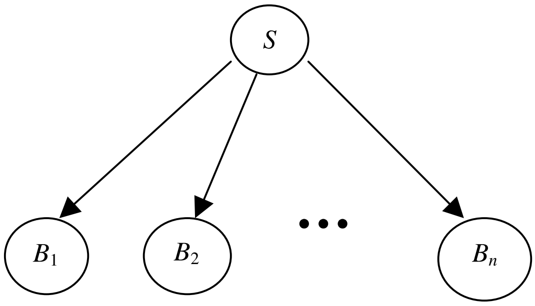


Fig. 1. Forward auction

Items are available in single units. A bid made by buyer B_i, $i = 1, 2, \ldots, n$, is defined as

$$b_i = \{C, p_i(C)\},$$

where

$C \subseteq R$, is a combination of items,

$p_i(C)$, is the offered price by buyer B_i for the combination of items C.

The objective is to maximize the revenue of the seller given the bids made by buyers. Constraints establish that no single item is allocated to more than one buyer.

Bivalent variables are introduced for model formulation:

$x_i(C)$ is a bivalent variable specifying if the combination C is assigned to buyer B_i ($x_i(C) = 1$).

The forward auction can be formulated as follows

$$\sum_{i=1}^{n} \sum_{C \subseteq R} p_i(C)\, x_i(C) \rightarrow \quad \max$$

subject to

$$\sum_{i=1}^{n} \sum_{C \subseteq R} x_i(C) \leq 1, \quad \forall j \in R, \tag{1}$$

$$x_i(C) \in \{0, 1\}, \ \forall\ C \subseteq R, \ \forall i, \ i = 1, 2, \ldots, n.$$

The objective function expresses the revenue. The constraints ensure that overlapping sets of items are never assigned.

Reverse auction

We present a reverse auction of indivisible items with one buyer and several sellers. This type of auction is important for supplier selection problem. Let us suppose that m potential sellers S_1, S_2, ..., S_m offer a set R of r items, $j = 1, 2, \ldots, r$, to one buyer B (see Fig. 2).

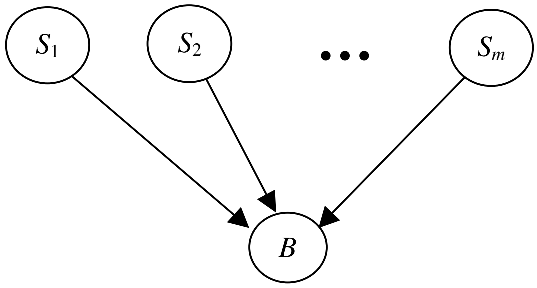


Fig. 2: Reverse auction

A bid made by seller S_h, $h = 1, 2, \ldots, m$, is defined as

$$b_h = \{C, c_h(C)\},$$

where

$C \subseteq R$, is a combination of items,

$c_h(C)$, is the offered price by seller S_h for the combination of items C.

The objective is to minimize the cost of the buyer given the bids made by sellers. Constraints establish that the procurement provides at least set of all items.

Bivalent variables are introduced for model formulation:

$y_h(C)$ is a bivalent variable specifying if the combination C is bought from seller S_h ($y_h(C) = 1$).

The reverse auction can be formulated as follows

$$\sum_{h=1}^{m} \sum_{C \subseteq R} c_h(C)\, y_h(C) \rightarrow \quad \min$$

subject to

$$\sum_{h=1}^{m} \sum_{C \subseteq R} y_h(C) \geq 1, \ \forall j \in R, \qquad (2)$$

$$y_h(C) \in \{0, 1\}, \ \forall \ C \subseteq R, \ \forall h, \ h = 1, 2, \ldots, m.$$

The objective function expresses the cost. The constraints ensure that the procurement provides at least set of all items.

3. Double auctions

Auctions with multiple buyers and multiple sellers are becoming increasing popular in electronic commerce. The numerous applications in electronic commerce, including stock exchanges, business-to-business commerce, bandwidth allocation, etc. have led to a great deal of interest in double auctions. Double auctions are not so often studied in the literature as single-sided auctions (see [6]).

For double auctions, the auctioneer is faced with the task of matching up a subset of the buyers with a subset of the sellers. The profit of the auctioneer is the difference between the prices paid by the buyers and the prices paid to the sellers. The objective is to maximize the profit of the auctioneer given the bids made by sellers and buyers. Constraints establish the same conditions as in single-sided auctions.

We present a double auction problem of indivisible items with multiple sellers and multiple buyers. Let us suppose that m potential sellers $S_1, S_2, \ldots, S_m$ offer a set R of r items, $j = 1, 2, \ldots, r$, to n potential buyers $B_1, B_2, \ldots, B_n$ (see Fig. 3).

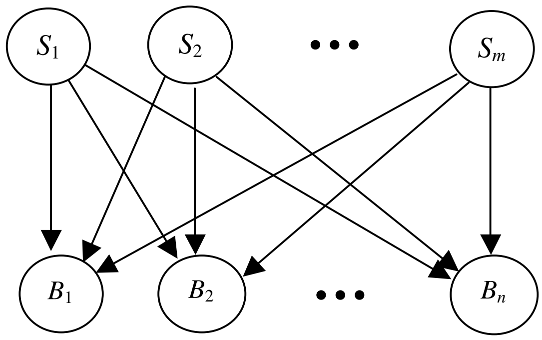


Fig. 3: Double auction

A bid made by seller S_h, $h = 1, 2, \ldots, m$, is defined as $b_h = \{C, c_h(C)\}$,

a bid made by buyer B_i, $i = 1, 2, \ldots, n$, is defined as $b_i = \{C, p_i(C)\}$,

where

$C \subseteq R$, is a combination of items,

$c_h(C)$, is the offered price by seller S_h for the combination of items C,

$p_i(C)$, is the offered price by buyer B_i for the combination of items C.

Bivalent variables are introduced for model formulation:

$x_i(C)$ is a bivalent variable specifying if the combination C is assigned to buyer B_i ($x_i(C) = 1$),

$y_h(C)$ is a bivalent variable specifying if the combination C is bought from seller S_h ($y_h(C) = 1$).

$$\sum_{i=1}^{n} \sum_{C \subseteq R} p_i(C)\, x_i(C) - \sum_{h=1}^{m} \sum_{C \subseteq R} c_h(C)\, y_h(C) \rightarrow \max$$

subject to

$$\sum_{i=1}^{n} \sum_{C \subseteq R} x_i(C) \leq \sum_{h=1}^{m} \sum_{C \subseteq R} y_h(C), \quad \forall j \in R, \tag{3}$$

$$x_i(C) \in \{0, 1\}, \ \forall\ C \subseteq R, \ \forall\, i,\ i = 1, 2, \ldots, n,$$

$$y_h(C) \in \{0, 1\}, \ \forall\ C \subseteq R, \ \forall\, h,\ h = 1, 2, \ldots, m.$$

The objective function expresses the profit of the auctioneer. The constraints ensures for buyers to purchase a required item and that the item must be offered by sellers.

Special case of double auction for one seller is the forward auction and special case of double auction for one buyer is the reverse auction.

4. Solving of double auctions

The formulated combinatorial double auction can be transformed to a combinatorial single-sided auction. Substituting $y_h(C)$, $h = 1, 2, \ldots, m$, with $1 - x_i(C)$, $i = n+1, n+2, \ldots, n+m$, and substituting $c_h(C)$, $h = 1, 2, \ldots, m$, with $p_i(C)$, $i = n+1, n+2, \ldots, n+m$, we get a model of a combinatorial single-sided auction.

$$\sum_{i=1}^{n+m} \sum_{C \subseteq R} p_i(C)\, x_i(C) - \sum_{i=n+1}^{n+m} \sum_{C \subseteq R} p_i(C) \rightarrow \max$$

subject to

$$\sum_{i=1}^{n+m} \sum_{C \subseteq R} x_i(C) \leq m, \quad \forall j \in R, \tag{4}$$

$$x_i(C) \in \{0, 1\}, \ \forall\ C \subseteq R, \ \forall\, i,\ i = 1, 2, \ldots, n+m.$$

The model (4) can be solved by methods for single-sided combinatorial auctions. Complexity is a fundamental question in combinatorial auction design. The algorithms proposed for solving combinatorial auctions are exact algorithms and approximate ones. Many researchers consider iterative auctions as an alternative.

One way of reducing some of the computational burden in solving the problem is to set up a fictitious market that will determine an allocation and prices in a decentralized way. In the iterative approach, there are multiple rounds of bidding and allocation and the problem is solved in an iterative and incremental way. Iterative combinatorial auctions are attractive to bidders because they learn about their rivals' valuations through the bidding process, which could help them to adjust their own bids (see [4]). The key challenge in the iterative combinatorial auctions design is to provide information feedback to the bidders after each iteration (see [5]).

We propose to use an iterative approach for combinatorial auctions (see [3]). The primal-dual approach is used for solving. For the problem (4) we will formulate the LP relaxation and its dual. The scheme can be outlined as follows:

1. Choose initial prices for sellers and buyers.
2. Announce current prices and collect bids.
3. Compute the current dual solution by interpreting the prices as dual variables. Try to find a feasible allocation, an integer primal solution that satisfies the stopping rule. If such solution is found, stop and use it as the final allocation. Otherwise update prices and go back to 2.

5. Conclusions

The paper is devoted to modeling and solving double auctions. The numerous applications in electronic commerce have led to a great deal of interest in double auctions. Modeling and solving of combinatorial double auctions is a promising subject for research and for practical exploitations. The paper presents a model of double auctions. The proposed combinatorial double auction gives us an opportunity to design modifications of the auction. The formulated model can be transformed to a combinatorial single-sided auction and solved by methods for single-sided combinatorial auctions. We propose to use an iterative approach to solving combinatorial double auctions. The primal-dual algorithm can be taken as a decentralized and dynamic method to determine equilibrium.

Acknowledgements

The research project was supported by Grant No. 402/07/0166 from the Grant Agency of Czech Republic „Combinatorial Auctions - Modeling and Analysis“.

References

[1] Cramton, P., Shoham, Y. and Steinberg, R. (eds.) (2006). *Combinatorial Auctions.* MIT Press, Cambridge.

[2] de Vries, S., Vohra, R. V. (2003). Combinatorial auctions: A survey. *INFORMS Journal of Computing*, 15 (1): 284-309.

[3] Flusserová, L., Fiala, P. (2008). Iterative Combinatorial Auctions. In: *Mathematical Methods in Economics.* Liberec : Faculty of Economics, Technical University, 115–124. ISBN 978-80-7372-387-3.

[4] Parkes, D. C. (2001): *Iterative Combinatorial Auctions: Achieving Economic and Computational Efficiency*. PhD thesis, University of Pennsylvania.

[5] Pikovsky, A, Bichler, M. (2005): Information Feedback in Iterative Combinatorial Auctions. In *Conference proceedings Wirtschaftsinformatik*, Bamberg, Springer.

[6] Xia, M., Stallaert,J., Whinston, A. B. (2005). Solving the combinatorial double auction problem. *European Journal of Operational Research*, 164, 239–251.

MODIFIED DEA MODELS AND THEIR USE IN OPTIMIZATION OF PRODUCTION STRUCTURE IN AGRICULTURAL ENTERPRISES IN THE CZECH REPUBLIC

Ludvík Friebel, Jana Friebelová

University of South Bohemia
Faculty of Economics, Department of Applied Mathematics and Informatics,
friebel@ef.jcu.cz

Abstract. The contribution is focused on adjustments of Data Envelopment Analysis with the aim to improve the production structure of agricultural enterprises on the basis of their production capabilities and current market conditions. Two specific methods for evaluation of enterprises and subsequent recommendations are discussed: Weighted Slack-Based Measures (WSBM) and Generic Directional Distance Model (GDDM). A common feature of these methods is the possibility to include weights and/or directional vectors in the recommended changes. In addition, both models can be modified to allow for negative outputs that are often seen in practical situations. Another goal of our work is to propose a specific procedure for weight determination in each particular model; this is relevant as the same weight vector might lead to different recommendations - and efficiency scores - in each model. In the applications part, we examine farming enterprises located in the Czech potato-growing region - the largest part of arable land in the Czech Republic. An attractive feature of this choice is the homogeneity of the grain production structure throughout this region.
Keywords. Efficiency, data envelopment analysis, performance, slack-based measure of efficiency, generic directional distance model, standardization

1. Introduction

Recently many authors have devoted their work to the application of Data Envelopment Analysis (DEA) in agricultural context. See, for instance, Ahmad and Bravo-Ureta [1], Wilson, Hadley and Asby [2], Iráizoz, Rapún and Zabaleta [3], Lansink and Reinhard [4], among others.

Data envelopment analysis (DEA) is instrumental to the technical efficiency evaluation of decision making units (DMUs) based on the size of inputs and outputs. DEA is used to measure the relative efficiency of comparable units. In particular, while these units use the same kind inputs to produce the same outputs, there may be differences in their performance. The number of units must not be too small, because many units seem to be efficient with a small amount of units in the group and a big criteria number.

One item of primary concern is the choice of evaluation criteria. Attention must be paid to the fact that the criteria must not be correlated with performance units. DEA models give a set of recommendations in order to improve the efficiency of non-efficient units, by either increasing outputs or decreasing outputs. However, in agricultural context, some recommendations given by basic DEA models may be unrealistic for the following reasons: smaller production units are preferred, but the classical DEA models do not take into account the unique character of particular units. Therefore other (modified) DEA models are proposed for applications in agriculture.

As groundwork, Weighted Slack-Based Measures of efficiency (WSBM) proposed by Tone [5] and Generic Directional Distance Model (GDDM) proposed by Chambers [6] and [7] are used for evaluation of enterprises. An advantage of these modified DEA methods is the possibility of affecting recommended changes by weights or directional vectors. Further, WSBM and GDDM can be improved by the inclusion of negative data [8], [10].

The production conditions and utilization of agricultural land in the Czech Republic are characterized by farming areas and sub-areas depending on the climate and soil conditions. There are five production areas (maize growing region, sugar beet growing region, grain growing region, potato growing region and forage growing region), and each of them is further split into sub areas (21 total).

This article is focused on the potato growing region because that area constitutes the largest part of arable land in the Czech Republic. The area is characterized by typical altitudes of 400-650 m above the sea level, its total cultivated land is 60-80% and the average slope is 5-12%. Another reason for this choice was the homogeneity from the point of view of grain production structure.

2. Methodology

The above mentioned Weighted Slack-Based Measure of efficiency (WSBM) and Generic directional distance model (GDDM) seemed very suitable.

2.1. Weighted Slack-Based Measure of Efficiency (WSBM) model

The WSBM can be applied to any situation with a priori defined weights of inputs and outputs. This model was derived from the SBM model, see Tone [9] and was further improved see Tone and Tsutsui [10] in order to permit negative data in the dataset.

In case VRS the WSBM model considering weights of inputs and outputs is introduced as follows:

$$\min\left\{\frac{1-\frac{1}{m}\sum_{i=1}^{m} w_{i0}s_i^-/x_{i0}}{1+\frac{1}{s}\sum_{k=1}^{s} r_{k0}s_k^+/y_{k0}}\middle|\ \mathbf{x_0}=\mathbf{X\lambda}+\mathbf{s}^-,\ \mathbf{y_0}=\mathbf{Y\lambda}-\mathbf{s}^+,\ \boldsymbol{\lambda}\geq\mathbf{o},\mathbf{s}^-\geq\mathbf{o},\mathbf{s}^+\geq\mathbf{o},\ \sum_{j=1}^{n}\lambda_j=1\right\}, \quad (1)$$

where w_{i0}, r_{k0} are weight for input i and output k of DMU_0. For a detailed description see Tone (2001).

2.2. Generic Directional Distance Model (GDDM)

The other possibility of influencing changes in the size of inputs and outputs is represented with the Generic Directional Distance Model (GDDM).

This model was proposed by Chambers [6] and [7]. The main advantage of this model is the ability to project inefficient units on the efficiency frontier with a selected direction. This feature is particularly attractive applicable in our situation, because the model provides each unit with its own way to efficiency, respecting its improvement potential. The GDDM is as follows:
For the set of unit $j=1,2,...,n$, with input levels x_{ij}, $i=1,2,...,m$ and output levels $y_{kj}, k=1,2,...,s$ and unit $o\in j$ which is to be assessed is the generic directional distance model given as follows:

$$\max\left\{\begin{array}{l}\beta_o \mid \sum_{j=1}^{n}\lambda_j y_{kj}\geq y_{ko}+\beta_o g_{y_k}, k=1,2,...,s, \sum_{j=1}^{n}\lambda_j x_{ij}\leq x_{io}-\beta_o g_{x_i}, i=1,2,...,m, \sum_{j=1}^{n}\lambda_j=1,\\ \lambda_j,\beta_o,g_{x_i},g_{y_k}\geq 0\end{array}\right\}, \quad (2)$$

where the vector $\mathbf{g}_x$ $(\mathbf{g}_y)$ represents possible changes of input (output).

The model (2) is valid in the case of variable returns to scale (VRS). Target values of inputs (outputs) were obtained as the product of **X (Y)** and **λ**. If negative data occurs among observed inputs (outputs), the last constraint of model (2) would be violated.

GDDM was further modified by Portela et al. [8] in order to handle negative data for the Range Directional Model (RDM+) and Inverse Range Directional Model (RDM-). These models substitute the vector of desired changes with the vector of inputs and output ranges in the case of RDM+ and the vector of the inverted value of the ranges in the case of RDM-. Both types of RDM use different ways to compute efficiency scores similarly to radial efficiencies traditionally used in DEA (see Portela et al. [8]).

Efficiency scores for the standardized model can be calculated as follows.

$$\varphi_o = 1 - \frac{\sqrt{\sum_{i=1}^{m}\left(x_{io}^{*} - x_{io}\right)^2 + \sum_{k=1}^{s}\left(y_{ko}^{*} - y_{ko}\right)^2}}{\sqrt{\sum_{i=1}^{m} R_{io}^{\ 2} + \sum_{k=1}^{s} R_{ko}^{\ 2}}}, \tag{3}$$

where φ_o is efficiency for assessed unit o,

x_{io}^{*} is the target value of i-th input projected on the efficiency frontier,

y_{ko}^{*} is the target value of k-th output projected on the efficiency frontier,

$R_{ko} = \max_j \{y_{kj}\} - y_{ko}, k = 1,2,..., s$,

$R_{io} = x_{io} - \min_j \{x_{ij}\}, i = 1,2,..., m.$

The main advantage of this model is the ability to project inefficient units on the efficiency frontier with the selected direction and to obtain a most realistic recommendation. This feature is well applicable in our application, because the model provides each unit with its own way to efficiency that respects its improvement potential.

3. Data and application

We considered a group of 28 farms of similar characteristics (in the potato growing region). As inputs, we took the ratio of land area designated for growing wheat, barley and the rape. As outputs, we took yield of the afore-mentioned crops in tons per hectar (see Table 1).

Symbol	Description
x_1	ratio of soil for growing wheat
x_2	ratio of soil for barley
x_3	ratio of soil for rape
y_1	yield per hectare of growing wheat [t]
y_2	yield per hectare barley [t]
y_3	yield per hectare rape [t]

Table 1: Description of inputs and outputs

The first step to successful use of the above-mentioned models is to determine the direction of the potential improvement described by vectors of weights for the WSBM and the directional vector for the GDDM.

For the outputs represented by per-hectare yields, we adjust the weights according to the price level of crop plants produced by particular farmer. Values of per hectare profits are proposed for the input weights. But weights for inputs are established to the contrary to weights for outputs. Consequently for the outputs the biggest emphasis is placed on the crop with the highest price. For the inputs the biggest emphasis is placed on the reduction of the crop with the lowest per hectare profit. In the special case of negative profit of two particular crops and one positive profit, the vector of potential improvement was set to (0.5; 0.5; 0). This situation occurred for unit 3 and many other units, see Table 2. Likewise this adjustment ensures binding of inputs and outputs which are typical in agriculture.

Input data are presented in Table 2.

	Inputs			Outputs				Inputs			Outputs		
j	x_1	x_2	x_3	y_1	y_2	y_3	j	x_1	x_2	x_3	y_1	y_2	y_3
1	0.264	0.106	0.191	6.30	5.30	3.50	15	0.307	0.156	0.086	7.56	6.29	4.21
2	0.235	0.237	0.150	5.29	4.11	2.49	16	0.341	0.377	0.000	5.87	4.49	0.00
3	0.461	0.259	0.278	5.39	3.82	3.04	17	0.338	0.307	0.131	5.71	4.13	3.04
4	0.300	0.216	0.138	3.89	3.81	2.76	18	0.315	0.241	0.082	5.25	4.65	3.49
5	0.288	0.036	0.099	6.43	4.82	2.80	19	0.445	0.000	0.131	4.94	0.00	4.54
6	0.257	0.225	0.067	7.07	6.74	3.76	20	0.254	0.045	0.179	4.40	2.98	2.36
7	0.303	0.356	0.058	5.81	4.72	3.24	21	0.370	0.075	0.063	5.52	5.53	3.46
8	0.323	0.141	0.039	7.80	6.71	3.60	22	0.318	0.091	0.140	6.53	5.27	3.70
9	0.330	0.128	0.111	6.39	5.15	3.26	23	0.288	0.216	0.072	6.33	5.04	3.82
10	0.392	0.171	0.163	5.54	3.82	3.82	24	0.379	0.098	0.183	5.79	5.63	3.64
11	0.353	0.231	0.076	5.33	4.65	2.89	25	0.343	0.077	0.113	5.33	4.55	3.20
12	0.270	0.239	0.126	6.92	5.52	3.43	26	0.208	0.160	0.143	5.52	4.79	3.65
13	0.326	0.020	0.150	5.65	5.77	3.44	27	0.231	0.246	0.155	5.55	4.28	3.64
14	0.264	0.106	0.191	6.30	5.30	3.50	28	0.144	0.122	0.058	3.91	5.00	3.71

Table 2: Inputs and outputs

Determination of weights for evaluation models

	Inputs			Outputs				Inputs			Outputs		
j	x_1	x_2	x_3	y_1	y_2	y_3	j	x_1	x_2	x_3	y_1	y_2	y_3
1	0.367	0.372	0.260	0.26	0.22	0.53	15	0.448	0.052	0.500	0.26	0.21	0.53
2	0.478	0.457	0.065	0.23	0.24	0.53	16	0.000	0.500	0.500	0.22	0.32	0.47
3	0.500	0.500	0.000	0.21	0.21	0.58	17	0.407	0.358	0.235	0.52	0.48	0.00
4	0.500	0.500	0.000	0.21	0.25	0.54	18	0.248	0.500	0.252	0.24	0.25	0.51
5	0.500	0.500	0.000	0.24	0.20	0.56	19	0.500	0.500	0.000	0.22	0.24	0.54
6	0.365	0.500	0.135	0.26	0.25	0.49	20	0.428	0.275	0.297	0.25	0.00	0.75
7	0.337	0.467	0.195	0.24	0.27	0.48	21	0.471	0.500	0.029	0.21	0.29	0.50
8	0.442	0.417	0.141	0.24	0.25	0.51	22	0.286	0.379	0.335	0.22	0.20	0.58
9	0.490	0.500	0.010	0.23	0.24	0.53	23	0.332	0.500	0.168	0.26	0.23	0.51
10	0.500	0.500	0.000	0.21	0.22	0.57	24	0.120	0.500	0.380	0.25	0.24	0.51
11	0.014	0.500	0.486	0.25	0.22	0.53	25	0.382	0.463	0.155	0.22	0.25	0.54
12	0.500	0.500	0.000	0.23	0.25	0.52	26	0.500	0.000	0.500	0.24	0.20	0.56
13	0.194	0.306	0.500	0.26	0.22	0.53	27	0.500	0.500	0.000	0.22	0.26	0.52
14	0.367	0.372	0.260	0.23	0.24	0.53	28	0.324	0.453	0.223	0.21	0.21	0.58

Table 3: Weight (directional) vectors

4. Results

Results for WSBM and GDDM are depicted in Tables 4 and 5. There is clear that the particular model gives different recommendation.

	Inputs			Outputs				Inputs			Outputs		
j	x_1	x_2	x_3	y_1	y_2	y_3	j	x_1	x_2	x_3	y_1	y_2	y_3
1	0.261	0.103	0.088	6.30	5.65	3.50	16	0.210	0.183	0.057	5.77	5.61	3.21
2	0.193	0.169	0.050	5.30	5.05	2.82	17	0.276	0.134	0.042	6.76	6.20	3.57
3	0.261	0.060	0.108	5.39	5.38	3.44	19	0.223	0.026	0.091	4.41	3.98	2.38
4	0.165	0.081	0.066	3.89	4.36	3.03	21	0.300	0.067	0.114	6.56	5.92	3.72
7	0.268	0.126	0.038	6.54	5.87	3.30	22	0.252	0.150	0.053	6.37	6.32	3.90
9	0.275	0.072	0.095	6.39	5.43	3.37	23	0.359	0.033	0.135	5.89	5.63	3.67
10	0.289	0.068	0.127	5.54	5.96	4.03	24	0.313	0.041	0.101	5.34	4.56	3.25
11	0.324	0.185	0.031	7.38	6.24	2.89	25	0.206	0.160	0.065	5.52	5.74	3.65
12	0.252	0.220	0.066	6.92	6.60	3.68	26	0.203	0.176	0.062	5.55	5.88	3.70
14	0.304	0.156	0.083	7.56	6.99	4.21	28	0.149	0.129	0.046	4.07	4.35	2.77

Table 4: Results for WSBM model for inefficient units

	Inputs			Outputs				Inputs			Outputs		
j	x_1	x_2	x_3	y_1	y_2	y_3	j	x_1	x_2	x_3	y_1	y_2	y_3
1	0.264	0.094	0.096	6.30	5.93	3.92	16	0.210	0.178	0.085	5.71	7.30	5.42
2	0.195	0.165	0.078	5.29	6.76	5.02	17	0.313	0.137	0.038	7.56	6.50	3.49
3	0.311	0.019	0.143	5.39	5.50	3.28	19	0.254	0.016	0.117	4.40	4.49	2.68
4	0.143	0.121	0.058	3.89	4.97	3.69	21	0.318	0.067	0.112	7.28	6.00	3.70
7	0.227	0.144	0.058	5.81	6.23	4.12	22	0.243	0.168	0.072	6.33	7.16	4.92
9	0.297	0.047	0.111	6.39	5.39	3.26	23	0.379	0.020	0.164	6.19	5.63	3.98
10	0.355	0.026	0.163	6.20	6.37	3.82	24	0.323	0.031	0.113	6.42	4.55	3.20
11	0.259	0.113	0.031	6.26	5.39	2.89	25	0.208	0.160	0.073	5.52	6.67	4.78
12	0.255	0.216	0.103	6.92	8.85	6.57	26	0.204	0.173	0.082	5.55	7.10	5.27
14	0.306	0.156	0.052	7.56	7.12	4.21	28	0.149	0.126	0.060	4.05	5.18	3.84

Table 5: Results for GDDM model for inefficient units

A comparison for two specific DMUs can be found in Table 6 and 7. Somewhat surprisingly neither model respects the given improvement vector in the obtained recommendations. This situation can be explained the following way: only eight units are efficient and many units are not and thus a significant part of units is not covered by efficiency envelopment. For those units the requested reduction of inputs and extension of outputs can not be fulfilled.

The efficiency scores for the both compared units are shown in tables 6 and 7. For the GDDM was computed score by the formula (3). The ranking of the units according their efficiency is more describing than the efficiency score. Rank of DMUs is not very different for compared units 6 and 7 as well as for the rest of units.

The changes proposed by GDDM for DMU 7 are obviously more considerable than the changes proposed by WSBM. In general we can say that the recommendation of the WSBM is more moderate than the recommendation of the GDDM, the average percentage of the change is 20.6 % for WSBM (31.2% – input changes, 10.0 % output changes) and 25.8 % for GDDM (29.2% – input changes, 22.3 % output changes). For these reasons, WSBM can be recommended if more moderate or balanced changes are needed.

	x_1	x_2	x_3	y_1	y_2	y_3	*eficiency*	*rank*
Weights / vector components	0.337	0.467	0.195	0.24	0.27	0.48	-	-
WSBM	12%	65%	35%	12%	24%	2%	0.523	21
GDDM	25%	60%	1%	0%	32%	27%	0.404	21

Table 6: Comparison relative differences between models for unit 7

	x1	x2	x3	y1	y2	y3	*eficiency*	*rank*
Weights / vector components	0.500	0.000	0.500	0.24	0.20	0.56	-	-
WSBM	12%	28%	60%	0%	37%	2%	0.569	17
GDDM	12%	29%	47%	0%	66%	45%	0.464	19

Table 7: Comparison relative differences between models for unit 26

5. Conclusion

In this paper we tested two additive DEA models with the aim to find realistic recommendations for agricultural enterprises where classical DEA models fail to respect the particular improvement which takes into account existing results of farming and market conditions. As this projection on the efficiency frontier is generally non-radial, we propose a new method to compute an efficient score too.

Acknowledgments

The following paper was supported by a grant project of FRVŠ 1820/2009.

References

[1] Ahmad, M., Bravo-Ureta, B. *Technical efficiency measures for dairy farms using panel data: a comparison of alternative model specifications*. The Journal of Productivity Analysis 7, p. 399 -415, 1996

[2] Wilson, P., Hadley, D., Asby, C. *The influence of management characteristics on the technical efficiency of wheat farmers in Eastern England*. Agricultural Economics 24, p. 329-338. 2001

[3] Iráizoz, B., Rapún, M., Zabaleta, I. Assessing the technical efficiency of horticultural production in Navarra, Spain. Agricultural Systems 78, p. 387-403. 2003

[4] Lansink, A. O., Reinhard, S. Investigating technical efficiency and potential technological change in Dutch *pig farming*. Agricultural Systems, 79, p. 353-367. 2004

[5] Tone, K. A slacks-based measure of efficiency in data envelopment analysis. European Journal of Operational Research 130, p. 498-509. 2001

[6] Chambers, R. G., Chung, Y., Färe, R. Benefit and functions. Journal of Economic Theory 70, p. p. 407 -419. 1996

[7] Chambers, R. G., Chung, Y., Färe, R. Profit, directional distance functions and Nerlovian efficiency. Journal of Optimization Theory and Applications 98(2), p. 351 -364, 1998

[8] Portela, M.C.A.S., Thanassoulis, E. and Simpson, G.P.M. Negative data in DEA: A directional distance approach applied to bank branches. Journal of the Operational Research Society, 55, p. 1111-1121. 2004

[9] Tone, K. (2002). A slack-based measure of super-efficiency in data envelopment analysis. European Journal of Operation Research 143, 32-41.

[10] Tone, K., Tsutsui M.: Network DEA: A slacks-based measure approach. Discussion Paper: 07-08. GRIPS Policy Information Center. 2008

APPLICATION OF PARAMETRIC BENCHMARKING METHOD – STOCHASTIC FRONTIER ANALYSIS IN COST EFFICIENCY ESTIMATION OF THE BANKING SECTOR

Andrea Furková

Department of Operations Research and Econometrics
Faculty of Economic Informatics, University of Economics in Bratislava
furkova@euba.sk

Abstract. This paper deals with the application of parametric benchmarking method – Stochastic Frontier Analysis (SFA) in measuring cost efficiency of Slovak banks. The Slovak banking environment has significantly transformed as a result of various economic changes in recent years. Due to these extensive changes banking sector becomes very competitive branch whereby increase the significance of its performance evaluation and its financial conditions monitoring. The efficiency measuring and relative efficiency comparison of banks are crucial questions for analysts as well as for economic policy creators. Having applied Stochastic Frontier Analysis to an unbalanced panel of Slovak banks we focus on cost efficiency estimation over period 2000 - 2007. In our analysis have been applied two versions of stochastic frontier cost function models. The differences in estimated scores, parameters and ranking of banks are compared across different panel data models.
Keywords. Stochastic Frontier Analysis, Cost Efficiency, Panel Data Models, Banking sector

1. Introduction

Since the beginning of the 1990s, environment of banking sectors in transition countries of Central and Eastern Europe has been significantly influenced by many different regulatory, technological and economic changes. Also in Slovak banking sector has been realized large reorganization for purpose of creation of modern banking system as one assumption of shipshape economy. Due to these extensive changes, the banking industry has become very competitive industry and increases its performance evaluation and its financial conditions monitoring. Over the last years large attention was dedicated to banking sector efficiency measuring especially of transforming countries and its comparison with old EU member states banking sectors. The efficiency measuring and relative efficiency comparison of banks are crucial questions for analysts as well as for economic policy creators.

2. Methodology for analyzing banking efficiency

In banking efficiency analysis may be used financial ratios or parametric and nonparametric frontier techniques. Historically, the financial ratios have been the standard technique of banking efficiency measuring and examining of their performance. These traditional ratio measures are attractive because of simplicity of understanding, but they fail to consider the multidimensional input and output process, and are unable to identify the best performers in a group of units. These limitations have led to the development of more complex and adequate tools for the units performance. These tools are based on benchmarking, that is, measuring a unit's efficiency compared with a reference performance (so-called efficient frontier). The frontier analyses are able to identify the best performers and separate them from their inefficient counterparts. Inefficiency can results from technological deficiencies or non-optimal allocation of resources into production. Both technical and allocative inefficiencies are included in cost inefficiency, which is by definition, the deviation from minimum costs to produce a given level of output with given input prices. Some of benchmarking methods can also identify the sources of inefficiency. The efficient frontier (cost or production) is unknown and must be empirically estimated from the real data set by parametric and nonparametric techniques.

Generally, there are two families of methods based on efficient frontier:

- Non-parametric methods, like Data Envelopment Analysis (DEA) or Free Disposal Hull (FDH). These methods originate from operations research and use linear programming to calculate an efficient deterministic frontier against which units are compared. Detail presentation of DEA and FDH models we can find in [6] and [5].

- Parametric methods, like Stochastic Frontier Analysis (SFA), Thick Frontier approach (TFA) and Distribution Free Approach (DFA). Econometric theory is used to estimate pre-specified functional form and inefficiency is modeled as an additional stochastic term. Detail presentation of these models we can find in [13] and [5].

There are many studies dealing with bank efficiency estimation based on efficient frontier however most of them use nonparametric methods especially models of DEA. Because DEA nor other nonparametric approach methods do not enable the separation of the inefficiency effect from the statistical noise we decided to exploit only parametric method, namely SFA and only panel data models will be used due to the fact that they provide information on the same units over several periods that is not possible with cross section data.

3. Stochastic Frontier Analysis – parametric benchmarking technique

A frontier cost function defines minimum costs with given output level, input prices and the existing production technology. It is unlikely that all units will operate at the frontier. Failure to attain the cost frontier implies the existence of technical and allocative inefficiency. Different approaches can be used to estimate a frontier cost function with panel data. The main goal of this paper is to compare different models of SFA with respect to the estimated cost function parameters, estimated efficiency scores and ranking of observed units. In SFA models is used econometric theory to estimate pre-specified functional form and inefficiency is modeled as an additional stochastic term.

The stochastic frontier cost function (single Cobb-Douglas form) for panel data can be formulated as:

$$\ln C_{it} = \beta_0 + \beta_y \ln y_{it} + \sum_n \beta_n \ln w_{nit} + v_{it} + u_i \qquad i = 1,2,\dots,N \text{ and } t = 1,2,\dots,T \qquad (1)$$

where

C_{it} are observed total costs of the *i*-th unit in year *t*,

y_{it} is a vector of outputs of the *i*-th unit' in year *t*,

w_{it} is an input price vector of the *i*-th unit in year *t*,

u_i are non negative time-invariant random variables assumed to be half normal distributed (u_i~iid $N^+(0, \sigma_u^2)$),

v_{it} are random variables which are assumed to be normally distributed (v_{it} ~ iid $N(0, \sigma_v^2)$).

In this specification the error term is composed of two uncorrelated parts. The first part u_i is capturing the effect of inefficiency (including both allocative and technical inefficiencies) and the second part v_{it} is reflecting effect of statistical noise. This random effect model can be estimated using Maximum Likelihood Estimation (MLE) method or Method of Moments. The next step is to obtain estimates of the cost efficiency of each unit. The problem is to extract the information that ε_i contains on u_i (we have estimates of $\varepsilon_i = u_i + v_i$, which obviously contain information on u_i). A solution to the problem is obtained from the conditional distribution of u_i given ε_i, which contains whatever information ε_i contains concerning u_i. This procedure is known as JLMS decomposition (for more details see [11]). For separation the inefficiency effect from the statistical noise can be also used an alternative minimum squared error predictor estimator (for more details see [13]). Once the point estimates of u_i are obtained, estimates of the cost efficiency of each unit can be obtained by substituting them into equation (2). If the cost frontier is specified as being stochastic, the appropriate measure of cost efficiency becomes:

$$CE_i = \frac{c(\mathbf{y}_{it}, \mathbf{w}_{it}, \beta)\exp\{v_{it}\}}{C_{it}} = \exp\{-u_i\} \qquad i = 1,\dots,N \quad t = 1,\dots,T \qquad (2)$$

which defines cost efficiency as the ratio of minimum cost attainable in an environment characterized by $\exp\{v_{it}\}$ to observed expenditures. The main advantage of the stochastic cost frontier approach is the separation of the inefficiency effect from the statistical noise.

Another random effect model was proposed by Schmidt and Sickles [13] to overcome the problem of specifying a particular distribution for the inefficiency by rewriting equation (1):

$$\ln C_{it} = \beta_0 + \beta_y \ln y_{it} + \sum_n \beta_n \ln w_{nit} + v_{it} + \alpha_i \qquad i = 1,2,\dots,N \text{ and } t = 1,2,\dots,T \qquad (3)$$

where $\alpha_i = \alpha + u_i$ and α is an intercept. Now, the Generalized Least Squares (GLS) method can be used. The estimate of inefficiency component is defined as the distance from the firm specific intercept to the minimal intercept in the sample:

$$u_i = \alpha_i - \min_i(\alpha_i) \qquad (4)$$

Subsequently the estimates of the cost efficiency of each unit can be obtained by substituting (4) into equation (2).

There are no required assumptions about distribution for inefficiency in this model. The remaining restrictive assumption is that two random components must not be correlated with each of the explanatory variables.

Till now we maintained the assumption that cost efficiency is constant through the time. This assumption is a strong one, particularly if the operating environment is competitive, it is hard to accept the notation that the cost efficiency remains constant through many time periods. The longer the panel, the more desirable it is to relax this assumption.

If we allow efficiency changes in time, inefficiency component will consist of two parts, namely cross-section component (u_i) and time component (β_t):

$$u_{it} = u_i + \beta_t \qquad i = 1,\dots,N \quad t = 1,\dots,T \qquad (5)$$

The time-invariant cost efficiency model given by equation (1) we reformulate as follows:

$$\ln C_{it} = \beta_{0t} + \beta_y \ln y_{it} + \sum_n \beta_n \ln w_{nit} + v_{it} + u_{it} \qquad i = 1,\dots,N \quad t = 1,\dots,T \qquad (6)$$

where β_{0t} is the cost frontier intercept common to all units in period t and all other variables are as previously defined. There are proposed various approaches to estimated time – varying cost frontier model given by equation (6).

Battese and Coelli [5], [13] presented a model where they model the inefficiency component in (6) according to:

$$u_{it} = \delta \mathbf{z}_{it} + w_{it} \qquad (7)$$

where

v_{it} ~ iid $N(0, \sigma_v^2)$,

z_{it} is a vector of variables which may influence the efficiency of unit,

u_{it} are non-negative random variables which are assumed to account for cost inefficiency and are assumed to be independently distributed as truncations at zero of the $N^+(\delta \mathbf{z}_{it}, \sigma_u^2)$ distribution,

δ is a vector of unknowns parameters to be estimated,

w_{it} is random variable defined by the truncation of the normal distribution with zero mean and constant variance w_{it} ~ $N(0, \sigma_w^2)$, since $u_{it} \geq 0$, the truncation point is $-\delta \mathbf{z}_{it}$.

In this model the observable environmental variables are to allow directly influence the stochastic component of the cost function. The likelihood function of this model is a generalization of the likelihood function for the conventional model (for more details see [5]).

4. Model specification and Data

The above-mentioned models have been applied to an unbalanced panel of 9 (Tatra banka, a.s., Všeobecná úverová banka, a.s., Slovenská sporiteľňa, a.s., Dexia banka Slovensko, a.s., OTP banka Slovensko, a.s., Istrobanka, a.s., Poštová banka, a.s., Unibanka, a.s., Ľudová banka, a.s) Slovak banks over a period from 2000 to 2007. In order to create homogeneous data set in term of production technologies, we consider only with commercial banks, our data set does not contain specialized banks. The sample includes 66 observations. All data are based on information from the annual reports of the banks.

The first part of analysis was based on the assumption of time invariant cost efficiency. We applied SFA panel data model with time invariant cost efficiency assumption (REM – GLS) and the analysis was based on the estimation of the model given by equation (8).

$$\ln\left(\frac{C}{P_F}\right)_{it} = \beta_0 + \beta_L \ln\left(\frac{P_L}{P_F}\right)_{it} + \beta_K \ln\left(\frac{P_K}{P_F}\right)_{it} + \beta_U \ln U_{it} + \beta_V \ln V_{it} + \beta_Z \ln Z_{it} + \beta_T T + v_{it} + u_i$$

$$i = 1,\ldots,N \quad t = 1,\ldots,T \tag{8}$$

In order to estimate the model with time varying cost efficiency (REM – ML) was formulated model given by equation (9):

$$\ln\left(\frac{C}{P_F}\right)_{it} = \beta_0 + \beta_L \ln\left(\frac{P_L}{P_F}\right)_{it} + \beta_K \ln\left(\frac{P_K}{P_F}\right)_{it} + \beta_U \ln U_{it} + \beta_V \ln V_{it} + \beta_T T + v_{it} + u_{it}$$

$$i = 1,\ldots,N \quad t = 1,\ldots,T \tag{9}$$

where C represents total costs, P_K, P_L, P_F are the prices of capital, labor and funds respectively. Three inputs (physical capital, labor, funds) are used to produce two outputs total loans U and total deposits V and both models include additional variables Z – equity capital and T – trend. v_{it} is reflecting effect of statistical noise and u_i or u_{it} are random variables described in equations (4) and (7). The dependent variable, total costs (C) is the sum of total expenses, price of capital (P_K) was constructed as the depreciation over fixed assets, price of labor (P_L) was defined as the ratio of personnel expenses over total assets, price of funds (P_F) was defined as the ratio of interest expenses over the sum of deposits. Equity capital (Z) is the amount of bank equity that reflects the size of banking operations. The condition of linear homogeneity in input prices was imposed by dividing total costs and input prices by the price of funds. The traditional Cobb-Douglas cost function has been considered in both models.

In the case of REM GLS model we used Generalized Least Squares method for estimation equation given by (8) and there are no required assumptions about distribution for inefficiency component in this model. In the case of REM MLE model, the cost function in (9) has been estimated by Maximum Likelihood Estimation method and u_{it} are non-negative random variables which are assumed to account for cost inefficiency and are assumed to be independently distributed as truncations at zero of the $N^+(\delta Z_{it}, \sigma_u^2)$ distribution. Having excluded variable Z from the deterministic core of the cost frontier and its including to the inefficiency component ($u_{it} = \delta_0 + \delta_1 Z_{it} + w_{it}$) we tried to model heterogeneity in production environment of banks. In both models for separation the inefficiency effect from the statistical noise was used Battese and Coelli point estimator [3]. The final estimates of the parameters both cost frontier models are listed in table 1.

	REM GLS		**REM ML**	
	Estimate	**Standard Error**	**Estimate**	**Standard Error**
constant	3.0432**	0.2603	1.8369**	0.4098
lnP_L/P_F	0.0710*	0.0318	0.3528**	0.0545
lnP_K/P_F	0.4034**	0.0389	0.0994**	0.0409
ln$U$	0.0567	0.0377	0.0577	0.0465
ln$V$	0.6809**	0.0343	0.9178**	0.0425
ln$Z$	0.2047**	0.0404		
T	0.0088	0.0080	0.0180*	0.0098
δ_0			8.5792	8.3307
δ_1			0.8556	0.8105
σ^2			0.5772	0.4768
γ			0.9883	0.0115
R^2	0.9859			
logLF			37.7088	

Table 1: *Parameters of the Cost Function*
Source: Self calculations (Eviews5 and Frontier 4.1)
* significant at $\alpha = 0{,}05$
** significant at $\alpha = 0{,}1$

Table 2 and 3 provide efficiency ranking for the banks and efficiency scores according to REM GLS and REM ML models.

Bank	REM GLS
Istrobanka, a.s.	1(1.0000)
Unibanka, a.s.	2(0.9865)
OTP banka, a.s.	3(0.9674)
Slovenská sporiteľňa, a.s.	4(0.9086)
Ľudová banka, a.s.	5(0.9062)
Všeob. úverová banka, a.s.	6(0.8986)
Dexia banka, a.s.	7(0.8895)
Poštová banka, a.s.	8(0.8553)
Tatra banka, a.s.	9(0.8497)
Average	0.9180

Table 2: Efficiency ranking for the banks and efficiency scores – REM GLS model
Source: Self calculations

Banka	2000	2001	2002	2003	2004	2005	2006	2007
Tatra banka, a.s.	2(0.9336)	7(0.7242)	8(0.7488)	8(0.7468)	9(0.7144)	5(0.9101)	4(0.9319)	3(0.9330)
Všeob. úverová banka, a.s.	5(0.9232)	4(0.9180)	6(0.8964)	7(0.9008)	6(0.8919)	2(0.9484)	7(0.8782)	7(0.9115)
Slovenská sporiteľňa, a.s.	1(0.9599)	1(0.9443)	3(0.9460)	3(0.9215)	3(0.9168)	1(0.9705)	5(0.9297)	5(0.9298)
Dexia banka, a.s.	6(0.9181)	6(0.9135)	2(0.9586)	1(0.9640)	4(0.9135)	6(0.8913)	3(0.9378)	2(0.9398)
OTP banka, a.s.		3(0.9216)	4(0.9395)	4(0.9157)	5(0.9023)	7(0.8194)	6(0.8929)	4(0.9329)
Istrobanka, a.s.	4(0.9301)	8(0.4601)	1(0.9652)	2(0.9612)	1(0.9334)	4(0.9340)	2(0.9446)	1(0.9544)
Poštová banka, a.s.	3(0.9306)	5(0.9155)	5(0.9280)	6(0.9079)	7(0.8433)	9(0.6347)	9(0.6236)	8(0.7468)
Unibanka, a.s.					2(0.9300)	3(0.9378)	1(0.9815)	6(0.9246)
Ľudová banka, a.s.	7(0.8822)	2(0.9328)	7(0.8924)	5(0.9145)	8(0.8369)	8(0.7812)	8(0.8754)	
Average	0.9254	0.8412	0.9094	0.9040	0.8758	0.8697	0.8884	0.9091

Table 3: Efficiency ranking for the banks and efficiency scores – REM ML model
Source: Self calculations

5. Conclusion

Table 1 shows that almost all the coefficients are significant and the coefficients are mildly different from one model to another. In addition in model REM ML are listed estimations of δ_0 and δ_1 by reason of different formulation of u_{it}, in this model we made assumption of time varying inefficiency and variable *Z* was excluded from the deterministic core of the cost frontier and included to the inefficiency component in order to model heterogeneity in production environment of banks. Non signification of parameter δ_1 implies that the equity capital is not variable which might capture heterogeneity in production environment of banks and account for inefficiency.

Tables 2 and 3 provide efficiency estimates and ordering of banks (the estimated efficiency scores are given in brackets). The scores can move between 0 and 1, where the highest value implies a perfectly efficient bank and the difference from 1 approximates the percentage of the total costs that the bank can potentially save. As the results suggest, the studied banks are on average about 90 % efficient according both models. The individual efficiency estimates are not stable across models by reason of different assumption about inefficiency component. In consequence of relaxation time invariant efficiency assumption we obtained efficiency scores for each observed year in REM ML model. In spite of this advantage and for reason of non signification of δ_1 we prefer REM GLS model in our data set.

References:

[1] Annual reports of banks 2000 - 2007

[2] APFELBECK, J., CULLMANN, A., VON HIRSCHHAUSEN, CH.: *Efficiency Analysis of East European Electricity Distribution Utilities* - Legacy of the Past? Working Paper, 2005, Department of Management and Economics, Dresden University of Technology.

[3] BATTESE, G. E., COELLI, T. J: *Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data.* In: Journal of Econometrics, 1988, č. 38, s. 387 - 399. ISSN 0304-4076

[4] COELLI, T. J: A Guide to Frontier Version 4.1: *A Computer Program for Stochastic Frontier Production and Cost Function Estimation*. CEPA Working Paper, 1996, Centre for Efficiency and Productivity Analysis, University of New England, Australia, ISSN 1832-4398

[5] COELLI, T. J., RAO PRASADA, D., BATTESE, G.: *An Introduction to Efficiency and Productivity Analysis*. Springer, 2005, 349 s. ISBN O387242661

[6] COOPER, W. W., SEIFORD, L. M., TONE, K.: *Data Envelopment Analysis*, Kluwer Academic Publishers, 2000

[7] FARSI, M., FILIPPINI, M., GREENE, W.: *Application of Panel Data Models in Benchmarking Analysis of the Electricity Distribution Sector*. CEPE Working Paper Nr.39, 2005, Centre for Energy Policy and Economics, Swiss Federal Institute of Technology.

[8] FARSI, M., FILIPPINY, M.: *Regulation and Measuring Cost Efficiency with Panel Data Models: Application to Electricity Distribution Utilities*. CEPE Working Paper Nr.19, 2003, Center for Energy Policy and Economics, Swiss Federal Institute of Technology.

[9] FILIPPINI, M., HROVATIN, N., ZORIČ, J.: *Regulation of the Slovenian Electricity Distribution Companies. In: Energy Policy*, 2004, č.32, s.335 - 344. ISSN 0301-4215

[10] FURKOVÁ, A.: *Analýza nákladovej efektívnosti slovenských a českých distribučných podnikov elektrickej energie*, dizertačná práca, 2007, 180 s. Bratislava

[11] JONDROW, J., LOVELL, C. A. K., MATEROV, I. S, SCHMIDT, P.: *On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model*. In: Journal of Econometrics, 1982, č. 19 (2-3), s.233 - 238. ISSN 0304-4076

[12] KUENZLE, M.: *Cost Efficiency in Network Industries: Application of Stochastic Frontier Analysis*. 2005, Dissertation Thesis, Swiss Federal Institute of Technology Zurich, s.133.

[13] KUMBHAKAR, S. C., LOVELL, C. A. K.: *Stochastic Frontier Analysis.* Cambridge University Press, 2000, 333 s. ISBN 05216666635

[14] KOŠAK, M., ZAJC, P.: *Bank consolidation and bank efficiency in Europe*, Working Paper, miha.ef.uni-lj.si/_dokumenti/wp/Kosak__Zajc_Bank_Consolidation_Efficiency_WP.doc

[15] ŽABKA, M.: *DATA ENVELOPMENT ANALYSIS a jej využitie pri vyhodnocovaní efektívnosti bankového sektora*, 2007, 143 s., Bratislava

EIGENPROBLEM FOR CIRCULANT MATRICES IN EXTREMAL ALGEBRAS

Martin Gavalec and Hana Tomášková

University of Hradec Králové, Faculty of Informatics and Management,
martin.gavalec@uhk.cz, hana.tomaskova@uhk.cz

Abstract. The eigenproblem in extremal algebras, such as max-plus or max-min algebra, is investigated with restriction to circulant matrices. For the max-plus case, an algorithm working in linear time is proposed which computes the dimension of the eigenspace of a given circulant max-plus matrix. For max-min circulant matrices, the structure of the eigenspace is described, in contrast to the fact that the description of the eigenspace structure is not known for general max-min matrices.
Keywords: max-min algebra, max-plus algebra, eigenproblem, circulant matrices, eigenvector

1 Introduction

Models in many applications use operations of maximum and minimum, together with further arithmetical operations. Such algebras are often referred to as extremal. Max-plus algebra is useful for investigation of discrete events systems and the sequence of states in discrete time corresponds to powers of matrices in max-plus algebra. The eigenvectors of max-plus matrices describe the steady states of the system. The eigenproblem in max-min algebra is important for applications connected with system reliability, or with fuzzy relations. For these reasons, the structure of the eigenspace (the set of all eigenvectors) of a given extremal matrix has been intensively studied by many authors, see e.g. [1–4, 8, 9]. The theoretical results were found and polynomial algorithms were proposed for computing the eigenvalue and eigenvectors of given max-plus or max-min matrices.

For special types of matrices, the computation can often be performed in a simpler way than in the general case, hence the investigation of special cases is important from the computational point of view. In this contribution the eigenproblem for circulant max-plus and max-min matrices is investigated. A square matrix is called circulant, if the input values in every row of the matrix are the same as the values in the previous row, but they are cyclically shifted by one position to the right. In applications, some types of computer nets can be represented by circulant matrix, e.g. tokenring is a simple circulant matrix.

Concerning the max-plus case, an algorithm for computation of the eigenspace dimension is proposed in this paper. The algorithm computes the eigenspace dimension of a given circulant max-plus matrix in linear time, in contrast to the computational complexity in the general case, which is $O(n^3)$. For max-min circulant matrices, the structure of the eigenspace is completely described. Such description of the eigenspace structure is not known for general max-min matrices.

2 Notions and notation

By a max-plus algebra we understand a triple $(\mathcal{R}, \oplus, \otimes)$, where $\mathcal{R}$ is the set of all real numbers and $\oplus, \otimes$ are binary operations on $\mathcal{R}$ defined as

$$a \oplus b = \max(a, b), \quad a \otimes b = a + b$$

for all $a, b \in \mathcal{R}$. Similarly, a max-min algebra is a triple $(\mathcal{R}, \oplus, \otimes)$, where $\mathcal{R}$ is the set of all real numbers and $\oplus, \otimes$ are binary operations on $\mathcal{R}$ defined as

$$a \oplus b = \max(a, b), \quad a \otimes b = \min(a, b)$$

for all $a, b \in \mathcal{R}$. In max-min algebra, the additive structure of real numbers is neglected, and only the linear ordering on $\mathcal{R}$ is used.

The operations $\oplus, \otimes$ are extended to matrices and vectors in the same way as in conventional linear algebra. Hence, in both types of extremal algebras, the matrix product $A \otimes B$ is defined for matrices $A \in \mathcal{R}^{m\times p}, B \in \mathcal{R}^{p\times n}$ as a matrix $C \in \mathcal{R}^{m\times n}$ by formula

$$c_{ij} = \bigoplus_{k=1}^{p} (a_{ik} \otimes b_{kj})$$

for $i = 1, \dots m$, $j = 1, \dots n$. The kth power of a square matrix $A \in \mathcal{R}^{n\times n}$ is denoted by $A^{(k)}$ and defined by recursion on $k = 2, 3, \dots$

$$A^{(k)} = A \otimes A^{(k-1)} \ .$$

In this paper we consider a special form of matrices, called *circulant* matrices. Square matrix A of type $n \times n$ is circulant if

$$a_{ij} = a_{i'j'}$$

whenever

$$i - i' \equiv j - j' \pmod{n} \ .$$

Hence, circulant matrix A is fully determined by its inputs in the first row, denoted as $a_0, a_1, \dots, a_{n-1}$ (a_0 is the common value of all diagonal inputs, and similarly each a_i is the common value of all inputs on a line parallel to the matrix diagonal)

$$\begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \dots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \dots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \dots & a_0 \end{pmatrix} \ .$$

3 Eigenproblem in max-plus algebra

The eigenproblem in max-plus algebra is formulated as follows. For given $A \in \mathcal{R}^{n\times n}$, find $\lambda \in \mathcal{R}$ and $x \in \mathcal{R}^n$ satisfying

$$A \otimes x = \lambda \otimes x \ .$$

The value λ and the vector x fulfilling the above equation are called the *eigenvalue* and the *eigenvector* of matrix A. The set of all eigenvectors is called the *eigenspace* of A. It has been shown in [1] that the eigenvalue of a given max-plus matrix can be efficiently described by considering cycles in specific digraphs.

The *associated digraph* D_A of a matrix $A \in \mathcal{R}^{n\times n}$ is defined as a complete arc-weighted digraph with the node set $V = N = \{1, 2, \dots, n\}$, and with the arc weights $w(i, j) = a_{ij}$ for every $(i, j) \in N \times N$. If p is a path or a cycle in D_A, of length $r = |p|$, then the weight $w(p)$ is defined as the sum of all weights of the arcs in p. If $r > 0$, then the mean weight of p is defined as $w(p)/r$. The maximal mean weight of a cycle in D_A is denoted by $\lambda(A)$.

Lemma 1. *[1] If $A \in \mathcal{R}^{n\times n}$, then $\lambda(A)$ is the unique eigenvalue of A.*

Eigenvectors of a given max-plus matrix can be found using the following procedure: for any $B \in \mathcal{R}^{n\times n}$ we denote by $\Delta(B)$ the matrix $B \oplus B^{(2)} \oplus \dots \oplus B^{(n)}$. Further, we denote $A_\lambda = -\lambda(A) \otimes A$ (in the formal product of a scalar value $-\lambda(A)$ and a matrix A we have $[A_\lambda]_{ij} = -\lambda(A) + a_{ij}$ for any $(i, j) \in N \times N$). It is shown in [1] that matrix $\Delta(A_\lambda)$ contains at least one column, the diagonal element of which is 0. Every such a column is an eigenvector (so called: *fundamental eigenvector*) of the matrix A. Moreover, every eigenvector of A can be expressed as a linear combination of fundamental eigenvectors.

Let us denote by $g_1, g_2, \dots, g_n$ all columns of $\Delta(A_\lambda)$. We shall say that vectors g_j, g_k are equivalent, if there is $\alpha \in G$ such that $g_j = \alpha \otimes g_k$. It has been shown in [1] that vectors g_j, g_k are equivalent if and only if the vertices j, k are contained in a common cycle c with $w(c) = \lambda(A)$ in D_A. The *eigenspace dimension* of matrix A is defined as the maximal number of non-equivalent fundamental eigenvectors. In general case, the eigenspace dimension can be found by computing matrix $\Delta(A_\lambda)$ in $O(n^3)$ time. In this section we show that for circulant matrices the eigenspace dimension can be computed in $O(n)$ time.

The eigenproblem for circulant max-plus matrices was studied in [4] and following lemmas were presented.

Lemma 2. *[4] If A and B are circulant max-plus matrices of the same type $n \times n$, then $A \otimes B$ is also a circulant matrix of type $n \times n$.*

Lemma 3. *[4] For a circulant matrix A the eigenvalue $\lambda(A)$ is equal to the maximal value of the elements in the first row of A.*

Theorem 1. *The eigenspace dimension of a given circulant matrix A is equal to the greatest common divisor of all positions of the maximal value $\lambda(A)$ in the first row of matrix A and the size n of A.*

3.1 Proof of Theorem 1

Let $A \in \mathcal{R}^{n\times n}$ be a fixed circulant matrix. Without any loss of generality we can assume that the maximal value in the first row of A is equal to zero. Then $\lambda(A) = 0$, $A_\lambda = A$, and any two fundamental eigenvectors g_j, g_k are equivalent if and only if vertices j, k are contained in a common cycle c with $w(c) = 0$ in D_A. Such a cycle will be called *zero-cycle*. For reader's convenience, the proof is divided into three cases.

Case 1. In the first case we assume that the maximal value 0 in the first row of matrix A is contained on the position 0 and nowhere else. In other words, $a_0 = 0$ and $a_i < 0$ for $i = 1, 2, \ldots, n-1$. Then matrix A contains zeros on the main diagonal and negative values on other positions. Hence, the only zero-arcs in the associated digraph D_A are the loops on all vertices, which are also the only zero-cycles. Hence, and all fundamental eigenvectors are pairwise non-equivalent. As a consequence, the eigenspace dimension is $n = \gcd(0, n)$, and the assertion of Theorem 1 in Case 1 is true.

Case 2. In this case we assume that the maximal value 0 in the first row is situated on a single position $p > 0$. Then the zero-arcs in the associated digraph D_A are exactly the arcs of span p, i.e. the arcs connecting every vertex i with vertex $i + p$. Hence, every zero-cycle is composed of arcs of span p and, therefore, the length of every zero-cycle is a multiple of p. At the same time, the length of any cycle is a multiple of n, and a multiple of the greatest common divisor $d = \gcd(p, n)$, as well. By a well-known theorem from number theory, every sufficiently large multiple of d can be expressed as a linear combination of numbers p, n with positive coefficients. As a consequence, any two vertices $i, j \in D_A$ are contained in a common zero-cycle if and only if $i - j \equiv d \pmod n$. By this, there are exactly d classes of eigenvector equivalence. In other words, the eigenspace dimension is equal to $d = \gcd(p, n)$.

Case 3. In fact, the last case is the most general and covers also the cases 1 and 2. We assume that the maximal value 0 in the first row of A is situated on k positions $p_1, p_2, \ldots, p_k$. Let us denote $d = \gcd(p_1, p_2, \ldots, p_k, n)$. By similar arguments as in the previous two cases, we get that any two vertices $i, j \in D_A$ are contained in a common zero-cycle if and only if $i - j \equiv d \pmod n$, i.e. the number of equivalence classes in the eigenspace of A is exactly d. By this, the proof of Theorem 1 is complete.

4 Eigenproblem in max-min algebra

It is well-known fact that the eigenvalue can be eliminated in the formulation of the eigenproblem in max-min algebra. Hence, the max-min eigenproblem stands as follows: for given max-min matrix $A \in \mathcal{R}^{n\times n}$, find vector $x \in \mathcal{R}^n$ satisfying the equation

$$A \otimes x = x \ .$$

The structure of the max-min eigenspace for circulant matrices is described in the following theorem. We use the notation $M(A) = \max(a_0, a_1, \ldots, a_{n-1})$ and $m(A) = \max(a_1, \ldots, a_{n-1})$, where the row vector $(a_0, a_1, \ldots, a_{n-1})$ is the first row of a given circulant matrix A.

Theorem 2. *Let matrix $A \in \mathcal{R}^{n\times n}$ be circulant. Then vector $x = (x_1, x_2, \ldots, n)$ is a max-min eigenvector of A if one of the following conditions is satisfied*

(i) $x_1 = x_2 = \ldots = x_n \le M(A)$,

(ii) $m(A) \le x_i \le M(A)$ *for every* $i = 1, 2, \ldots, n$.

4.1 Proof of Theorem 2

Let $A \in \mathcal{R}^{n\times n}$ be a given circulant matrix with elements $(a_0, a_1, \ldots, a_{n-1})$ in the first row, let $x = (x_1, x_2, \ldots, n) \in \mathcal{R}^n$. We shall prove that any of the conditions (i), (ii) is sufficient for the max-min equation $A \otimes x = x$.

Condition (i). Let us assume that the condition $x_1 = x_2 = \ldots = x_n \leq M(A)$ is satisfied. By definition of $M(A)$, there exists $p \in N_0 = \{0, 1, 2, \ldots, n-1\}$ such that $a_p = M(A)$. Using cyclic indexation modulo n of elements in A, we have for any $i \in N = \{1, 2, \ldots, n\}$

$$(A \otimes x)_i = \bigoplus_{k \in N} (a_{k-i} \otimes x_k) \geq a_p \otimes x_{p+i} = M(A) \otimes x_{p+i} = x_i \ .$$

On the other hand, we have

$$(A \otimes x)_i = \bigoplus_{k \in N} (a_{k-i} \otimes x_k) \leq \bigoplus_{k \in N} (M(A) \otimes x_k) = \bigoplus_{k \in N} x_k = x_i$$

As a consequence, we have the equation $A \otimes x = x$, i.e. x is eigenvector of A

Condition (ii). Let us assume that the inequalities $m(A) \leq x_i \leq M(A)$ hold true for every $i \in N$. By definition of $m(A)$, there exists $p \neq 0$ such that $a_p = m(A)$.

We shall distinguish two cases: (a) $a_0 = M(A)$, (b) $a_0 < M(A) = m(A)$. In case (a) we have, for any $i \in N$

$$(A \otimes x)_i = \bigoplus_{k \in N} (a_{k-i} \otimes x_k) \geq a_0 \otimes x_i = M(A) \otimes x_i = x_i \ ,$$

$$\begin{aligned}(A \otimes x)_i &= \bigoplus_{k \in N} (a_{k-i} \otimes x_k) = (a_0 \otimes x_i) \oplus \bigoplus_{k \in N \setminus \{i\}} (a_{k-i} \otimes x_k) \\ &\leq (M(A) \otimes x_i) \oplus \bigoplus_{k \in N \setminus \{i\}} (m(A) \otimes x_k) = x_i \oplus m(A) = x_i \ .\end{aligned}$$

In case (b) we have $M(A) = m(A)$, and condition (ii) reduces to condition (i). Hence, in both cases (a) and (b) the equation $A \otimes x = x$ holds true, i.e. x is eigenvector of A.

Note 1. There exist circulant matrices with max-min eigenvectors not satisfying any of conditions in Theorem 2. The paper describing the form of such eigenvectors is in preparation.

5 Acknowledgements

The support of Czech Science Foundation #402/09/0405 is kindly acknowledged.

References

1. R. A. Cuninghame-Green, *Minimax algebra*, Lecture Notes in Econ. and Math. Systems 166, Springer-Verlag, Berlin, 1979.
2. R. A. Cuninghame-Green, *Minimax algebra and Application.* In: Advances in Imaging and Electron Physics 90, (P. W. Hawkes, Ed.). Academic Press, New York, 1995.
3. M. Gavalec, J. Plavka, *Eigenproblem in extremal algebras*, Proc. of the 9th Int. Symp. Oper. Res. '07, Nova Gorica, Slovenia (2007), 15-21.
4. J. Plavka, *Eigenproblem for circulant matrices in max-algebra*, Optimization 50 (2001), 477-483.
5. J. Plavka, *l-parametric eigenproblem in max algebra*, Discrete Applied Mathematics 150 (2005), 16-28.
6. H. Tomášková, *Eigenvalue problem for parametrized circulant matrices*, Abstracts of the Conf. MME 2008, Prague (2008).
7. H. Tomášková, *Eigenproblem for circulant matrices in max-plus algebra*, Abstracts of the Conf. IMEA 2009, Hradec Králové (2009).
8. K. Zimmermann, *Extremal Algebra* (in Czech), Ekon. ústav ČSAV Praha, 1976.
9. U. Zimmermann, *Linear and Combinatorial Optimization in Ordered Algebraic Structure*, Ann. Discrete Math., No 10, North Holland, Amsterdam, 1981.

OPTIMIZATION PROBLEMS UNDER (max, min)-LINEAR EQUATIONS - THEORY, METHODS, APPLICATIONS*

Martin Gavalec, Karel Zimmermann

University of Hradec Kralove, Faculty of Informatics and Management
Charles University Prague, Faculty of Mathematics and Physics
martin.gavalec@uhk.cz, karel.zimmermann@mff.cuni.cz

Abstract. An iteration method for solving optimization problems with constraints described by a system of (max, min)-linear equations is proposed. Some theoretical properties of the problems are proved. Some prossibilities of applications of such problems are presented.
Keywords. Nonconvex Optimization Methods, (max,min)-linear equations.

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. [5], [10]). A systematic theory of such algebraic structures was published probably for the first time in [5]. In these publications among other problems systems of the so called $(\max, +)$- or $(\max, \min)$-linear equations with variables on only one side of the equations were investigated. Since operation max replacing addition is no more a group, but only a semigroup operation, it is a substantial difference between solving systems with variables on one side and systems with variables occuring on both sides of the equations. The former systems will be called "one-sided" and the latter systems "two-sided". Special two-sided systems were studied in [4], [5], [8], [10] in connection with the so called $(\max, +)$- or $(\max, \min)$-eigenvalue problem. General two-sided $(\max, +)$-linear systems were studied in [2], [3]. Two-sided systems with a more general structure, in which on both sides of the equations residual functions occur were investigated in [6], where a general iteration method for solving such systems was proposed. The method can be applied also to $(\max, +)$- or $(\max, \min)$-linear systems. In [7] a polynomial method for finding the maximum solution of the $(\max, \min)$-linear system was proposed. Using this method we will propose an iteration method for solving optimization problems with max-separable objective functions and the set of feasible solutions described by a system of $(\max, \min)$-linear equations.

2 Notations, Problem Formulation, Theoretical Properties.

Let us introduce the following notations:

$J = \{1, \ldots, n\}$, $I = \{1, \ldots, m\}$, $R = (-\infty, \infty)$, $\overline{R} = [-\infty, \infty]$,

$R^n = R \times \ldots \times R$ (n-times), similarly $\overline{R}^n = \overline{R} \times \ldots \times \overline{R}$, $x = (x_1, \ldots, x_n) \in \overline{R}^n$, $\alpha \wedge \beta \equiv \min\{\alpha, \beta\}$ for any $\alpha, \beta \in \overline{R}$, we set per definition $-\infty \wedge \infty = -\infty$, $x^{(0)} \equiv (\infty\ , \ldots,\ \infty)$,

$a_{ij}, b_{ij} \in \overline{R}$, $\forall\ i \in I$, $j \in J$ are given numbers,

$$a_i(x) \equiv \max_{j \in J}(a_{ij} \wedge x_j) \quad \text{for all } \ i \in I\,,$$
$$b_i(x) \equiv \max_{j \in J}(b_{ij} \wedge x_j) \quad \text{for all } \ i \in I\,,$$

We will consider the following system of $(\max, \min)$-linear equations

$$a_i(x) = b_i(x) \quad \forall i \in I. \tag{1}$$

The set of all solutions of system (1) will be denoted by M. We define further sets $M(\overline{x})$, $M_i(\overline{x})$ for any $\overline{x} \in \overline{R}^n$, $i \in I$ as follows:

* Under the support of GA ČR #402/09/0405, #202/03/2060982

$$M(\overline{x}) \equiv \{x | x \in M \ \& \ x \leq \overline{x}\} \tag{2}$$

$$M_i(\overline{x}) \equiv \{x | a_i(x) = b_i(x) \ \& \ x \leq \overline{x}\} \tag{3}$$

Let us note that sets $M(\overline{x})$, $M_i(\overline{x})$ are always nonempty, since e.g. $x(\alpha) \equiv (\alpha, \ \ldots, \ \alpha) \in M(\overline{x})$, if $\alpha \leq \min_{(i,j) \in I \times J}(a_{ij} \wedge b_{ij} \wedge \overline{x}_j)$. Let us note further that if $\overline{x} = x^{(0)} \equiv (\infty, \ \ldots, \ \infty)$, then evidently $M(\overline{x}) = M$, and if $\underline{x} = (-\infty, \ \ldots, \ -\infty)$, then $\underline{x} \leq x$ for any $x \in M$.

Definition 1. Let $L \subseteq \overline{R}^n$, $\tilde{x} \in L$. Let the following implication holds: $x \in L \Rightarrow x \leq \tilde{x}$. Then $\tilde{x}$ is called the maximum element of L.

Remark 1. The algorithm presented in [7] finds the maximum element of $M(\overline{x})$ for any given $\overline{x}$, i.e. element $x^{max} \in M(\overline{x})$ such that $x \leq x^{max} \ \forall \ x \in M(\overline{x})$.

In this paper, we will solve the following minimization problem:

$$f(x) \ \longrightarrow \ \min \ \ s.t. \ x \in M(\underline{x}, \overline{x}) \equiv \{x | x \in M \ \& \ \underline{x} \leq x \leq \overline{x}\}, \tag{4}$$

where $\underline{x}, \ \overline{x} \in R^n$ and the objective function $f(x) = f(x_1, \ \ldots, \ x_n)$ is defined as follows:

$$f(x_1, \ \ldots, \ x_n) \equiv \max\{f_1(x_1), \ \ldots, \ f_n(x_n)\}.$$

We will assume that $f_j(x_j), \ j \in J$ are continuous strictly increasing functions and $\underline{f} \equiv f(\underline{x}), \ \overline{f} \equiv f(\overline{x})$. In the next section we will propose an iteration algorithm for solving minimization problem (4).

3 Solution Method

Lemma 1. Let $\underline{y} \leq \overline{y}$, let y^{max} be the maximum element of M($\overline{y}$). Then $M(\underline{y}, \overline{y}) \neq \emptyset$ if an only if $\underline{y} \leq y^{max}$

Proof:
If $\underline{y} \leq y^{max}$, then $y^{max} \in M(\underline{y}, \overline{y})$ and thus $M(\underline{y}, \overline{y}) \neq \emptyset$.
If $\underline{y} \not\leq y^{max}$, then $M(\underline{y}, \overline{y}) = \emptyset$, since any element of $M(\underline{y}, \overline{y})$ would have to satisfy the inequalities $\underline{y} \leq y \leq y^{max}$, which is under the assumption that $\underline{y} \not\leq y^{max}$ impossible. □

ALGORITHM

$\boxed{1}$ Input $\underline{x}, \ \overline{x}, \ \epsilon > 0, \ \underline{f} := f(\underline{x}), \ \overline{f} := f(\overline{x})$;
$\boxed{2}$ Find the maximum element $\tilde{x}$ of $M(\overline{x})$ using the method from [7];
$\boxed{3}$ If $\underline{x} \not\leq \tilde{x}$, then $M(\underline{x}, \overline{x}) = \emptyset$, STOP.
$\boxed{4}$ $\overline{f} := f(\tilde{x})$, $f^{(1)} := \underline{f} + (\overline{f} - \underline{f})/2$, $\overline{x}_j^{(1)} := f_j^{-1}(f^{(1)}) \wedge \tilde{x}_j \ \forall j \in J$, $\overline{x}^{(1)} := (\overline{x}_1^{(1)}, \ \ldots, \ \overline{x}_n^{(1)})$ (determining the new upper bound);
$\boxed{5}$ Find the maximum element $\tilde{x}^{(1)}$ of $M(\overline{x}^{(1)})$ using the method from [7];
$\boxed{6}$ If $\underline{x} \leq \tilde{x}^{(1)}$, then set $\tilde{x} := \tilde{x}^{(1)}$ and go to $\boxed{8}$;
$\boxed{7}$ Set $\underline{f} := f(\overline{x}^{(1)})$ (determining new lower bound for f);
$\boxed{8}$ If $(f(\tilde{x}) - \underline{f}) > \epsilon$, then go to $\boxed{4}$;
$\boxed{9}$ $\tilde{x}$ is the ϵ-approximation of the optimal solution STOP.

4 Applications, Numerical Examples.

Example 1. Let us consider a situation, in which transportation means of different size provide transporting goods from in places $i \in I$ to a terminal T. The goods are unloaded in T and the transportation means (possibly with other goods uploaded in T) have to return to i. We assume that the connection between i and T is possible only via one of the placies (e.g. cities) $j \in J$ and the capacities of the roads between i and j is equal to a_{ij}. We have to join placies j with T by a road with a capacity x_j so that the total capacity of the connection between i and T is equal to $a_{ij} \wedge x_j$. The transport from T to i is carried out via possibly different places $k \in K$ with capacities between k and i equal to b_{ik}. We have to join placies k with T by a road with a capacity y_k so that the total capacity of the connection between T and i is equal to $b_{ik} \wedge y_k$. We assume that the transportation means can pass only through some roads with a capacity corresponding to the size of the transportation mean and our task is to choose appropriate capacities $x_j,\ y_k$. In order that each of the transportation means may return to i, it is natural to require for each i that the maximal attainable capacity of connections between i and T via j is equal to maximal attainable capacity of connections between T and i via k. In other words, we have to choose x_j and y_k in such a way that

$$\max_{j\in J}(a_{ij}\wedge x_j)=\max_{k\in K}(b_{ik}\wedge y_k)\ \forall\ i\in I. \tag{5}$$

We see that this problem can be transformed to a system having form (1) (we can namely consider vector (x,y) and introduce coefficients at y on the left hand sides and at x on the right hand sides equal to zero). Note that the model is flexible enough to include different real situations. If for instance the road between i and j does not exist, we set simply $a_{ij}=0$, if the road is a two way road with equal capacity in both directions, we set $b_{ik}=a_{ij}$, if we do not want to connect j with T, we set the lower and upper bound on x_j equal to 0. Let us assume further that the choice of $x_j,\ y_k$ is connected with penalties $f_j(x_j),\ g_k(y_k)$ respectively. Tne penalties may be connected e.g. with some economic or ecologic requirements (costs, air pollution) so that it is quite natural to accept that $f_j,\ g_k\ \forall j\in J,k\in K$ are continuous strictly increasing functions. The problem consisting in minimizing the maximum penalty under the constraints given by (5) and some lower and upper bounds on x,y can be easily transformed to an optimization problem having form (4).

Example 2. Let us consider the following optimization problem:
Minimize $f(x)\equiv\max(2x_1,3x_2,x_3)$ subject to

$$\max(4\wedge x_1,5\wedge x_2,0\wedge x_3)=\max(0\wedge x_1,4\wedge x_2,5\wedge x_3),$$

$$\underline{x}\le x\le\overline{x},$$

where $\underline{x}=(5,3,4),\ \overline{x}=(6,6,6)$.
We proceed in accordance with the AGORITHM as follows:

$\boxed{1}$ $\underline{x}=(5,3,4),\ \overline{x}=(6,6,6),\ \underline{f}:=10,\ \overline{f}:=18,\ \epsilon=1;$
$\boxed{2}$ $\tilde{x}=\overline{x}=(6,6,6);$
$\boxed{3}$ $\underline{x}\le\tilde{x};$
$\boxed{4}$ $\overline{f}:=f(\tilde{x})=18,\ f^{(1)}:=10+(18-10)/2=14,\ \overline{x}^{(1)}:=(6,14/3,6);$
$\boxed{5}$ $\tilde{x}=(6,14/3,14/3);$
$\boxed{6}$ $\underline{x}\le\tilde{x};$
$\boxed{8}$ $f(\tilde{x})-\underline{f}=14-10=4>\epsilon=1;$
$\boxed{4}$ $\overline{f}:=f(\tilde{x})=14,\ f^{(1)}=10+(14-10)/2=12,\ \overline{x}^{(1)}:=(6,4,14/3);$
$\boxed{5}$ $\tilde{x}=(6,4,4)$
$\boxed{6}$ $\underline{x}\le\tilde{x};$
$\boxed{8}$ $f(\tilde{x})-\underline{f}=12-10=2>\epsilon=1;$
$\boxed{4}$ $\overline{f}:=f(\tilde{x})=12,\ f^{(1)}:=10+(12-10)/2=11;$
$\boxed{5}$ $\tilde{x}=(11/2,11/3,4);$
$\boxed{6}$ $\underline{x}\le\tilde{x};$
$\boxed{8}$ $f(\tilde{x})-\underline{f}=11-10=1=\epsilon=1;$
$\boxed{9}$ $\tilde{x}$ is the ϵ-optimal solution, STOP;

Example 3. Let us consider the following optimization problem:
Minimize $f(x)\equiv\max(2x_1,3x_2,x_3)$ subject to

$$\max(4\wedge x_1,5\wedge x_2,0\wedge x_3)=\max(0\wedge x_1,4\wedge x_2,5\wedge x_3),$$

$$\underline{x}\le x\le\overline{x},$$

where $\underline{x}=(5,3,5),\ \overline{x}=(6,6,6)$.
We proceed in accordance with the AGORITHM as follows:

$\boxed{1}$ $\underline{x}=(5,3,5),\ \overline{x}=(6,6,6),\ \underline{f}:=10,\ \overline{f}:=18,\ \epsilon=1;$
$\boxed{2}$ $\tilde{x}=\overline{x}=(6,6,6);$
$\boxed{3}$ $\underline{x}\le\tilde{x};$
$\boxed{4}$ $\overline{f}:=f(\tilde{x})=18,\ f^{(1)}:=10+(18-10)/2=14,\ \overline{x}^{(1)}:=(6,14/3,6);$

$\boxed{5}$ $\tilde{x} \;=\; (6, 14/3, 14/3)$;
$\boxed{6}$ $\underline{x} \;\not\leq\; \tilde{x}$;
$\boxed{7}$ $\underline{f} := f(\overline{x}^{(1)}) = 14,\; \tilde{x} := \overline{y}$, go to $\boxed{4}$;
$\boxed{4}$ $\overline{f} := f(\tilde{x}) = 18,\; f^{(1)} := 14 + (18 \;-\; 14)/2 = 16,\; \overline{x}^{(1)} := (6, 16/3, 6)$;
$\boxed{5}$ $\tilde{x} = (6, 16/3, 5)$;
$\boxed{6}$ $\underline{x} \;\leq\; \tilde{x}^{(1)},\; \tilde{x} \;:=\; \tilde{x}^{(1)} = (6, 16/3, 5)$;
$\boxed{8}$ $f(\tilde{x}) \;-\; \underline{f} = 16 - 14 = 2 > \epsilon$;
$\boxed{4}$ $\overline{f} := f(\tilde{x}) = 16,\; f^{(1)} := 14 + (16 - 14)/2 = 15,\; \overline{x}^{(1)} := (6, 5, 5)$;
$\boxed{5}$ $\tilde{x} = (6, 5, 5)$;
$\boxed{6}$ $\underline{x} \;\leq\; \tilde{x}$;
$\boxed{8}$ $f(\tilde{x}) \;-\; \underline{f} = 15 - 14 = 1 = \epsilon$;
$\boxed{9}$ $\tilde{x}$ is the ϵ-optimal solution, STOP;

References

1. Baccelli, F.L., Cohen, G., Olsder, G.J., Quadrat, J.P.: Synchronization and Linearity, An Algebra for Discrete Event Systems, Wiley, Chichester, 1992.
2. Butkovič, P., Hegedüs, G.: An Elimination Method for Finding All Solutions of the System of Linear Equations over an Extremal Algebra, Ekonomicko - matematický obzor 20 (1984), No. 2, pp. 203-215.
3. Butkovič, P., Zimmermann, K.: A Strongly Polynomial Algorithm for Solving Two-sided Linear Systems in Max-algebra, Discrete Applied Mathematics 154(2006), pp. 437-446.
4. Cechlárová, K.: Efficient Computation of the Greatest Eigenvector in Fuzzy Algebra, Tatra Mt. Math. Publications, 1997, pp. 73-79.
5. Cuninghame-Green, R.A.: Minimax Algebra, Lecture Notes in Economics and Mathematical Systems, 166, Springer Verlag, Berlin, 1979.
6. Cuninghame-Green, R.A., Zimmermann, K.: Equation with Residual Functions, Comment. Math. Univ. Carolinae 42(2001), 4, pp. 729-740.
7. Gavalec, M., Zimmermann, K.: Solving Systems of Two-Sided $(\max, \min)$-Linear Equations, to appear.
8. Sanchez, E.: Resolution of Eigen Fuzzy Sets Equations, Fuzzy Sets and Systems, vol. 1, No 1, 1978, pp. 69 - 74.
9. Sanchez, E.: Inverses of Fuzzy Relations. Applications to Possibility Distributions and Medical Diagnosis, Fuzzy Sets and Systems, vol. 2, No 1, 1979, pp. 75 - 86.
10. Vorobjov, N.N.: Extremal Algebra of positive Matrices, Datenverarbeitung und Kybernetik, 3, 1967, pp. 39-71 (in Russian).

ESTIMATION OF PD OF FINANCIAL INSTITUTIONS WITHIN LINEAR DISCRIMINANT ANALYSIS

Petr Gurný, Martin Gurný

VŠB – TU Ostrava
Faculty of Economics, Department of Finance,
petr.gurny.ekf@vsb.cz

Abstract. Correct determination of probability of default (PD) of particular financial subjects plays a necessary role in lot of crucial financial fields such are risk management, requests for loans, determination of rating and particularly for pricing of credit derivatives. Among the most widely used models to forecast a company's default, is a class of statistical models, generally known as credit-scoring models. In this paper a possibility of determination of financial institution's PD according to the linear discriminant analysis (one of the categories of credit-scoring models) is discussed. The main part of the paper is devoted to the estimation of the score via discriminant function and to the direct determination of the probability of default associated with the individual companies analyzed. Further we introduce index, which is possible to use to measure the success of a model. Linear discriminant analysis is applied and verified on sample of 14 financial institutions. There are also discussed restrictions of introduced model a possibilities of its improvement in the paper.
Keywords. Probability of default, linear discriminant analysis, economic indicators of financial institutions, Wilk's Lambda.

1. Introduction

Estimating the borrower's risk level, namely the probability of default, by assigning a different PD to each borrower is now widely employed in many banks. The false estimation of PD led to unreasonable rating, incorrect pricing of financial instruments and thereby it was one of the causes of the financial crisis. One of the possibilities how to estimate the probability of default is to use so-called credit-scoring models. Although the techniques underlying these models were devised in the 1930s by authors such as Fischer [5] and Durand [4], the decisive boost to the development and spread of these models came in the 1960s, with studies by Beaver [3], Altman [1] and others. Presently it is tendency to analyze and to simulate the time series of the particular variables by means of Lévy processes [7] as it could by helpful to better predictor ability of the models.

The goal of the paper is the application and verification of usage possibility of linear discriminant analysis which is one of the credit-scoring models. Within this method we arrange a model for determination of probability of default of financial institutions.

In the theoretical part of the paper we expand in more detail the linear discriminant analysis at first. This part also involves the way of calculation of PD and the possibility of measuring the success rate of a model. There will be introduced a sample of 14 financial institutions, that are to be divided in healthy companies and default ones, including their financial indicators in the next part. Further we will apply the linear discriminant analysis on the chosen sample of the banks and determine the particular PDs. In conclusion we will discuss the results and also the restrictions and the drawbacks of the model.

2. Review of credit-scoring models

There are a lot of models to estimate the probability of default in financial world. Among the most widely used ones are models, generally known as credit-scoring models. These are multivariate models which use the main financial indicators of a company as input, attributing a weight to each of them that reflects its relative importance in forecasting default. We can group these models into three categories: a) linear discriminant analysis; b) regression models; c) inductive models. Linear discriminant analysis (will be presented in more detail bellow) is based on the estimation of the discriminant function which make it possible to discriminate between healthy companies and abnormal ones. In regression models (linear, logit and probit) we identify the variables that lead to the default of company, and their weights, with a simple linear regression. Both above mentioned models are based on a deductive approach designed to explain the economic causes of default. By contrast, inductive models (neural networks, genetic algorithm) follow a purely empirical inductive. These models are often “black-boxes”, which can be used to generate results rapidly, but whose logic may not be fully understood. We will be concerned with linear discriminant analysis in more detail in the next part of the paper.

3. Linear discriminant analysis

Basically, discriminant analysis is a classification technique which uses data obtained from a sample of companies to draw a boundary that separates the group of reliable ones from the group of insolvent ones. This means (how we mentioned above) to find the discriminant function. To do that we have to initially identify the variables, which typically are economic and financial indicators taken from financial statements. Afterwards the discriminant function estimates so-called z-score, which is a linear combination of the chosen independent variables. For i-th company and n independent variables, the score will be calculated as follows:

$$z_i = \sum_{j=1}^{n} \gamma_j x_{i,j} \,. \tag{1}$$

The coefficients γ_j of this linear combination are chosen as to obtain a score z which discriminates abnormal and healthy companies. The value z must be obtained in such a way that it maximizes the distance between centroids (means of z_A and z_B) of the two groups of healthy and insolvent companies. The vector of gamma coefficients is calculated as follows:

$$\gamma = \Sigma^{-1}\left(x_A - x_B\right), \tag{2}$$

where x_A and x_B are vectors containing the mean values of the n independent variables for the group of healthy companies and the group of insolvent ones. Σ is the matrix of variances and covariances between the n independent variables and can be estimated as the “average” Σ of the matrixes of variances and covariances for each group of companies, weighted by the number companies $\left(n_A, n_B\right)$ present in each group:

$$\Sigma = \frac{n_A - 1}{n_A + n_B - 2}\Sigma_A + \frac{n_B - 1}{n_A + n_B - 2}\Sigma_B \,.$$

A very important task is to identify the indicators which will be used to estimate the model. In this paper the selection of discriminant variables will follow the stepwise method. It means the initially including all the variables in the list and subsequently removing those with lower discriminating power.

In the following it is necessary to estimate so-called cut-off point (α), which facilitates us to distinguish between relatively healthy companies and too risk ones. We can do that in the following way:

$$\alpha = \frac{1}{2}\gamma'(x_A + x_B)\,. \tag{3}$$

3.1. The estimation of the probability of default

Discriminant analysis can be used to produce a direct estimate of the probability of default. It can by shown (Altman 1981), that if the independent variables are distributed according to a multivariate normal distribution, the company’s probability of default is given by:

$$PD = p\left(B \middle| x_i\right) = \frac{1}{1 + \frac{1 - \pi_B}{\pi_B} e^{z_i - \alpha}} \,, \tag{4}$$

where z_i and α are quantities defined in (1) and (3) and π_B represents the prior probability of default, which depends on the general characteristic of the market.[1]

3.2. Wilk's Lambda

Wilk's Lambda is a widely used index to measure the success rate of a model. It is given by

$$\Lambda = \frac{\sum_{i \in A} \left(z_i - z_A\right)^2 + \sum_{i \in B} \left(z_i - z_B\right)^2}{\sum_{i=1}^{n} \left(z_i - \bar{z}\right)^2}, \tag{5}$$

where $\bar{z}$ represents the mean of z_i in the entire sample of healthy and insolvent companies. If the scores for the individual companies from particular group are very similar to one another, the numerator (as will the entire Wilk's Lambda) will be close to zero and the model is very effective. However, if the two centroids are very similar to one another, the sum of the within deviances will be close to the total deviance, and the ratio will be close to 1.

4. Data description

In our paper we will work with sample of 14 financial institutions. As a first step we had to divide this sample of banks into two groups – the healthy companies and the abnormal (or defaulted) ones. Defaulted companies can be defined in a variety of ways. In this paper the defaulted banks are thought the financial institutions which have gone into liquidation or undergone financial restructuring processes (e.g. take-over by another company or by government). The samples of the financial institutions for both of these groups were chosen randomly pursuant to the publicly available information. As a second step we identified the financial indicators from financial statements. Table 1 and table 2 show the chosen financial indicators for both groups of companies including their mean values.[2] All indicators were calculated from data in year 2007.

	x_1: TA	x_2: LTA	x_3: EQ	x_4: YAEA	x_5: CIBL	x_6: NIM	x_7: ROAA	x_8: ROAE	x_9: IE II	x_{10}: CIR	x_{11}: PE OI	x_{12}: EQ TA
Bank of America	1715746	14,355	146803	7,16%	4,32%	34433	0,87%	14,25%	60,56%	75,41%	15,73%	8,56%
JPMorgan Chase	1562147	14,262	123221	10,31%	4,28%	26406	0,98%	18,51%	63,01%	74,50%	19,50%	7,89%
M&T Bank	64876	11,08	6485	7,49%	3,53%	1850	1,01%	14,87%	47,81%	74,21%	20,28%	10,00%
National City	150374	11,921	13408	8,03%	4,16%	4396	0,21%	2,77%	52,14%	85,61%	21,88%	8,92%
PNC Bank	138920	11,842	14854	8,78%	3,45%	2915	1,06%	14,10%	52,72%	75,80%	21,49%	10,69%
SunTrust Bank	179574	12,098	18053	8,20%	4,03%	4720	0,91%	12,46%	52,97%	78,35%	20,57%	10,05%
Wells Fargo & Comp.	575442	13,263	47628	9,27%	3,88%	19142	1,40%	24,41%	40,38%	69,09%	24,94%	8,28%
Zionsbancorporation	52947	10,877	5293	7,99%	3,56%	1882	0,93%	13,92%	41,28%	75,42%	22,12%	10,00%
Mean values:	***555003***	***12,46***	***46968***	***8,40%***	***3,90%***	***11968***	***0,92%***	***14,41%***	***51,36%***	***76,05%***	***20,81%***	***9,30%***

Table 1: Financial indicators of healthy companies

	x_1: TA	x_2: LTA	x_3: EQ	x_4: YAEA	x_5: CIBL	x_6: NIM	x_7: ROAA	x_8: ROAE	x_9: IE II	x_{10}: CIR	x_{11}: PE OI	x_{12}: EQ TA
Wachovia	782896	13,571	76872	8,87%	4,02%	18130	0,81%	11,41%	57,07%	79,10%	21,95%	9,82%
Merrill Lynch	1020050	13,835	31932	59,98%	13,02%	5549	-0,76%	-35,81%	90,26%	120,47%	25,37%	3,13%
The Bear Stearns	395362	12,888	11793	13,27%	3,38%	1350	0,06%	1,64%	88,32%	98,81%	21,21%	2,98%
Lehman Brothers	691063	13,446	22490	20,25%	9,32%	1947	0,61%	26,74%	95,33%	89,81%	16,09%	3,25%
Landsbanki	43692	10,685	2629	7,93%	5,84%	772	1,31%	24,76%	73,27%	93,98%	17,25%	6,02%
Glitnir banki	42140	10,649	2429	7,41%	5,88%	558	0,94%	19,93%	79,18%	100,36%	24,57%	5,76%
Mean values:	***495867***	***12,51***	***24691***	***19,62%***	***6,91%***	***4718***	***0,50%***	***8,11%***	***80,57%***	***97,09%***	***21,07%***	***5,16%***

Table 2: Financial indicators of default companies

5. Application and results

Application process will be divided into next steps:

i. According to (2) we will estimate particular weights of the variables and subsequently remove those with lower discriminating power.

[1] Because of lack of information, we will use the proportion of insolvent companies in sample, n_B / n_A.

[2] The more detail description of these financial indicators is shown in appendix A.

ii. We will write down the discriminant function and estimate z-score for every company.
iii. We will calculate cut-off point (α) and estimate PDs according to (3) and (4).
iv. By way of Wilk's Lambda, equation (5), we will measure the success rate of a model.
v. We will discuss the results and restrictions of the model.

The vector γ (weights of financial indicators) is according to (2) estimated as follows:

$$\gamma = \{0 \quad -0{,}8 \quad 0 \quad -6{,}2 \quad 0 \quad 0{,}0002 \quad -18 \quad 0 \quad 0 \quad 0 \quad 12 \quad 0\}.$$

Here we can see that only 5 indicators with the most significant discriminating power remained. Now we can plug this vector in equation (1) and we get the general discriminant function for estimating z-score:

$$z_i = -0{,}8x_{2,i} - 6{,}2x_{4,i} + 0{,}0002x_{6,i} - 18x_{7,i} + 12x_{11,i} \ ,$$

where x_2 is indicator *LTA*, x_4 *YAEA*, x_6 *NIM*, x_7 *ROAA*, x_{11} *PE OI.*

Now we can calculate the *z*-score for every particular company. The results are shown in table 3. After calculation of the cut-off point (α), $\alpha = -6{,}987$, we can directly estimate the probability of default for analyzed companies according to (4). The results are shown in table 3 again.

Non - default companies	z_i	**PD**	***Default companies***	z_i	**PD**
Bank of America	-4,436	5,6%	*Wachovia*	-5,811	18,9%
JPMorgan Chase	-5,426	13,7%	*Merrill Lynch*	-10,537	96,3%
M&T Bank	-6,656	35,2%	*The Bear Stearns*	-8,251	72,8%
National City	-6,601	34,0%	*Lehman Brothers*	-9,762	92,4%
PNC Bank	-7,027	44,1%	*Landsbanki*	-6,975	42,8%
SunTrust Bank	-6,987	43,1%	*Glitnir banki*	-5,976	21,6%
Wells Fargo & Company	-5,159	10,8%			
Zionsbancorporation	-6,278	27,1%			
Mean values:	***-6,071***	***26,7%***	***Mean values:***	***-7,885***	***57,5%***

Table 3: Z-score and PD

As we can see from results in table 3, our model is not very effective. The two groups are rather clustered around their respective centroids (equal to -6,071 for healthy companies and -7,885 for default ones), however, they are not perfectly separate, as some healthy companies have relative low scores, and especially some default companies (Wachovia and Glitnir banki) have relative high ones, which leads to the too low PD.

As we mentioned in the theoretical part, model efficiency can be also measured through the use of Wilk's Lambda. So by means of (5) we can measure the success rate of the model as

$$1 - \Lambda = 1 - \frac{\sum_{i \in A} (z_i - z_A)^2 + \sum_{i \in B} (z_i - z_B)^2}{\sum_{i=1}^{n} (z_i - \bar{z})^2} = 1 - \frac{6{,}34 + 19{,}46}{37{,}09} = 30{,}42\% \ .$$

How we can see from equation above, the deviance in the group of default companies is relative high, hence the success rate of the model is low. So what can we do to improve the model? Generally, when we use the credit-scoring models to estimate the PD of the particular company, we have to pay attention to some problems. First problem relates to the insufficient amount of the companies in the sample. Great numbers of observations are needed to obtain robust results. From this point of view the sample of 14 institutions is too low to guarantee model with sufficient accuracy. The second problem is associated with the definition of the default company. The definition used to break down the estimation sample obviously affects the results of the model. A third problem is due to the fact that credit-scoring models ignore numerous qualitative factors, which can be highly significant in determining the insolvency of a company (such as company's reputation, the stage of the economic cycle, the quality of the management etc.). And last but not least we have to observe that success of the model is given by characteristic of the financial markets.

6. Conclusion

We applied and verified the possibility of usage of linear discriminant analysis in determination of PD of financial institutions in the paper. First we introduced theoretical background of credit-scoring models, particularly linear discriminant analysis, including the way of measuring the success rate of a model. In the next part of the paper we applied this method on a sample of 14 financial institutions and so we developed the model for forecasting a financial company's default. Next we estimated the PD of these financial institutions and success rate of the model. However, we faced some serious restrictions of the model, which caused that its efficiency was by no means perfect. First limit was amount of companies in the sample. It would be advisable to work with much more larger samples in order to estimate the model with sufficient accuracy. Next problems were associated with the definition of the default company and with ignore qualitative factors of companies. Elimination of these restrictions and usage of the Lévy processes to simulation of the particular financial indicators will be part of the next research.

Acknowledgement

The paper is based on research activities sponsored through GAČR (Czech Science Foundation – Grantová Agentura České Republiky) project No. 402/08/1237. The support is greatly acknowledged.

References

[1] ALTMAN, E. I. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, September, 1968, 589-609.

[2] BALTHAZAR, L. *From Basel 1 to Basel 3: The Integration of State-of-the-Art Risk Modeling in Banking Regulation*. New York: Palgrave Macmillan, 2006. 294 p.

[3] BEAVER, W. Financial ratios as predictors of failures. Empirical Research in Accounting: Selected Studies – 1966, supplement to *Journal of Accounting Research*, 4, 1967, 71-111.

[4] DURAND, B. Risk elements in consumer installments financing. Working paper, 1941, NBER.

[5] FISHER, R. The use of multiple measurements in taxonomic problems. *Annals of Eugenics,* 7, 1936, 179-188.

[6] RESTI, A., SIRONI, A. *Risk management and Shareholders' value in banking*. Chichester: Wiley, 2007, 782 p.

[7] TICHÝ, T. Posouzení vybraných možností zefektivnění simulace Monte Carlo při opčním oceňování. *Politická ekonomie* 56, 6, 2008, 772-794.

Appendix A

Indicator	***Description of Indicator***	***Indicator's group***
TA	Total assets ($, mln)	*Size*
LTA	Logarithm of total assets	*Size*
EQ	Shareholders' Equity ($, mln)	*Size*
YAEA	Interest Income / Average Interest Earning Assets (%)	*Profitability*
CIBL	Interest Expense / Average Interest Bearing Liabilities (%)	*Profitability*
NIM	Net Interest Margin	*Profitability*
ROAA	Return on Average Assets (%)	*Profitability*
ROAE	Return on Average Equity (%)	*Profitability*
IE II	Interest Expense / Interest Income (%)	*Profitability*
CIR	Cost to Income Ratio (%)	*Efficiency*
PE OI	Personnel Expenses / Operating Income (%)	*Efficiency*
EQ TA	Shareholders' Equity / Total Assets (%)	*Capital adequacy*

Appendix A: Description of financial indicators

THE EMPIRICAL ANALYSIS OF THE EFFECTS OF NATIONAL MONETARY POLICIES IN THE VISEGRAD COUNTRIES

Jana Hančlová
VŠB-Technical University of Ostrava
Faculty of Economics, Department of Mathematical Methods in Economics, jana.hanclova@vsb.cz

Ivana Kubicová[1]
VŠB-Technical University of Ostrava, Faculty of Economics
Czech National Bank, ivana.kubicova@cnb.cz

Extended abstract. This paper analyzes the effects of national monetary policy in countries of the Visegrad Group V4 comprising the Czech Republic, Hungary, Poland and Slovakia. We investigate these effects using VAR and the structural VAR (SVAR) models. The empirical analysis covers the period 1999Q1 - 2009Q1. The variables included in the VAR (SVAR) model can be divided into two groups. The first group of exogenous variables contains the GDP and the consumer price index of the euro area and the EURIBOR nominal interest rate. The second group of endogenous domestic variables comprises the output gap, the consumer price index, the three-month interbank interest rate and the nominal NAC/EUR exchange rate. We analyze whether tightening of monetary policy leads to a fall in the output gap and prices using impulse response functions and we also investigate the behavior (persistence) of national monetary policy shocks on economy in the individual countries of the Visegrad countries.
Keywords. Visegrad countries, monetary policy, VAR, structural VAR, quarterly data.

1. Introduction

This paper is a follow-up paper focused on effects of monetary policy on the economy in Central and Eastern European countries. Our interest lies on features of national monetary policy of countries Visegrad Group 4. We investigate reactions of inflation and output on domestic interest rate shock and from this point of view we deal with each country separately. The motivation for this paper also arises from the future adhesion of Visegrad Group 4[2] to the euro area and thus there exists the interest in possible differences in monetary policy transmition mechanisms see [2]. The paper is concentrated on the the Czech Republic, Hungary, Poland and Slovakia.

The questions to be answered:

- Empirical verification of appropriate model from the group of Vector autoregression model (VAR) and structural VAR to fullfill the objective of the paper.
- Whether monetary policy impulses have different affects on economic activity in countries of the Visegrad Group V4.

In Section 2 we briefly summarize VAR and structural VAR (SVAR) methodology. We would like to investigate the effects of national monetary policy using vector approach to time series modeling according to economic theory. In Section 3 we provide a description of the construction VAR or SVAR models for V4 countries including time series analysis of 3 exogenous and also 4 endogenous time series, specification tests, models estimation, diagnostic checking of estimated models and identifying the relevant innovations and monetary policy impulses in four countries. Final Section is pursued to the summary of the empirical results and addressing of the questions above mentioned.

2. Methodology

The vector approach to time series modeling uses economic theory to model the relationship among the variables of interest. Unfortunately, economic theory is often not rich enough to provide a dynamic specification that identifies all of these relationships. Furthermore, estimation and inference are complicated by the fact that endogenous variables may appear on both the left and right sides of equations. These problems lead to

[1] The views expressed in this conference paper are not necessarily those of the Czech National Bank.
[2] The Slovak Republic has already entered to the euro area 1st January 2009.

alternative, non-structural approaches to modeling the relationship among several variables. This part of our paper shortly describes the estimation and analysis of vector autoregression (VAR) and structural vector autoregression (SVAR).

A structural form of the VAR model see [4] for an (*k* x *1)* vector of endogenous variables y_t *is given by:*

$$Ay_t = A_1 y_{t-1} + \ldots + A_p y_{t-p} + B_0 x_t + B_0 x_{t-1} + \cdots + B_p x_{t-p} + \varepsilon_t, \text{ for } t=1, 2, \ldots, T, \qquad \text{(E1)}$$

where x_t is an (*n x 1)* vector of exogenous variables, ε_t is an (*k* x *1)* vector of serially uncorrelated errors distributed independently of x_t with a zero mean and a constant possitive definite variance-covariance matrix $\Omega = (\omega_{ij})$. For given values of y_t and x_t the above dynamic system is stable if all the roots of the determinantal equation

$$\left|A - A_1\lambda - A_2\lambda^2 - \cdots - A_p\lambda^p\right| = 0$$

lie strictly outside the unite circle. This stability condition ensures the existence of long-run relationships between y_t and x_t, which will be cointegrating when one or more elements of x_t are integrated (i.e. they contain unit roots).

The reduced form of the VAR model (E1), which expresses the endogenous variables in terms of the predetermined and exogenous variables, is given by:

$$y_t = \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + \Psi_0 x_t + \Psi_1 x_{t-1} + \cdots + \Psi_p x_{t-p} + u_t, \qquad \text{(E2)}$$

where $\Phi_1 = A^{-1}A_i$, $\Psi_i = A^{-1}B_i$, $u_t = A^{-1}\varepsilon_t$.

One of the main features of the traditional macromodels was their dynamic multipliers, which measured the effect of a shock to an exogenous variable, e.g. a policy change, or shock to one of the structural errors ε_t on the expected future values of the endogenous variables. The standard approach to deriving impulse response functions is to start form the moving average representaions of the final form. Structural VAR (SVAR) estimation is to obtain non-recursive orthogonalization of the error terms for impulse analysis.

3. Estimation and results a VAR model for the Czech Republic, Hungary, Poland and Slovakia

In this part, we describe the construction of a quarterly SVAR model of the Czech Republic, Slovakia, Poland and Hungary using above mentioned techniques. The SVAR models will be estimated over the period Q1'1998 – Q1'2009. This is a sample of data which is relatively reliable.

3.1. Data

A vector x_t of 3 **exogenous** variables comprising 3 variables: *lgdp_emu* is measured as the natural logarithm of GDP EMU (2000=100); *lcpi_emu* is the logarithm of the CPI EMU (2000=100); *ir3m_eu* is the 3M_Euribor. y_t is a vector of 4 **endogenous** variables: *outputgap_xx* is measured as the logarithm of GDP_xx (2000=100) from its trend (estimated by the HP filter) and multiplied by 100; *lcpi_xx* is the logarithm of the CPI for each country; *ir3m_xx* are the domestic nominal interest rate for *CZ, HU, PO and SK*. We notice *xx* as a country, where *xx* is ***CZ*** for the Czech Republic, ***HU*** for Hungary, ***PO*** for Poland and ***SK*** for Slovakia. The last *lexr_xx* variable is natural logarithm of the bilateral exchange rate NAC/EUR.

Country	**Stacionarity test (ADF, PP, KPSS)**				**Lag structure Wald test for equation**				**Lag structure LR test**			**model**
	outputgap	lcpi	ir3m	lexr	1	2	3	4	AIC	SC	HQ	
CZ	I(0)	I(1)	I(0)	I(1)	1	2	**4**	0	**4**	1	**4**	VAR(4)
HU	I(0)	I(1)	I(0)	I(0)	1	**4**	1	0	1	1	1	VAR(4)
PO	I(0)	I(1)	I(1)	I(1)	0	**4**	**4**	0	**4**	1	**4**	VAR(4)
SK	I(0)	I(0)	I(0)	I(1)	**4**	1	1	0	1	1	1	VAR(4)
EU	*lgdp_emu I(1)*	I(1)	I(1)	-	-	-	-	-	-	-	-	-

Table 1: Stacionarity and lag structure tests for CZ, HU, PO, SK

It is important that the variables used in the empirical analysis were tested for presence of a unit root. We choose the Augmented Dickey-Fuller (ADF), Phillips-Peron (PP) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests. The results of these tests for all above mentioned variables are described in the Table 1. We can conclude that all *outpugap_xx* variables are stationary at 5% level of significance and also *ir3m_xx* with the exception of *PO*. l*cpi_xx* and *lexr_xx* variables can be reasonably argued to be I(1) with the exception of logarithm of the exchange rate for Hungary and the consumer price index for Slovakia. We also provide these unit roots tests for all exogenous variables and one can see in the last row of the Table 1 that these time series are integrated I(1).

3.2. Specification and specification tests

In this part we describe the estimation and testing of the core VAR (SVAR) model for analysing the effects of national monetary policy in countries of the Visegrad Group V4. To derive the monetary policy shocks we estimate the VAR(4) model for each the Visegrad Group country:

$$Ay_t = A_1 y_{t-1} + \ldots + A_p y_{t-4} + B_0 x_t + \varepsilon_t, \tag{E3}$$

where y_t = *(outputgap_xx*$_t$*, lcpi_xx*$_t$*, ir3m_xx*$_t$*, lexr_xx*$_t$*)* and x_t = *(c, lgdp_emu*$_t$*, lcpi_emu*$_t$*, ir3m_eu*$_t$*)*. The variables are ordered in according to the assumption that the monetary authorities choose: the interest rate taking into account the current level of prices and ouput (see [10]). We choose for our VAR model *outputgapxx* variable instead the logarithm of GDP that was discussed in papers [1] and [5] for monthly data from 1998:1 to 2006:5.

We have applied Wald test and also LR test with three criteria AIC, SC and HQ to set a lag structure of our VAR models. The results are shown in Table 1. The Wald test conducted for each equation recomends lag of four quartals. LR test results according to AIC and HQ recommends the order 4 with the exception of Slovakia, while SC criterion suggests 1 lagged differences to include into our VAR model. We have choosen to go with four lagged differences i.e. it seems to be valid to use the VAR(4) model for each country. We also made this decision based on the residual autocorrelations. However this lag sturcture of our VAR(4) model influences the degrees of freedom in the estimated models.

Diagnostic cheking is also an important stage of our VAR or SVAR modelling. We provide particularly AR roots test for stacionary VAR processes, tests for residual serial correlation and also the multivariate extensions of the Jarque-Bera residual normality test.

There are many of the tests for model adequacy. We apply AR roots graphs and tables with AR roots and modulus for detecting stationary VAR processes (see Figure 1). The estimated VAR is stable (stationary) if all roots have modulus less than one and lie inside the unit circle. The Figure 1 reports that the 16 AR roots of the VAR(4) system are in the unit cirle for all countries exluding Slovakia (0.43-1.0007), where two roots are outside the unit cirle. For this reason we do not analyse further VAR model for Slovakia, but the important results are reported. The results of all 16 modulus for the Czech Republic are inside of the interval, in the concrete range of 0.14–0.98, for Hungary modulus are 0.59 – 0.94 and for Poland modulus are 0.44 – 0.89. If modulus of highest root is next to one there should be cointegrating relations between endogenous variables. To conclude this part of our analysis, the VAR(4) models for Hungary, Poland and the Czech Republic are stable because the highest modulus lies inside the unit circle and that means VAR is stationary. Thus, the results for CZ, HU and PO allow to continue estimating SVAR model or cointegrating SVAR.

Next, all models are estimated as structural VAR (SVAR) model using AB-model in [9], which is defined as follows in a reduced form:

$$y_t = \Phi(L) y_{t-p} + \Psi(L) x_t + u_t, \tag{E4}$$

$u_t = A^{-1} B e_t, \quad e_t \sim i.i.d(0, I_k)$, where I_k is identity matrix, k is the number of variables, A and B are *(kxk)* matrices to be estimated:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a_{21} & 1 & 0 & 0 \\ a_{31} & 0 & 1 & a_{34} \\ a_{41} & a_{42} & a_{43} & 1 \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 \\ 0 & 0 & b_{33} & 0 \\ 0 & 0 & 0 & b_{44} \end{bmatrix} \tag{E5}$$

A matrix A represent that the monetary authorities: do not consider contemporaneous prices while deciding on the monetary policy ($a_{32}=0$); are likely to react to contemporaneous output (a_{31}, i.e an excess demand pressure indicator) and also exchange rate shocks (a_{34}), which is a reasonable assumption for small open economies (see [5]).

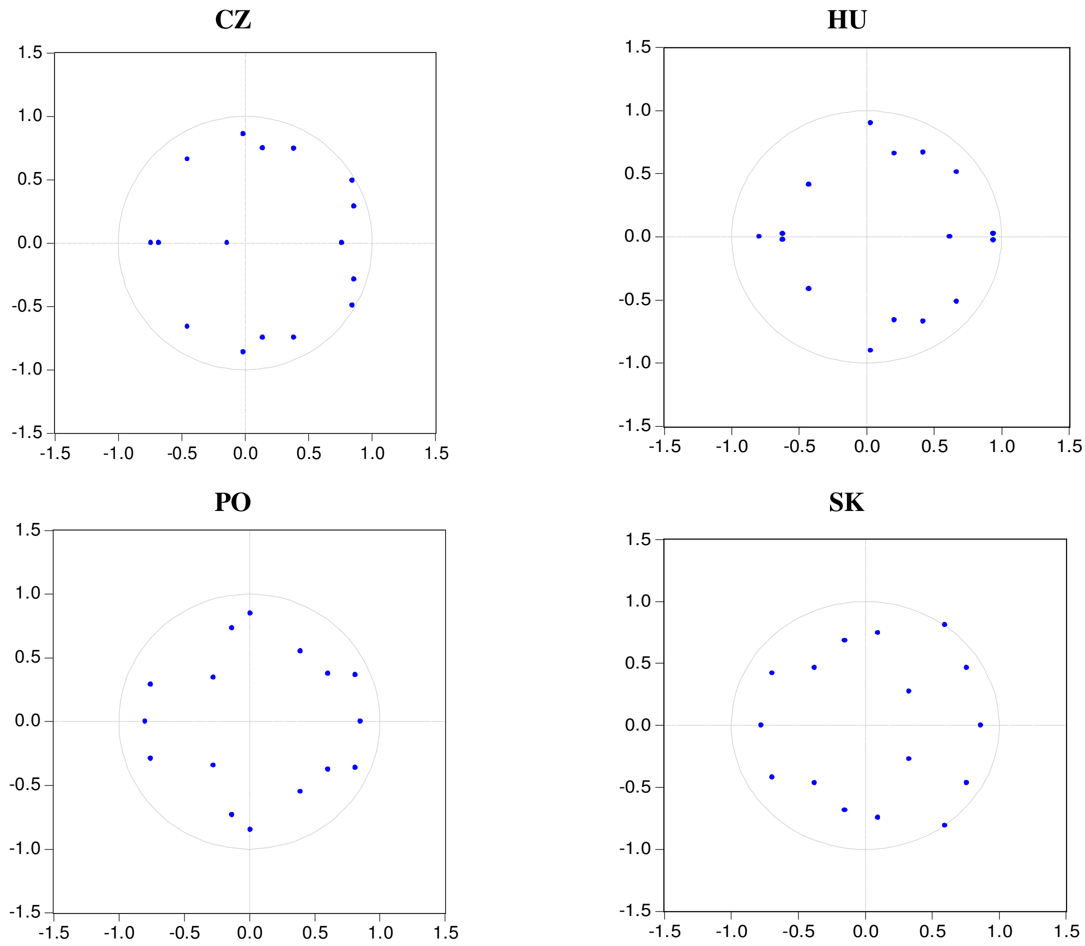


Figure 1: The inverse roots of the AR characteristic polynominal

model	AR roots	Autocorrelation tests		Normality J_B test	R^2_{adj}
		Correlograms	LM test (lag)		
CZ SVAR(4)	ok	ok	1., 3.	no no no no	0,888 0,994 0,993 0,984
SK SVAR(4)	**NO**	cpi-cpi(-2) ir3m-ir3m(-2)	ok	ok no no no	0,242 0,993 0,937 0,973
PO SVAR(4)	ok	ok	ok	no no no no	0,760 0,998 0,994 0,843
HU SVAR(4)	ok	ok	ok	no no no no	0,613 0,998 0,907 0,451

Table 2: Selected specification tests

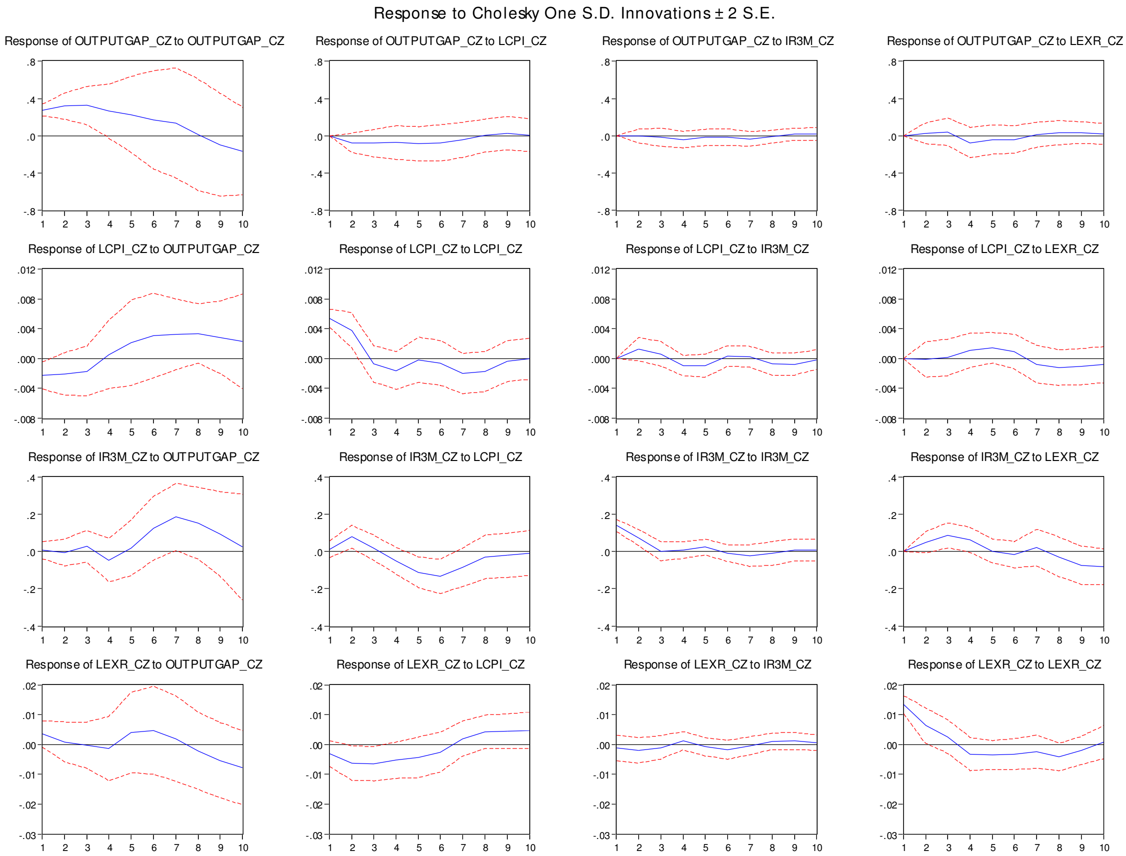
The next specification tests include presence of residual autocorrelation using the pairwise cross-correlograms for the estimated residuals in the SVAR(4) and also using autocorrelation LM test. It appears residual autororrelations is statistically significant fit order 1 and also order 3 for the Czech Republic (see Table 2). Next tests are for residual normality. Reports of the multivariate extensions of the Jarque-Bera residual normality test, which compares the third and fourth moments of the residuals to those from the normal distribution, are shown in the Table 2. We analyse VAR residual and we reject normality distribution for all residuals with the exception of 1st ouputgap-equation for the Slovak Republic. Non-normality of residuals is caused by issue relating to the kurtosis.

3.3. Identification of the shocks

Impulse responses are an important tool to discover relations between the variables in a SVAR and also tool for analysis of policy effectiveness. VARs can be reduced form models and strutural restriction are required to identify the relevant innovations and impulse responses. We deal with structural restriction relevant for VARs with integrated variables. Our general modelling strategy is to specify and estimate reduced form of the VARs model and then focus on the structural parameters and resulting structural impulse responses. We also consider both types of restrictions - the so-called AB model (see (E5)). In this case, a simultaneous equations system is formulated for the errors of the reduced form model rather than the observable variables directly. Thereby model accounts for shift from specifying direct relations for the observable variables to formulate relations for innovations. The impulse response functions of the SVAR(4) models with four endogenous variables for the Czech Republic, Hungary and Poland are shown in, respectively, Graphs 1 – 3.

From the graphs 1 – 3 one can conclude that effects of all the shocks gradually vanish although with differing intensity. This signals stability and stationarity of SVAR(4) models. However the effects of the shocks are limited as to their size and significance. The issue of the significance of responses to shocks is apparent also in other studies in this area.

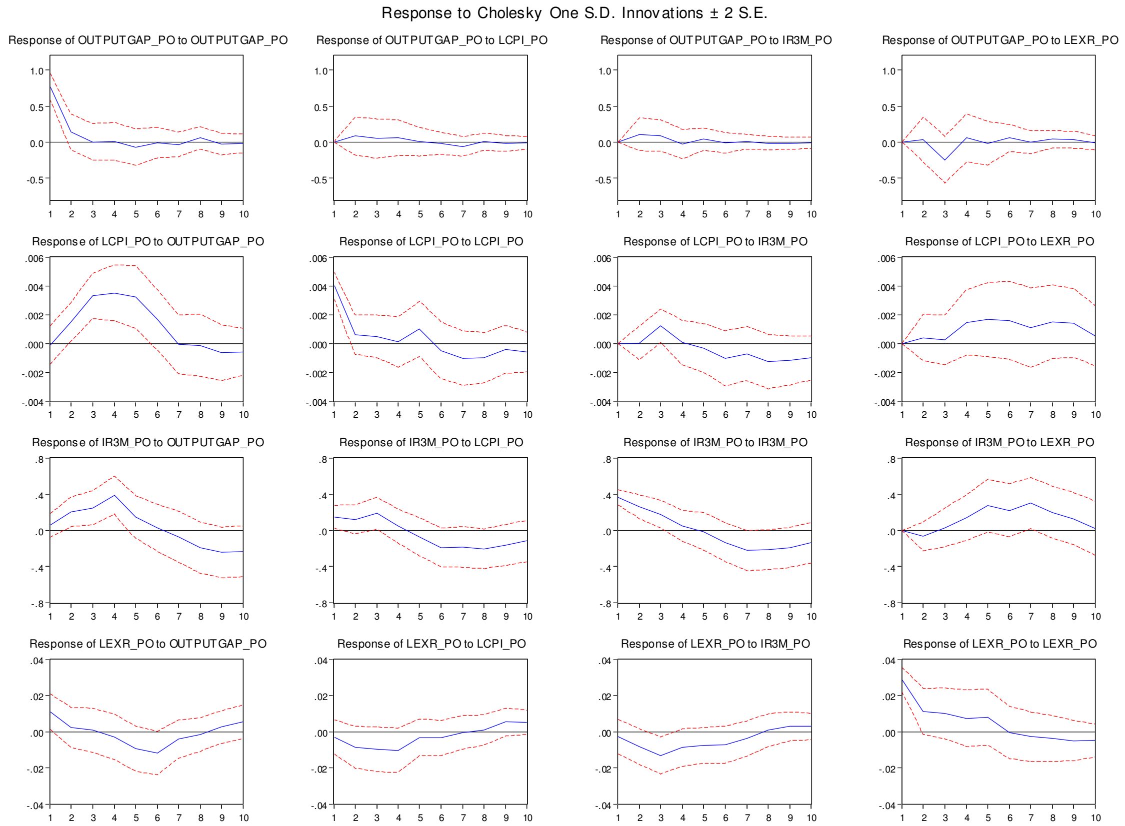
According to the graph 1 commenting the Czech Republic we can not confirm that tightening of monetary policy influences real economic activity and the exchange rate which should than pass through to inflation. The behaviour of exchange rate confirms the observed evidence that after increasing inflation markets expect increase of interest rates by the central bank and the exchange rate appreciates.



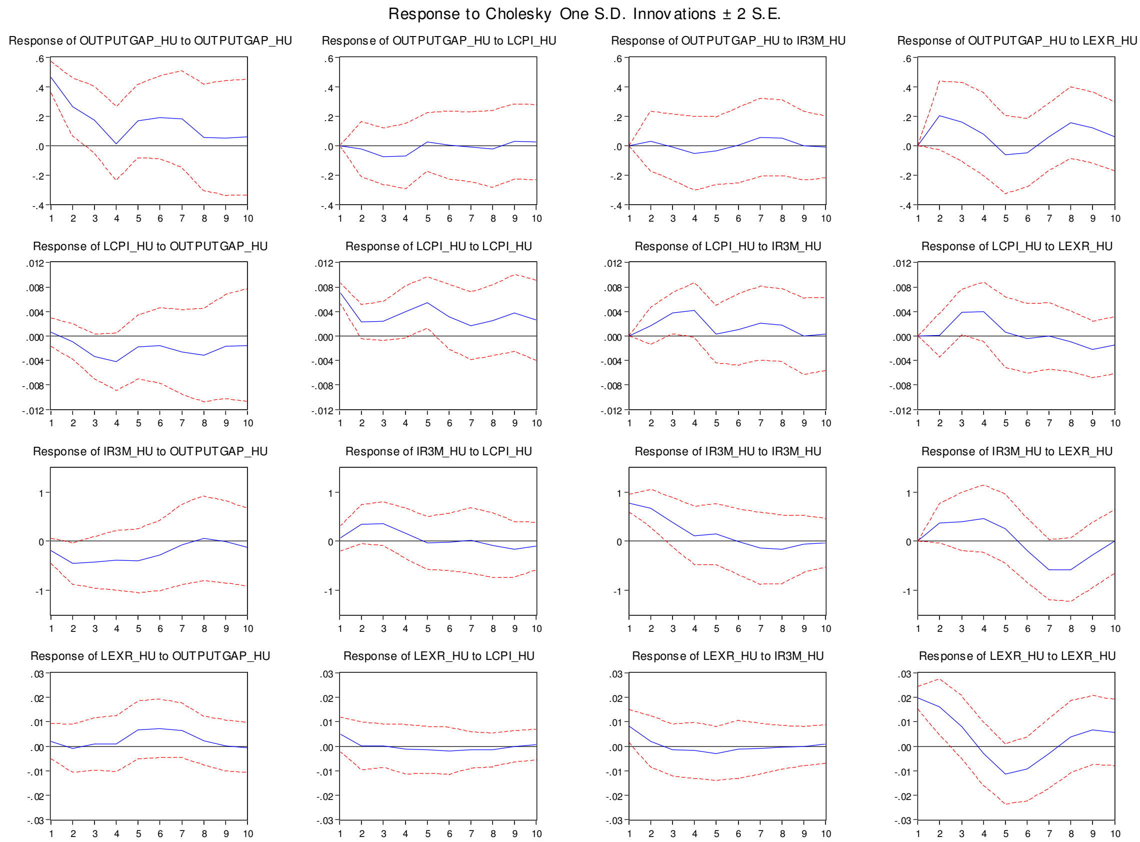
Graph 1: Estimated orthogonalized responses to error impulse for the SVAR model – CZ

The graph 2 confirms the existence of price puzzle in Poland which would indicate SVAR(4) mis-specification in terms of ommiting monetary policy relevant variable according to [12]. The behaviour of exchange rate response on interest rate shock is in line with so-called delayed overshooting [3]. The exchange rate appreciates 3Q after interest rate shock and some effect persists into to the second year after shock. Unfortunately the results do not confirm effect of interest rate shock on output gap. The reason would be due to data, especially ex-post revised output gap which could be not relevant as real-time output gap in the moment of monetary policy decision.

The graph 3 reports the similar results regarding the response of real economic activity and inflation to monetary policy in Hungary as in Poland. However we find different behaviour in the exchange rate. The exchange rate respond in a manner of immediate depreciation after monetary policy tightening and which is absorbed aftewards. But this effect is subdued.



Graph 2: Estimated orthogonalized responses to error impulse for the SVAR model – PO



Graph 3: Estimated orthogonalized responses to error impulse for the SVAR model – HU

4. Conlusions

In this paper we have investigated whether monetary policy impulses (interest rate shock) have effects on the output gap and consumer prices in four countries of the Visegrad Group V4. The empirical analysis covers the period 1999Q1 - 2009Q1 and we can sumarize obtained results into answers at the beginning of paper pre-defined two questions:

- Empirical results provide evidence that approximately half of all endogenous variables are nonstationary, integrated of the order one I(1) and the second half of variables is stationary. Stability tests of VAR (4) model recommend to investigate structural model SVAR(4) and further deal with the existence of cointegration relations according to the results of cointegration tests (2 cointegrations relations for the Czech Republic, 3 for Poland and 1 for Hungary). R^2_{adj} of SVAR (4) models ranges 0,451 – 0,998 without equations of Slovakia. Diagnostic tests have prooved in almost all cases non-normality of residuals, especially from the point of the kurtosis problem.
- Monetary policy tightening according to our results do not influence the output gap. We have found the price puzzle for Poland and Hungary. Further, the behaviour and responses of exchange rates on shocks differ across countries substantially.

Sugestions for further research

- To consider using real-time output gap instead of ex-post revised HDP (output gap).
- To consider including of the comodity price index as a proxy of supply side (with respect to degree of freedom issue).
- To investigate long run relations between variables, thus go for deep analysis of cointegration relations where relevant.

Acknowledgement

The research work on this paper was supported by the Czech Science Foundation, grant number 402/08/1015 - Macroeconomic Models of the Czech Economy and Economies of the other EU Countries.

References

[1] BORYS, M.M – HORVÁTH, R.- FRANTA, M. The Effects of Monetary Policy in the Czech Republic: an empirical study. *An Empirical Study*, 2009, forthcoming .

[2] CREEL, J. – LEVASSEUR, S.- Monetary Policy Transmission Mechanisms in the CEECs: How Important Are the Differences eith EURO Area?. WP OFCE no. 2/2005.

[3] EICHENBAUM, M. – EVANS, C: Some Empirical Evidence on the Effects of Monetary Policy Shocks on Exchange Rates, Quarterly Journal of Economics 110, 1995, 975-1010.

[4] GARRATT, A. – LEE, K. – PESARAN, M.H. – SHIN, Y. *Global and National Macroeconometric Modelling (A long.Run Strutural Approach).* Oxford: Oxford University Press, 1st Ed., 2006, 380 pp. ISBN 978-0-19-929685-9.

[5] HORVATH, R. Time-Varying Policy Neutral Rate in Real Time: A Predictor for Future Inflation?, *Czech National Bank Working Paper* Series No 4/2007.

[6] JUSELIUS, K. The *Cointegrated VAR Model: Methodology and Applications*. New York: Oxford University Press. 2007, ISBN 978-0-19-928566-2.

[7] LUKÁČIK, M. - SZOMOLÁNYI, K. - LUKÁČIKOVÁ, A. Analýza hospodárskych cyklov SR a EMÚ. In *AIESA - budovanie spoločnosti založenej na vedomostiach*: 12. medzinárodná vedecká konferencia. Bratislava: Iura Edition, 2008. ISBN 978-80-8078-233-7.

[8] LUKÁČIK, M. Štrukturálne VAR modely a Mundellov-Flemingov model. In *INFOSTAT,* Inštitút informatiky a štatistiky : Cevavstat. - Bratislava. Dostupné na internetu: http://www.infostat.sk/cevavstat/modelovanie/prispevky/Lukacik_prispevok.pdf.

[9] LÜTKEPOHL, H. *New Introduction to Multiple Time Series Analysis.* 2nd ed., Berlin: Springer, 2006. ISBN 3-540-40172-5.

[10]MOJON, B.- PEERSMAN, G. A VAR Description of the Effects of Monetary Policy in the Individual Countries of the Euro Area. *ECB Working Paper* No.92, December 2001.

[11] NIMARK, K.P. A Structural Model of Australia as a Small Open Economy. *Australian Economic Review.* 42(1): 24-41 MAR 2009.

[12] SIMS, C. Macroeconomics and Reality. *Econometrica*, 48, 1980, 1-48.

[13] IVANIČOVÁ, Z .- RUBLÍKOVÁ, E: Plnenie Maastrichtských kritérií v krajinách V4. (Za koľko rokov dosiahnu krajiny V4 priemernú životnú úroveň v krajinách EÚ-25?). *Ekonomické rozhľady,* Bratislava: EU v Bratislave. 2/2006, Volume XXX,. s. 137-146. ISSN 0323-262X.

EFFECTS OF SCHEDULED VERSUS UNEXPECTED NEWS IN INTRADAY PRICE MOVEMENTS: THE EVIDENCE FROM NEW EUROPEAN STOCK MARKETS[1]

Jan Hanousek, Evzen Kocenda

Charles University and the Academy of Sciences, CERGE-EI, Prague
jan.hanousek@cerge-ei.cz

Abstract. The goal of this paper is to study real time behavior on three emerging EU stock markets—in the Czech Republic, Hungary, and Poland—taking into account interactions with developed markets and the influence of macroeconomic news originating in the EU and in the U.S. We characterize the price discovery in these three emerging EU stock markets by employing high-frequency five-minute intraday data on stock market index returns and four classes of EU and U.S. macroeconomic announcements during 2004–2007. We account for the difference of each announcement from its market expectation and we jointly model the volatility of the returns accounting for intra-day movements and day-of-the-week effects. Our findings show that real-time interactions on the new EU markets are strongly determined by matured stock markets as well as the macroeconomic news originating thereby. In general, differences in results across markets are driven by differences in key market participants, mostly by the presence of foreign investors. Our findings yield insights into the process of the development of new European capital markets and stock market integration in the EU.
Keywords. Price discovery, stock markets, intra-day data, macroeconomic news, European Union, volatility, excess impact of news

1. Introduction, Motivation and Related Literature

Developments in emerging capital markets have long been characterized by less developed financial institutions and higher returns associated with higher volatility. While higher volatility is still present to a large extent [34], since the late 1990s the development of institutions in these markets attracted a large number of foreign institutional investors from developed countries who begun to consider emerging markets as a risk-sharing domain.[2] In addition, financial globalization brought emerging markets under the increasing influence of spillovers from developed markets, that themselves demonstrate degree of dependency [8], as well as impact of economic information flows. General consensus in the literature is that macroeconomic announcements have significant effects on financial markets, both in terms of asset returns and their volatility, but their impact, specifically on volatility, is generally ambiguous [5]. While the effect of macroeconomic news is quite well documented in developed markets, emerging counterparts are lagging behind and new markets in Europe are still under-researched. At the same time the presence of foreign institutional investors from developed markets in new European markets is heavy. [7] document that financial liberalizations that open emerging capital markets to foreign investors are associated with significant increases in real economic growth and [47] persuasively shows that competition and financial openness are beneficial for stock market development. Hence, knowing how new European markets process economic information is important from a development perspective.

We approach our analysis by using the data in an intra-day frequency in order to capture information flowing from developed markets and to illustrate its power on price formation in emerging markets. Modern research draws attention to the use of intraday data that are able to reveal the effect of macroeconomic announcements on stock market movements [11, 42, 32, 20, 48]. In our paper we contribute to the related literature in several ways. Most of the literature targets the developed capital markets in the U.S. and Europe, while European emerging markets are still under-researched. Therefore we investigate new EU members: the Czech Republic, Hungary, and Poland. Further, as an extension to the above literature, we use stock price data based on 5-minute intervals

[1] We would like to thank John Banko, John Brinkman, Dana Hájková, T.A. Chola, Petr Koblic, Magdalena Malinowska and participants at the FMA (Prague, 2008), Monetary and Financial Transformations in the CEECs (Paris, 2008), XVII International Tor Vergata Conference on Banking and Finance (Rome, 2008), and BESI (Acapulco, 2009) conferences for helpful comments. GAČR grant (402/08/1376) support is gratefully acknowledged. The usual disclaimer applies.

[2] Devereux and Sutherland (2009) note unprecedented improvements in the financial environment of the emerging markets during the last decade and analyze the determinants of an optimal risk-sharing portfolio for an emerging market economy and an advanced economy.

to provide more robust estimates of public information on stock returns in the new EU markets, which is not covered in the literature.[3]

2. Data and Methodology

We analyze the price discovery on the new EU stock markets and concentrate on the stock exchanges in Budapest, Prague, and Warsaw in particular. These markets are the largest European emerging markets in terms of market capitalization as well as the extent of liquidity [17].

We analyze the impact of macroeconomic announcements by employing an augmented version of the generalized autoregressive conditional heteroskedasticity (GARCH) model [10]. This approach allows us to assess the impact of news on stock returns and assess market volatility, as well as to account for the fact that errors from the mean equation are heteroskedastic. We deviate from the standard sequencing and introduce our data prior to describing the model since a description of the news announcements is needed to better describe our model.[4]

2.1. Data Set: Stocks and News

We constructed our dataset from intraday data on three emerging EU markets recorded by Bloomberg. Stock exchange index quotes ($I_{i,t}$) for market i are available in five-minute intervals at time t (ticks) for the stock markets in Budapest (BUX), Prague (PX-50), and Warsaw (WIG-20). In addition to these markets we also employ data from the Frankfurt stock exchange (the German DAX index) and the U.S. Dow Jones Industrial Average of 30 stocks index. Based on these quotes we construct a five-minute stock market index return $R_{i,t}$ ($R_{i,t} = \ln(I_{i,t} / I_{i,t-1})$) for each market i from time t-1 to time t. The time period of our data starts on 1 June 2004 at 9:00 and ends on 30 December 2007 at 16:30 Central European Daylight Time (CEDT). The beginning of our sample intentionally starts after the entry of the four countries to the European Union in May 2004. After accounting for weekends and public holidays, the time span gives the following numbers of trading days for each of the three new EU markets: 878 (Budapest), 880 (Prague), and 879 (Warsaw).

The Budapest index BUX consists of 16 titles, with four forming the bulk of the index (91.5%). The Prague index (PX-50) consists of 13 titles and 82.7% of it is formed by four titles. The Warsaw index WIG-20 contains 20 titles and five titles form a majority (64.0%). None of the companies that are included in the three indices are exposed to foreign economic conditions in a different way in terms of reporting activities as they are all obliged to report under international accounting standards. The energy, banking and telecom industries dominate all three indices and specifically the banking industry is represented in similar proportions in each of the three markets. The index composition is then to a large extent representative of each country's economy without any strong concentration in a specific industry. If there is any bias towards banking, the index composition hints that at least it is consistent across the three countries. In the same spirit all three countries exhibit a similarly consistent trading pattern with respect to the U.S. and the old EU-15.

Further, we compiled an extensive data set on 15 different macroeconomic announcements (news) that are divided into four categories. These are announcements on prices, real economy (GDP, current account, production, sales, trade balance, unemployment, etc.), monetary policy (monetary aggregate and interest rate), and economic confidence (consumer and industry confidence, business climate, etc.).

The macroeconomic announcements we employ are surveyed by Bloomberg and Reuters with a clearly defined calendar and timing of the news releases; as publication schedules of the releases is publicly available we do not report it for the sake of space.

We analyze the effect of the news from its *excess impact* perspective. Because markets form expectations about scheduled important news, it is not the news itself that matters but its difference from what the market expects it to be (market consensus). The news deviation, or its excess, has then an impending impact on stock prices. Following this logic, we construct a data set of announcements. There is news associated with indicator i in the form of various macroeconomic releases or announcements that are known ahead of time to materialize on specific dates t.[5] The extent of such news is not known but expectations on the market form a forecast. The excess impact news announcement is then defined as a deviation of the news from the market expectation formed earlier. Further, announcements are often reported in different units and therefore they are standardized to allow their meaningful comparison (see e.g. [4]). Formally, the excess impact news variable is labeled as xn_{it} and defined as $(sn_{it} - \mathrm{E}_{t-1}[sn_{it}]) / \sigma_i$, where sn_{it} stands for the value or extent of the scheduled announcement i at

[3] Other literature deals with emerging markets in Europe but on a lower frequency and without the specific effect of macroeconomic announcements (see e.g. [37, 49] among others).

[4] The theoretical framework linking macro announcements to stock returns is underdeveloped. We refer readers to the account of bond pricing with announcement effects of [44], the equities modeling framework with announcements' effect of [38], and an equilibrium asset pricing model with public announcements [15].

[5] There is also news in the form of an unexpected announcement that can be understood as a truly exogenous shock or surprise. The number of such news that is recorded is negligible and we do not consider them in the present study.

time t and $\mathrm{E}_{t-1}[sn_{it}]$ is the value of the announcement for time t expected by the market at time t-1, and σ_i is the sample standard deviation of the announcement i. The standardization does not affect the properties of the coefficients' estimates as the sample standard deviation σ_i is constant for any announcement indicator i.

From a practical perspective, we consider the immediate effect of each new announcement at the time of its release and account for its impact for 5 minutes as extension of the interval does not yield an improvement because the impact of the scheduled announcements dissipates very quickly.

2.2. Estimation Methodology

We employ the augmented generalized autoregressive conditional heteroskedasticity (GARCH) model [10] to empirically test for the effect of macroeconomic announcements on stocks and to assess stock market volatility. We augment the mean specification by parameters to account for the effect of macroeconomic news in the form of deviations of scheduled releases from market expectations and the effects of spillovers from neighboring emerging markets as well as two major developed markets (Germany and the U.S.). The volatility equation is augmented by a set of dummy variables to capture intraday and daily effects. Thus, our model effectively captures the effect of news and market spillovers on stock returns and the effect of trading patterns on stock volatility. The baseline model is specified in the following form:

$$R_{i,t}^{E} = \sum_{y=2004}^{2007} \lambda_y + \sum_{k\in\{EU,US\}} \sum_{j=1}^{p} \pi_k R_{k,t-j}^{M} + \sum_{i=1}^{2}\sum_{j=1}^{q} \gamma_i R_{i,t-j}^{E} + \sum_{j=1}^{n}\sum_{l=1}^{3} \delta_{l,j} xn_{EU}^{j} + \sum_{j=1}^{n}\sum_{l=1}^{3} \kappa_{l,j} xn_{US}^{j} + \varepsilon_t \quad (1)$$

$$h_{i,t} = \omega + \sum\nolimits_{m=1}^{r} \alpha_m \varepsilon_{t-m}^{2} + \sum\nolimits_{m=1}^{s} \beta_m h_{i,t-m} + \sum_{\tau\in T} \mu_\tau D_\tau + \sum\nolimits_{d=1}^{4} \psi_d W_d \ . \quad (2)$$

The variables in the mean equation (1) are coded as follows. Our dependent variable $R_{i,t}^{E}$ is the return on a specific emerging (E) market stock index i (Budapest, Prague, Warsaw) at time t. The parameter $R_{k,t-j}^{M}$ is the lagged return on a specific mature and developed (M) stock market index in the European Union (EU) and the United States (US). As a proxy for the Eurozone we employ the German DAX index from the Frankfurt stock exchange and for the U.S.A. we employ the Dow Jones Industrial Average of 30 stocks index.[6] Coefficients π_k capture the effects of market spillovers from the two developed markets. The parameter $R_{i,t-j}^{E}$ is the lagged return on a specific emerging market stock index other than that employed as a dependent variable and coefficients γ_i capture the effects of spillovers from emerging markets (e.g., in the case of the Prague index being the dependent variable, lagged indices from Budapest and Warsaw are right-hand side variables). Coefficients λ represent a set of year-specific dummy variables that provide information on stock index returns in a specific year during the period 2004–2007.

3. Results and concluding remarks

Despite varied effects inferred from estimates we can draw some generalizations specific to all three countries. The effects of other stock markets are dominated by spillovers from Frankfurt stock exchange while reaction to the New York market is smaller. The findings are sensible given the ongoing process of European integration that also affects financial markets and the narrow time window during which trading at the U.S. and European markets overlap. Spillovers from the neighboring markets are smaller or comparable in cumulative magnitudes to the effect of New York. Among them the Budapest stock market produces the strongest spillover effects, followed by Prague, and the smallest effect is from Warsaw.

The general finding is that the three markets are not efficient in the sense of efficient market theory because numerous significant coefficients associated with the impact of the news testify that news is not absorbed by the market immediately and reflected instantaneously in prices. Announcements originating in the Eurozone exhibit more effects than U.S. news. In terms of specific news, EU current account, consumer confidence and PMI affect all three markets while U.S. prices are the only news of the same reach. The volatility of the returns is accounted for at the beginning and end of the trading session and it declines dramatically during the rest of the day. The differences in the extent of volatility at the three markets should be credited to differences in trading hours on these markets.

The effects of macroeconomic announcements need more detail to summarize. Among the four classes of macroeconomic announcements, monetary news has virtually no impact on stock returns. The reason might rest in the relative detachment of monetary policy figures from stock market developments. [48] claim that the "detachment" of monetary policy expectations and asset prices from incoming economic news is partly related to

[6] Germany is the most important trading partner for the three new EU countries under research. Using a composite Stoxx 50 or EuroStoxx 50 index is not feasible as these are not available historically at the desired intra-day frequencies.

the difficulties associated with measuring the surprise component of that news. Since we account for the surprise component our findings show that the detachment might be due to the low value stock markets place on monetary announcements.

Prices on the other hand affect all three markets, mostly in a very intuitive manner: worse (better) than expected results bring negative (positive) effects on stock returns. This result upholds the market's ability to effectively incorporate inflation into stock prices. The interesting trait in the price effect findings is the dominating influence of U.S. prices while the Eurozone announcements pass nearly unnoticed. The possible and sensible explanation might be credited to the well-mapped expectations of the European Central Bank's operations that in the integrating Europe pose little challenge to financial market assessment.

The real economy class of announcements offers varied results from which the news on the EU current account stand out as it affects all three markets in the same manner without exception: better-than-expected results prompt a positive reaction and worse-than-expected results prompt a negative one. This finding should be paired with the heavy dependency of the three economies on foreign trade with other EU countries, the presence of EU firms in these markets, the similarity of supply and demand shocks [22], and a relatively high degree of business cycle correlation [23] between the old and new EU members. Needless to say that the most important companies present in the new EU economies as owners or co-owners of the major local firms and banks are also often quoted on the local stock markets. Other real economy announcements are limited in their reach to one or two markets. Industrial production influences Prague and Warsaw, while announcements on trade balance and unemployment are echoed in Prague and Budapest. Announcements on factory orders and retail sales do not provoke any market reactions. Real economy announcements originating in the U.S. bring only scarce evidence of their effects on stock returns. Many announcements are simply not available during the Europe-U.S.A. trading window. The Prague stock market is not affected by U.S. news at all while Budapest and Warsaw are only sparingly.

Finally, business climate and confidence announcements provide valuable insights to the previous categories. Practically no effect of the U.S. survey announcements has been found in any of the three markets and the effect of those originating in the Eurozone is limited. Only the news on consumer confidence and the Purchasing Manager's Index (PMI) impact all three markets in an intuitive manner common to developed markets: a worse-than-expected outcome provokes a negative effect on stock returns and better-than-expected results prompt a positive one. All the above results thus validate the *excess impact* approach that highly reduces difficulties in measuring "news" correctly.

News affects the volatility of the stock return indices in a similar manner but specific features vary across the three markets. The volatility of the Prague index is affected by the past announcements most but in no market is the effect destabilizing. The Budapest index exhibits the highest persistence of volatility. The volatility of the Warsaw index shows the slowest convergence to the steady state. In terms of the intra-day features the Budapest market exhibits the highest volatility at the beginning and end of the trading sessions while Prague records the lowest volatility during the two periods. Volatility declines dramatically on the three markets during the rest of the trading day and its extent is comparable across the markets. All three markets also show a decrease in volatility by the middle of the business week.

Our findings show that real-time interactions on the new EU markets are strongly determined by matured stock markets as well as the macroeconomic news originating thereby. The differences in results across the three markets seem to be driven by differences in the composition and origin of the key market participants. The discovered detailed effects are complemented by a characterization of market volatility. Our findings yield insights into the process of the development of new European capital markets and stock market integration in the EU. As these new markets become more mature and globally integrated they become more important for the real economy of specific countries. Information flows related to developed economies, their ups and downs, were shown to affect emerging European capital markets, whose developments then affect the real economy. This is because smaller emerging markets are quite sensitive to changes in the economic situation in developed markets as well as to changes in perceptions that are conveyed via macroeconomic news. As the recent financial problems unfold globally, we can expect increases in sensitivity to the information flows reaching emerging markets and the need to understand them better.

References

[1] ALBUQUERQUE, R., VEGA, C. *Asymmetric Information in the Stock Market: Economic News and Co-movement.* CEPR Discussion Paper 5598, 2006.

[2] ALEGRIA, C., MCKENZIE, G., WOLFE, S. *Earnings announcements by UK companies: Evidence of extreme events*? 15(2), European Journal of Finance, 2009, 137-156.

[3] ANDERSEN, T.G., BOLLERSLEV, T. *Deutsche mar-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies*. Journal of Finance 53, 1998, 219-65.

[4] ANDERSEN, T.G., BOLLERSLEV, T., DIEBOLD, F.X., VEGA, C. *Real-time price discovery in global stock, bond and foreign exchange markets*. Journal of International Economics 73, 2007, 251-277.

[5] ANDRITZKY, J.R., BANNISTER,G.J., TAMIRISA, N.T. *The Impact of Macroeconomic Announcements on Emerging Market Bonds*. Emerging Markets Review, 8(1), 2007, 20-37.

[6] BANKO, J., FLANNERY, M. *The Relation between Trading Time and Volatility: What is a "Day"?* Working Paper, Department of Finance, Graduate School of Business Administration, University of Florida, 2008.

[7] BEKAERT, G, HARVEY, C.R., LUNDBLAD, C. *Emerging equity markets and economic development*. Journal of Development Economics, 66(2), 2001, 465–504.

[8] BEINE, M., CAPELLE-BLANCARD, G., RAYMOND, H. *International nonlinear causality between stock markets*. European Journal of Finance, 14(8), 2008, 663-686.

[9] BERNDT, E. K., HALL, B. H., HALL, R. E., HAUSMAN, J.A. *Estimation of inference in nonlinear structural models*. Annals of Economic and Social Measurement 4, 1974, 653-665.

[10] BOLLERSLEV, T. *Generalized autoregressive conditional heteroscedasticity*. Journal of Econometrics 31, 1986, 307-327.

[11] BOLLERSLEV, A. T., CAI, J. *Intraday and Interday Volatility in the Japanese Stock Market*. Journal of International Financial Markets, Institutions, and Money 10, 2000, 107–130.

[12] BOMFIM, A. N. *Pre-announcement effects, News Effects, and Volatility: Monetary Policy and the Stock Market*. Journal of Banking and Finance 27, 2001, 133-151.

[13] BOYD, J. H., HU, J., JAGANNATHAN, R. *The Stock Market's Reaction to Unemployment News: Why Bad News Is Usually Good for Stocks*. Journal of Finance, 60(2), 2005.

[14] BREDIN, D., HYDE, S., O'REILLY, G. *European Monetary Policy Surprises: The Aggregate and Sectoral Stock Market*. Economic Analysis and Research Department Working Paper 10/RT/2005, Central Bank and Financial Services Authority of Ireland, 2005.

[15] CENESIZOGLU, T. *Risk and Return Reaction of the Stock Market to Public Announcements about Fundamentals: Theory and Evidence*. HEC Montreal, Department of Finance Working Paper, 2007.

[16] CHANG, E., PINEGAR, J., RAVICHANDRAN, R. *International Evidence on the Robustness of the Day-of-the-Week Effect*. Journal of Financial and Quantitative Analysis 12(1), 1993, 497-513.

[17] ÉGERT, B., KOČENDA, E. *Interdependence between Eastern and Western European Stock Markets: Evidence from Intraday Data*. Economic Systems, 31(2), 2007, 184-203.

[18] EHRMANN, M., FRATZSCHER, M. *Taking Stock: Monetary Policy Transmission to Equity Markets*. Journal of Money, Credit and Banking 36(4), 2004, 719-37.

[19] EHRMANN, M., FRATZSCHER M. *Global Financial Transmission of Monetary Policy Shocks*, European Central Bank Working Paper No. 616, 2006.

[20] ERENBURG, G., KUROV A., LASSER D. J. *Trading Around Macroeconomic Announcements: Are All Traders Created Equal*?, Journal of Financial Intermediation 15, 2005, 470-493.

[21] FAN, Y-J., LAI, H-N. *The Intraday Effect and the Extension of Trading Hours for Taiwanese Securities*. International Review of Financial Analysis 15, 2006, 328-347.

[22] FIDRMUC, J., KORHONEN, I. *Similarity of Supply and Demand Shocks between the Euro Area and the CEECs*. Economic Systems 27(3), 2003, 313-334.

[23] FIDRMUC, J., KORHONEN, I. *Meta-Analysis of the Business Cycle Correlation between the Euro Area and the CEECs*. Journal of Comparative Economics 34(3), 2006, 518-537.

[24] FLANNERY, M., PROTOPAPADAKIS A. *Macroeconomic Factors Do Influence Aggregate Stock Returns*. Review of Financial Studies 15, 2002, 751-782.

[25] FRENCH, K.R., ROLL, R. *Stock return variances: the arrival of information and the reaction of traders*. Journal of Financial Economics 17, 1986, 5-26.

[26] FRIJNS, B., MARGARITIS, D. *Forecasting daily volatility with intraday data*. European Journal of Finance, 14(6), 2008, 523-540.

[27] GREENE, J. AND WATTS, S. *Price discovery on the NYSE and the NASDAQ: The case of overnight and daytime news releases*. Financial Management, 25(1), 1996, 19-42.

[28] GURKAYNAK, R., SACK, B., SWANSON E. *Do Actions Speak Louder than Words?: The Response of Asset Prices to Monetary Policy Actions and Statements*. Finance and Economics Discussion Series 2004-66, Federal Reserve Board, Washington, D.C., 2004.

[29] HE, L. *Variations in Effects of Monetary Policy on Stock Market Returns in the past Four Decades*. Review of Financial Economics 15, 2006, 331-349.

[30] JENSEN, G., JOHNSON, R. *Discount Rate Changes and Security Returns in the US, 1962-1991*. Journal of Banking and Finance 19, 1995, 79-95.

[31] JENSEN, G., MERCER, J., JOHNSON, R. *Business Conditions, Monetary Policy, and Expected Security Returns.* Journal of Financial Economics 40, 1996, 213-237.

[32] JONES, B., LIN, CHIENG-TING, L., MASIH M. M. *Macroeconomic Announcements, Volatility, and Interrelationships: An Examination of the UK Interest Rate and Equity Markets.* International Review of Financial Analysis 14, 2005, 356-375.

[33] KAMINSKY, G., SCHMUKLER, S. *What triggers market jitters?: A chronicle of the Asian crisis*. Journal of International Money and Finance, 18(4), 1999, 537-560.

[34] KAMINSKY, G., REINHART, C. *Financial markets in times of stress.* Journal of Development Economics, 69(2), 2002, 451– 470.

[35] SUK-JOONG, K., MCKENZIE, M., FAFF, R. *Macroeconomic news announcements and the role of expectations: evidence for US bond, stock and foreign exchange markets*. Journal of Multinational Financial Management, 14(3), 2004, 217-232.

[36] KIYMAZA, H., BERUMENT, H. *The day of the week effect on stock market volatility and volume: International evidence*. Review of Financial Economics 12, 2003, 363–380.

[37] KORCZAK, P, BOHL, M. *Empirical Evidence on Cross-listed Stocks of Central and Eastern European Companies.* Emerging Markets Review 6(2), 2005, 121-137.

[38] MAMAYSKY, H. A Model For Pricing Stocks and Bonds. Yale ICF Working Paper No. 02-10, 2002.

[39] MANN, T., ATRA, R., DOWEN, R. *U.S. Monetary Policy Indicators and International Stock Returns.* International Review of Financial Analysis 13, 2004, 543-558.

[40] MCMILLAN, D., SPEIGHT, A. *Intra-day Periodicity and Long-Run Volatility.* Econometrica 59, 2002, 347–370.

[41] NELSON, D. *Conditional heteroskedasticity in asset returns: A new approach. Macroeconomic News and Stock Valuation in Europe.* Journal of Multinational Financial Management 14, 1991, 201-215.

[42] NIKKINEN, J., OMRAN, M., SAHLSTRÖM, M., ÄIJÖ, A. *Global stock market reactions to scheduled U.S. macroeconomic news announcements.* Global Finance Journal 17 (1), 2006, 92-104.

[43] NIKKINEN, J., SAHLSTRÖM, P. *Scheduled domestic and U.S. macroeconomic news and stock valuation in Europe*. Journal of Multinational Financial Management 14, 2004, 201–215.

[44] PIAZZESI, M. *An Econometric Model of the Yield Curve with Macroeconomic Jump Effects*. NBER Working Paper No. W8246, 2001.

[45] PATELIS, A. *Stock Return Predictability and the Role of Monetary Policy.* Journal of Finance 52(5), 1997, 1951-1972.

[46] RAMCHANDER, S., SIMPSON, M., THIEWES H. *The Effects of Macroeconomic News on German Closed-End Funds.* Quarterly Review of Economics and Finance, In Press, 2006.

[47] RAMOS, S. *Competition and stock market development*. European Journal of Finance, 15(2), 2009, 231-247.

[48] RIGOBON, R., SACK, B. *Noisy Macroeconomic Announcements, Monetary Policy, and Asset Prices*, NBER Working Paper No. W12420, 2006.

[49] SERWA, D., BOHL, M. *Financial Contagion Vulnerability and Resistance: A Comparison of European Stock Markets*. Economic Systems 29, 2005, 344-362.

[50] SIKLOS, P., ANUSIEWICZ, J. *The Effect of Canadian and U.S. M1 Announcements on Canadian Financial Markets: The Crow Years.* Journal of Economics and Business 50, 1998, 49-65.

[51] THORBECKE, W. *On Stock Market Returns and Monetary Policy*, Journal of Finance 52, 1997, 635-654.

[52] TONCHEV, D., TAE-HWAN, K. *Calendar effects in Eastern European financial markets: Evidence from the Czech Republic, Slovakia and Slovenia.* Applied Financial Economics, 14, 2004, 1035–1043.

[53] WANG, J. *Foreign equity trading and emerging market volatility: Evidence from Indonesia and Thailand.* Journal of Development Economics 84(2), 2007, 798–811.

[54] WONGSWAN, J. *Transmission of Information across International Equity Markets.* Review of Financial Studies 19(4), 2006, 1157-1189.

APPLICATION OF ADVANCED APPROACHES FOR CALCULATION OF OPERATIONAL RISK WITHIN FINANCIAL INSTITUTIONS

Jiří Havlický

VŠB – TU Ostrava
Faculty of Economics, Department of Finance
jiri.havlicky@seznam.cz

Abstract. The aim of this paper is to describe and apply advanced approaches for calculation of operational risk in line with the Loss Distribution Approach. At first, classical statistical distributions for modeling the number of losses and their severities are described together with special techniques how to aggregate these distributions together in order to get the probability distribution of the total loss. Then an extreme value theory is presented for cases when empirical loss data have a heavier right tail in comparison with standard probability distributions. In the application part both concepts (classical and extreme value theory) of the Loss distribution approach are applied on operational risk loss databases of Czech financial institution. Obtained results have shown the importance of the application of extreme value theory in the field of operational risk modeling.
Keywords. Operational risk, Loss distribution approach, Extreme value theory

1. Introduction

This paper deals with the calculation of the operational risk according to an advanced method under the Basel 2 called Loss distribution approach ("LDA"). The aim of the paper is to describe the standard LDA and its extended version with the Extreme value theory and apply both concepts on an operational risk events database of a financial institution.

First part of the paper contains a short presentation of the main idea of LDA followed by a description of its standard version which is afterwards extended through an application of Extreme value theory (Peak over threshold). Second part deals with a description of the application of both concepts on an operational risk loss database of a financial institution. Obtained results of both concepts are compared. Conclusion summarizes important findings for practical implementation of LDA within financial institutions.

2. Loss Distribution Approach ("LDA")

Following subchapters contain a short presentation of the main idea of LDA followed by a description of its standard version which is afterwards extended through an application of Extreme value theory (Peak over threshold).

2.1. Main idea and goal of the LDA

LDA belongs to a statistical bottom-up approach for the calculation of the operational risk loss and is based on the same idea as the Collective model commonly used in the insurance modeling (see for example Daykin, 1996). The idea is that the total operational risk loss is a simple sum of individual operational risk losses whereas the number of losses and the severity of each individual loss are random variables.

The goal of LDA is to get the probability distribution of the total loss for a given time period and calculate risk measures like expected loss, operational Value-at-risk and unexpected loss (see
figure 1). The operational Value-at-risk (VAR) corresponds to the value of the total loss for a given significance level p according to the following formula

$$VAR_p = F_L^{-1}(1-p), \tag{1}$$

where F_L^{-1} denotes the inverse function of the distribution function of the total loss. Expected loss (EL) is given by the mean of the total loss and unexpected loss (UL) can be expressed as

$$UL = VAR_p - EL. \tag{2}$$

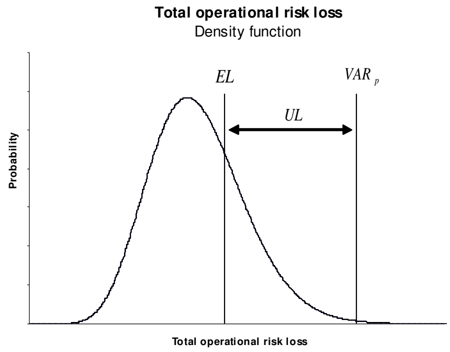


Figure 1 Example of the probability distribution of the total operational risk loss

2.2. Standard LDA

The total loss L of a given business line i and event type j for a given time period is given by the following formula

$$L(i,j) = \sum_{n=1}^{N(i,j)} S_n(i,j), \tag{3}$$

where $N(i,j)$ represents the number of risk events and $S(i,j)$ stands for severity (financial impact) of a single event. In case that we assume the independence between the number of losses and their severities and between each individual severities as well then the distribution function of the total loss F_L corresponds to the following formula (for more convenience indices i and j are excluded)

$$F_L(x) = \begin{cases} p(0) + \sum_{n=1}^{\infty} p_N(n) F_S^{n*}(x) & \text{for } x > 0 \\ p_N(0) & x = 0, \end{cases} \tag{4}$$

where p_N represents the density function of $N(i,j)$, F_S stands for the distribution function of $S(i,j)$ and $F_S^{n^*}$ means n-fold convolution of F_S with itself. The amount of the capital (TC) that a financial institution should have in order to cover the operational risk is usually given by

$$TC = \sum_i \sum_j (EL + UL)_{i,j} \tag{5}$$

for a significance level of 0,1% and a time period of one year.

The application of LDA consists of operational risk loss data collecting, building a model for the number of losses (frequency model), building a model for the financial impact of a single loss event (severity model) and aggregation of the respective frequency and severity models (details about each step of LDA application can be found for example in Panjer, 2006 or Cruz, 2002).

Operational risk loss data colleting

The goal of the operational risk events data collecting is to build up and maintain a database about the operational risk events of the whole financial institution at every level of business units. The data structure differs around institutions and is based on its modeling and managing needs.

Frequency and severity model

Building a model for the number of losses and a model for the financial impact of a single loss consist of choosing an arbitrary probability distribution, estimating its parameters, performing goodness of fit tests and repeating this algorithm until the appropriate distribution is found. For frequency model is usually used Poisson or Binomial distribution. In case of severity models mostly lognormal, gamma or Weibull distribution is applied, i.e., we generally assume a particular representative of broad Lévy models family, (for details of large set of probability distribution see Panjer, 2006).

Aggregation of frequency and severity model

The aggregation of the model for the number of losses and the model for the severity of a single event is usually done via Monte Carlo simulation or Panjer recursion method. Unfortunately there is no analytical solution for the distribution function of the total loss (4) because of problematic calculations of convolutions.

Monte Carlo simulation (see e.g. Tichý, 2008) is based on generation of a large number of total loss scenarios. At first a discrete random variable is generated which represents the number of losses for a given time period. Afterwards a respective number of trials of a continuous random variable is generated. These trials represent the financial impact of every individual loss and their sum is one scenario of the total loss. This algorithm is repeated many times in order to get the probability distribution of the total loss.

Panjer recursion method is the second most frequently used method for the aggregation of frequency and severity model and represents an analytical approximation of the total loss F_L. If the chosen probability distribution for modeling the number of losses p_N is a member of the group *(a,b,0)* then it satisfies the following equation

$$p_N(n) = \left(a + \frac{b}{n}\right) p_N(n-1) \text{ for } n > 1 \tag{6}$$

and then the density function of the total loss f_L can be expressed with the following formula

$$f_L(i) = \frac{1}{(1 - a \times f_s(0))} \left(\sum_{k=1}^{i} \left(a + \frac{b \times k}{i} \right) f_s(k) f_L(i-k) \right) \tag{7}$$

with the initial condition

$$f_L(0) = p_N(0) + \sum_{n=1}^{\infty} p_N(n) f_s^{*n}(0), \tag{8}$$

where f stands for the density function of the chosen probability distribution for the severity model. The coefficients a and b depends on the chosen probability distribution of p_N and can be found for example in Panjer, 2006. The distribution function of the total loss can be then expressed with the following formula

$$F_L = F(i \cdot C) = \sum_{j=0}^{i} f_L(j) \tag{9}$$

where C represents a given monetary unit of the financial impact of a single loss (the application of the Panjer recursion formula requires a discretization of the continuous severity model - for details of mid point discretization method see Daykin, 1996).

2.3. LDA with Extreme value theory

The empirical operational risk loss data for some event types have a right heavy tail of the probability distribution. Because of this fact, it is useful to implement Extreme value theory (especially the version called Peak over threshold - POT) in LDA framework.

According to respective theorem that was stated by Pickans (1975), Balkema and de Haan (1974), the conditional probability distribution of a single loss above a high threshold can be very well approximated by generalized Pareto distribution (for details of the Peak over Threshold approach to EVT see for example Panjer, 2006).

The application of LDA together with the Peak over threshold methodology consists of same steps as the application of standard LDA. But this time the building of a model for the severity of a single event is a little bit more complicated and consists of the following steps:

1) determination of threshold,
2) building a model for the severity of a single loss that don't exceed the defined threshold,
3) building a model for the severity of a single loss that is higher than defined threshold.

There are several techniques to determine an appropriate threshold, for example a graphic representation of a mean excess loss function (round about the selected threshold should be linear with the positive slope), an expert opinion or a given percentile. Determination of the threshold is a trade-off between its sufficiently high value and enough data for the parameter estimation of the generalized Pareto distribution.

Steps for building a model for the severity of a single loss that don't exceed the defined threshold are the same as in the case of standard LDA. Model for losses above the defined threshold consists of the utilization of generalized Pareto distribution according to the above defined theorem. For details of application of Extreme value theory see for example Panjer, 2006 or Cruz, 2002.

3. Application of LDA on operational risk loss database

Following subchapters contain a description of an application of LDA framework within a financial institution. At first standard LDA according to the subchapter 1.2 is applied. Then an application of LDA with the Peak over threshold methodology is presented. Afterwards both concepts and their result are compared.

3.1. Operational risk loss database

The calculation of the operational risk was done with the use of operational risk loss database containing 186 cases of external frauds from retail banking. The figure 2 shows the empirical probability function obtained from the mentioned database.

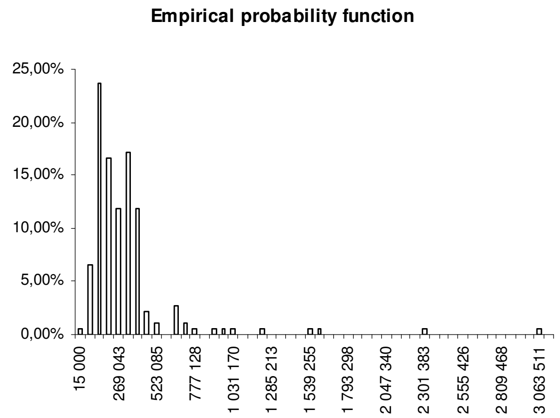


Figure 2 Empirical probability function of the database

3.2. Model for the number of losses

In order to model the number of losses the Poisson distribution with the parameter λ amounting to 47 was used (average number of loss events per year). Goodness of fit tests between the chosen Poisson distribution and the empirical data was not performed because of the short data series.

3.3. Severity model of a single loss

The model for the severity of a single loss within standard LDA was done for lognormal, gamma and Weibull distribution. Results of the parameter estimation are presented in the Table 1 (for further calculations the results of the maximum likelihood estimation are used).

Distribution Parameters	Maximum Likelihood Estimation	Method of Moments
Lognormal distribution - μ	12,28	12,17
Lognormal distribution - σ	0,75	0,91
Gamma distribution - α	1,76	0,79
Gamma distribution - β	165 858	370 333
Weibull distribution - K	1,20	1,62
Weibull distribution - λ	313 827	306 078

Table 1 Results of the parameter estimation within severity models of standard LDA

The analytical goodness of fit tests Kolgomorov-Smirnov and Anderson-Darling were calculated on the confidence level of 5%. The following figures display the results of the graphical goodness of fit tests QQ Plots. According to the results of the goodness of fit tests can be stated that none of the analyzed probability distribution is suitable for the modeling of the right tail of the distribution of financial impact of a single loss. Because of poor results of the goodness of fit tests LDA with the Extreme value theory was also applied.

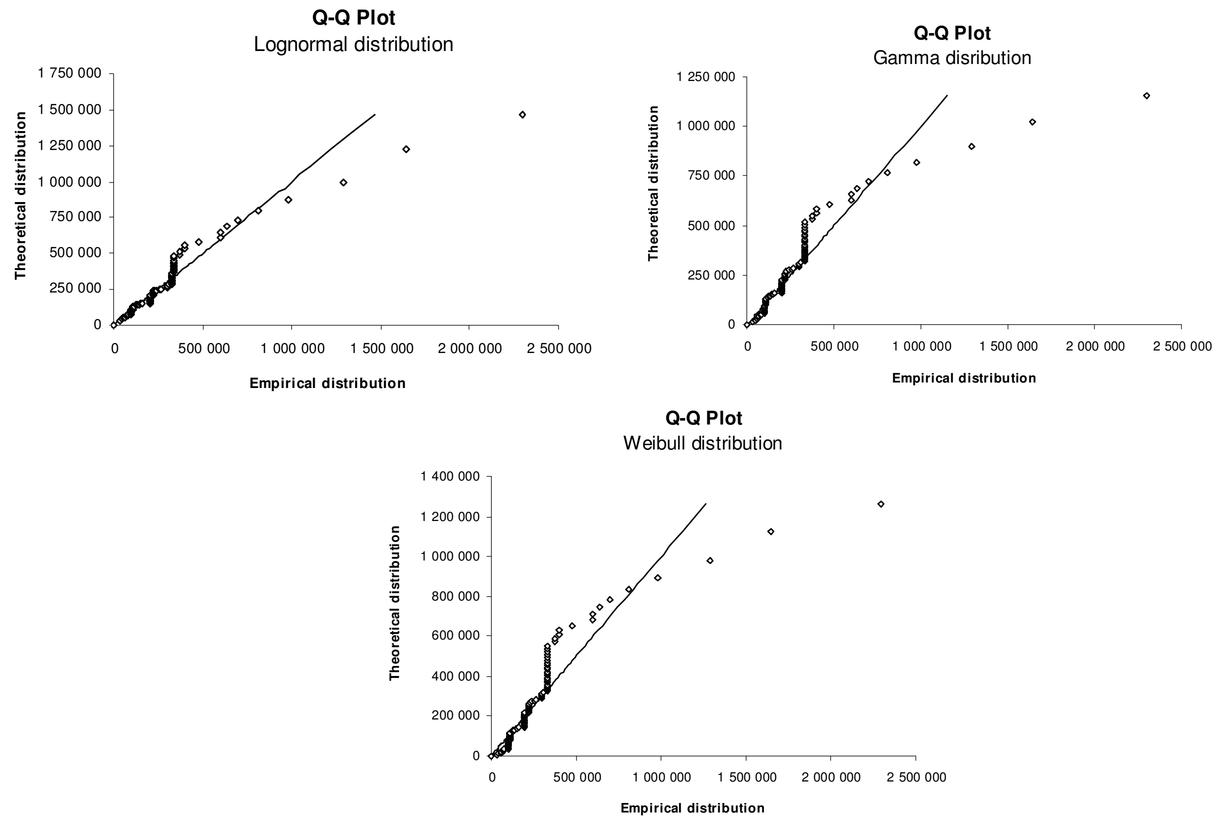


Figure 3 QQ Plots for severity models of standard LDA

As the first step within LDA framework with the Peak over threshold methodology a threshold has to be determined. For this purpose a graphical representation of the mean excess loss function was used. The threshold should be set in the area where the mean excess loss function is approximately linear with the positive slope (determination of the threshold is a trade-off between its sufficiently high value and enough data for the parameter estimation of the generalized Pareto distribution). The threshold was set out at 650 000 CZK (see figure 4).

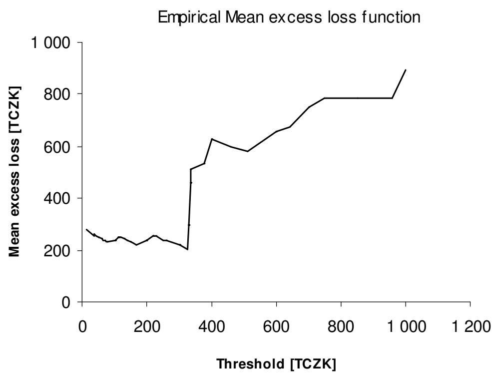


Figure 4 Empirical mean excess loss function

Severity models for losses that don't exceed the stated threshold were done for the same distribution as in the previous case (lognormal, gamma and Weibull). Results of the parameter estimation are presented in the Table 2 (for further calculations results of the maximum likelihood estimation are used).

Lognormal distribution	Maximum Likelihood Estimation	Method of Moments
Lognormal distribution - μ	12,17	12,21
Lognormal distribution - σ	0,62	0,50
Gamma distribution - α	3,14	3,47
Gamma distribution - β	72 442	65 582
Weibull distribution - K	1,96	2,00
Weibull distribution - λ	257 327	256 950

Table 2 Results of the parameter estimation within the severity models of LDA with the POT

The analytical goodness of fit tests Kolgomorov-Smirnov and Anderson-Darling were calculated on the confidence level of 5%. The following figures display the results of graphical goodness of fit tests QQ Plots. According to both type of goodness of fit test can be stated that the results are much better in this case compared to the previous standard LDA where all data was used. It is obvious that the performed database split according to the threshold of 650 000 leads to better severity models for losses not higher than 650 000.

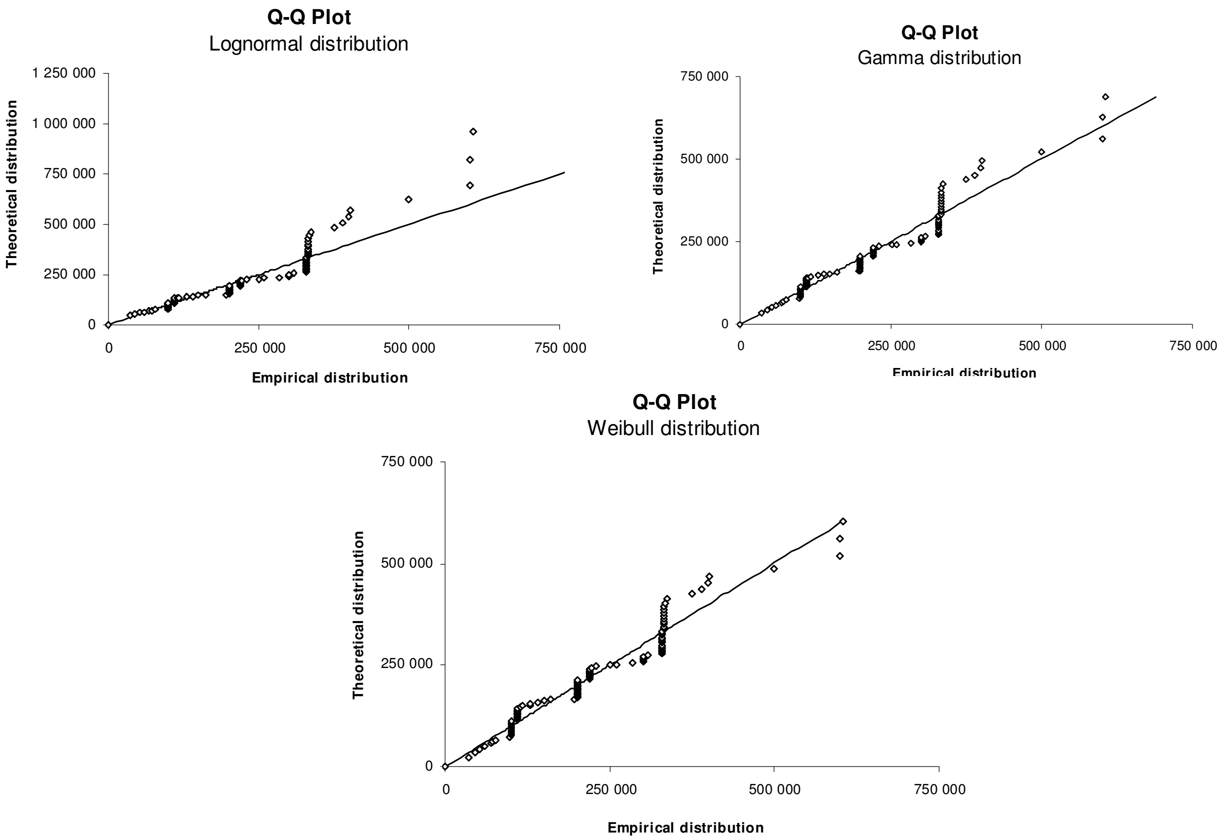


Figure 5 QQ Plots for severity models of LDA with the Peak over threshold methodology

The empirical data above the stated threshold (11 records) was used for the parameter estimation of the generalized Pareto distribution. The result is shown in the Table 3.

Generalized Pareto distribution	Value
Parameter - σ	599 258
Parameter - ξ	0,1

Table 3 Result of the parameter estimation of Generalized Pareto distribution

The next figure 6 displays the goodness of fit test QQ Plot for the generalized Pareto distribution. According to the performed QQ Plot can be stated that the generalized Pareto distribution describes the right tail of the empirical data very well (this was not the case for the standard LDA).

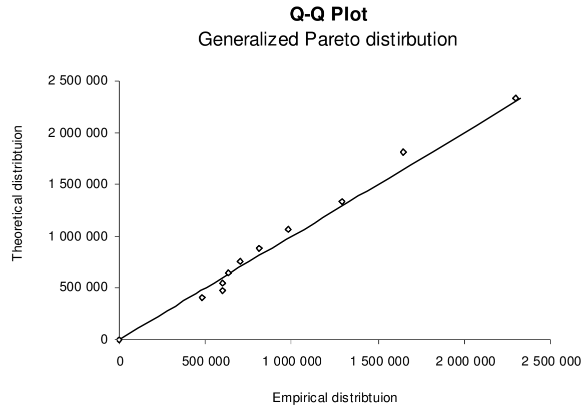


Figure 6 QQ Plots for the generalized Pareto distribution

3.4. Aggregation of the model for the number of losses and the severity model

In the case of standard LDA the Poisson distribution was aggregated with the gamma distribution with the use of the Panjer recursion formula (7). Discretization of the gamma distribution was done according to the mid point method with the money unit *C* amounting to 35 000.

In the case of LDA with the Peak over threshold methodology the Poisson distribution was aggregated with the gamma distribution respectively the generalized Pareto distribution via Monte Carlo simulation with 500 000 scenarios of the total loss. Obtained distributions are displayed in the figure 7.

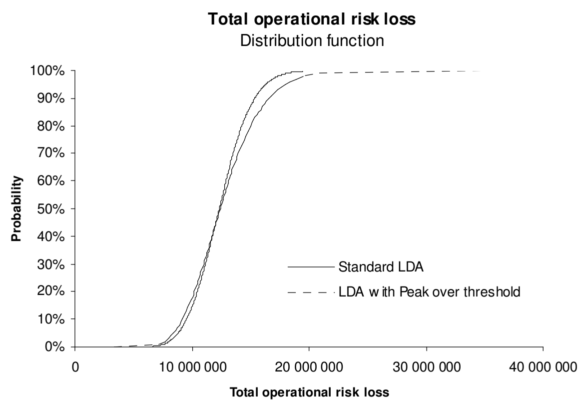


Figure 7 Distributions of the total operational risk loss

3.5. Results

From the distribution of the total operational risk loss were derived usual risk measures like operational Value-at-risk, expected and unexpected loss for different significance levels. Results are presented in the Table 4. The difference between both approaches is increasing with the higher percentiles of the distribution which is in line with the expectations because of the Extreme value theory. The amount of the capital for covering operational risk loss from analyzed business line and event type is equal to 24.8 million for LDA with the Peak over threshold (sum of expected and unexpected loss for significance level of 0.1%) and is higher by 23% in comparison with standard LDA.

Risk measure	Standard LDA	LDA with Peak over threshold	Absolute difference	Relative difference
Expected loss	12 413 826	12 689 234	275 408	+ 2 %
Value-at-Risk (p=5%)	16 275 000	17 962 998	1 687 998	+ 10 %
Value-at-Risk (p=1%)	18 060 000	20 854 899	2 794 899	+ 15 %
Value-at-Risk (p=0,1%)	20 230 000	24 812 373	4 582 373	+ 23 %
Unexpected loss (p=0,1%)	7 816 174	12 123 139	4 306 965	+ 55 %

Table 4 Obtained risk measures for the standard LDA and LDA with the POT methodology

Obtained results have proven that the Extreme value theory (in this case the version called Peak over threshold) plays an import role in the calculation of the operational risk loss when the empirical data have a heavier right tail compared to classical distribution functions.

4. Conclusion

This paper deals with the quantification of the operational risk according to the standard LDA framework and its extended version with the Extreme value theory. First part of the paper consists of a description of both concepts which are afterwards applied on the database of a financial institution.

According to the obtained results can be stated that the empirical operational risk data have a heavy right tail of the distribution and because of this fact plays the Extreme value theory important role in the measurement of operational risk. Difference between the capital requirement in case of the application of standard LDA and LDA with the Extreme Value Theory (Peak over Threshold) has been 23%.

Acknowledgement

The research is due to the support provided by GAČR (Czech Science Foundation - Grantová Agentura České Republiky) under the project No. 402/08/1237. All support is greatly acknowledged

References to literature

[1] AKKIDIZ, I., S., *Guide to optimal operational risk and Basek II*, Auerbach Publications, 2006.
[2] BASEL COMMITTEE ON BAKING SUPERVISION, *International Convergence of Capital Measurement and Capital Standards – A revised framework,* 2005, Available on www.bis.org.
[3] CHERNOBAI, A., RACHEV, S., Applying Robust Methods to Operational Risk Modeling, *Working Paper*, 2007.
[4] CRUZ, M. G., *Modeling, Measuring and Hedging Operational Risk*, John Wiley & Sons Ltd., 2002.
[5] DAYKIN, C.D., PENTIKÄINEN, T., PESONEN, M. *Practical Risk Theory for Actuaries*, Chapman&Hall, 1996.
[6] DEGEN, M., EMBRECHTS, P., LAMBRIGGER, D., The Quantitative Modeling of Operational Risk: Between g-and-h and EVT, *Working Paper*, 2006.
[7] GIESECKE, K., SCHMIDT, T., WEBER, S., *Measuring the Risk of Extreme Events,* 2005. Available on http://www.stanford.edu/dept/MSandE/people/faculty/giesecke/kts.pdf.
[8] GIESECKE, K., SCHMIDT, T., WEBER, S., Measuring the Risk of Extreme Events, *Working Paper,* Cornell University – School of Operations Research nad Industrial Ingineering, 2005.
[9] GILII, M., KELLEZI, E., An application of Extreme Value Theory for Measuring Risk, *Working Paper,* University of Geneva, 2003.
[10] McNEIL, J., *Extreme Value Theory for Risk Managers*, 1999. Available on http://www.math.ethz.ch/~mcneil/ftp/cad.pdf.
[11] PANJER, H. H., *Operational Risk – Modeling Analytics*, John Wiley & Sons, 2006.
[12] PETERS, G., W., JOHANSEN, A., M., DOUCET, A., Simulation of the Annual Loss Distribution in Operational Risk via Panjer Recursions and Volterra Integral Equations for Value at Risk and Expected Shortfall Estimation, *Working Paper*, 2006.
[13] TICHÝ, T. Posouzení vybraných možnost zefektivnění simulace Monte Carlo při opční oceňování. *Politická ekonomie* 56, (pp. 772-794), 2008.
[14] ZECH, G., ASLAN, B., Comparison of different Goodness-of-fit Test, *Working Paper*, Universität Siegen, 2005.

ESTIMATING OUTPUT GAP IN THE CZECH REPUBLIC: DSGE APPROACH

Pavel Herber[1] and Daniel Němec[2]

[1] Masaryk University, Faculty of Economics and Administrations
Department of Economics
Lipová 41a, 602 00 Brno, Czech Republic
pherber@seznam.cz
[2] Masaryk University, Faculty of Economics and Administrations
Department of Economics
Lipová 41a, 602 00 Brno, Czech Republic
nemecd@econ.muni.cz

Abstract. In our contribution, we estimate the output gap that is consistent with a fully specified DSGE model. The output gap is defined in this framework as a deviation of actual output from its flexible-price equilibrium level. The flexible-price equilibrium corresponds to the state of the economy with more efficient allocation. These estimates are thus useful indicators for monetary policy. Our output gap illustrates Czech business cycles which are rather different to other estimates (e.g. HP filter). This result may be typical for economies in transition. Moreover, our results for the Czech economy show that the turning points in the gaps are accompanied by government changes.
Keywords. Output gap, DSGE models, Bayesian estimation.

1 Introduction

There are various methods and models to estimate the potential output and output gap. One way is to use pure statistical estimates (e.g. HP filter). Such an approach lacks of economic content. Production function method estimates an aggregate production function with equilibrium inputs. "More economic" approaches estimate potential output within the framework of the Phillips curve. These models exploit reduced-form equations and the resulting output gaps are non-accelerating inflation measure. In our contribution, we estimate the output gap (using Bayesian techniques) that is consistent with a fully specified DSGE model. The output gap is defined in this framework as a deviation of actual output from its flexible-price equilibrium level. The flexible-price equilibrium corresponds to the state of the economy with more efficient allocation. These estimates are thus useful indicators for monetary policy.

Our contribution is structured as follows. The next section provides a simple variant of the New Keynesian DSGE model which is used for our analysis. Section 3 explains used data and estimation techniques. In Section 4, our baseline estimation results are presented and robustness of our analysis is checked. Section 5 concludes this contribution.

2 The model

The model is a simple variant of the standard New Keynesian (NK) DSGE model used by Hirose and Naganuma [4]. The model consists of optimizing households and monopolistic firms facing price stickiness. Monetary policy follows an interest-rate feedback rule. All real variables are detrended by a non-stationary trend component of the productivity process $\bar{A}_t$ with the constant growth rate γ^*. This assumption guarantees stationarity of the model. We will present only the main equations and log-linearized version of the model. Derivation of first-order conditions and details to the log-linearization of corresponding equations are provided by Herber [3]. For further references and extensions of basic NK DSGE model see Galí [2].

2.1 The representative household

The representative household is infinitely lived. Its utility is derived from composite consumption good C_t, real money balances M_t/P_t and leisure $1 - N_t$. We assume a consumption habit formation. The household maximize the following expected utility function:

$$E_t \sum_{i=0}^{\infty} \beta^i D_{t+i} \left[\frac{1}{1-\tau} \left(\frac{C_{t+i}}{C^h_{t+i-1}} \right)^{1-\tau} + \frac{\mu}{1-b} \left(\frac{M_{t+i}}{P_{t+1}} \right)^{1-b} - \chi \frac{N^{1+\eta}_{t+i}}{1+\eta} \right], \tag{1}$$

where D_i is a preference shock, C^h_{t-1} represents habit stock with the habit persistence parameter $0 < h < 1$.Parameter β is the discount factor ($0 < \beta < 1$), $\tau^{-1} > 0$ is the intertemporal elasticity of substitution, $b^{-1} > 0$ is elasticity of substitution between consumption and real money balances and $\eta^{-1} > 0$ is elasticity of substitution between consumption and labor. Parameters $\mu > 0$ and $\chi > 0$ are scale factors.

Given the aggregate price index, we have following budget constraint:

$$C_t + \frac{M_t}{P_t} + \frac{B_t}{P_t} = \frac{W_t}{P_t}N_t + \frac{M_{t-1}}{P_t} + R_{t-1}\frac{B_{t-1}}{P_t} + \Pi_t, \tag{2}$$

where B_t are nominal government bonds with nominal interest rate R_t, W_t/P_t is the real wage and Π_t is real profits of the firms (households are the owners).

2.2 The firms

The final (composite) consumption good Y_t is produced using differentiated intermediate goods $Y_t(j)$, $k \in [0,1]$ which are produced by monopolistic firms. Final good is produced (aggregated) using following technology (CES index):

$$Y_t = \left(\int_0^1 Y_t(j)^{\frac{\lambda_t - 1}{\lambda_t}} \mathrm{d}j\right)^{\frac{\lambda_t}{\lambda_t - 1}}, \tag{3}$$

where $\lambda_t > 0$ is time-varying elasticity of demand for each intermediate good $Y_t(j)$. The cost minimization problem of the final good producer gives the demand function for good j and aggregate price index:

$$Y_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\lambda_t} Y_t \quad P_t = \left(\int_0^1 P_t(j)^{1-\lambda_t}\mathrm{d}j\right)^{\frac{1}{\lambda_t - 1}} \tag{4}$$

Each monopolistic firm faces demand curve 4. We assume a linear production function

$$Y_t(j) = A_t N_t(j), \tag{5}$$

where A_t is exogenous productivity shock and $N_t(j)$ is labor input. Each firm minimizes its total costs TC_t which are expressed as a product of real wages $\frac{W_t}{P_t}$ and total hours worked $N_t = \frac{Y_t}{A_t}$:

$$TC_t = \frac{W_t}{P_t}\frac{Y_t}{A_t}. \tag{6}$$

Real marginal cost Φ_t must satisfy the condition

$$\Phi_t = \frac{\partial TC_t}{\partial Y_t} = \frac{W_t/P_t}{A_t}. \tag{7}$$

The firms have an opportunity to change their prices in a given period with probability $1 - \omega$ (see Calvo [1]). Each firm j chooses the price $P_t(j)$ to maximize expected discounted profits

$$E_t \sum_{i=0}^{\infty} \omega^i Q_{t,t+i}\left[\left(\frac{P_t(j)}{P_{t+i}}\right)Y_{t+i}(j) - \Phi_{t+i}Y_{t+i}(j)\right], \tag{8}$$

where $Q_{t,t+i} = \beta^i \frac{U^*_{C,t+i}/C_{t+i}}{U^*_{C,t}/C_t}$ is the discount factor. Equilibrium condition $C_t(j) = Y_t(j)$ holds.

2.3 Flexible-price equilibrium and monetary policy

We consider the case where all of the firms adjust their prices every period. The flexible prices setting is characterize when $\omega = 0$, optimal price $P^*_t = P_t$ and markup $Z_t = \bar{Z}$. The definition of real marginal cost in (7) implies:

$$\frac{W_t}{P_t} = \frac{A_t}{\bar{Z}}. \tag{9}$$

This relationship is combined with the first order conditions of the households and equilibrium conditions of the whole economy. We may thus define optimal consumption, hours worked and flexible-price equilibrium output with the superscript f. The output gap is defined as:

$$gap_t = y_t - y^f_t, \tag{10}$$

where y_t is a log of the actual output and y^f_t is a log of the flexible-price equilibrium output. The gap measures the percentage deviation of the actual output from its flexible-price equilibrium.

The model is closed by specifying a monetary policy rule. We use the standard Taylor-type rule which links the nominal interest rate with the movement of inflation and the output gap from its target levels. Equilibrium dynamics is forced by demand shock d_t, cost shock z_t and productivity shock a_t. These shocks follow the stationary $AR(1)$ processes.

2.4 Log-linearized equations

Our system consists from following log-linearized equations:

$$
\begin{aligned}
u^*_{C,t} - y_t &= E_t u^*_{C,t+1} - E_t y_{t+1} + r_t - E_t \pi_{t+1} && (11)\\
w_t - p_t &= d_t + \eta n_t - u^*_{C,t} + y_t && (12)\\
u^*_{C,t} &= \frac{1}{1-\beta h}\Bigg[(1-\tau)\Big((1+\beta h^2)y_t - h y_{t-1} - \beta h E_t y_{t+1}\Big) \\
&\quad + d_t - \beta h E_t d_{t+1}\Bigg] && (13)\\
\pi_t &= \beta E_t \pi_{t+1} + \frac{(1-\beta\omega)(1-\omega)}{\omega}\varphi_t + \frac{1-\omega}{\omega}(z_t - \beta\omega E_t z_{t+1}) && (14)\\
y^f_t &= a_t + \frac{1}{1+\eta} u^{*f}_{c,t} - \frac{1}{1+\eta} d_t && (15)\\
u^{*f}_{C,t} &= \frac{1}{1-\beta h}\Bigg[(1-\tau)\Big((1+\beta h^2)y^f_t - h y^f_{t-1} - \beta h E_t y^f_{t+1}\Big) \\
&\quad + d_t - \beta h E_t d_{t+1}\Bigg] && (16)\\
r_t &= \rho_r r_{t-1} + (1-\rho_r)\Big[\psi_\pi \pi_t + \psi_y (y_t - y^f_t)\Big] + \epsilon_{r,t} && (17)\\
d_t &= \rho_d d_{t-1} + \epsilon_{d,t} && (18)\\
z_t &= \rho_z z_{t-1} + \epsilon_{z,t} && (19)\\
a_t &= \rho_a a_{t-1} + \epsilon_{a,t} && (20)
\end{aligned}
$$

Lower case letters are logarithms of the corresponding upper cases. We have seven equations for endogenous variables and three equations which describe exogenous processes. The system includes four shocks – $\epsilon_{r,t}, \epsilon_{d,t}, \epsilon_{z,t}$ a $\epsilon_{a,t}$. Previous equations may be read as follows: Equation (11) is a New Keynesian IS curve (Euler equation for the consumption). Equation (12) represents labor supply curve of the households. Equation (13) express marginal utility of consumption. Equation (14) is a New Keynesian Phillips curve. Equation (15) defines potential output. Equation (16) reflects marginal utility of consumption in the case of flexible-price framework. Equation (17) is a modified Taylor rule. Equation (18) is the $AR(1)$ process of demand shock. Equation (19) is the $AR(1)$ process of cost shock. Equation (20) is the $AR(1)$ process of productivity shock.

We will use the log-linearized versions of both production function and real marginal costs:

$$
\begin{aligned}
y_t &= a_t + n_t, && (21)\\
\varphi_t &= w_t - p_t - a_t. && (22)
\end{aligned}
$$

We estimate 15 parameters. These parameters are described in the Table 1.

3 Estimation

We use Bayesian techniques (Random-Walk metropolis Hasting algorithm) and Kalman filter algorithm in order to estimate both the parameters and unobserved potential output. The model is thus identified by using the Dynare toolbox for Matlab [5]. We use quarterly seasonally adjusted macroeconomic data of the Czech Republic from the second quarter 1996 to the fourth quarter 2008. Observable variables are quarter-on-quarter productivity growth, inflation and nominal interest rate. Productivity is obtained from the output growth of real GDP divided by the labor force. The inflation rate is based on CPI and interest rate is the 3M PRIBOR. Inflation and nominal interest rate are annualized. One of the alternative models uses data of the inflation target. The variables in the model are expressed as deviation from their steady-states. We have added measurement equations that relate the model variables to the data (instead of prefiltering the observed data):

$$
\begin{aligned}
Y^{obs}_t &= \gamma^{ss} + y_t - y_{t-1}, && (23)\\
\Pi^{obs}_t &= \pi^{ss} + 4\pi_t, && (24)\\
R^{obs}_t &= rr^{ss} + \pi^{ss} + 4r_t, && (25)
\end{aligned}
$$

Table 1. Parameters

Parameter	Description	Range	Density
τ^{-1}	Intertemporal elasticity in consumption	$\mathbb{R}^+$	Gamma
β	Discount factor	$(0,1)$	Beta
h	Habit persistence	$(0,1)$	Beta
η^{-1}	Elasticity of substitution between labor and consumption	$\mathbb{R}^+$	Gamma
ω	The share of non-optimizing firms	$[0,1]$	Beta
ψ_π	Elasticity of the interest rate with respect to inflation	$\mathbb{R}^+$	Gamma
ψ_y	Elasticity of the interest rate with respect to output gap	$\mathbb{R}^+$	Gamma
ρ_r	Interest rate smoothness	$[0,1)$	Beta
ρ_d	Persistence in the demand shock	$[0,1)$	Beta
ρ_z	Persistence in the supply shock	$[0,1)$	Beta
ρ_a	Persistence in the productivity shock	$[0,1)$	Beta
σ_r	Standard deviation of the monetary shock	$\mathbb{R}^+$	Inv. Gam.
σ_d	Standard deviation of the demand shock	$\mathbb{R}^+$	Inv. Gam.
σ_z	Standard deviation of the supply shock	$\mathbb{R}^+$	Inv. Gam.
σ_a	Standard deviation of the productivity shock	$\mathbb{R}^+$	Inv. Gam.

where γ^{ss} is equilibrium growth rate of a non stationary technology shock, π^{ss} is annualized inflation at the steady-state and rr^{ss} is real interest rate at the steady-state.

Priors are presented in the Table 2. They are mostly in line with the estimates of similar estimates for the Czech Republic (see Musil and Vašíček [6] or Remo and Vašíček [7]). Prior hyperparameters of π^{ss} and rr^{ss} are obtained from the historical averages of inflation and real interest rate.

4 Results

Table 2 reports posterior distributions of the structural parameters. The posterior mean of the parameter τ is 2.31, intertemporal elasticity of substitution τ^{-1} is thus 0.43. Households prefer actual consumption to future consumption. The parameter h is 0.97, i.e.past consumption plays crucial role in current decisions. The parameter η is 1.54. Elasticity for labor supply is therefore 0.65. This low elasticity is typical for the Czech labor market because of low flexibility of the labor force. Price rigidity is related to the parameter ω. The mean value 0.8 means that only 20% of firms change their price every quarter. The mean contract duration is thus five quarters. The reaction function of the Czech National Bank is based

Table 2. Estimates - parameters and shocks

Parameter	Prior mean	Prior s.d.	Post. mean	90% HPDI	
τ	1.10	0.50	2.3128	1.5128	3.0692
h	0.80	0.10	0.9661	0.9317	0.9982
η	1.50	0.30	1.5442	1.0440	2.0247
ω	0.60	0.20	0.7944	0.7163	0.8774
ψ_π	1.50	0.30	1.3157	0.9985	1.5823
ψ_y	0.30	0.15	0.2706	0.0627	0.4794
ρ_r	0.65	0.25	0.7290	0.6603	0.7990
ρ_d	0.50	0.25	0.8308	0.7253	0.9414
ρ_z	0.50	0.25	0.5724	0.2029	0.9779
ρ_a	0.70	0.25	0.9336	0.8790	0.9953
γ^{ss}	0.50	0.30	0.6725	0.5067	0.8467
π^{ss}	3.50	0.50	3.6368	2.8588	4.3909
rr^{ss}	1.80	0.30	1.7559	1.2991	2.1924
σ_r	0.30	∞	0.3355	0.2771	0.3946
σ_d	0.50	∞	1.0901	0.4321	1.7915
σ_z	0.50	∞	0.4230	0.1050	0.9410
σ_a	0.80	∞	1.4366	0.8301	2.0816

on the parameters ρ_r, ψ_π a ψ_y. Interest rate rule may be written as

$$r_t = 0.73 r_{t-1} + 0.36\pi_t + 0.07 gap_t + \epsilon_{r,t}. \tag{26}$$

The smoothing parameter is relative high and plays the most important role in monetary decisions. Central bank places five times less weight on the output gap relative to inflation variations. The autoregressive coefficients for the shocks represent persistency of the shocks.

Figure 1 plots the posterior mean of smoothed estimate of the output gap and its 90% confidence intervals. This trajectory may be compared with the gap obtained using Hodrick-Prescott filter. Using

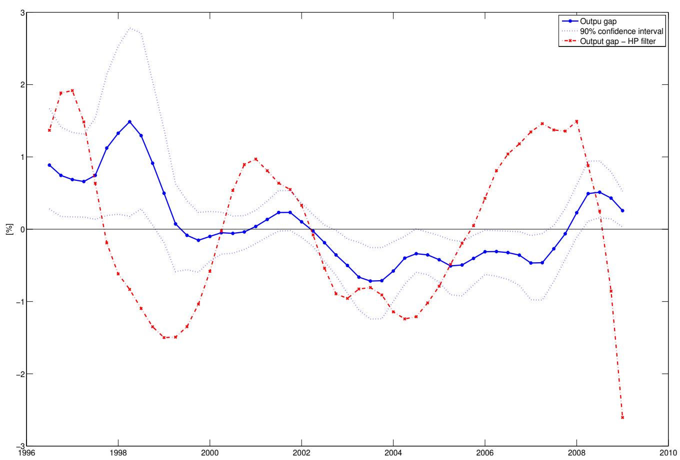


Fig. 1. Output gap and confidence intervals

U.S. data we were able to replicate the original results presented by Hirose and Naganuma [4] which are similar compared with other commonly used estimates. Our output gap, on the other hand, illustrates Czech business cycles which are rather different to the estimate using HP filter.

Our approach shows the potential of the economy with flexible-prices that would prevail in the absence of cost shocks. We may present alternative estimates which measures output gap as the deviation of the actual output from its equilibrium (trend) level. In this case, the dynamics of the gap should be in accordance with HP filter estimates. This may be shown on the Figure 2. Resulting trajectories differ in levels because our data are detrended by steady-state growth rate of a technology shock (parameter γ_{ss}). As argued by Woodford [8], an optimal monetary policy replicates flexible-price equilibrium, which can

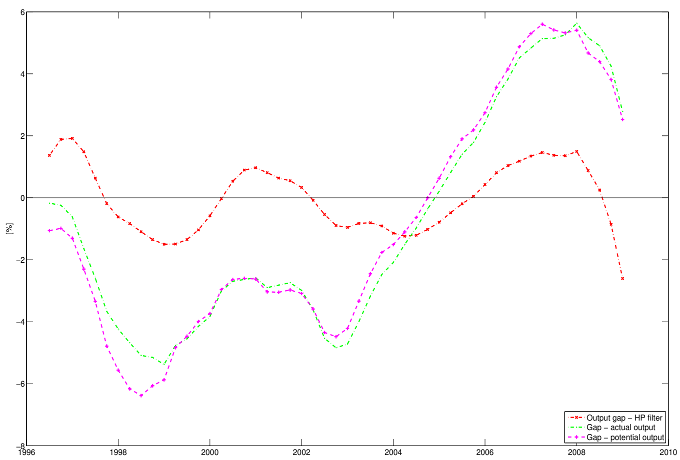


Fig. 2. Actual output gap and potential output gap

be the best outcome of the economy where government offsets monopolistics distortions by appropriate transfers. The output gap that is defined here should be a useful measure for monetary policy makers (from a welfare perspective). Monetary policy is able to stabilize inflation and welfare of the households (approximated by the output gap). These goals are conformable with each other.

New Keynesian approach to the output gap offers relevant economic background compared to the "traditional" estimates. Output gap is linked to real marginal costs which influence inflation within the Phillips curve framework. There is a strong relationship between output gap and inflation which may be used by monetary authority.

What says our output gap in the Figure 1? Its trajectory indicates the existence of rigid prices in the Czech economy. On the one hand, these rigidities have damped impact of recession in the period since 1997 and 2001 (positive output gap). On the other hand they restrain economic growth in the periods of the booms (negative output gap since 2005). Our results for the Czech economy show that the turning points in the gaps are accompanied by government changes (1998, 2002 and 2006).

4.1 Prior sensitivity and model alternatives

Our results may be influenced by the choice of prior hyperparameters. We have specified two alternative prior sets based on tight and loose prior. Our choice concerns following parameters: τ, h, η, ω, ψ_π, ψ_y, γ^{ss}, π^{ss} and rr^{ss}. In the case of "tight prior" the prior standard deviation equals a half of the base one. In the case of "loose prior" the standard deviations has been doubled. As expected, for most of the parameters, deviations of the posterior estimates from the prior means were remarkable under the loose prior. The tight prior leads the posterior estimates close to the prior. Figure 3 depicts the corresponding

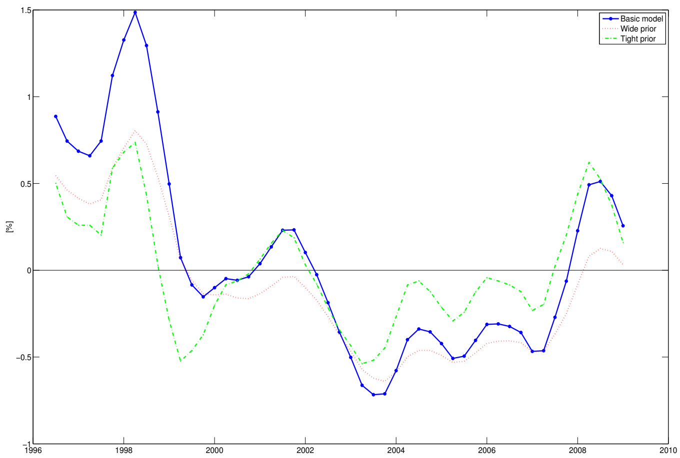


Fig. 3. Output gaps - prior sensitivity

estimates of the output gap. We can see that the output gap is similar to the gap under the loose prior at the beginning of the period and to the gap under loose prior in the second half of the period. The output gap under the tight prior exhibits smaller volatility due to the smaller values of h and ω. The dynamic of all trajectories is almost the same within the whole period. Our baseline estimate may be taken as a relevant "compromise" estimate. For further discussion regarding sensitivity analysis see Herber [3].

We have made three alternative specifications of the model which may be more appropriate to describe the Czech economy. The first model (Model 1) assumes time-varying steady-state growth rate of technology shocks:

$$\gamma_t^{ss} = \gamma_{t-1}^{ss} + \epsilon_{\gamma,t}, \qquad \epsilon_{\gamma,t} \sim N(0, \sigma_\gamma^2). \tag{27}$$

The second model (Model 2) assumes time-varying steady-state inflation:

$$\pi_t^{ss} = \pi_{t-1}^{ss} + \epsilon_{\pi,t}, \qquad \epsilon_{\pi,t} \sim N(0, \sigma_\pi^2). \tag{28}$$

This assumption affects the dynamic of nominal interest rate that should be rewritten as follows:

$$R_t^{obs} = rr^{ss} + \pi_{t+1}^{ss} + 4r_t. \tag{29}$$

The last model (Model 3) takes into account inflation targeting. Inflation gap will be defined as the difference between actual inflation and its target level

$$\Pi_t^{obs} - \Pi_t^{target} = 4\pi_t. \tag{30}$$

Nominal interest rate is thus determined by

$$R_t^{obs} = r^{ss} + 4r_t, \tag{31}$$

where the parameter r^{ss} is annualized steady-state nominal interest rate. The relevant prior is based on historical averages.

The results are not sensitive to the changes in measurement equations. Obtained output gaps (see Fig. 4) have similar dynamics in comparison with the basic model. They differ only in levels of the gaps and in the confidence intervals which are larger. Interesting results are offered by Model 3 which takes into account inflation targeting. Central bank was not able to achieve its inflation target at the beginning of the period of inflation targeting. This fact leads to higher volatility of the inflation gap in our model. The Phillips curve and the dynamics of the real marginal costs are the mechanisms through which the volatility of the output gap increases. For more detailed discussion about our alternative models (including the trajectories of time-varying steady-state values) see Herber [3].

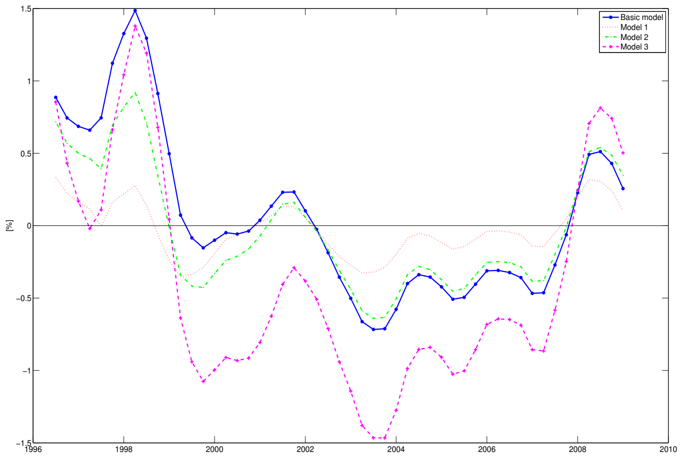


Fig. 4. Output gaps - alternative models

5 Conclusion

In our contribution, we estimated the output gap which is consistent with the fully specified DSGE model. This gap is a useful measure for welfare of the economy. The output gap is a plausible indicator for monetary policy. Structural parameters may provide an economic interpretation for movements of the estimated output gap. Using U.S. data we were able to replicate the original results presented by Hirose and Naganuma [4] which are similar compared with other commonly used estimates. Our output gap, on the other hand, illustrates Czech business cycles which are rather different to other estimates (e.g. HP filter). This result may be typical for economies in transition. Rigidities in the Czech economy have damped the fall of economy in the periods of recession but they have hindered the potential growth in the periods of booms. Moreover, our results for the Czech economy show that the turning points in the gaps are accompanied by government changes.

Acknowledgement This work was supported by MŠMT project Research centers 1M0524, and funding of specific research at ESF MU.

References

1. Calvo, G. A.: Staggered Prices in a Utility-Maximizing Framework. Journal of Monetary Economics, 12(3), 383-398 (1983)
2. Galí, J.: Monetary Policy, Inflation, and the Business Cycle. Princeton University Press (2008)
3. Herber, P.: Odhady potenciálního produktu - DSGE přístup. Master Thesis, Masaryk University, Brno (2009)
4. Hirose, Y., Naganuma, S.: Structural Estimation of the Output Gap: A Bayesian DSGE Approach for the U.S. Economy. Bank of Japan Working Paper Series, no. 7 (2007)
5. Juillard, M.: DYNARE Toolbox for Matlab. ver. 3.65. `www.cepremap.cnrs.fr/dynare/`
6. Musil, K., Vašíček, O.: The Czech Economy with Inflation Targeting Represented by DSGE Model: Analysis of Behaviour. Working papers CVKS ESF MU, no. 7. (2005)
7. Remo, A., Vašíček, O.: Monetary policy and stability of Czech economy: optimal commitment policy in NOEM DSGE framework. In *Tomeš, Z., Žídek, L.: Determinanty hospodářského růstu v zemích východní Evropy.* Brno, CVKS ESF MU, 409-460 (2008)
8. Woodford, M.: Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press, Princeton, New Jersey (2003)

EXTENSIVE MARGIN IN INTERNATIONAL TRADE: EMPIRICAL ANALYSIS FOR VISEGRAD COUNTRIES

Miroslav Hloušek[1]

Faculty of Economics and Administration, Masaryk University
Department of Economics, Lipová 41a, 602 00 Brno
hlousek@econ.muni.cz

Abstract. This paper deals with empirical analysis of international trade between Visegrad countries and EU-15 during past two decades. The goal of the paper is to find out if the growth in export is of intensive or extensive type. We follow methodology of Kehoe and Ruhl (2002) and use detailed trade statistics on the value of trade flows by commodity according to Standard International Trade Classification (SITC) codes. We find out that the goods that were traded the least in the benchmark year account for disproportionate share in trade after liberalization and reduction of trade barriers. The most significant increase was found in Hungary. The set of goods which accounted for only ten percent of trade in 1993 accounts for about forty percent of trade following the liberalization. Similar patterns were identified also in other Visegrad countries. The countries thus began to export goods that they had not been previously trading. This is growth on the extensive margin and it should be reflected in models of international trade.
Keywords. International trade, trade barriers, liberalization, export growth, intensive and extensive margin, Visegrad countries

1. Introduction

This paper analyses international trade between Visegrad countries and EU-15. The goal of the paper is to find out if the growth in export is of intensive or extensive type. We follow methodology of Kehoe and Ruhl (2002) and use detailed trade statistics on the value of trade flows by commodity according to Standard International Trade Classification (SITC). The results show that the goods that were traded the least in the benchmark year account for disproportionate share in trade after liberalization and reduction of trade barriers. In other words the countries began to export goods that they had not been previously trading. This is growth on the extensive margin and should be reflected in models of international trade and also in decisions of policymakers that usually focuses only on supporting of traditional export industries.

The rest of the paper is organized as follows. Section 2 briefly describes data used for analysis and section 3 focuses on measures of extensive margin growth. The results are presented and discussed in Section 4. The stress is put on analysis of Hungary and the Czech Republic. Section 5 deals with sensitivity analysis of the measure of extensive margin and its influence on results. Final section concludes with prospects for further research.

2. Data

The data are obtained from OECD database. The measure is annual flow values of exports of particular country into EU-15. The data are disaggregated by commodities according to four-digit Standard International Trade Classification (SITC), revision 2. The data are quoted in thousands of U.S. dollars, however we are interested in relative quantities and thus the units are not important. The data sample for each country is determined by availability of data. The data sample starts in 1992 for Hungary and Poland, in 1993 for the Czech Republic and in 1997 for Slovakia. The last available period is 2006 for all countries. It must be mentioned that the process of trade liberalization was gradual and in some cases already started before period that we study. However, main reduction of trade barriers happened before accessions of Visegrad countries to European Union in 2004 which is covered in our analysis.

3. Measures of the extensive margin

For each country's exports into EU15, the SITC codes are ordered by their value of trade in the first three years of the sample.[2] Then the cumulative sets of the ordered codes are constructed. Every set includes one-tenth of

[1] This work is supported by MŠMT project Research centers 1M0524 and funding of specific research at ESF MU.
[2] Average value in three years is used for the sake of robustness of the ordering. Implications of the way of ordering (one year or three years average) are discussed in Section 5.

total export. The first set starts with the smallest codes[3] and other codes are added to this set until the sum of their values reaches one-tenth of total export value. The next set is formed in similar way by summing the remaining smallest codes until the value of the set reaches one-tenth of total trade. This procedure produces ten sets of codes where each set represents one-tenth of total trade. The first set consists of the "least-traded" commodities – they have the smallest export value. Subsequent set contains less codes then the previous set as the (relative) trade value of the codes increases. Since the set comprises exactly one-tenth of total export, some codes are split into two sets. Therefore the number of codes in the set need not be integer number.

Given this partitioning of the SITC codes, two measures of export growth are considered. First measure corresponds to change in each set's share of trade over the sample period, second measure focuses on the time evolution of the least-traded set of codes with the aim to capture timing of the export growth of these goods.

The first measure is constructed by calculating the share of total exports for each of the ten sets of codes in the last year of the sample period. The interpretation of this measure is as follows. If the growth in trade is driven only by proportional increase in the value of goods already traded, each set of codes would retain its one-tenth share in trade. On the other hand, if the trade liberalization leads only to trade of goods that were previously untraded, the first set of codes would gain trade share, while share of other set would decline. The first case is intensive margin, the second case is extensive margin in trade growth.

The second measure uses the same partition of SITC codes but looks only on the share of least-traded goods in total export. This share is calculated for each year of the sample period. If the lowering of trade barriers leads to trade of goods not previously traded, there should be an increase in the share of trade accounted for by this set of goods. This measure should show the timing of any change it the trade of new goods. If an increase in the share of exports coincides with the implementation of trade reforms, we can think of it as an evidence of the link between trade liberalization and growth in the extensive margin.

4. Results

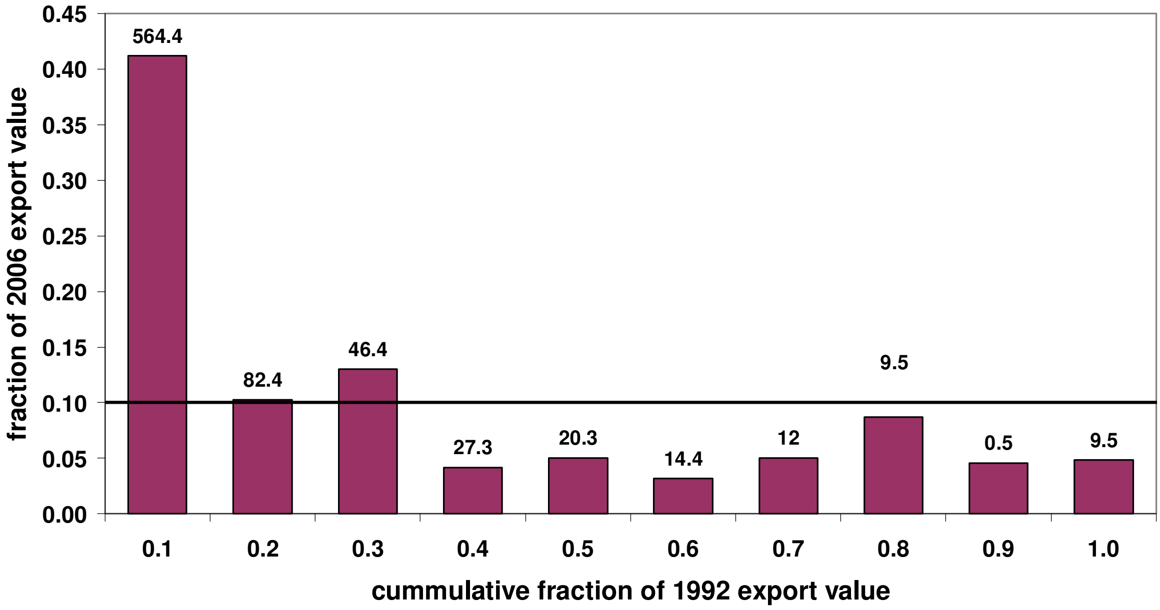
Overall results indicate significant export growth on the extensive margin for all Visegrad countries. Table 1 shows the end of sample export shares of the least-traded goods. The most significant increase in the extensive margin is observed in Hungary. Least-traded goods comprise 41 % of the total export share in 2006. This is more than four times its original trade share. The shares of least-traded goods in Poland and the Czech Republic account for 29 and 27 percent, respectively, which is also quite high. The share in Slovakia is only 19 % but this fact is probably influenced by shorter sample.

Country	Benchmark year	Share of least-traded goods
Czech Republic	1993	0,27
Hungary	1992	0,41
Poland	1992	0,29
Slovakia	1997	0,19

Table 1: Share of export value in 2006: least-traded goods in benchmark year

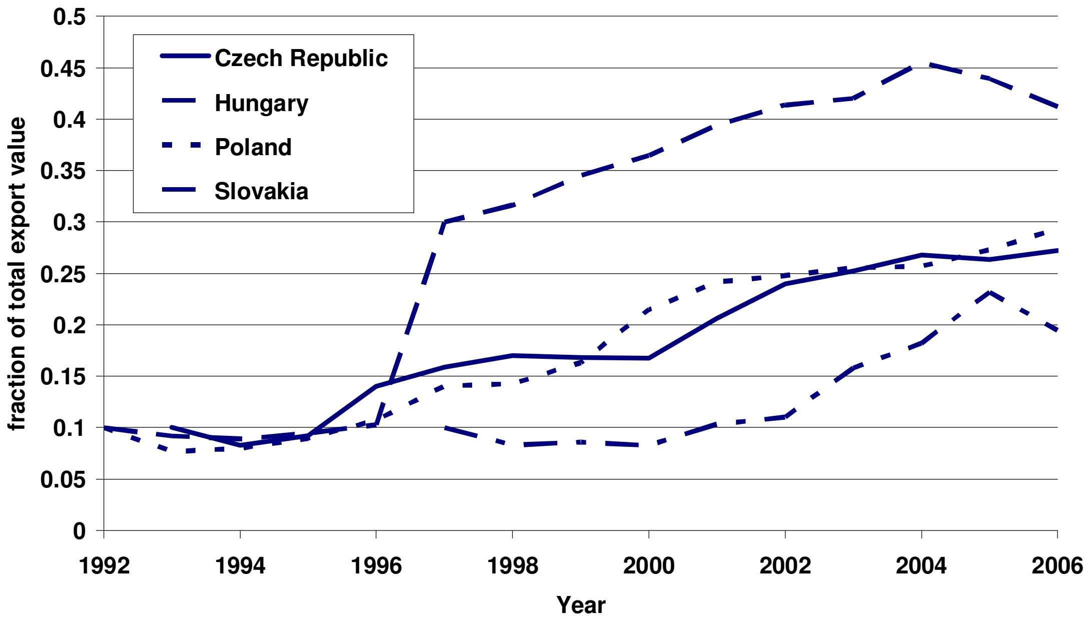
The most evident example of extensive margin growth is Hungary. Graph 1 shows the decomposition of trade into individual sets of goods for this country. The set of least-traded goods that account for 41 % trade share in 2006 includes 564 SITC codes. Given the detailed structure of the measure, we can find that the most significant increase in trade share has code "7132: Internal combustion piston engines for propelling vehicles" – more than 10 % -- followed by code "7643: Radiotelegraphic & radiotelephonic transmitters" with nearly 6 % increase.

[3] The term 'smallest codes' means codes that account for the smallest value of total trade.



Graph 1: Composition of Exports: Hungary into EU 15

Trade growth in Hungary also shows interesting timing. Graph 2 plots the trade share of the least-traded goods over the sample period for all countries. In case of Hungary (dashed line), there is abrupt and large increase in the trade share between years 1996 and 1997 – from 10 % up to 30 %. What are possible causes of such change?

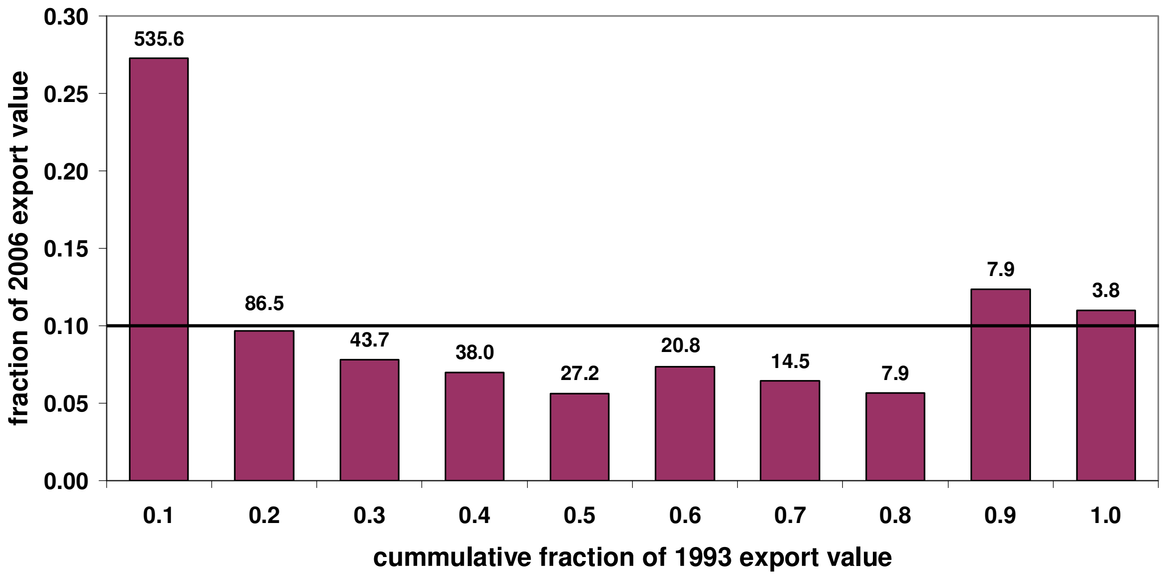


Graph 2: Time evolution of least traded goods: Hungary

Hungary's Trade Association Agreement with the EU became effective in February of 1994 and provided for an asymmetrical liberalization of trade over the next 5 years. The agreement immediately removed EU duties on

70% of Hungary's industrial exports to the EU and lifted quotas on 60% of its total exports.[4] However, there is not any change in trade share following this event. The argument that some time is needed for setup of production and managing export probably does not hold. We should see (at least gradual) increase in least-traded share in the first or second year after liberalization, not abrupt jump in the third year. In sum, this event does not coincide with our measure showed in Graph 2. Probably one of the most important developments influencing trade between the EU and Hungary has been the so-called Pan-European Cumulation System (PECS). This system comprises a set of EU bilateral agreements, which aims to harmonize standards and rules-of-origin laws between the EU, EFTA, CEFTA, and other countries. The Government of Hungary joined the PECS in December 1996, with an effective date of July 1, 1997. Increase in the least-traded share surprisingly coincides with implementation of this system.[5]

Graph 3 shows trade decomposition in the Czech Republic. The set of least-traded goods includes 535 codes and comprise 27 % of export share in 2006. There are not many single codes (among least-traded goods) that would achieve large share over the analyzed period.[6] However, it is worth to note that goods in two last deciles (most-traded goods) account for more than proportional increase in their sets. No wonder that these sets include codes "7810: Passenger motor cars for transport of pass. & goods" and „7849: Other parts & accessories of motor vehicles" – traditional branches of Czech industry.[7]



Graph 3: Composition of Exports: Czech Republic into EU 15

For time evolution of least-traded goods in the Czech Republic we look again on Graph 2. Compared to Hungary there is not such sudden increase of the share. Regarding trade liberalization, the free trade agreement between EFTA and Czechoslovakia was arranged on 1 July 1992.[8] EFTA states abolished the customs and quantity limits unilaterally as the agreement came into force. Since we have data from 1993, the effect of this liberalization is partly distorted. However, we can see two waves (accruals) in the share of least-traded goods: one from 1995 to 1998 and the second from 2000 to 2004. But it is hard to find some concrete causes. The Czech Republic also joined the PECS, but this event does not have such important effect as in Hungary. Why the evolution in the Czech Republic and Hungary is so different deserves further study.

Detailed analysis of the composition of exports in Poland and Slovakia can be found in Hloušek (2009). Regarding time evolution of least-traded goods, development in Poland is quite similar to the Czech Republic.

[4] See: http://www.factbook.net/countryreports/hu/HuTradeRegs.htm

[5] However, there could be objection that half a year of functioning of this system can not be capable to produce such large change.

[6] The largest increase – by 4 % – was recorded for code „7523: Complete digital central processing units".

[7] These two codes increased their share by 5 % and 7 %, respectively.

[8] After splitting of Czechoslovakia it was automatically related to the Czech Republic.

Comparison with Slovakia is difficult due to shorter data sample but the extensive margin growth is also considerable.

5. Sensitivity analysis

This section focuses on sensitivity analysis of our empirical measures. First, we check how the results depend on the choice of cutoff level. Table 2 reports the extensive margin growth rates for 5%, 10% and 20% cutoffs. Each column shows percentage growth rate of the least traded goods between benchmark year and year 2006. All countries exhibit the same pattern. If we consider smaller cutoff, the least-traded goods grow more. It again support the idea, that goods with very small trade shares drive the extensive margin growth and our measure that uses 10% cutoff can even underestimate size of the extensive margin growth. The measure calculated with 20% cutoff shows smaller growth and it is roughly one half of the measure using 10% cutoff. Even if larger cutoff makes the set of least-traded goods quite big, the increase is still significant.

		Cutoff		
Country	**Benchmark year**	**5%**	**10%**	**20%**
Czech Republic	1993	267,4	172,4	84,5
Hungary	1992	560,4	312,0	157,4
Poland	1992	281,0	193,2	92,1
Slovakia	1997	151,2	94,5	32,7

Table 2: Results under different cutoff values

Second sensitivity check deals with ordering of the goods at the beginning of the sample period. Primary analysis made in the paper used the average of export value in first three years. The goods were then ordered according to this measure and the first year was chosen to calculate the deciles. Alternative way is to order the goods according to their value in the first year. The least-traded goods thus contain little different set of codes (usually more of them). The results of this procedure are reported in Table 3. The extensive margin growth is higher for all countries and thus our analysis understated the importance of this growth.[9] The largest difference is for Slovakia. Contrary to previous measure, the least-traded goods now comprise 36 % of total export value. It is even more than in Poland and the Czech Republic. Detailed analysis reveals that code „7611: Television receivers, color“ is responsible for this change.[10] Time evolution of least-traded goods for Visegrad states using alternative ordering is shown in Graph 4. What is different? We see that the share is growing already from the beginning – there is no initial decrease in first years that was present in Figure 2. Main trends are very similar for all countries except of Slovakia. Here, the share of least-traded goods exhibits rapid growth from 2002 that finally overtook the share in other countries, as discussed above.

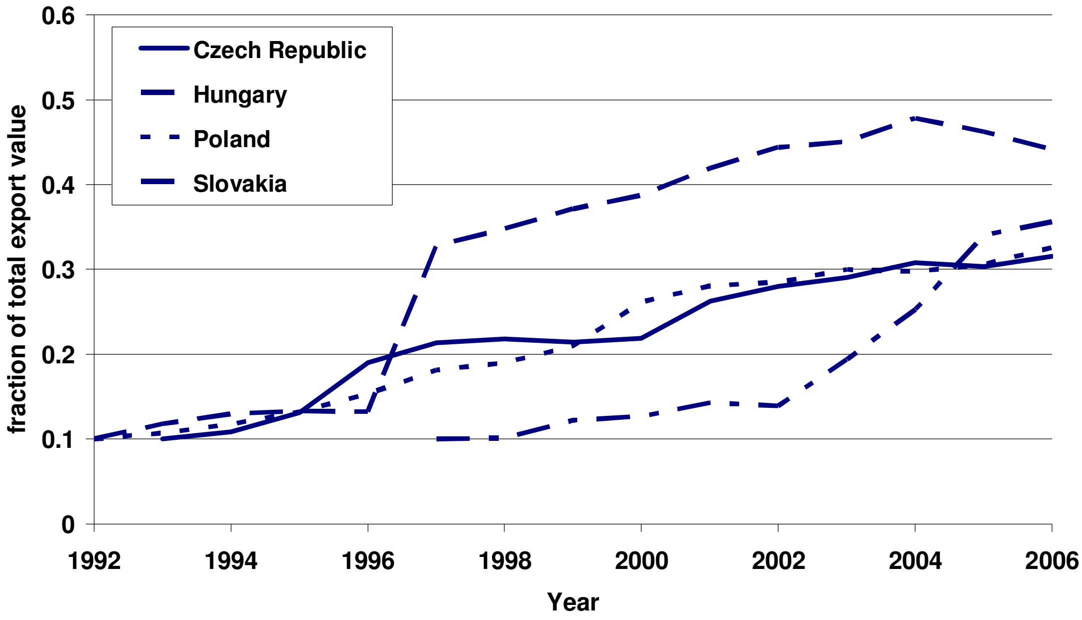
Country	Benchmark year	Share of least-traded goods
Czech Republic	1993	0,32
Hungary	1992	0,44
Poland	1992	0,33
Slovakia	1997	0,36

Table 3: Share of export value in 2006: least-traded goods in benchmark year

To summarize it, using smaller cutoff implies more significant extensive margin growth. Similarly, ordering of the goods according to the first year export value (instead of the average of three years) increases the importance of new goods margin in international trade.

[9] Compare to Table 1.

[10] This code was included in the third decil when previous ordering was used. (according to the average of first three years.)



Graph 4: Time evolution of least traded goods:Visgrad countries

6. Conclusion

This paper analysed international trade between Visegrad countries and EU15. Results show that there is clear evidence of extensive margin in trade growth. Countries are exporting goods that they had not been previously trading. It is sometimes hard to find if trade liberalization was the main cause of this change because reduction of trade barriers was gradual. This issue deserves more attention. The implications for economic theory are clearcut. International trade models should focus on modeling of the extensive margin. The modified Ricardian model used in Kehoe and Ruhl (2002) is one of the examples. Calibration of this model using data on intra-industry trade is topic for further research.

7. References

[1] HLOUŠEK, M. (2009) Importance of Extensive Margin in International Trade: empirical analysis for Visegrad countries. *Working paper CVKCE*, forthcoming.

[2] KEHOE, T.J. (2002) *An Evaluation of the Performance of Applied General Equilibrium Models of the Impact of NAFTA*, Federal Reserve Bank of Minneapolis. Available at: <http://www.econ.umn.edu/~tkehoe/papers/NAFTAevaluation.pdf>.

[3] KEHOE, T.J. and RUHL, K.J. (2002) *How Important is the New Goods Margin in International Trade?* Available at: < http://www.econ.umn.edu/~tkehoe/papers/Kim.pdf >.

[4] *Disaggregated Trade data*. U.S.-China and China-U.S. Available at: <http://www.econ.umn.edu/~tkehoe/classes/8401-07.html>, [cit. 2009-05-12].

[5] *EFTA – Czech Republic Free Trade Agreement*. Available at: <http://intl.econ.cuhk.edu.hk/rta/index.php?did=14>, [cit. 2009-05-12].

[6] *Hungary –Trade Regulations and Standards*, Available at: <http://www.factbook.net/countryreports/hu/HuTradeRegs.htm>, [cit. 2009-05-12].

[7] *ITCS International Trade by Commodity Database – SITC, Revision 2*. Available at: <http://www.sourceoecd.org>

DENSITY FUNCTION SMOOTHING USING DISCRETE F-TRANSFORM*

Michal Holčapek and Tomáš Tichý

Dep. of Mathematical Methods in Economics, Dep. of Finance
Faculty of Economics VŠB-TU Ostrava
{michal.holcapek,tomas.tichy}@vsb.cz

Abstract. There exist many approaches to the estimation of probability distribution function. A general principal is to reduce a noise (more precisely, a white noise) in data by means of e.g. stochastic processes, kernel regressions, integral transforms, wavelet transforms, or even fuzzy filters. Without applying such a method the data set would be deformed by a noise so that the resulting interpretation might be problematic or, what would be worse, misleading. Moreover, a construction of valuable models from such data is also complicated. In this paper we propose and apply a relatively new and more or less simple approach to filter a noise from data – the fuzzy transform (F-transform) originally introduced in Perfilieva [6]. More particularly, we will introduce a filter using the direct and inverse F-transform and show that the filtered data have a smaller noise, i.e., the variance of the random variable describing a filtered data noise is smaller than the variance of the random variable expressing an original data noise. Finally, the proposed filter is illustrated with the aim to estimate the probability distribution function of financial market returns.
Keywords. Fuzzy transform, probability distribution, financial returns.

1 Introduction

Without loss of generality we can say that any economic variable or, more correctly – its state (e.g. the price), follows a stochastic process, i.e. it moves randomly in time. In order to describe such a randomness, a suitable probability distribution can be used. Doing that, either characteristic function, or cumulative distribution function, or density function can be used. Since the data are generally observable only in a discrete time, some approach of statistical inference must be applied in order to get a continuous function, e.g., a density function, and remove a white noise from the data.

A basic principal is to reduce a white noise in data by means of e.g. stochastic processes, kernel regressions, integral transforms, wavelet transforms, or even fuzzy filters, see e.g. [1–4] for more details. For example, a relatively new approach of smoothing by fuzzy filters utilizes a discrete fuzzy-transform, which is based on fuzzy partitioning of the universe into fuzzy subsets.

The aim of this paper is to introduce the fuzzy transform and develop a filter based on it that would allow us efficient smoothing of probability distribution of financial market returns.

We proceed as follows. The intention of the following two sections is to set the preliminaries for stochastic processes and probability distribution estimation (Section 2) and fuzzy transform (Section 3). Next, the filter to smooth the probability distribution function is introduced (Section 4). Finally, the proposed model is applied in order to estimate a smooth distribution function of financial data (Section 5).

2 Probability density estimation for financial data

Assume a time series of financial data $\{x_t \mid t = 1, \ldots, T\}$. A general model for x_t can look as follows:

$$x_t = f(x_{t-1}) + \varepsilon_t, \tag{1}$$

where $f(x_{t-1})$ can be any function (linear, non-linear) of preceding values $x_1, \ldots, x_{t-1}$ or simply a given constant and ε_t describes a white noise, i.e., ε_t is a value of random variable with zero mean and no autocorrelation.

Further, assume that $\{x_t \mid t = 1, \ldots, T\}$ is a set of values of a continuous random variable X. To make a deeper analysis of data we are often interested in the unknown probability density function of

* The support of GAČR (Czech Science Foundation – Grantová Agentura České Republiky) under the project No. 402/08/1237 is kindly announced.

the random variable X. In practice, we usually use some of methods for a probability density function estimation. Perhaps, the simplest method how to approximate the probability density function (shortly PDF) is a method called a *histogram*. Let the suitable interval be divided in M mutually disjoint bins B_j of binwidth h covering the considered interval. The PDF is then given by

$$\hat{f}_h(x) = \frac{1}{Th} \sum_{t=1}^{T} \sum_{j=1}^{M} I(x_t \in B_j) I(x \in B_j), \tag{2}$$

where T is the length of the series of data and $I(\cdot)$ is an indicator function

$$I(A) = \begin{cases} 1, \text{ if } A \text{ holds}, \\ 0, \text{ otherwise.} \end{cases} \tag{3}$$

It is easy to see that the histogram appears to be strongly dependent on h and the PDF is neither smooth nor continuous. Therefore, the most standard approach how to smooth empirical observations of random variable is to apply a *Parzen windows estimator* based on a kernel function $K(y)$:

$$\hat{f}(x) = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{h} K\left(\frac{x_t - x}{h}\right), \tag{4}$$

where h is a bandwith parameter for the window size description. A natural choice for kernel function $K(y)$ is the Gaussian kernel or any kernel from the symmetric Beta family. Note that this approach is a special case of the *Kernel regression* (see [1]). For a comparative study of various probability density estimation methods as well as references to the substantial literature we refer to [5].

3 Discrete F-transform

In this section, we will describe the direct discrete fuzzy transform (shortly, F-transform). The key idea on which the F-transform is based is a fuzzy partition of the universe into fuzzy subsets (factors, clusters, granules etc.). A fuzzy partition may be understood as a system of neighborhoods of some chosen nodes. A sufficient representation of a function we may consider its average values over fuzzy subsets from the partition. Then, a function can be associated with a mapping from a set of fuzzy subsets to the set of thus obtained average function values. We take an interval [a, b] as a universe. That is, all (real-valued) functions considered in this section have this interval as a common domain. The fuzzy partition of the universe is given by fuzzy subsets of the universe [a, b] (determined by their membership functions) which must have properties described in the following definition.

Definition 1 ([6]). Let $x_1 < \cdots < x_n$ be fixed nodes within $[a, b]$, such that $x_1 = a$, $x_n = b$ and $n \geq 2$. We say that fuzzy sets $A_1, \ldots, A_n$, identified with their membership functions $A_1(x), \ldots, A_n(x)$ defined on $[a, b]$, form a fuzzy partition of $[a, b]$ if they fulfill the following conditions for $k = 1, \ldots, n$:

(1) $A_k : [a, b] \to [0, 1]$, $A_k(x_k) = 1$;
(2) $A_k(x) = 0$ if $x \notin (x_{k-1}, x_{k+1})$ where for the uniformity of denotation, we put $x_0 = a$ and $x_{n+1} = b$;
(3) $A_k(x)$ is continuous;
(4) $A_k(x)$, $k = 2, \ldots, n$, strictly increases on $[x_{k-1}, x_k]$ and $A_k(x)$, $k = 1, \ldots, n-1$, strictly decreases on $[x_k, x_{k+1}]$;
(5) for all $x \in [a, b]$

$$\sum_{k=1}^{n} A_k(x) = 1. \tag{5}$$

The fuzzy sets $A_1, \ldots, A_n$ are called *basic functions*.

A set of basic functions which form a fuzzy partition of $[a, b]$ will be denoted by $\mathcal{A}_a^b$ or simply $\mathcal{A}$, when we need not to consider a concrete interval $[a, b]$.

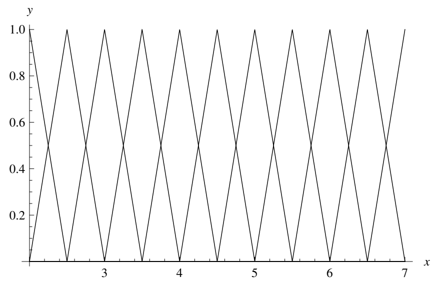
Let $\mathcal{A}_a^b = \{A_1, \ldots, A_n\}$ be a set of basic functions. We say that $\mathcal{A}_a^b$ defines a *uniform* fuzzy partition of $[a, b]$, if $x_{i+1} - x_i = h$ for any $i = 1, \ldots, n-1$, where $h > 0$ is a constant, and there is a symmetric fuzzy set[1] $A : [-h, h] \to [0, 1]$ such that $A(x) = A_i(x_i - x)$ for any $x \in [a, b]$ and $i = 1, \ldots, n$. Obviously, $A(0) = 1$. Since all basic functions may be determined from a one fuzzy set, the uniform fuzzy partitions seem to be very profitable from the practical point of view. Let us show two examples of uniform fuzzy partitions.

[1] It means an even function.

Example 1. Let $A, B : [-h, h] \to [0, 1]$ be fuzzy sets defined as follows:

$$A(x) = \begin{cases} \frac{h+x}{h}, & x \in [-h, 0], \\ \frac{h-x}{h}, & x \in [0, h], \\ 0, & \text{otherwise}, \end{cases} \qquad B(x) = \begin{cases} 0.5\left(\cos\frac{\pi}{h}(x - x_i) + 1\right), & x \in [-h, h], \\ 0, & \text{otherwise}. \end{cases}$$

If $x_1 < \cdots < x_n$ are nodes within $[a, b]$ such that $x_1 = a$, $x_n = b$ and $x_i = x_1 + (i - 1)h$, then $\mathcal{A}_a^b = \{A_i \mid A(x) = A_i(x_i - x),\ i = 1, \ldots, n\}$ and $\mathcal{B}_a^b = \{B_i \mid B(x) = B_i(x_i - x),\ i = 1, \ldots, n\}$ define a *triangle* and a *cosine uniform fuzzy partition of* $[a, b]$, respectively. On fig. 1 we can see the triangle and cosine uniform fuzzy partitions of the interval $[2, 7]$ for $h = 0.5$.



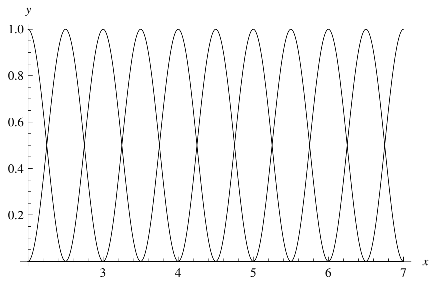


Fig. 1. Triangle and cosine uniform fuzzy partitions

The following lemma shows a useful property holding for the uniform fuzzy partitions.

Lemma 1. *Let $\mathcal{A}_a^b$ be a uniform partition of $[a, b]$ such that $n \geq 3$. Then*

$$\int_a^b A_1(x)dx = \int_a^b A_n(x)dx = \frac{h}{2}, \tag{6}$$

$$\int_a^b A_i(x)dx = h \tag{7}$$

for any $i = 2, \ldots, n - 1$.

Proof. See [6].

Now, we can introduce the definition of the discrete fuzzy transform (shortly discrete F-transform) which assigns, using basic functions, to a discrete function f a vector of real numbers representing the discrete function f. Let $x_1, \ldots, x_l \in [a, b]$ be nodes and $\mathcal{A}_a^b$ be a set of basic functions. We say that the set of nodes $x_1, \ldots, x_l$ is *sufficiently dense with respect to* $\mathcal{A}_a^b$, if for each $A_k \in \mathcal{A}_a^b$ there is a node x_j such that $A_k(x_j) > 0$.

Definition 2 ([6]). Let a function f be given at nodes $x_1, \ldots, x_l \in [a, b]$ and $\mathcal{A}_a^b = \{A_1, \ldots, A_n\}$, $n < l$, be basic functions which form a fuzzy partition of $[a, b]$. We say that the n-tuple of real numbers $[F_{\mathcal{A}_a^b,1}, \ldots, F_{\mathcal{A}_a^b,n}]$ is the discrete F-transform of f with respect to $\mathcal{A}_a^b$ if

$$F_{\mathcal{A}_a^b,k} = \frac{\sum_{j=1}^{l} f(x_j)A_k(x_j)}{\sum_{j=1}^{l} A_k(x_j)}. \tag{8}$$

For more information about the discrete F-transform we refer to [6].

4 FT-smoothing filters

In the following part, we will propose a smoothing filter which is based on the discrete F-transform. Let $\mathbb{R}$ denote the set of all real numbers and X be a subset of $\mathbb{R}$. We say that a function $f : X \to \mathbb{R}$ is *discrete*, if X is a finite set. Let $\mathcal{A}_a^b$ be a set of basic functions which form a fuzzy partition of $[a, b]$. We denote $\mathcal{D}_{\mathcal{A}_a^b}$ the set of all discrete functions f such that $\min(\mathrm{Dom}(f)) = a$ and $\max(\mathrm{Dom}(f)) = b$, $\mathrm{Dom}(f)$ is sufficiently dense with respect to $\mathcal{A}_a^b$ and $|\mathcal{A}_a^b| < |\mathrm{Dom}(f)|$. Obviously, the set $\mathcal{D}_{\mathcal{A}_a^b}$ contains all functions on which the discrete F-transform may be applied. Finally, let $\mathcal{C}([a, b])$ denote the set of all continuous functions.

Definition 3. Let $\mathcal{A}_a^b = \{A_1, \dots, A_n\}$ be a set of basic functions. An *FT-smoothing filter determined by* $\mathcal{A}_a^b$ is a mapping $\mathcal{F}_{\mathcal{A}_a^b} : \mathcal{D}_{\mathcal{A}_a^b} \to \mathcal{C}([a,b])$ defined by

$$\mathcal{F}_{\mathcal{A}_a^b}(f)(x) = \sum_{k=1}^{n} F_{\mathcal{A}_a^b,k} A_k(x) \tag{9}$$

for any $x \in [a,b]$, where $F_{\mathcal{A}_a^b,k}$, $k = 1, \dots, n$, are the components of the discrete F-transform.

Remark 1. One can check easily that the linear combination of continuous functions is a continuous function. Hence, our definition is correct and $\mathcal{F}_{\mathcal{A}_a^b}$ is really a mapping to the set of all continuous functions.

Remark 2. Let us note that our definition of FT-smoothing filter is the inverse F-transform for the continuous case. For the discrete case the resulted function of the inverse F-transform is again a discrete function, more precisely, if f is a discrete function with $\mathrm{Dom}(f) = \{x_1, \dots, x_l\}$, then

$$f_{F_{\mathcal{A}_a^b}}(x_i) = \mathcal{F}_{\mathcal{A}_a^b}(f)(x_i) \tag{10}$$

defines the discrete function which is the inverse F-transform of f in the sense of Definition 5 in [6].

Now, we may ask, if the mapping $\mathcal{F}_{\mathcal{A}_a^b}$ may be considered as a smoothing filter, that means, if it reduces the white noise of the original discrete functions. To give an answer we will investigate the properties of random functions expressing the original functions, where a white noise is supposed. First, however, let us recall a definition of random function. Let $(\Omega, \mathcal{G}, P)$ be a probability space. A function $f : X \times \Omega \to \mathbb{R}$ is called a *random function*, if $\{\omega \mid f(x,\omega) \le y\} \in \mathcal{G}$ holds for any $y \in \mathbb{R}$ and $x \in X$. Obviously, f is a random function if and only if $f(x)$ is a random variable. We say that a random function $f : X \times \Omega \to \mathbb{R}$ is discrete, if X is a finite set. Now, suppose that

$$g(x) = f(x) + \varepsilon(x) \tag{11}$$

is a discrete random function such that f is a discrete function and ε is a discrete random function describing the white noise such that

(i) $\varepsilon(x)$ is a random variable with the probability distribution $N(0, \sigma^2)$ for any $x \in \mathrm{Dom}(f)$,
(ii) $\varepsilon(x), \varepsilon(y)$ are independent random variable for any $x, y \in \mathrm{Dom}(f)$ with $x \ne y$.

Let us note that f is often an unknown discrete function and g is a discrete random function with the probability distribution $N(f(x), \sigma^2)$ for each $g(x)$. Now, we may consider $\mathcal{F}_{\mathcal{A}_a^b}$ as a smoothing filter, if the variance of the probability distribution of $\mathcal{F}_{\mathcal{A}_a^b}(g)(x)$ is smaller than the variance of the probability distribution of the random variable $g(x)$ for any $x \in \mathrm{Dom}(g)$ and $g \in \mathcal{D}$. In another words, the function after applying an FT-smoothing filter has no such outliers that can deform it as the original one.

Theorem 1. *Let $g(x) = f(x) + \varepsilon(x)$ be a discrete random function from (11) such that $\varepsilon(x)$ is a random variable satisfying (i) and (ii), $\mathcal{F}_{\mathcal{A}_a^b}$ be an FT-smoothing filter determined by $\mathcal{A}_a^b$ and $f \in \mathcal{D}_{\mathcal{A}}$. Then*

$$\mathcal{F}_{\mathcal{A}_a^b}(g)(x) = \mathcal{F}_{\mathcal{A}_a^b}(f)(x) + \mathcal{F}_{\mathcal{A}_a^b}(\varepsilon)(x) \tag{12}$$

is a random function defined on $[a,b]$ such that $\mathsf{Var}[\mathcal{F}_{\mathcal{A}_a^b}(\varepsilon)(x)] \le \sigma^2$ for any $x \in [a,b]$.

Proof. Let $g(x) = f(x) + \varepsilon(x)$ be a discrete random function such that ε satisfies (i) and (ii), $\mathrm{Dom}(g) = \{x_1, \dots, x_l\}$, $\mathcal{F}_{\mathcal{A}_a^b}$ be an FT-smoothing filter determined by $\mathcal{A}_a^b = \{A_1, \dots, A_n\}$ and $f \in \mathcal{D}_{\mathcal{A}_a^b}$. First, we will describe the components $G_{\mathcal{A}_a^b,k}$, $k = 1, \dots, n$, of the direct F-transform. It is easy to see that $G_{\mathcal{A}_a^b,k}$, $k = 1, \dots, n$, are random variables. In fact, we can write

$$G_{\mathcal{A}_a^b,k} = \frac{\sum_{i=1}^{l} (f(x_i) + \varepsilon(x_i)) A_k(x_i)}{\sum_{i=1}^{l} A_k(x_i)} = \frac{\sum_{i=1}^{l} f(x_i) A_k(x_i)}{\sum_{i=1}^{l} A_k(x_i)} + \frac{\sum_{i=1}^{l} \varepsilon(x_i) A_k(x_i)}{\sum_{i=1}^{l} A_k(x_i)} = F_{\mathcal{A}_a^b,k} + \mathcal{E}_{\mathcal{A}_a^b,k},$$

where $F_{\mathcal{A}_a^b,k}$ and $\mathcal{E}_{\mathcal{A}_a^b,k}$ are the k-th component of F-transform, where clearly $\mathcal{E}_{\mathcal{A}_a^b,k}$ is a random variable. Hence, $G_{\mathcal{A}_a^b,k}$ is a random variable. Further, the expected value of the random variable $\mathcal{E}_{\mathcal{A}_a^b,k}$ is 0 for any $k = 1, \dots, n$, since

$$\mathsf{E}[\mathcal{E}_{\mathcal{A}_a^b,k}] = \mathsf{E}\left[\frac{\sum_{i=1}^{l} \varepsilon(x_i) A_k(x_i)}{\sum_{i=1}^{l} A_k(x_i)}\right] = \frac{\sum_{i=1}^{l} E[\varepsilon(x_i)] A_k(x_i)}{\sum_{i=1}^{l} A_k(x_i)} = \frac{\sum_{i=1}^{l} 0 \cdot A_k(x_i)}{\sum_{i=1}^{l} A_k(x_i)} = 0,$$

where we use the well-known equality $\mathsf{E}[a_1X_1+\cdots+a_nX_n] = a_1\mathsf{E}[X_1]+\cdots+a_n\mathsf{E}[X_n]$ for any real number $a_1,\ldots,a_n$ and random variable $X_1,\ldots,X_n$. For the covariance of $\mathcal{E}_{\mathcal{A}_a^b,k}$ and $\mathcal{E}_{\mathcal{A}_a^b,k'}$ we have

$$\mathsf{E}[(\mathcal{E}_{\mathcal{A}_a^b,k}-0)(\mathcal{E}_{\mathcal{A}_a^b,k'}-0)] = \mathsf{E}[\mathcal{E}_{\mathcal{A}_a^b,k}\mathcal{E}_{\mathcal{A}_a^b,k'}] = \mathsf{E}\left[\frac{\sum_{i=1}^{l}\sum_{j=1}^{l}\varepsilon(x_i)\varepsilon(x_j)A_k(x_i)A_{k'}(x_j)}{\sum_{i=1}^{l}A_k(x_i)\sum_{j=1}^{l}A_{k'}(x_j)}\right] =$$
$$\frac{\sum_{i=1}^{l}\sum_{j=1}^{l}\mathsf{E}[\varepsilon(x_i)\varepsilon(x_j)]A_k(x_i)A_{k'}(x_j)}{\sum_{i=1}^{l}A_k(x_i)\sum_{j=1}^{l}A_{k'}(x_j)} = \frac{\sum_{i=1}^{l}\mathsf{E}[\varepsilon(x_i)^2]A_k(x_i)A_{k'}(x_i)}{\sum_{i=1}^{l}A_k(x_i)\sum_{j=1}^{l}A_{k'}(x_j)} =$$
$$\frac{\sum_{i=1}^{l}A_k(x_i)A_{k'}(x_j)}{\sum_{i=1}^{l}A_k(x_i)\sum_{j=1}^{l}A_{k'}(x_j)}\sigma^2$$

where the independence of random variables $\varepsilon(x_1),\ldots,\varepsilon(x_l)$ is applied, i.e., $\mathsf{E}[\varepsilon(x_i)\varepsilon(x_j)] = 0$ for $i \neq j$. We can see that the random variables $\mathcal{E}_{\mathcal{A}_a^b,1},\ldots,\mathcal{E}_{\mathcal{A}_a^b,n}$ are not mutually independent in general. Put

$$a_{kk'} = \frac{\sum_{i=1}^{l}A_k(x_i)A_{k'}(x_i)}{\sum_{i=1}^{l}A_k(x_i)\sum_{j=1}^{l}A_{k'}(x_i)}. \tag{13}$$

It is easy to see that $a_{kk'} \leq 1$ and thus the covariance of $\mathcal{E}_{\mathcal{A}_a^b,k}$ and $\mathcal{E}_{\mathcal{A}_a^b,k'}$ (i.e. $\sigma_{kk'} = a_{kk'}\sigma^2$) is smaller than or equal to σ^2. Now, let us apply the FT-smoothing filter $\mathcal{F}_{\mathcal{A}_a^b}$ to g, i.e.,

$$\mathcal{F}_{\mathcal{A}_a^b}(g)(x) = \sum_{k=1}^{n}G_{\mathcal{A}_a^b,k}A_k(x) = \sum_{k=1}^{n}(F_{\mathcal{A}_a^b,k}+\mathcal{E}_{\mathcal{A}_a^b,k})A_k(x) =$$
$$\sum_{k=1}^{n}F_{\mathcal{A}_a^b,k}A_k(x) + \sum_{k=1}^{n}\mathcal{E}_{\mathcal{A}_a^b,k}A_k(x) = \mathcal{F}_{\mathcal{A}_a^b}(f)(x) + \mathcal{F}_{\mathcal{A}_a^b}(\varepsilon)(x).$$

Since $\mathcal{E}_{\mathcal{A}_a^b,k}$, $k = 1,\ldots,n$ is a random variable, then $\mathcal{F}_{\mathcal{A}_a^b}(e)(x)$ is also a random variable for any $x \in [a,b]$ and $\mathcal{F}_{\mathcal{A}_a^b}(g)(x)$ is a random function. Further, it is easy to prove (analogously as in the previous part) that $E(\mathcal{F}_{\mathcal{A}_a^b}(\varepsilon)(x)] = 0$ for any $x \in [a,b]$. For the variance of $\mathcal{F}_{\mathcal{A}_a^b}(\varepsilon)(x)$ we can write

$$\mathsf{Var}[\mathcal{F}_{\mathcal{A}_a^b}(\varepsilon)(x)] = \mathsf{E}[(\mathcal{F}_{\mathcal{A}_a^b}(\varepsilon)(x))^2] = \mathsf{E}[(\sum_{k=1}^{n}\mathcal{E}_{\mathcal{A}_a^b,k}A_k(x))^2] = \mathsf{E}[\sum_{k=1}^{n}\sum_{k'=1}^{n}\mathcal{E}_{\mathcal{A}_a^b,k}A_k(x)\mathcal{E}_{\mathcal{A}_a^b,k'}A_{k'}(x)] =$$
$$\sum_{k=1}^{n}\sum_{k'=1}^{n}\mathsf{E}[\mathcal{E}_{\mathcal{A}_a^b,k}\mathcal{E}_{\mathcal{A}_a^b,k'}]A_k(x)A_{k'}(x) = \sum_{k=1}^{n}\sum_{k'=1}^{n}a_{kk'}\sigma^2A_k(x)A_{k'}(x) =$$
$$\left(\sum_{k=1}^{n}\sum_{k'=1}^{n}a_{kk'}A_k(x)A_{k'}(x)\right)\sigma^2 \leq \left(\sum_{k=1}^{n}\sum_{k'=1}^{n}A_k(x)A_{k'}(x)\right)\sigma^2 = \left(\sum_{k=1}^{n}A_k(x)\right)^2\sigma^2 = \sigma^2,$$

since $\sum_{k=1}^{n}A_k(x) = 1$ (see the presumptions on the basic functions). □

Remark 3. As we proved the components $\mathcal{E}_{\mathcal{A}_a^b,1},\ldots,\mathcal{E}_{\mathcal{A}_a^b,n}$ are mutually dependent in general. Note that this fact is a natural consequence of the procedure to obtain a smooth shape of the resulted function.

5 Illustrative example

As in similar studies, a data set of continuous financial returns (ie. a natural logarithm of discretely observed prices) is considered here. We choose continuous financial returns since they are (relatively) scale-free and since their statistical properties allow us easy handling. Thus, the data set consist of 2017 daily returns of CZK/EUR exchange rate over last eight years.

Now, we can use the FT-smoothing filter to estimate the unknown PDF as follows. First, we create a discrete function $\{(x_i, f(x_i)) \mid i = 1,\ldots,M\}$, where x_i are the centers of bins and

$$f(x_i) = \frac{1}{T}\sum_{t=1}^{T}I(x_t \in B_i). \tag{14}$$

Further, we apply FT-smoothing filter determined by a uniform fuzzy partition on the discrete function $f(x_i)$. On the left side of Fig. 2 we can see the smooth functions $\mathcal{F}_{\mathcal{A}_a^b}(f)(x)$ for the triangle and cosine uniform fuzzy partitions. Since the integral is not equal to 1, then, according to Lemma 1, the normalization is defined by

$$\widehat{\mathcal{F}_{\mathcal{A}_a^b}(f)}(x) = \frac{\mathcal{F}_{\mathcal{A}_a^b}(f)(x)}{\frac{h}{2}(F_{\mathcal{A}_a^b,1} + F_{\mathcal{A}_a^b,n}) + h\sum_{i=2}^{n-1} F_{\mathcal{A}_a^b,i}}. \qquad (15)$$

The PDF estimation for the financial returns can be seen on the right side of Fig. 2.

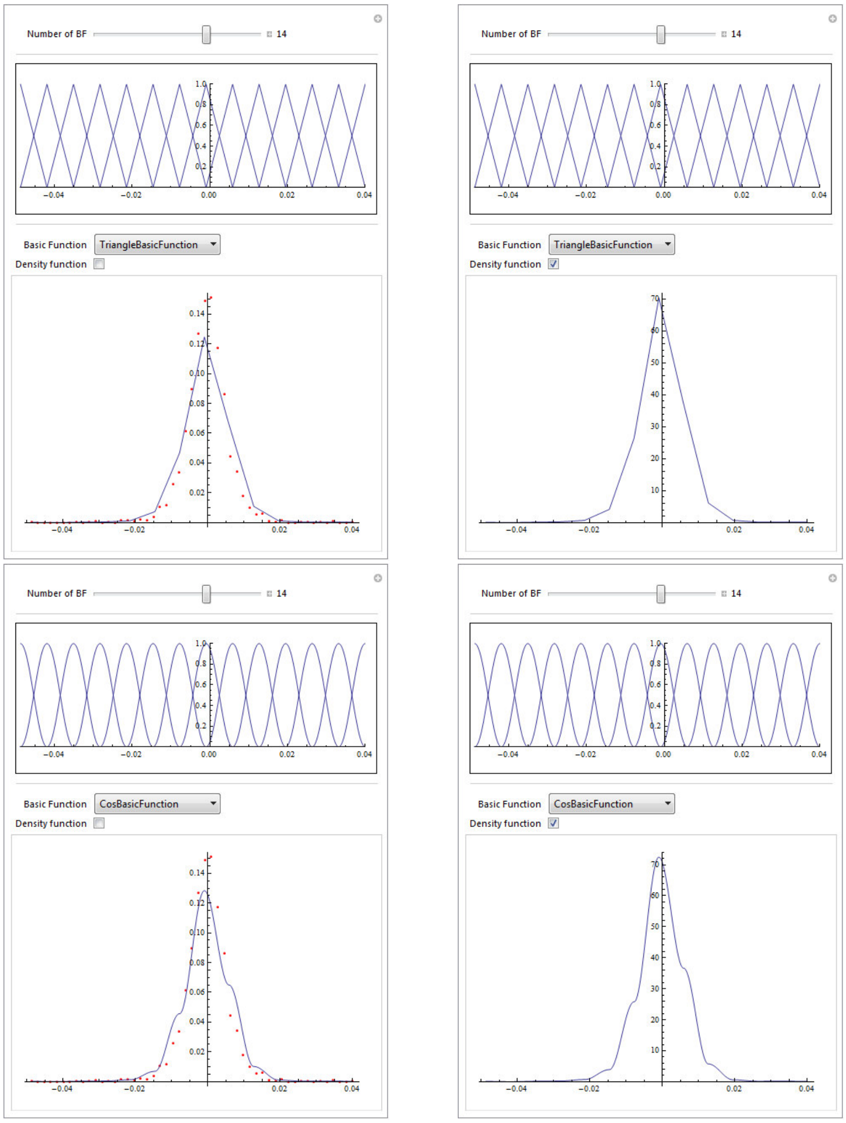


Fig. 2. FT-smoothing of PDF for financial returns based on a triangle and fuzzy partition

6 Conclusions

In this paper we have proposed an alternative approach to density function estimation. The application on financial market returns was presented. The approach based of discrete fuzzy-transform filtering can be of great value for financial modeling and forecasting. The future research should be devoted to comparison of the efficiency for various data, as well as the time costs of this method as compared to more standard approaches.

References

1. J. Fan and Q. Yao. *Nonlinear Time Series: Nonparametric and Parametric Methods.* Springer Series in Statistics. Springer-Verlag, 2005.
2. E.J. Kostelich and J.A. Yorke. Noise reduction: Finding the simplest dynamical system consistent with the data. *Physica*, D 41:183–196, 1990.
3. Hui-Hsiung Kuo. *White noise distribution theory.* CRC Press, 1996.
4. M. Nachtegael, D. Van der Weken, D. Van De Ville, and Kerre E.E., editors. *Fuzzy Filters for Image Processing.* Studies in Fuzziness and Soft Computing. Springer-Verlag, 2003.
5. A. Assenza, M. Valle, and M. Verleysen. A comparative study of various probabilty density estimation methods for data analysis. *Int. J. of Computational Intelligence Systems*, 1(2):188–201, 2009.
6. I. Perfilieva. Fuzzy transforms: Theory and applications. *Fuzzy sets and systems*, 157(8):993–1023, 2006.

AN EXAMPLE OF A NON-WORKING BOSS INCREASING EXPECTED OUTOUT

Milan Horniaček

Comenius University in Bratislava
Faculty of Social and Economic Sciences, Institute of Economics,
milan.horniacek@fses.uniba.sk

Abstract. We are motivated by the following feature of equilibria of infinitely countably repeated partnership games with discounting of future payoffs: the expected output is not maximized after some non-terminal histories that occur with a positive probability in equilibrium. This raises the question whether subordinating the partners to a non-working principal (i.e. transforming the partnership into a corporation) can change the outcome. We give an example of a game between the principal and two agents (former partners) in which the answer to the above question is affirmative. A weakly renegotiation-proof public equilibrium is the solution concept that we apply. It results from the application of Farrell's and Maskin's concept of a weakly renegotiation-proof equilibrium to a perfect public equilibrium. We show that for discount factors close enough to one the game has a weakly renegotiation-proof public equilibrium with the following properties: the sum of players' equilibrium payoffs equals the maximal sum of their feasible payoffs and the expected output is maximized after each non-terminal history that occurs with a positive probability in equilibrium and the agents choose their actions after it. Thus, the equilibrium payoff vector is strictly Pareto efficient with respect to the set of all feasible payoff vectors in the repeated game.
Keywords. Repeated game, principal-agent game with two agents, perfect public equilibrium, weakly renegotiation-proof equilibrium, maximal expected output.

1. Introduction

Equilibria of infinitely countably repeated partnership games (of two partners) with discounting of future payoffs have an undesirable property. The expected output is not maximized after some non-terminal histories that occur with a positive probability in equilibrium. (This holds unless the stage game has a strategy profile that maximizes the expected output and at most one partner can increase his stage game payoff by a unilateral deviation).

A perfect public equilibrium (henceforth PPE), which the authors of [2] analyze, is the best known type of equilibrium in these games. Nevertheless, sequential equilibria, in which players' actions can depend on their private histories, also suffer from the above mentioned shortcoming. Kandori's result [3] is not relevant here because it is formulated for games with at least four players.

This undesirable property of equilibria stems from their two features resulting from the fact that the partnership game is a game with two-sided moral hazard. First, as the identity of a deviator cannot be established, a punishment has to harm both partners. Thus, continuation equilibrium payoff vectors in subgames, in which a suspected deviation is being punished, are not (even weakly) Pareto efficient. Moreover, the partners' expected joint output is not maximized when a punishment takes place. Second, an effect of a deviation cannot be distinguished from an effect of random factors on the partners' joint output. Therefore, the punishment has to be triggered with a positive probability even when nobody deviates. Thus, non-terminal histories, after which it is triggered, occur with a positive probability in equilibrium. The latter feature cannot be eliminated when the moral hazard applies to at least one player. The former feature cannot be eliminated when – as in partnership games - the moral hazard applies to all players. An interesting question is whether it can be eliminated when there is at least one player to whom the moral hazard does not apply. In particular, with respect to the partnership game, a natural question is whether subordinating the former partners to a non-working principal – i.e. transforming the partnership into a corporation – can eliminate the former feature. We give an example of a game between the principal and two agents (former partners) in which the answer to the above question is affirmative.

We apply the solution concept called a 'weakly renegotiation-proof public equilibrium' (henceforth WRPPE) to the analyzed example**.** It results from the application of Farrell's and Maskin's [1] concept of a weakly renegotiation-proof equilibrium to a PPE. We take into account that in our model the stage game is an extensive form non-cooperative game and all subgames in the repeated game are not identical.

We show that, for discount factors close enough to one, the analyzed example has a WRPPE that has two additional desirable properties. First, the sum of equilibrium payoffs equals the maximal sum of feasible payoffs in the repeated game. Thus, the equilibrium payoff vector is strictly Pareto efficient. Second, the expected joint output of the agents equals the maximal possible one after each non-terminal history that occurs with a positive probability in equilibrium and the agents choose their actions after it.

2. The analyzed example

2.1.1. The stage game

The stage game G is an extensive form non-cooperative game with two phases. There are three players: the principal - player 0 and two agents – players 1 and 2. The sets of agents' feasible actions when they work for the principal are $A_1 = A_2 = \{0,1\}$. Action 1 corresponds to working and 0 to shirking. The costs of these actions (in terms of disutility caused by taking them) are $c_1(1) = c_2(1) = 1.1$ and $c_1(0) = c_2(0) = 0$. The agents can also take the outside option (i.e. not to work for the principal), denoted by -1. It gives them zero expected payoff. We let $\bar{A}_i = \{-1\} \cup A_i$ for each $i \in \{1,2\}$, $A = A_1 \times A_2$, and $\bar{A} = \bar{A}_1 \times \bar{A}_2$. An agent's action when he works for the principal can be observed neither by the principal nor by the other agent. The taking of the outside option by an agent is observed both by the principal and by the other agent.

The set of feasible levels of a joint output of the agents, expressed in pecuniary terms, is $Y = \{10,5,0\}$. The probability distributions on Y conditional on a profile of agents' actions are

$$\pi((1,1)) = (\pi(10|(1,1)), \pi(5|(1,1)), \pi(0|(1,1))) = (0.4,0.4,0.2),$$
$$\pi((1,0)) = \pi((1,-1)) = \pi((0,1)) = \pi((-1,1)) = (0.39,0.01,0.6),$$
$$\pi((0,0)) = \pi((-1,-1)) = \pi((0,-1)) = \pi((-1,0)) = (0,0,1).$$

The reward scheme is a function $q : Y \to [0,10]^2$ with the property that $q_1(y) + q_2(y) \le y$ for each $y \in Y$. Here $q_i(y)$ is the pecuniary reward that the principal pays to agent $i \in \{1,2\}$ when i does not take the outside option and the output is y. We denote the set of all reward schemes by Q.

In the first phase of G the principal announces the reward scheme. This announcement is observed by both agents. In the second phase each agent $i \in \{1,2\}$ chooses an action $a_i \in \bar{A}_i$. Let $L(a) = \{i \in \{1,2\} | a_i \neq -1\}$. The principal's expected payoff equals $\sum_{y \in Y} [y - \sum_{i \in L(a)} q_i(y)] \pi(y|a)$. The expected payoff of agent $i \in L(a)$ equals $\sum_{y \in Y} q_i(y) \pi(y|a) - c_i(a_i)$. The expected payoff of agent $i \notin L(a)$ equals zero.

Note that action profile $a^* = (1,1)$ is the only action profile that gives the maximal expected output (equal to 6) and the maximal difference $\sum_{y \in Y} y\pi(y|a) - c_1(a_1) - c_2(a_2)$ on A (equal to 3.8). The sum of players' expected payoffs in G equals $\sum_{y \in Y} y\pi(y|a) - \sum_{i \in L(a)} c_i(a_i)$. Taking of the outside option has the same effect on the expected output as shirking. Therefore, a^* maximizes the sum of expected payoffs in G.

2.2. The repeated game

The repeated game, $\Gamma(\delta)$, is an infinite countable repetition of G with discounting of future payoffs of all players by a discount factor $\delta \in (0,1)$. Payoffs in it are expressed as expected discounted average payoffs over its time horizon. Thus, the maximal sum of players' payoffs in $\Gamma(\delta)$ equals to the maximal sum of their payoffs in G.

A public history leading to phase $k \in \{1,2\}$ of period $t \in N$ contains reward scheme announced by the principal, identities of agents taking the outside option, and output from the beginning of the game up to (but excluding) phase k of period t. We denote the set of such public histories by $H^{(t,k)}$ and let $H = \bigcup_{t\in N}\left(H^{(t,1)} \cup H^{(t,2)}\right)$. $H^{(1,1)}$ contains only the initial (empty) public history. A private history of agent $i \in \{1,2\}$ leading to (the beginning of) period $t \in N$ contains his action in each period $\tau \in \{1,\ldots,t-1\}$ in which he did not take the outside option. We denote the set of such private histories by $H_i^{(t)}$. For each $i \in \{1,2\}$, $H_i^{(1)}$ contains only i's initial (empty) private history.

In our example we restrict attention to pure strategies in $\Gamma(\delta)$. Nevertheless, allowing behavioral strategies (conditioning actions on signals of a private or the public randomizing device) would not change our results. A principal's pure strategy is a function $s_0 : \bigcup_{t\in N} H^{(t,1)} \to Q$. We denote the set of her pure strategies by S_0. A pure strategy of agent $i \in \{1,2\}$ is a function $s_i : \bigcup_{t\in N}\left(H^{(t,2)} \times H_i^{(t)}\right) \to \bar{A}_i$. We denote the set of pure strategies of agent $i \in \{1,2\}$ by S_{i+} and let $S_+ = S_0 \times S_{1+} \times S_{2+}$. Function $\gamma_i : S_+ \to \Re$ assigns to each profile of pure strategies expected discounted average payoff of player $i \in \{0,1,2\}$ in $\Gamma(\delta)$. We define function $\gamma : S_+ \to \Re$ by $\gamma(s) = (\gamma_0(s), \gamma_1(s), \gamma_2(s))$.

A strategy $s_i \in S_{i+}$ of agent $i \in \{1,2\}$ is public if it does not depend on his private histories. That is, s_i is public if $s_i(h,h_i') = s_i(h,h_i'')$ for each $t \in N$, each $h \in H^{(t,2)}$, and each $(h_i', h_i'') \in \left(H_i^{(t)}\right)^2$. We denote the set of public strategies of agent $i \in \{1,2\}$ by S_i and let $S = S_0 \times S_1 \times S_2$.

For each $t \in N \setminus \{1\}$ and each $h \in H^{(t,1)}$ we denote by h^- the subhistory of h leading to the first phase of period $t-1$. For each $t \in N$ and each $h \in H^{(t,2)}$ we denote by h^- the subhistory of h leading to the first phase of period t. For each $h \in \left(\bigcup_{t\in N\setminus\{1\}} H^{(t,1)}\right) \cup \left(\bigcup_{t\in N} H^{(t,2)}\right)$ let $q(h) = (q_1(h), q_2(h))$ be the latest principal's announcement of a reward scheme contained in h. For each $h \in \bigcup_{t\in N\setminus\{1\}} H^{(t,1)}$ let $y(h)$ be the latest output contained in h. For each $t \in N \setminus \{1\}$ and each $h \in H^{(t,1)}$ let $I(h)$ be the set of agents who took the outside option in period $t-1$.

We denote by $\Gamma_{(h)}(\delta)$ the public subgame of $\Gamma(\delta)$ following non-terminal history h. It is a class of 'standard' subgames (that follow a profile of histories formed by a public history and a private history for each agent). When $h \in H^{(t,k)}$ then $\Gamma_{(h)}(\delta)$ contains the subgame following (h, h_1', h_2') for each $(h_1', h_2') \in H_1^{(t)} \times H_2^{(t)}$. We use the subscript (h) in symbols denoting restrictions of sets and functions already defined for $\Gamma(\delta)$ to $\Gamma_{(h)}(\delta)$.

2.3. The solution concept

DEFINITION 1: A profile of public strategies $s^* \in S$ is a *PPE of* $\Gamma(\delta)$ if there does not exist $i \in \{0,1,2\}, h \in H$, and $s_i \in S_{i(h)}$ such that $\gamma_{i(h)}\left(s_i, s^*_{-i(h)}\right) > \gamma_{i(h)}\left(s^*_{(h)}\right)$.

In Definition 1 we consider only deviations to public strategies. Nevertheless, it is easy to see that an agent, who can increase his payoff in (each subgame contained in) a public subgame by a deviation to a non-public strategy, can do so also by a deviation to a public strategy. (This follows from two facts. First, the other agent uses a public strategy in a PPE. Second, private histories are payoff irrelevant. They affect neither the sets of players' future feasible actions nor their payoff consequences.) Thus, a PPE is immune also to unilateral deviations to strategies that are not public.

DEFINITION 2: A strategy profile $s^* \in S$ is a *WRPPE of* $\Gamma(\delta)$ if it is a PPE of $\Gamma(\delta)$ and it satisfies the following conditions:

(i) There does not exist $h \in \bigcup_{t\in N} H^{(t,1)}$ and $h' \in \left(\bigcup_{t\in N} H^{(t,1)}\right) \setminus (h)$ such that $\gamma_{i(h)}\left(s^*_{(h)}\right) > \gamma_{i(h')}\left(s^*_{(h')}\right)$ for each $i \in \{0,1,2\}$.

(ii) There does not exist $h \in \bigcup_{t\in N} H^{(t,2)}$ and $h' \in \left(\bigcup_{t\in N} H^{(t,2)}\right) \setminus (h)$ such that $q(h) = q(h')$ and $\gamma_{i(h)}\left(s^*_{(h)}\right) > \gamma_{i(h')}\left(s^*_{(h')}\right)$ for each $i \in \{0,1,2\}$.

Farrell and Maskin [1] study an infinite countable repetition of a strategic form non-cooperative game with perfectly observable actions and with discounting of stage game payoffs. In such game all subgames are identical. They satisfy two conditions. First, their game forms coincide. Second, the same strategy profile leads to the same payoff vector in all of them. Therefore, they define a weakly renegotiation-proof equilibrium as a subgame perfect equilibrium in which no continuation equilibrium payoff vector is strictly Pareto dominated by another continuation equilibrium payoff vector. In our example the stage game is an extensive form non-cooperative game. Public subgames starting in different phases of a period have different game forms. In public subgames starting in the second phase of a period the second condition is satisfied only if the latest principal's announcement of a reward scheme preceding them is the same. This is reflected in the formulation of Definition 2.

3. The result

PROPOSITION: *There exists* $s^* \in S$ *with the following properties: (i) it is a WRPPE of* $\Gamma(\delta)$ *for each* $\delta \in [0.95,1)$*, (ii)* $\sum_{i=0}^{2} \gamma_i\left(s^*\right) = 3.8$*, and (iii) the expected joint output of the agents in the first period of* $\Gamma_{(h)}(\delta)$ *equals* 6 *for each* $h \in \bigcup_{t\in N} H^{(t,2)}$ *that occurs with a positive probability when the players follow* s^*.

PROOF: *Description of* s^*. Let $\bar{q}_1(y) = \bar{q}_2(y) = 0.5y$ for each $y \in Y$, $\underline{q}_1(y) = \underline{q}_2(y) = 0$ for each $y \in Y$, $\bar{q} = \left(\bar{q}_1, \bar{q}_2\right)$, and $\underline{q} = \left(\underline{q}_1, \underline{q}_2\right)$. We define s^* by induction on the length of public histories. We set $s^*_0(\varnothing) = \bar{q}$. For each $h \in H^{(1,2)}$ and each $i \in \{1,2\}$ we have $s^*_i(h) = 1$ if $q(h) = \bar{q}$, otherwise $s^*_i(h) = -1$. Next suppose that there is $\tau \in N \setminus \{1\}$ such that we have already defined $s^*_0(h)$ for each $h \in \bigcup_{t=1}^{\tau-1} H^{(t,1)}$ and $s^*_i(h)$ for each $i \in \{1,2\}$ and each $h \in \bigcup_{t=1}^{\tau-1} H^{(t,2)}$. Take $h \in H^{(\tau,1)}$. We set $s^*_0(h) = \bar{q}$ if $I(h) = \varnothing$, $y(h) \neq 0$, and either $q(h) = s^*_0(h^-)$ or $q(h) = \bar{q} \neq s^*_0(h^-)$. If $q(h) = s^*_0(h^-)$ or $q(h) = \bar{q} \neq s^*_0(h^-)$, and $I(h) \neq \varnothing$ then $s^*_{0i}(h) = \bar{q}_i$ for each $i \in \{1,2\} \setminus I(h)$ and $s^*_{0i}(h) = \underline{q}_i$ for each $i \in I(h)$. We set $s^*_0(h) = \underline{q}$ if $I(h) = \varnothing$, $y(h) = 0$, and either $q(h) = s^*_0(h^-)$ or $q(h) = \bar{q} \neq s^*_0(h^-)$. If $q(h) \neq s^*_0(h)$ and $q(h) \neq \bar{q}$ then $s^*_{0i}(h) = \bar{q}_i$ for each $i \in I(h)$ and $s^*_{0i}(h) = \underline{q}_i$ for each $i \in \{1,2\} \setminus I(h)$. For each $h \in H^{(\tau,2)}$ and each $i \in \{1,2\}$ we set $s^*_i(h) = 1$ if either $q(h) = s^*_0(h^-)$ or $q(h) = \bar{q} \neq s^*_0(h)$, otherwise $s^*_i(h) = -1$. Note that $s^*_1(h) = s^*_2(h) = 1$ for each $h \in \bigcup_{t\in N} H^{(t,2)}$ occurring with a positive probability when the players follow s^*. Therefore, the expected joint output of the agents in the first period of $\Gamma_{(h)}(\delta)$ equals 6.

Continuation payoffs. Let $v = \gamma(s^*)$. Denote by w_i be the continuation payoff of agent $i \in \{1,2\}$ in a public subgame in the first period of which he faces zero reward for each output. Let w_{02} (w_{01}) be the principal's continuation payoff in a public subgame in the first period of which both gents face (one agent faces) zero reward for each output. We have

$$w_1 = w_2 = (1-0.2\delta)^{-1}\left[-1.1(1-\delta)+0.8\delta v_j\right] \ \& \ w_{02} = (1-0.2\delta)^{-1}\left[6(1-\delta)+0.8\delta v_0\right], \tag{1}$$

$$v_1 = v_2 = 3(1-0.2\delta)-1.1 \ \& \ v_0 = 1.2\delta, \ w_{01} = 3(1-\delta)+0.8\delta v_0 + 0.2\delta w_{02}. \tag{2}$$

Note that

$$v_0 + v_1 + v_2 = w_{02} + w_1 + w_2 = w_{01} + v_1 + w_2 = w_{01} + w_1 + v_2 = 3.8. \tag{3}$$

Consider the principal single period expected payoff when the profile of agents' actions is $(1,1)$. It equals 6 for reward scheme $\underline{q}$, 3 for reward schemes $\left(\bar{q}_1, \underline{q}_{-2}\right)$ and $\left(\underline{q}_{-1}, \bar{q}_2\right)$, and zero for reward scheme $\bar{q}$. Therefore, we have $w_{02} > w_{01} > v_0$. Note that $\delta v_i > w_i$ for each $i \in \{1,2\}$ and each $\delta \in (0,1)$.

Unilateral single period deviations by agents. $\Gamma(\delta)$ is a game continuous at infinity. Thus, we can apply the single period deviation principle in the analysis of unilateral deviations. First suppose that $i \in \{1,2\}$ deviates from working to shirking when he faces zero reward for each output. Such deviation does not increase his continuation payoff if and only if

$$0.1.\delta(1-\delta)+\delta^2(0.8v_i+0.2w_i) \le -1.1(1-\delta)+1.3\delta(1-\delta)+\delta^2(0.8v_i+0.2w_i). \tag{4}$$

This inequality holds for each $\delta \in [11/12,1)$, so it holds for each $\delta \in [0.95,1)$. Next suppose that $i \in \{1,2\}$ deviates from working to shirking when he faces $\bar{q}_i$. Such deviation brings lower single period expected gain than the previously considered one. (It decreases the expected reward from the principal from 3 to 1.975.) Nevertheless, the probability of the punishment is the same. Thus, it cannot increase the deviator's continuation payoff. A deviation from working to taking of the outside option does not increase a single period expected gain in comparison with the deviation to shirking. However, it increases the probability of the punishment from 0.6 to 1. Thus, it cannot increase the deviator's continuation payoff. A deviation by $i \in \{1,2\}$ from taking of the outside option to working gives a single period expected gain bounded from above by $0.39 \times 10 + 0.01 \times 5 - 1.1 = 2.85$.

Therefore, it does not increase the deviators continuation payoff if and only if

$$2.85(1-\delta)-1.1\delta(1-\delta)+\delta^2(0.8v_i+0.2w_i) \le 1.9\delta(1-\delta)+\delta^2(0.8v_i+0.2w_i). \tag{5}$$

This inequality holds for each $\delta \in [0.95,1)$.

Principal's unilateral single period deviations. A deviation from $s_0^*(h)$ to $q \neq \bar{q}$ gives the principal continuation payoff $\delta v_0 < v_0 \leq \gamma_{0(h)}\left(s_{(h)}^*\right)$. A deviation from $s_0^*(h) \neq \bar{q}$ to $\bar{q}$ gives her continuation equilibrium payoff $v_0 < \gamma_{i(h)}\left(s_{(h)}^*\right)$. Thus, a deviation decreases her continuation payoff.

Strategy profile s^* is a WRPPE. For each $h \in \bigcup_{t\in N} H^{(t,1)}$ (each $h \in \bigcup_{t\in N} H^{(t,2)}$ with $q(h) = \bar{q}$) $\gamma_{(h)}\left(s_{(h)}^*\right)$ is strictly Pareto efficient. For each $h \in \bigcup_{t\in N} H^{(t,2)}$ with $q_i(h) \notin \left\{\bar{q}_i, \underset{-i}{q}\right\}$ for at least one $i \in \{1,2\}$ we have $\gamma_{(h)}\left(s_{(h)}^*\right) = \delta v$. Thus, it does not violate condition (ii) in Definition 2. Finally, consider $q \in Q \setminus \left\{\bar{q}\right\}$ with $q_i \in \left\{\bar{q}_i, \underset{-i}{q}\right\}$ for each $i \in \{1,2\}$, $h \in \bigcup_{t\in N} H^{(t,2)}$ with $q(h) = q = s_0^*(h^-)$, and $h' \in \bigcup_{t\in N} H^{(t,2)}$ with $q(h) = q \neq s_0^*((h')^-)$. Then there is $j \in \{1,2\}$ with $q_j = \underset{-j}{q}$, $\gamma_{0(h)}\left(s_{(h)}^*\right) > v_0 > \delta v_0 = \gamma_{0(h')}\left(s_{(h')}^*\right)$, and $\gamma_{j(h')}\left(s_{(h')}^*\right) = \delta v_j > w_j = \gamma_{j(h)}\left(s_{0(h)}^*\right)$. Thus, neither $\gamma_{(h)}\left(s_{(h)}^*\right)$ strictly Pareto dominates $\gamma_{(h')}\left(s_{(h')}^*\right)$ nor vice versa. *Q.E.D.*

Acknowledgement: Research reported in this paper was supported by the grant VEGA 1/0138/09.

References

[1] FARRELL, J. and E. MASKIN *Renegotiation in Repeated Games.* Games and Economic Behavior, 1, 1989, p. 327-360.

[2] FUDENBERG, D., D. LEVINE, and E. MASKIN *The Folk Theorem with Imperfect Public Information.* Econometrica, 62, 1994, p. 997-1039.

[3] KANDORI, M. *Randomization*, Communication, and Efficiency in Repeated Games with Imperfect Public Monitoring. Econometrica*,* 71, 2003, p. 345-353.

FACTORS DETERMINING STUDENTS' WORKLOAD WITHIN EUROPEAN TRANSFER AND ACCUMULATION SYSTEM (ECTS)

Vladimíra Hovorková Valentová, Jiří Rozkovec

Faculty of Economics
Department of Economic Statistics
Technical University of Liberec
Studentská 2, 46117 Liberec
email: vladimira.valentova@tul.cz,jiri.rozkovec@tul.cz

Abstract. European Credit Transfer and Accumulation System (ECTS) is known to be a credit system for higher education that was adopted within the Bologna Process in June 1999. Thirty one countries have adopted ECTS by law for their higher education systems. ECTS is a student-centred system for credit accumulation and transfer that aims to facilitate planning, delivery, evaluation, recognition and validation of qualifications and educational programmes. The whole ECTS consists of several parts. The authors of the paper have focused on the part which is concerned on allocation of credits to courses of study programmes. The aim of the research was to find factors influencing on students´ workload which is ECTS based on. Another aim of the research was to recognize the strength of the influence of these factors. The students´ workload is supposed to be from 1500 to 1800 hours for an academic year. At once the paper contains a list of factors that could have any influence on the students´ workload and gets in an example of application of these factors on the planned new study programme Economics and Management, specialization Economics and Management of Services of Faculty of Economics, Technical University of Liberec. The calculated credits are based on Saaty matrix of the workload factors and the weights derived from this matrix. The weights of factors together with the particular values for each taught subject produced relative difficulties of subjects, which were modified to give the total number of credits demanded by the study programme.
Keywords. Bologna Process, European Credit Transfer and Accumulation System (ECTS), factor, multicriteria decision making, Saaty matrix, workload.

1 Introduction

It is more than twenty years when the European Credit Transfer System was set up as a pilot scheme within the framework of the Erasmus programme. The first Czech universities have begun to implement this system to their study programmes since 1991. The process of implementation of the current version of European Credit Transfer and Accumulation System (ECTS) has been proceeding at the Technical University of Liberec (follows as TUL) as well. Activities concerning with ECTS are presently carrying out within the university development project focused on the implementation of ECTS into all study programmes of all the faculties of TUL.

1.1 What is ECTS ?

ECTS is one of the most important ideas of the Bologna process. The Bologna Declaration of 19 June 1999 is a joint declaration of the European Ministers of Education. The main points of the Bologna Declaration [9] are:

- adoption of a system of easily readable and comparable degrees,
- adoption of a system essentially based on the two main cycles, undergraduate and graduate,
- establishment of a system credits,
- promotion of mobility for students, for teachers, researches and administrative staff,
- promotion of European co-operation in quality assurance,
- promotion of the necessary European dimensions in higher education.

Today the Bologna process has 46 signatory countries. The Czech Republic belongs to them.

ETCS is defined as a student-centred system based on the student workload which is needed for achieving of expected outcomes of study. The outcomes of study describe what a student is expected to know, understand and be able to do after successful completion of a study of a certain subject. ETCS makes study programmes easy to read and compare and its advantage is that can be used for all types of programmes. It has benefits for the both, mobile and non-mobile students. It can be used for credits transfer between institutions and for credits accumulation within a home institution.

ECTS is based on the principle that every full-time student has to obtain 60 credits during one academic year. These 60 credits measure the student workload. The workload of a full-time student is supposed to be mostly around 1500 – 1800 hour per year. It results in the conclusion that one credit stands for around 20 to 30 working hours.

In ECTS students can obtain credits only after a successful completion of the learning activities required by their study programmes. They have to achieve appropriate learning outcomes. The use of learning outcomes facilitates students to understand what they have to do to obtain ETCS credits from any subject and makes the comparison of qualifications easier. Credits are allocated to study programmes as well as their educational components. The term educational components means courses, dissertation work, placement, laboratory work etc.

1.2 Student Workload

Student workload is one of the most important terms for our research presented in this paper. Student workload is expressed in time that students typically need to complete all the activities connected with a study of a certain subject. It includes e.g. lectures, seminars, practical work, private study, examinations, semester projects etc. The estimation of workload must not be based on contact hours only. And another wrong way how to allocate credits is to link it to the status of a course or the prestige of a teacher. According to main ECTS documents the allocation of credits should be based on an estimation of the real student workload that average student needs for achieving learning outcomes.

As ECTS users´ guide mentions [4] the university project "Tuning Educational Structures in Europe", supported by the European Commission in the framework of the Socrates programme, identifies a four steps approach for determining student workload:

1. Introducing modules/course units.
   The system can be modularized and non-modularized. The difference between these two systems consists in a different number of credits that each course can have. In a modularized system the course has a fixed workload. On the other hand, in a non-modularized system each course unit can have the different number of credits.
2. Estimating student workload.
   Teachers estimate the time that students need to complete their learning activities. The learning activities can be defined by considering the aspects such as types of courses (lecture, laboratory work, seminar, project work, etc.), types of learning activities (attending lectures, writing papers, practising special skills, etc.) and types of assessment (oral or written examination, thesis, oral presentation, etc.).
3. Checking the estimated workload through student evaluations.
   It is necessary to check if a teacher estimated student workload correctly. There are a few methods how to do it. The most common method is a questionnaire survey. Students are asked about the time spent on all the activities connected with the completion of the course.
4. Adjustment of workload and/or educational activities.
   Generally, it is necessary to adjust workload and/or educational activities anyway when the survey shows that estimated student workload does not correspond to the actual workload.

1.3 Base of the Research

In our opinion the current criteria for credit assignment mentioned above are very subjective and not able to describe a real difficulty of any course perfectly. Moreover, we spent a lot of time over the discussion about meaning of the terms "average student" or "typical student". How to recognise an "average student" ? Is it possible to measure it ? Is an "average student" same in 2000 like in 2009 ? The ECTS Users´ Guide [4] mentions: "However, basing a programme on a reasonable and realistic estimate of the time required by an average learner protects all students from unrealistic and overloaded programmes or from excessively light and undemanding ones." It sounds logically but nobody has showed how the "average student" looks like. It is a considerably abstract term. In reference to these facts we decided to supplement the current criteria for determining student workload with others which allow estimate the workload more exactly. We tried to find the most important items which influence the student workload but at the same time do not depend on the students´ quality.

2 Calculation Method

In March 2009 TUL, Faculty of Economics, proposed for accreditation a new study programme *Economics and Management*, the specialization *Economics and Management of Services*. The proposal included credits of all courses. These credits are usually derived from the data like amounts of hours of lectures, hours of practising and studying hours for written or oral exam, which are stated by tutors in a special question-form. At TUL the broad issue of credits, their correction respectively, has been already discussed for a long time. On that occasion we decided to develope a new and more objective calculation method. Our proceeding was composed of the following steps:

1. to set new criteria (factors) and create a new question-form
2. to compare the criteria and to set the weights of factors
3. to collect the data (filled question forms) from tutors
4. to calculate the relative difficulties of all the courses
5. to modify them to integers according the required total of the credits

2.1 Workload Factors

As mentioned above, we considered the current question-form not to be fully objective, so we prepaired a new form with new criteria - *workload factors*. The number of criteria we will further denote k.

These new factors are not dependent on students' quality or the estimates of amounts, so they should be more objective. We decided to evaluate the courses of the selected study programme from the following points of view:

1. amount of homeworks (number from [0; 1] - it means how often students get any homework, whereas it is assumed that a typical homework needs 30-60 minutes. For example, if it equals to 0.50, on average the homework is enjoined every second week of semester.),
2. the number of semestral works,
3. the number of tests,
4. whether there is a written exam,
5. whether there is an oral exam (obviously the both are also possible),
6. if some courses are required as precourses,
7. handling with any special tools (specialized software etc.),
8. the total time of lectures per week,
9. the total time of practising per week.

2.2 Weights of Factors

Now we estimated the weights of factors by *Saaty's method*. We constructed Saaty matrix $S_{(k\times k)}$ (Table 1) with elements s_{ij}, $i, j = 1, \dots, k$, where $s_{ji} = \frac{1}{s_{ij}}$ (details in [?], [?]), its values emerged from a discussion with our collagues and students.

Criteria	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	v_i	w_i (%)
Homeworks	**1**	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{8}$	$\frac{1}{9}$	1	3	3	2	0.61	4
Semestral work	4	**1**	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{7}$	3	2	5	4	1.12	8
Credit test	5	4	**1**	$\frac{1}{4}$	$\frac{1}{6}$	3	5	7	6	2.01	14
Written exam	8	6	4	**1**	$\frac{1}{4}$	5	6	7	6	3.40	23
Oral exam	9	7	6	4	**1**	8	9	9	8	5.83	39
Precourses	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{5}$	$\frac{1}{8}$	**1**	2	2	2	0.66	4
Special tools	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{9}$	$\frac{1}{2}$	**1**	2	1	0.44	3
Lectures amount	$\frac{1}{3}$	$\frac{1}{5}$	$\frac{1}{7}$	$\frac{1}{7}$	$\frac{1}{9}$	$\frac{1}{2}$	$\frac{1}{2}$	**1**	$\frac{1}{3}$	0.29	2
Practising amount	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{2}$	1	3	**1**	0.44	3

Table 1

From this matrix we calculated the weights v_i of all criteria. To simplify we used a geometric mean of the values on every row

$$v_i = (\prod_{j=1}^{k} s_{ij})^{\frac{1}{k}}. \tag{1}$$

Then we standardized them when $w_i = \dfrac{v_i}{\sum_{j=1}^{k} v_j}$

2.3 Data from Tutors

Further, we needed the data, i. e. to know values of all criteria for each course of the surveyed study programme. Therefore we distributed our question-form to the tutors of all the courses and having their feedbacks we created a matrix $X_{(k \times n)}$ with elements x_{ij}, where n denotes the number of all the courses and k the number of the criteria. As we dealt with more than 50 courses within this study programme, we present for example only three columns of the matrix X here.

Course	**Microeconomics**	**Statistics**	**Accounting I**
Homeworks	0.50	0.30	1.00
Semestral work	0	0	0
Credit test	2	1	0
Written exam	0	1	0
Oral exam	1	0	0
Precourses	0	1	1
Special tools	0	1	0
Lectures amount	2	2	2
Practising amount	2	2	1

Table 2

2.4 Relative Difficulties of Courses

We decided to calculate the relative difficulty r_j, $j = 1, \dots, n$ of a particular course j as sumproduct

$$r_j = \boldsymbol{w}^T \boldsymbol{x}_j \tag{2}$$

where the vector $\boldsymbol{w}$ contains the weights of the factors (see the last column of Table 1) and the vector x_j is the j-th column of matrix X, i. e. the values of the criteria of j-th subject. In Table 3 there are relative difficulties of the selected subjects from Table 2.

Course	Microeconomics	Statistics	Accounting I
Relative difficulty	0.78	0.55	0.29

Table 3

2.5 Credits

The values like in Table 3 cannot be naturally used as ECTS credits, as they should be integers, so any transformation is necessary. The only constraint, which we took into account, is the total of credits - according [?] it is 180 credits within the first three years of the study. So we sumed all the relative difficulties $r = \sum_{i=1}^{n} r_i$ ($r = 31.61$) and multiplied all r_j by $\frac{180}{r}$ to get the credits c_j for all the courses $j = 1, \dots, n$

$$c_j = r_j \frac{180}{r}. \tag{3}$$

Note that $\sum_{i=1}^{n} c_j = 180$. But still c_j are real numbers, not integers, so we rounded them keeping at the same time the demanded total of 180 credits. After the rounding we got the final integer values of credits C_j

Course	Microeconomics	Statistics	Accounting I
Credits c_j	4.44	3.13	1.65
Integer credits C_j	4	3	2

Table 4

3 Conclusion

Our attempt does not change the ECTS grounds at all. It is only a local modification of credit calculation and it has served for TUL needs. Although we managed to appreciate the difficulties of the courses, it is nearly impossible to create any comprehensive and comparable system applicable for all concerned universities (within ECTS). The reason is, among others, that various universities can teach the same course according its name but with different syllabi and different requirements (i. e. different criteria values). Thus, such a course would be assigned to different credits at different schools. Obviously, it would be better and more comfortable in general, if the same courses have the same credits and be compatible.

References

1. *Boloňský proces.* MMT, [online]. [cit. 2009-05-05]. Available from: `http://www.bologna.msmt.cz`
2. *ECTS System – incoming and outgoing students are assessed through the European Community Course Credit Transfer System.* [online]. [cit. 2009-05-05]. Available from: `http://www.enscm.fr/ects_system.htm`
3. *ECTS USERS´ GUIDE.* [online]. [cit. 2009-05-05]. Available from: `http://www.hrk.de/bologna/de/download/dateien/ECTS_guide_en.pdf`
4. *ECTS USERS´ GUIDE – European Credit Transfer and Accumulation System and the Diploma Supplement.* [online]. [cit. 2009-05-05]. Available from: `http://www.eua.be/eua/jsp/en/upload/ECTS_DS_Users_guide_en.1094119167134.pdf`
5. *European Credit Tranfer and Accumulation System (ECTS).* [online]. [cit. 2009-05-05]. Available from: `http://ec.europa.eu/education/lifelong-learning-policy/doc48_en.htm#doc1`
6. JABLONSKÝ, J., DLOUHÝ, M. *Modely hodnocení efektivnosti produkčních jednotek.* První vydání. Professional Publishing, Praha, 2004. ISBN 80-86419-49-5
7. JABLONSKÝ, J. *Operační výzkum – Kvantitativní modely pro ekonomické rozhodování.* Druhé vydání. Professional Publishing, Praha, 2002. ISBN 80-86419-42-8
8. *Příručka pro uivatele ECTS – Evropského kreditního transfer systému.* Z úvodu k publikaci „European Credit Transfer System, Users Guide, Erasmus Bureau, CECA-CE-CEEA, Brussels, Luxembourg, 1995, ISBN 92-826-9020-2. [online]. [cit. 2009-05-05]. Available from: `http://www.natur.cuni.cz/www/en/data/study/ects.html`
9. *The Bologna Declaration of 19 June 1999.* [online]. [cit. 2009-05-05]. Available from: `http://www.bologna-bergen2005.no/Docs/00-Main_doc/990719BOLOGNA_DECLARATION.PDF`

DO STOCK MARKET PRICES GRANGER CAUSE ECONOMIC GROWTH IN CR?

Roman Hušek, Filip Tichý[1]

University of Economics, Department of Econometrics
husek@vse.cz, tichyf@vse.cz

Abstract. One of the good features of VAR models is that they allow us to test for the direction of causality. Suppose the relationship between two stationary variables, say y_t (real GDP) and x_t (stock market prices), is captured by a VAR model. We will adopt an appropriate procedure – **Granger causality test** that allows us to test and statistically detect the cause and relationship between those variables.
The dataset used in estimation and testing consist of quarterly observations of CZ GDP and CZ PX index from the first quarter of 1996 to the third quarter of 2006. The results show that the causality direction runs from the CZ PX index to the real GDP, while the opposite hypothesis that real GDP Granger causes stock market prices is strongly rejected. Thus, one can suppose that the PX index is **leading indicator** of GDP dynamics.
Keywords: Granger causality**,** VAR model, leading indicator, efficient market hypothesis

1. Introduction

The overall movement in stock market prices is regarded as a **leading indicator** of future economic growth dynamics. It is assumed that an aggregated decline in stock market prices (bear market) is sooner or later followed by a recession and vice versa: aggregated increase in stock market prices (bull market) is a signal for improved future economic situation.

Granger causality (GC) test is used to fulfil the objective of this paper: empirical verification of the effect of changes in stock market prices towards the anticipated economic growth in the Czech Republic (CR). The relation between stock market index PX and Czech real GDP was tested upon quarterly data from 1.Q.1996 to 3.Q.2006, which were used for estimation of the base two-equation VAR(3) model. The actual GC test is built around a chi-square (χ^2) statistics, as included in the PcGive software we have applied.

2. Empirical VAR model

Granger causality test for two stationary variables y_t (real HDP) and x_t (PX index) is based on the estimated two-equation VAR(3) model

$$y_t = a_1 + \sum_{i=1}^{3} \alpha_i y_{t-i} + \sum_{i=1}^{3} \beta_i x_{t-i} + u_t, \tag{1}$$

$$x_t = a_2 + \sum_{i=1}^{3} \gamma_i y_{t-i} + \sum_{i=1}^{3} \delta_i x_{t-i} + v_t, \tag{2}$$

where y_t is the second difference of seasonally adjusted real GDP, using prices from the year 2000,
x_t – second difference of the PX index,
u_t, v_t – spherical random factors.

[1] Financial support from the GAČR – No. 402/09/0273 is appreciated.

Unit root tests as performed using the ADF criteria indicated that both time series are integrated of order 2, i.e. $I(2)$. Both time series were found trend stationary in their second differences at a 1 percent significance level, hence standard procedures of statistical induction (e.g. Hušek, 2007) may be applied to the VAR(3) model. Optimal lag length of 3 quarters was chosen using the Akaike information criteria (AIC).

Least squares estimation (LSE) of equations from the model (1) and (2) provided us with regression parameters as shown in table 1.

Table 1: Estimates of VAR(3) model parameters

GDP equation	Point estimate	Standard error	*t* probability
a_1	1070,91	411,400	0,02
α_1	- 0,69082	0,155	0,00
α_2	- 0,70961	0,135	0,00
α_3	- 0,74680	0,147	0,00
β_1	19,48090	5,176	0,01
β_2	17,45830	6,730	0,02
β_3	7,13234	5,434	0,20
PX equation			
a_2	- 8,87685	15,570	0,57
γ_1	0,00068	0,006	0,91
γ_2	0,00337	0,005	0,52
γ_3	0,00839	0,006	0,14
δ_1	- 0,80759	0,196	0,00
δ_2	- 0,20961	0,255	0,42
δ_3	0,12632	0,206	0,54

F test values confirm that the model as a whole is statistically significant at 1 a percent significance level.

3. Statistical Granger causality test

The substance of GC testing (Granger, 1969) lies in validating whether changes in one variable precede changes in other variable, i. e. whether the inclusion of an additional variable into a regression model provides significantly improved explanatory power. Hence, in Granger's conception, variable y_t stipulates x_t if the inclusion of lagged y_t values improves x_t forecasts, provided other explanatory variables in the regression remain unchanged. Also, an alternative test for causality in econometric models was set up afterwards by Sims (1972).

GC test is used to verify the null hypothesis that x_t does not Granger-cause y_t within the two estimated regressions, which for our VAR(3) model are outlined as

$$y_t = a_1 + \sum_{i=1}^{3} \alpha_i y_{t-i} + \sum_{i=1}^{3} \beta_i x_{t-i} + u_t, \tag{3}$$

$$y_t = a_1 + \sum_{i=1}^{3} \alpha_i y_{t-i} + u_t. \tag{4}$$

Given the asymptotical normality of parameters from the correct specified VAR model estimated using LSE method, we are able to use the F test values in order to verify statistical significance of lagged values of the variables x_{t-i} from equation (3), as follows:

$$F = \frac{RSS_R - RSS_U}{RSS_U} \cdot \frac{T-K}{m}, \qquad (5)$$

where RSS_R, RSS_U are the sums of squared residuals from the restricted (4) and unrestricted (3) regressions,
T – number of observations,
$m = 3$ – number of parameter restrictions in equation (4),
$k = 2m + 1$ – number of parameters estimates from equation (3).

For F values calculated from (5) where $F > F^*_{m,T-k}$, we reject the null hypothesis in favour of the alternative hypothesis that variable x_t stipulates y_t. Gujarati (1995) pointed out that testing for Granger non-causality might be susceptible to the choice of lag length m.

Our VAR(3) model (3) and (4) may lead to different GC conclusions.

- Coefficients of lagged x_{t-i} values in (3) significantly differ from zero and lagged y_{t-i} values in (4) do not have coefficients significantly different from zero. Hence, x_t stipulates y_t.
- Coefficients of lagged y_{t-i} values in (4) are statistically different from zero, but coefficients of variables x_{t-i} from (3) do not significantly differ from zero. Then, conversely, y_t stipulates x_t.
- Coefficients of lagged values of both variables from equations (3) and (4) are statistically different from zero, so both variables are mutually stipulating each other (bi-directional causality).
- Coefficients of lagged values of both variables from (3) and (4) do not statistically differ from zero hence x_t and y_t are mutually independent.

When testing for GC, we found out that all β_i parameters from equation (3) differ from zero at a one percent significance level, therefore we rejected the null hypothesis that variable x_t does not Granger-cause the variable y_t. Next step consisted of testing the null hypothesis that variable y_t does not Granger-cause the variable x_t. For this purpose, we swapped both variables from the regression equations (3) and (4). In this second step of our GC testing, we have accepted the null hypothesis that y_t does not stipulate x_t. Therefore, we have concluded that x_t does Granger-cause y_t, i.e. PX index lagged values contribute to increased accuracy of real GDP forecasts. This conclusion, however, may not be interpreted as if variables x_t and y_t were in a cause and effect type of relation.

Additionally, it should be noted that GC tests for two variables may be performed even with VAR models consisting of more than two equations (e.g. Lütkepohl, 1993). Also, Sims (1972) identifies the existence of Granger non-causality with the exogeneity of variables. In reality, however, it's only related to the so called **weak exogeneity**, given Granger non-causality provides only for the necessary condition of the so called **strong exogeneity** (Maddala, 1992). VAR models, though more typically deal with long time-series, can also be applied to panel data. See e.g. Pánková (2005) with an application to Czech machinery data.

4. Conclusion

The test results of testing real GDP and PX index for GC have shown us that the PX index may be regarded as the **leading indicator** of economic growth for the period 1996-2006. However, the **efficient market hypothesis** (EMH) precondition must be met for PX index to be the leading indicator of GDP. EMH ensures that stock prices quickly and accurately reflect any new relevant price changing information so that there are no discrepancies between share market prices and their internal values. Also, market investors are assumed to form rational expectations and therefore forecasts made by economic subjects do not suffer from systematic error

(Fama, 1965). Preconditions being met, then in accordance with economic theory we may expect that decline on share market shall be followed by a somewhat delayed economic recession and vice versa: increase in share prices stipulates improved future economic activity.

We have found that GC between GDP and PX index is statistically significant at a one percent significance level, which is a surprising conclusion to some extent, given the actual level of market efficiency of Czech share market.

References

[1] Fama, E. (1965): The Behavior of Stock Market Prices. Journal of Business, 38, pp. 34-105.

[2] Granger, C.W.J. (1969): Investigating Causal Relations by Econometric Models and Cross Spectral Methods”. Econometrica, 35, pp. 424 – 438.

[3] Gujarati, D.N. (1995): Basic Econometrics. McGraw Hill, New York.

[4] Hušek, R. (2007): Ekonometrická analýza, 2nd ed. Oeconomica, Praha.

[5] Lütkepohl, H. (1993): Testing for Causation Between Two Variables in Higher Dimensional VAR Models. In H. Schneeweiss and K.F. Zimmermann (eds.): Studies in Applied Econometrics. Physica Verlag, Heidelberg.

[6] Maddala, G.S. (1992), Introduction to Econometrics, 2nd ed. Macmillan, New York.

[7] Pánková, V. (2005): Tobinovo Q – teorie a aplikace. Politická ekonomie, roč. LIII, 5, pp. 601-608. ISSN 0032-3233.

[8] Sims, C.A. (1972): Money, Income and Causality. American Economic Review, 62, pp. 540-552.

Internet sources

http://www.czso.cz

http://www.pse.cz

VALIDITY OF THE PURCHASING POWER PARITY IN THE SELECTED EUROPEAN UNION'S COUNTRIES: STATIONARITY AND COINTEGRATION APPROACH[1]

Michaela Chocholatá

University of Economics Bratislava
Faculty of Economic Informatics, Department of Operations Research and Econometrics
chocholatam@yahoo.com

Abstract. This paper deals with the analysis of the purchasing power parity (PPP) in the selected European Union's countries (Czech Republic, Hungary, Latvia, Poland, Romania and Slovakia) during the period January 1999 – November 2008 (119 observations) using the euro as a base currency. In the first step we examine the stationarity of the individual real exchange rates using the ADF (Augmented Dickey – Fuller) and PP (Phillips – Perron) unit root tests. All the real exchange rates were identified as to be non-stationary, so the PPP can not hold. Further analysis was done using the Engle-Granger cointegration approach and Johansen cointegration approach. Both approaches provide support for the validity of the PPP only for Hungary - the values of estimated parameters, however, are not consistent with the PPP theory. Finally we present the Larson et al. panel cointegration technique which enables to overcome the problems with the low power of the unit root and cointegration tests in case of short sample sizes. The absence of cointegration in majority of analysed countries makes it impossible to apply this modern technique.
Keywords: Exchange rate, purchasing power parity (PPP), stationarity, cointegration.

1. Introduction

Purchasing power parity (PPP) represents an important key building block in international macroeconomics. Although the origins of this theory can be found in the sixteenth century in Spain, the term purchasing power parity (expressing the link between exchange rates and national price levels) was first used by the Swedish economist G. Cassel in 1920s (see e.g. [8]). The PPP theory has become extremely popular since the advent of flexible exchange rates in 1970s. The studies dealing with testing of the PPP validity use various approaches. We can divide these studies into three groups based on methods the authors used for analysis (see e.g. [7], [15]). The early studies testing the PPP didn't take into account possible non-stationarities in the analysed series. The second group build studies focused on testing for stationarity of the real exchange rates using various unit root tests (see e.g. [4]). The third group includes the various cointegration - based studies analysing the trivariate relationship among the nominal exchange rate, the domestic price level and the foreign price level or the bivariate relationship between exchange rate and relative price. The main disadvantage of the studies using the concept of stationarity and cointegration is in case of short sample sizes the low power of the tests. One of the possibilities how to solve this problem is to analyse the time series data from a large number of countries using the panel data techniques (see e.g. [1], [15], [16]).

The main aim of this paper is to analyse the validity of the PPP using the stationarity and cointegration concept and panel cointegration approach in case of some European Union's (EU's) countries during the period January 1999 – November 2008. We wanted to do the analysis for exchange rates of all "new" member states which became EU members in 2004 and later and didn't adopt the euro during the above mentioned period (i.e. Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia). Since in Estonia and in Lithuania, the Monetary Authorities committed themselves to retain the exchange rate of Estonian kroon against the euro and Lithuanian litas against the euro, respectively, unchanged at the central parity rate during the Exchange Rate Mechanism II (ERM II) membership in order to maintain the stability and other benefits of a fixed exchange rate regime, these countries were excluded from the analysis. One more currency excluded from the analysis was the Bulgarian lev, which has been unilaterally tied to the euro since January 1999.

2. The Methodological Issues

[1] Acknowledgements: This research was supported by the grant project VEGA No. 1/4652/07 „Analysis of the current problems of the Slovak economy development before the entrance into the European Monetary Union – econometrical approach".

2.1. Testing for Unit Roots

The theory of PPP links the national price levels and exchange rates. In case that the PPP relationship holds, the national price levels and exchange rates should form an equilibrium relationship. If we denote p_t and p_t^* logarithms of the price levels in domestic and foreign[2] country, and s_t the logarithm of the nominal exchange rate defined as the domestic price of a foreign currency (euro), the logarithm of the real exchange rate r_t can be then calculated as follows:

$$r_t = s_t + p_t^* - p_t \tag{1}$$

The PPP can be considered to be valid if the logarithm of the real exchange rate r_t is stationary. The stationarity or non-stationarity (i.e. the existence of the unit roots) can be tested by various unit root tests, the most famous of which are the ADF (Augmented Dickey – Fuller) test and PP (Phillips – Perron) test[3]. The non-stationarity of the r_t indicates, that there is no tendency of the real exchange rate to return to its equilibrium value, so the PPP could not hold. The rejection of the unit root hypothesis on the other hand means, that the real exchange rate has a character of the stationary stochastic process with a certain ability to return to its equilibrium value.

2.2. Cointegration Tests

Since the typical feature of the majority of macroeconomic variables is non-stationarity (see e.g. [14]), the another possibility how to test the PPP validity is the cointegration concept. In case of PPP validity, the sequence f_t formed by the sum $s_t + p_t^*$, expressing the logarithm of the domestic value of the foreign price level, should be cointegrated with the p_t sequence. We can also analyse the following equation:

$$f_t = \phi_0 + \phi_1 p_t + \varepsilon_t \tag{2}$$

where ϕ_0 and ϕ_1 are unknown parameters and ε_t is an error term reflecting any short-run deviations from the long-run equilibrium caused by stochastic shocks.

To test for cointegration between two variables (f_t, p_t) the Engle-Granger method [5] and Johansen method [10] can be used. The cointegration concept concentrates on analysis whether the linear combination of non-stationary variables is stationary. Assuming that a pair of series has the same order of integration, it is useful to examine the equilibrium relations further with the cointegration tests.

The Engle-Granger method is the OLS (Ordinary Least Squares) based cointegration approach for a bivariate system defined e.g. by equation (2). In the first step we need to determine the order of integration of both the variables f_t and p_t. If they are both integrated of the same order[4] – usually I(1), we can conclude that they are cointegrated if the residuals from the equation (2) are stationary. It is not appropriate to use the "classic" ADF critical values, the adequate critical values are more negative (see e.g. [4]). In order to say that the PPP concept is valid, it is necessary to confirm that the cointegrating vector is such that $\phi_1 = 1$. In case that the residual sequence $\hat{\varepsilon}_t$ contains a unit root, the variables f_t and p_t are not cointegrated and the PPP relationship does not hold.

The Johansen method is based on maximum-likelihood estimation procedure which enables to capture the feedback effects between variables and is independent of the choice of endogenous variable. The whole procedure is based on the following VAR (Vector Autoregression Representation) of the vector of N stationary variables $\mathbf{x_t}$ ($t = 1,2,...,T$):

$$\mathbf{x_t} = \mathbf{\Pi_1 x_{t-1}} + \mathbf{\Pi_2 x_{t-2}} + \mathbf{...} + \mathbf{\Pi_k x_{t-k}} + \mathbf{u_t}, \tag{3}$$

---

[2] As a foreign country was used the Economic and Monetary Union (EMU) in its changing composition.

[3] For more information about various unit root tests see e.g. [4], [6].

[4] With differing orders of integration, it would have been possible to immediately conclude that long-run PPP failed [4].

where $u_1, u_2, ..., u_T$ are N-dimensional independent identically distributed normal variables and $\mathbf{x_t}$ is a vector of all endogenous variables in system. In case that all variables in $\mathbf{x_t}$ are non-stationary and achieve stationarity after being differenced once, the VAR (3) can be rewritten in the VECM (Vector Error Correction Model) form as follows:

$$\mathbf{\Delta x_t} = \mathbf{\Gamma_1 \Delta x_{t-1}} + \mathbf{\Gamma_2 \Delta x_{t-2}} + \mathbf{...} + \mathbf{\Gamma_{k-1} \Delta x_{t-k+1}} + \mathbf{\Pi x_{t-1}} + \mathbf{u_t} \tag{4}$$

$$\text{where } \mathbf{\Gamma_i} = -\sum_{j=i+1}^{k} \mathbf{\Pi_j} \quad (i = 1,2,...,k-1) \text{ a } \mathbf{\Pi} = -\left(\mathbf{I} - \sum_{j=1}^{k} \mathbf{\Pi_j} \right). \tag{5}$$

The order of the VAR, k, is assumed to be finite. $\mathbf{\Pi}$ ($N \times N$) represents the long-run or cointegration matrix and if there exist r cointegrating vectors, is of reduced rank $r < N$. In this case, the matrix $\mathbf{\Pi}$ can be rewritten as $\mathbf{\Pi} = \boldsymbol{\alpha\beta}'$, where $\boldsymbol{\alpha}$ ($N \times r$) includes the speed of adjustment coefficients and $\boldsymbol{\beta}$ ($N \times r$) the r cointegrating vectors. To test the hypothesis that there are at most r cointegrating vectors, two test statistics - maximum eigenvalue statistic $\lambda_{\max}$ and trace statistic λ_{trace}, are used. These statistics have the χ^2 - distribution with $(N-r)$ degrees of freedom. The null hypotheses in these tests are accepted if the estimated values are less than the critical values at the appropriate significance level. Furthermore it is also important to mention that the Johansen method also enables testing for the possible linear restrictions regarding coefficients of the matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ (see e.g. [1], [4], [12]).

2.2.1. Panel Cointegration Test

In order to overcome the problems with the low power of the above mentioned tests in case of short sample sizes, the concept of the panel cointegration can be used. There are different possible tests which enable to test the cointegration in panels (see e.g. [1], [16]). Although the majority of them is based on the Engle – Granger cointegration relationship, we will concentrate on Larsson et al. test based on Johansen's approach. Larsson et al. suggest the estimation of the ECM (Error Correction Model) separately for each country (N countries with time dimension T and a set of p variables are assumed) using the maximal likelihood method for the calculation of the trace statistic LR_{iT}. The panel rank trace statistic LR_{NT} can be calculated as an average of the trace statistics obtained for each of the N analysed countries. The null hypothesis is that the number of cointegrating vectors (r_i) among the p variables is the same in all N countries, i.e. $H_0 : rank(\mathbf{\Pi}_i) = r_i \le r$ for all $i = 1,...,N$, against the alternative hypothesis $H_1 : rank(\mathbf{\Pi}_i) = p$ for all $i = 1,...,N$. The panel cointegration rank trace test statistic Y_{LR} can be obtained by calculating the average of the trace statistics LR_{NT} received for individual countries. After standardising it we will receive:

$$Y_{LR} = \frac{\left(LR_{NT} - E\left[Z_k\right]\right)\sqrt{N}}{\sqrt{Var\left[Z_k\right]}} \tag{6}$$

where $E\left[Z_k\right]$ and $Var\left[Z_k\right]$ are the mean and variance of the asymptotic trace statistic (see e.g. [1], [15]).

3. Empirical Results

In the analysis of the PPP validity we use the monthly exchange rates of domestic currencies of the selected EU's countries (Czech Republic, Hungary, Latvia, Poland, Romania and Slovakia) against the euro[5]. The price variables are measured by the harmonized indices of consumer prices (HICP). The sample data covers the period January 1999 – November 2008 (119 observations). All variables are expressed in logs and are in the form of indices relative to a base month (January 1999 = 1,00). The data for analysis were obtained from the Eurostat web-page [18] and the European Central Bank web-page [17], the whole analysis was done in econometric software EViews 5.1.

[5] On 1 July 2005 the Romanian Leu (ROL) was replaced by the new leu (RON), with a conversion factor of 1 RON = 10000 ROL. In order to receive consistent results for Romania, we used for analysis in January 1999 – June 2005 the actual exchange rates divided by 10000.

In the first step of analysis we tested the real exchange rates (1) for stationarity using the ADF and PP test. From the results in table 1[6] it is obvious, that all analysed real exchange rates are non-stationary (have one unit root), which means that the PPP could not hold. The real exchange rate in this case can be characterised as a sequence of real shocks which permanently influence the level of the real exchange rate, i.e. it doesn't exist the tendency of the real exchange rate to return to its equilibrium value or trend.

Table 1 – The real exchange rates – unit root test results

	Level			*1. difference*
	trend and intercept	*intercept*	*neither trend nor intercept*	*trend and intercept*
CZE	-1,8265	0,0725	1,8699	-9,5180*
HUN	-2,9454	-1,7947	0,2776	-8,1699*
LAT	-0,8733	-0,7445	0,5725	-7,7984*
POL	-2,4255	-2,4467	-0,8169	-6,9295*
ROM	-1,9496	-1,4376	0,4035	-8,3755*
SVK	-3,6737** -3,0142	-0,1671	2,7106	-8,1069*

*Note: The symbols *, ** denote the rejection of the null hypothesis at the 0,01 and 0,05 significance level respectively.*

In the next step of our analysis we tested the cointegration between variables f_t and p_t (see equation (2)) using the Engle – Granger method and Johansen method. The existence of cointegration in both methods requires the same order of integration in case of both analysed variables f_t and p_t. The majority of variables was according to ADF and PP test identified as to be integrated I(1)[7], the only exception was Romania, where the variable p_t was stationary, and variable f_t integrated of order 1. It means that in case of Romania there is no cointegration between variables from equation (2) and therefore the PPP could not hold.

Assuming the Engle – Granger cointegration concept, we estimated equation (2) for the remaining countries and in order to check on the robustness also the regression with the reversed order of variables of the form $p_t = \varphi_0 + \varphi_1 f_t + \xi_t$ (see e.g. [1], [4], [13]). The both calculated residuals for individual countries were tested for stationarity using the ADF test with neither trend nor intercept. The results are in table 2 from which it is clear that the existence of cointegration was confirmed only in case of Hungary since both the residuals $\hat{\varepsilon}_t$ and $\hat{\xi}_t$ were stationary. The estimated values of parameters $\hat{\phi}_1 = 0,36$ and $\hat{\varphi}_1 = 2,09$ (for Hungary), however, were far away from unity. One of the possible reasons presented in [13] is the existence of transportation charges. Taking into account the large discrepancies from a value of 1 in case of Hungary and absence of cointegration for the remaining countries, we can conclude, that according to the presented results the PPP can not hold.

Table 2 – Unit root test results of residuals $\hat{\varepsilon}_t$ and $\hat{\xi}_t$

	CZE		HUN		LAT		POL		SVK	
	$\hat{\varepsilon}_t$	$\hat{\xi}_t$	$\hat{\varepsilon}_t$	$\hat{\xi}_t$	$\hat{\varepsilon}_t$	$\hat{\xi}_t$	$\hat{\varepsilon}_t$	$\hat{\xi}_t$	$\hat{\varepsilon}_t$	$\hat{\xi}_t$
level	-3,04	-3,06	-3,54**	-3,3***	-1,25	-0,74	-0,63	-2,26	-2,63	-2,9
1.diff.	-11,1*	-11,2*			-7,8*	-7,7*	-7,6*	-6,1*	-7,7*	-7,8*

*Note: The symbols *, **, *** denote the rejection of the null hypothesis at the 0,01; 0,05 and 0,1 significance level respectively.*

To investigate the existence of cointegration based on Johansen method - after identification of the same order of integration of variables f_t and p_t (all countries with exception of Romania) - it is necessary to set the appropriate lag length of the VAR model (3). Based on FPE (Final Prediction Error) and AIC (Akaike

[6] The results obtained by both mentioned tests were identical in all cases with exception of Slovakia (in table 1 there are results of the ADF test, in case of Slovakia of both tests), where the ADF test identified the real exchange rate to be stationary at the significance level 0,05 (first row SVK), but according to the PP test it is clearly non-stationary (second row SVK). Since the PP test identified the Slovak real exchange rate to be non-stationary, we considered it to be non-stationary for further analysis.

[7] In order to save space, the results of ADF and PP tests for individual f_t and p_t are not presented here, but are available upon request.

Information Criterion) we received the following results for the lag lengths: CZE – 2, HUN – 9, LAT – 3, POL – 2 and SVK – 2 lags. The appropriate form of the model regarding the deterministic components was identified according to the Pantula principle (see [1]) and the existence of cointegration was tested using the statistics λ_{trace} and $\lambda_{\max}$. The results are in table 3[8].

Table 3 – The Pantula principle

	Model 2				Model 3				Model 4			
	λ_{trace}		$\lambda_{\max}$		λ_{trace}		$\lambda_{\max}$		λ_{trace}		$\lambda_{\max}$	
	$r=0$	$r=1$	$r=0$	$r=1$	$r=0$	$r=1$	$r=0$	$r=1$	$r=0$	$r=1$	$r=0$	$r=1$
CZE	23,9	5,3	18,6	5,3	11,2*	0,3	11,0*	0,3	17,8	4,7	13,1	4,7
HUN	31,4	8,0*	23,5	8,0*	20,4	1,4	19,0	1,4	29,9	10,7	19,2	10,7
LAT	21,1	0,9	20,1	0,9	10,3*	0,3	10,0*	0,3	28,2	9,0	19,2	9,0
POL	30,8	4,1	26,8	4,1	13,3*	4,0	9,3*	4,0	20,2	4,0	16,2	4,0
SVK	30,5	3,1	27,4	3,1	13,5*	0,4	13,1*	0,4	29,4	11,8	17,7	11,8

*Note: The symbol * indicates the first time that the null hypothesis can not be rejected.*

The results in table 3 indicate no cointegration in all cases with exception of Hungary, where it was identified one cointegrating vector (Model 2). This means that the validity of the PPP was clearly not confirmed for the Czech Republic, Latvia, Poland and the Slovak Republic. For Hungary we can estimate the VECM with 9 lags[9]. The VECM synthesises the short-run dynamic relationships and long-run equilibrium relationships. The long-run information is included in the cointegrating vector $\boldsymbol{\beta} = \begin{pmatrix} 1 & \beta_0 & \beta_1 \end{pmatrix} = \begin{pmatrix} 1 & 0{,}15 & -0{,}49 \end{pmatrix}$, where β_0 corrensponds to the constant, β_1 corresponds to the variable p_t and both parameters are highly statistically significant. However the value of the parameter β_1 is not in accordance with the PPP theory. From the speed of adjustment coefficients $\alpha_1 = 0{,}081$ and $\alpha_2 = 0{,}075$ is the first one not statistically significant, but the second one is highly statistically significant.

The residuals of this VECM were tested for uncorrelatedness and normality using the *LM* (Lagrange multiplier) test and Doornik – Hansen normality test (see e.g. [9], [19]). The *LM* statistic for 9 lags, i.e. $LM(9) = 4{,}049$ indicates that the residuals are uncorrelated at the significance level 0,01 till the lag 9. Worse results were received by testing the normality using the multivariate extension of the Jarque – Bera residual normality test, the Doornik – Hansen test. The calculated statistics of the normality test were 9,82 and 97,36 which indicates the evidence of deviations from normality in case of both equations. Islam and Ahmed in [9] mention that the deviation from normality does not render the cointegration test invalid. Taking into account the economically not correct value of the parameter β_1, we can conclude that the PPP validity is also in case of Hungary problematic.

Since the existence of cointegration was rejected in almost all analysed countries, it was not possible to apply the modern panel cointegration test, because it supposes the existence of cointegration in individual countries.

4. Conclusion

The analysed real exchange rates were all identified as to be non-stationary, which is in contrast with the PPP validity. The possible validity of the PPP was tested based on the cointegration methods (Engle-Granger method and Johansen method) as well and was clearly rejected for all analysed countries with exception of Hungary. Although the existence of cointegration was confirmed by both mentioned methods in case of Hungary, the estimated parameters were not in coincidence with the theory. Some of the reasons why PPP does not hold are e.g. the existence of transportation costs, tariffs, non-tradable goods, preferences for goods in different countries etc. (see e.g. [16]). Boršič and Bekö in [3] analysing the PPP in Slovenia and Hungary present the opinion that the problematic validity of the PPP could be caused by faster growth of non-tradable to tradable prices in transition countries in comparison to relative prices of developed market economies.

[8] Model 2 is a model with intercept (no trend) in CE (Cointegrating Equation)– no intercept in VAR, Model 3: Intercept (no trend) in CE and VAR, Model 4: Intercept and trend in CE – no trend in VAR. The use of models 1 and 5 is from the economic theory point of view unusual (Model 1: No intercept or trend in CE or VAR and Model 5: Intercept and trend in CE – linear trend in VAR) and they were therefore not analysed.

[9] The whole VECM will not be presented here from the space reasons, but the results can be provided upon request.

References

[1] ASTERIOU, D. – HALL, S.G.: *Applied Econometrics. A Modern Approach using EViews and Microfit*. Houndmills, Basingstoke, Hampshire and New York, PALGRAVE MACMILLAN 2007.

[2] BAHMANI–OSKOOEE, M. – BARRY, M.: *The Purchasing Power Parity and the Russian Ruble*. Comparative Economic Studies (Association for Comparative Economic Studies) 39, Spring 1997, No. 1, p. 82 – 94.

[3] BORŠIČ, D. – BEKÖ, J.: *Empirical Evidence on Purchasing Power Parity in Two Transition Countries: The Case of Slovenia and Hungary*.
(http://evenements.univ-lille3.fr/recherche/jemb/programme/papiers/borsic_beko_lille06.pdf)

[4] ENDERS, W.: *Applied Econometric Time Series*. New York, John Wiley&Sons, Inc. 1995.

[5] ENGLE, R.F. – GRANGER, C.W.J.: Co-integration and error correction: representation, estimation, and testing. Econometrica 55, 1987, p. 251 – 276.

[6] FRANSES, P.H. – DIJK, D. van: *Non-Linear Time Series Models in Empirical Finance*. Cambridge, Cambridge University Press 2000.

[7] FROOT, K.A. – ROGOFF, K.: *Perspectives o PPP and long-run real exchange rates*. In: Grossman, G. – Rogoff, K.(eds.): The Handbook of International Economics. Vol. 3, Elsevier Press, Amsterdam, 1995, p. 1647-1688.

[8] ISARD, P.: *Exchange Rate Economics*. Cambridge, Cambridge University Press 1995.

[9] ISLAM, A.M. – AHMED, S.M.: *The Purchasing Power Parity Relationship: Causality and Cointegration Tests Using Korea-U.S. Exchange Rate and Prices*. Journal of Economic Development 24, December 1999, No. 2, p. 95-111.

[10] JOHANSEN, S.: *Statistical Analysis of Cointegration Vectors*. Journal of Economic Dynamics and Control 12, 1988, p. 231-254.

[11] LUKÁČIK, M. – LUKÁČIKOVÁ, A. – SZOMOLÁNYI, K.: *Ekonometrické prognózovanie importu Slovenskej republiky*. Ekonomické rozhľady XXXVI, No. 2, Bratislava: Ekonomická univerzita, 2007, p. 179-192.

[12] LUKÁČIK, M. – LUKÁČIKOVÁ, A. – SZOMOLÁNYI, K.: *The Long-Run Determinants of Money Demand in Slovakia*. Proceedings from the 13th International Conference - Quantitative Methods in Economics (Multiple Criteria Decision Making), Bratislava, 2006, p. 92-100.

[13] McNOWN, R. – WALLACE, M.S.: *Cointegration tests of purchasing power parity among four industrial countries: results for fixed and flexible rates*. Applied Economics 22, 1990, p. 1729 – 1737.

[14] NELSON, C. – PLOSSER, C.: *Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications*. Journal of Monetary Economics 10, 1982.

[15] SIDERIS, D.: *Purchasing Power Parity in economies in transition: evidence from Central and East European countries*. Working paper, University of Ioaninna, April 2005 (http://www.econ.uoi.gr/working_papers/sideris/sideris2.pdf).

[16] SOSVILLA-RIVERO, S. – GARCÍA, E.: *Purchasing Power Parity*. Documento de trabajo 2003-20, FEDEA, December 2003, 49 p.

[17] www.ecb.int

[18] epp.eurostat.ec.europa.eu

[19] EViews 5 User's Guide

A TWO-STAGE AHP/DEA MODEL FOR EVALUATION OF EFFICIENCY

Josef Jablonsky

Faculty of Informatics and Statistics, Department of Econometrics
University of Economics Prague
jablon@vse.cz, URL: http://nb.vse.cz/~jablon

Abstract. The paper presents a two-stage model for evaluation and ranking of decision making units based on the analytic hierarchy process (AHP) with interval paired judgments and data envelopment analysis (DEA) with weight restrictions. The first stage of the model consists in assessment of weights of input and outputs by means of the interval AHP model, i.e. the AHP model with interval judgments. The results of this stage are interval weights derived by optimization approach on the one hand and simulation on the other hand. The second stage of the model uses a DEA model with weight restrictions for evaluation of efficiency. The problem is that the DEA models with direct weight restrictions are often infeasible. That is why we use models with virtual weights restrictions where the infeasibility does not occur so often. The DEA model with virtual weight restrictions splits the evaluated units into two subsets – efficient and inefficient ones. Then, the super-efficiency models and the AHP model with interval weights are used for ranking of efficient units. The results of both models are compared and discussed. The proposed modeling approach is applied in assessing the efficiency of banks operating in the Czech Republic.
Keywords. Analytic hierarchy process; data envelopment analysis; efficiency, super-efficiency

1. Data envelopment analysis models

Minor parts of the text describing basic data envelopment analysis (DEA) models in this section and interval analytic hierarchy models in the following section are based on the paper Jablonsky [3].

One of the most popular modelling approaches for efficiency evaluation of decision making units is data envelopment analysis (DEA). DEA models measure relative efficiency of decision making units (DMU) that are usually described by several inputs spent for production of several outputs. Let us consider the set E of n decision making units E = {DMU_1, DMU_2, …, DMU_n}. Each of the units produces r outputs and for their production spent m inputs. Let us denote $\mathbf{x}^j = \{x_{ij}, i=1,2,\ldots,m\}$ the vector of inputs and $\mathbf{y}^j = \{y_{kj}, k=1,2,\ldots,r\}$ the vector of outputs of the DMU_j. Then **X** is the (m, n) matrix of inputs and **Y** the (r, n) matrix of outputs. DEA models evaluate efficiency of the DMU_q, $q\in\{1,2,\ldots,n\}$ by looking for a virtual unit with inputs and outputs defined as the weighted sum of inputs and outputs of the other units in the decision set - $\mathbf{X}\lambda$ a $\mathbf{Y}\lambda$, where $\lambda=(\lambda_1, \lambda_2,\ldots, \lambda_n)$, $\lambda>\mathbf{0}$ is the vector of weights of the DMUs. The problem of looking for a virtual unit can be generally formulated as a standard linear programming problem:

$$
\begin{array}{ll}
\text{minimize} & z = \theta - \varepsilon(\mathbf{e}^{\mathbf{T}}\mathbf{s}^{+} + \mathbf{e}^{\mathbf{T}}\mathbf{s}^{-}), \\
\text{subject to} & \mathbf{Y}\lambda - \mathbf{s}^{+} = \mathbf{y}^{q}, \\
& \mathbf{X}\lambda + \mathbf{s}^{-} = \theta\mathbf{x}^{q}, \\
& \lambda, \mathbf{s}^{+}, \mathbf{s}^{-} \geq \mathbf{0},
\end{array} \qquad (1)
$$

where $e^T = (1,1,\ldots,1)$ and ε is an infinitesimal constant. The variables s^+, s^- are just slack variables expressing the difference between virtual inputs/outputs and appropriate inputs/outputs of the DMU_q. Obviously, the virtual inputs/outputs can be computed using the optimal values of variables of the model (1) as follows:

$$
\mathbf{x}^{q\prime} = \mathbf{x}^{q}\theta^{*} - \mathbf{s}^{-},
$$
$$
\mathbf{y}^{q\prime} = \mathbf{y}^{q} + \mathbf{s}^{+}.
$$

The DMU_q is considered to be efficient if the virtual unit is identical with evaluated unit. In this case $\mathbf{Y}\lambda = \mathbf{y}^q$, $\mathbf{X}\lambda = \mathbf{x}^q$ and the minimum value of the objective function z = 1. Otherwise the DMU_q is not efficient and minimum value of $\theta<1$ can be interpreted as the need of proportional reduction of inputs in order to reach the efficient frontier.

Dual variables of the model (1) v_i, i=1,2,…,m, and u_k, k=1,2,…,r, express relative importance of the inputs and outputs respectively. In order to have the possibility to restrict values of dual variables the dual formulation to the model (1) can be simply used:

$$
\begin{array}{llll}
\text{maximize} & z = \sum_{k}^{r} u_k y_{kq}, & & \\
\text{subject to} & \sum_{i}^{r} u_k y_{kj} \le \sum_{i}^{m} v_i x_{ij}, & j = 1, 2, \ldots, n, & (2) \\
 & \sum_{i}^{m} v_i x_{iq} = 1, & & \\
 & u_k \ge \varepsilon, & k = 1, 2, \ldots, r, & \\
 & v_i \ge \varepsilon, & i = 1, 2, \ldots, m. &
\end{array}
$$

Variables of the model (2) can be restricted by lower and upper bounds in order to influence the relative importance of the inputs and outputs in the model. This means that the constraints in model (2)

$$v_i \ge \varepsilon, \; u_k \ge \varepsilon,$$

are replaced by the following ones:

$$d_i \le v_i \le h_i,$$
$$b_k \le u_k \le c_k,$$

where d_i a h_i, i = 1,2,...,m, are lower and upper bounds for i-th input and b_k a c_k, k = 1,2,...,r, are lower and upper bound for k-th output. The main problem is that the model (2) with weight restrictions is often infeasible. Many researchers published papers dealing with infeasibility of DEA models with weight restrictions. One of the approaches based on models in Wang and Beasley [6] and Helcio [2] is called virtual weight restriction approach. The main idea of this approach is not to restrict the values of variables v_i and u_k directly as above but restrict them through the virtual inputs and outputs as follows:

$$d_i \le \frac{v_i x_{iq}}{\sum_{s}^{m} v_s x_{sq}} \le h_i, \qquad i = 1,2,\ldots,m, \tag{3}$$

$$b_k \le \frac{u_k y_{kq}}{\sum_{s}^{m} u_s y_{sq}} \le c_k, \qquad k = 1,2,\ldots,r, \tag{4}$$

where $0 \le d_i \le h_i \le 1$ and $0 \le b_k \le c_k \le 1$.

The constraints (3) and (4) are added to the model (2) and two different approaches are proposed in paper Helcio [2] - to restrict in this way only the observed unit DMU$_q$ or to restrict all the DMUs of the model. We will use the first of them in our paper.

One of the main problems in DEA models with weight restrictions is to set up the bounds of the inputs and outputs. We propose to use interval analytic hierarchy process (IAHP) model for estimation of lower and upper bounds. This model is described in the next section of the paper. Section 3 describes numerical experiments with different sets of bounds on the case of evaluation of efficiency of bank branches. Last section of the paper contains conclusions with summarization of the main results of the paper.

2. Interval AHP model

The AHP is a powerful tool for analysis of complex decision problems. The AHP organizes the decision problem as a hierarchical structure containing always several levels. The first (topmost) level defines a main goal of the decision problem and the last (lowest) level describes usually the decision alternatives or scenarios. The levels between the first and the last level can contain secondary goals, criteria and subcriteria of the decision problem, in our case inputs and outputs of the model. The number of the levels is not limited, but in the typical case it does not exceed four or five.

The AHP model can be simply used for evaluation of efficiency of decision making units. One of the possible models presented in Jablonsky [3] is on Figure 1. The topmost level of the hierarchy contains the goal of the decision problem, i.e. the evaluation of efficiency of the DMUs. This goal depends on the set of criteria used in the model - on the inputs on the one hand and the outputs on the other hand. They are placed on the next two levels of the hierarchy. The last level of the hierarchy contains DMUs that have to be compared with respect to the elements of the preceding level.

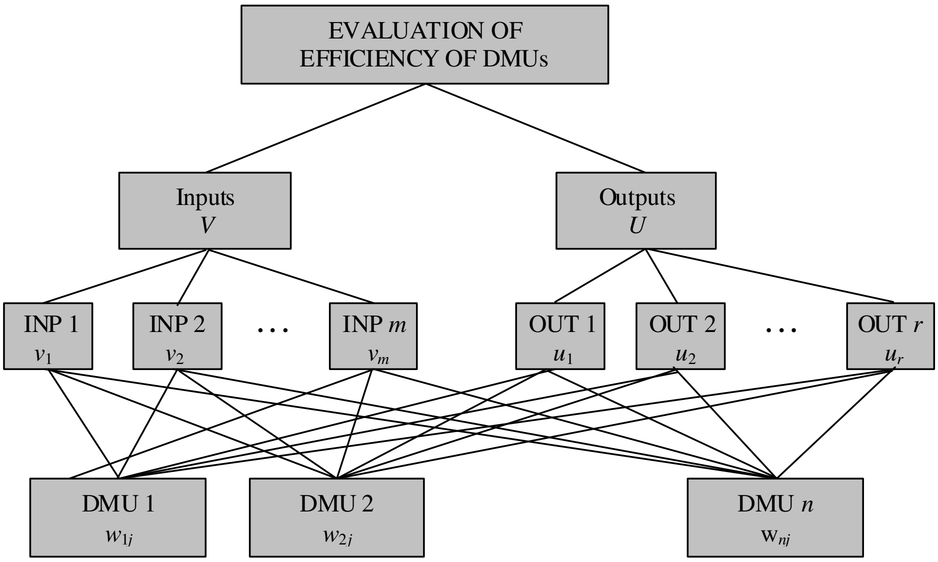


Figure 1: AHP model for efficiency evaluation.

By pairwise comparisons the decision maker can derive relative importance coefficients of the inputs and outputs – they are denoted as v_j, j=1,2,...,m, and u_k, j=1,2,...,r. The preference indices of the DMUs with respect to the given input or output are denoted as w_{ij}, i=1,2...,n, j=1,2,...,$m+r$. In the standard AHP model the decision maker judgements are organised into paired comparison matrices at each level of the hierarchy. The judgements are point estimates of the preference between two elements of the level. Let us denote the pairwise comparison matrix $\mathbf{A} = \{a_{ij} | a_{ji}=1/a_{ij}, a_{ij}>0, i,j=1,2,...,n\}$, where n is the number of elements of the particular level. Saaty [5] proposes to use a_{ij} integers in the range 1 through 9 for preference expression, where 1 means that the i-th and the j-th element are equally important and 9 that the i-th element is absolutely more important than the j-th element. The local priorities are derived by solving the eigenvector problem (5)

$$\mathbf{A.v} = \lambda_{\max}\mathbf{v}, \qquad \sum_{i=1}^{k} v_i = 1, \tag{5}$$

where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{A}$ and $\mathbf{v}$ is the normalised right eigenvector belonging to $\lambda_{\max}$.

In the deterministic AHP approach the decision maker always specifies point estimates that express his preference relations between two elements in the given hierarchical level. It can often be very difficult to fulfil this condition for decision makers. They feel much better and closer to have the possibility to express their preferences as interval estimates. The AHP model with interval decision maker judgements is usually called interval AHP (IAHP) model. It is characterised by interval pairwise comparison matrices given as follows:

$$\mathbf{A} = \begin{bmatrix} 1 & <p_{12},q_{12}> & ... & <p_{1k},q_{1k}> \\ <p_{21},q_{21}> & 1 & ... & <p_{2k},q_{2k}> \\ : & : & ... & : \\ <p_{k1},q_{k1}> & <p_{k2},q_{k2}> & ... & 1 \end{bmatrix} \tag{6}$$

where p_{ij} is lower bound and q_{ij} upper bound for preference relation (a_{ij}) between the i-th and j-th element. Due to the reciprocal nature of the pairwise comparison matrices the relation $p_{ij}.q_{ji} = 1$ holds for all i,j=1,2,...,k.

The judgements in the IAHP can be considered as random variables defined over the given interval. In this way the IAHP changes from the deterministic model to the model with some stochastic features. That is why it cannot be analysed in the traditional way – by solving the eigenvector problem (5). The random variables for description of interval judgements can be selected from the available probabilistic distributions. We will use uniform distribution defined over the interval $<p_{ij}, q_{ij}>$ and normal distribution with mean value standard deviation in our numerical experiments below. The preferences of elements derived form from matrix **A** are random variables. Their characteristics can be computed by several approaches – either Monte Carlo simulation that is very simple tool that offers lower and upper bounds for the preferences or one of the optimization approaches derived by different researchers. In our experiments presented in the next section we used Monte Carlo simulation for deriving interval weights inputs and outputs.

The interval AHP/DEA approach for evaluation of decision making units can be divided into two stages and described as follows:

1. Estimation of lower and upper bounds for weights of the inputs and outputs by means of the IAHP model. The weights of the inputs on the one hand and the weights of the outputs on the other hand are computed separately.
2. Solving a DEA model, e.g. the model (2) with additional constraints (3) and (4) where the lower and upper bounds are taken from step 1. Of course it is possible to experiment with different sets of weights not only given by interval AHP model. The efficiency scores of the new model are always worse or equal comparing to the efficiency scores of the original model. Some of the originally efficient units are not efficient in the model with virtual weights restrictions and new results can be taken as a basis for new ranking of the DMUs.

Bank	# of employees	operational expenses	# of branches	active debts	profit
A	84	43	1	6 760	247
B	893	141	11	86 498	895
C	267	15	5	21 186	390
D	116	204	1	5 911	110
E	10 897	575	636	588 526	12 148
F	616	80	155	138 048	1 220
G	239	27	7	20 350	759
H	10 357	16 965	222	504 294	10 837
I	2 299	204	214	61 112	2 366
J	131	20	39	29 407	155
K	426	20	30	1 111	805
L	230	235	1	27 109	218
M	7 764	13 558	386	54 023	11 225
N	319	10	29	162	121
O	362	41	246	64 051	432
P	118	36	1	20 898	693
Q	204	23	54	39 659	201
R	1 627	67	54	72 265	780
S	209	57	636	89 827	1 105
T	1 594	543	60	166 322	2 938
U	635	4	54	23 685	320
V	191	19	209	29 954	142

Table 1: Banks and their characteristics

3. Numerical experiments

This section contains simple example that demonstrates the approach described in the previous Section. The example is based on real data set but it is not a serious case study. Table 1 contains data about 22 DMUs, banks operating on the financial market in the Czech Republic – the units are denoted by letters A to V. Each of the banks is described by three input and two outputs:

- INP1 – number of employees,
- INP2 – operational expenses in 2007 in millions of CZK,
- INP3 – number of branches,
- OUT1 – active debts in 2007 in millions of CZK,
- OUT2 – gross profit in 2007 in millions of CZK.

The first stage of our approach is estimation of interval weights. The interval pairwise comparison matrices for inputs and outputs are below. We suppose uniform distribution for single comparisons. The random values as presented below are generated and the matrix is completed by appropriate reciprocal values. The logarithmic least square method is used for estimation of final priorities, i.e. weights of the inputs/outputs. The results from 100 simulations are presented in Table 2.

$$\begin{array}{c} \mathrm{INP1} \\ \mathrm{INP2} \\ \mathrm{INP3} \end{array} \begin{bmatrix} 1 & & <1{,}2> \\ <3{,}5> & 1 & <5{,}7> \\ & & 1 \end{bmatrix}$$

$$\begin{array}{c} \mathrm{OUT1} \\ \mathrm{OUT2} \end{array} \begin{bmatrix} 1 & \\ <5{,}7> & 1 \end{bmatrix}$$

Bounds	**# of employees**	**operational expenses**	**# of branches**	**active debts**	**Profit**
Lower	0.180	0.649	0.082	0.125	0.834
Upper	0.252	0.731	0.114	0.166	0.875

Table 2: Lower and upper bounds for inputs and outputs

The results of DEA models with virtual weights restrictions (VWR) and their comparison with standard models are presented in Table 3. It is clear that the efficiency scores of the DMUs are lower in VWR models comparing to standard DEA models. Under assumption of constant returns to scale 7 DMUs are recognized as efficient. We used one of the super-efficiency measures to rank the units - the final ranking is given in the second column of Table 3.

Bank	**CCR-I**		**CCR-I VWR**		**BCC-I**	**BCC-I VWR**
A	0.514	17	0.358	11	**1.401**	0.996
B	0.855	13	0.318	14	**2.413**	0.518
C	**1.771**	4	0.909	5	**1.791**	**1.466**
D	0.288	20	0.086	19	**1.000**	0.422
E	0.743	14	0.657	6	unbnd	unbnd
F	**1.204**	8	0.597	7	**2.348**	0.826
G	**1.209**	7	**1.157**	3	**1.307**	**1.211**
H	0.271	21	0.077	20	**2.453**	0.521
I	0.390	18	0.366	10	0.703	0.520
J	0.987	9	0.343	13	**1.334**	**1.011**
K	**1.236**	6	0.214	17	**1.423**	0.625
L	**1.297**	5	0.124	18	**1.947**	0.319
M	0.243	22	0.056	21	**1.306**	0.196
N	0.298	19	0.046	22	**1.176**	0.945
O	0.889	11	0.404	8	0.897	0.479
P	**3.159**	2	**1.642**	2	**6.264**	**1.691**
Q	0.981	10	0.348	12	**1.071**	0.746
R	0.685	15	0.372	9	0.967	0.406
S	**1.990**	3	0.986	4	**2.230**	**1.362**
T	0.571	16	0.292	15	**1.808**	0.770
U	**3.815**	1	**1.851**	1	**4.021**	**2.663**
V	0.880	12	0.285	16	0.987	0.799

Table 3: Results of efficiency analysis

The third a fourth columns contain efficiency scores and ranking of DMUs by the same model with VWR. Only three of the originally efficient units remain efficient in this restricted model. Ranking of other units is quite different in some cases, e.g. the unit L was originally efficient and now is inefficient with a very low efficiency score. The BCC model with variable returns to scale indicates 18 of 22 units as efficient without considering weights restrictions (the super-efficiency model for unit E has not optimal solution). The VWR model substantially reduces the number of efficient unit – only 7 of them remained efficient.

4. Conclusions

Both data envelopment analysis and analytic hierarchy process are one of the most often used decision making tools. Their combination can lead to interesting results in the analyzed problem. Our aim was to develop a procedure for ranking decision making units in DEA. The procedure is based on solving standard DEA models with virtual weights restrictions with the weights derived by interval AHP model. The given results allow better and simpler ranking of DMUs comparing to standard DEA models. What must be solved is the problem of infeasibility of the optimization problem in case the intervals for virtual weights are too close. It will be the topic of further research.

Acknowledgements

The research is supported by the Grant Agency of the Czech Republic – grant no. 402/09/0231.

References

[1] ESTELLITA LINS, M.P., MOREIRA da SILVA, A.C., LOVELL, C.A.K. *Avoiding infeasibility in DEA models with weight restrictions*. European Journal of Operational Research, 2007, 181, pp. 956-966.

[2] HELCIO, V.J. *Multicriteria approach to data envelopment analysis*. Pesquisa Operacional, 2008, 28(2), pp. 231-242.

[3] JABLONSKY, J.: *Measuring the Efficiency of Production Units by AHP Models*. Mathematical and Computer Modelling, 2007, 46, pp. 1091-1098.

[4] JABLONSKY, J., DLOUHY, M.: *Modely hodnocení efektivnosti produkčních jednotek* (*Models for evaluation of efficiency of production units*). Professional Publishing, Praha 2004.

[5] SAATY, T.L. *The Analytic Hierarchy Process*. RWS Publications, Pittsburgh 1990.

[6] WONG, Y., BEASLEY, J.E. *Restricting Weight Flexibility in Data Envelopment Analysis*. Journal of the Operational Research Society, 1990, 41(9), pp. 829-835.

ON TREATING THE GENERAL CONSTRAINTS IN AGRICULTURAL AND FORESTRY OPTIMIZATION PROBLEMS

Jitka Janová[1]

Mendel University of Agriculture and Forestry in Brno
janova@mendelu.cz

Abstract. The possibility of realistic modeling of the complex constrained optimization problems in Czech agriculture and forestry business is studied using the results of mathematical programming and dynamic programming. The real-life constraints arising in the crop planting and the logging decision making problems in Czech Republic are introduced and the optimization models presented are analyzed.
Keywords. Optimization problems, general constraints, agriculture, forestry, mathematical programming .

1 Introduction

Currently, the general mathematical programming models applicable in agriculture and forestry (AF) are under worldwide research. The linear and non-linear programming methods are most widely used in AF practice together with the methods of integer and mixed-integer programming (see e.g.[1], [2]). There is a broad research also on stochastic programming models for production and harvesting planning (see e.g. [3]–[7]) and some studies of dynamic programming utilization possibilities in AF business have been made (see [8]–[12]). The stochastic dynamic programming applications, when planning under uncertainty, can be found e.g. in [13], [14].

In Czech AF business mainly expert estimation is used for decision making and planning, especially in the medium- and small enterprises. Such approach was able to produce a good profit, but currently number of restrictions covering the soil preservation and sustainable development in natural resources utilization are arising, not yet fully taken into account in AF managerial decision making. These new restrictions make the identification of the optimal or even feasible solution using only expert estimation very difficult. Hence, there is increasing need of the decision support systems, which allow solving of optimization problems with upcoming real-world complex constraints. In this paper the initial study of the solution of typical Czech AF constrained optimization problems using mathematical programming methods is done. In each problem the respective constraint is constructed, the optimization problem for is designed, methods for its solution are presented and the applicability of the suggested approach and feasibility of its results in Czech AF practice is discussed.

2 Selected AF optimization problems

2.1 The crop planting design

A typical farmer's optimization problem is the planning of the sowing design for the next period. The goal is to maximize the profit from the harvests of crops (denote i, $1 \le i \le Q$), when considering existing restrictions on total arable land (X), capital, labour, etc. This optimization task can be formulated as a linear programming problem:

$$z^\star = \max\ \Sigma_{i=1}^{Q} c_i x_i \tag{1}$$
$$\Sigma_{i=1}^{Q} x_i \le X, \tag{2}$$
$$\Sigma_{i=1}^{Q} a_{ji} x_i \le b_j, \tag{3}$$
$$x_i \ge 0, \tag{4}$$

where decision variables x_i stand for areas of arable land planted with crop i. In the goal function (1) the parameters c_i are the random variables of total profit per 1 ha of planted crop i which means that (1-4) represents a stochastic programming problem (see e.g. [15]). Apart from the simple constraints describing above mentioned restriction, a complex problem arises when considering the restrictions on crop rotation. Each of the crops is allowed to be sowed again on the same place after distinct number r_i of periods (from one to six years according to particular crop). We introduce this restriction into model (1-4) by adding the set of linear inequality constraints (see [16])

$$1 \leq p \leq Q : \Sigma_{l=1}^{p} x_{i_l} \leq X - \bar{x}_{i_1 \cdots i_p}, \tag{5}$$

where $\{i_1, i_2, \cdots, i_p\}$ are combinations from $\{1, \ldots, Q\}$. The restriction on arable land contained in (5) replaces constraint (2) and ensures also that the crop will not be sowed on the area recently sowed by the same crop. The parameter $\bar{x}_{i_1 \ldots i_p}$ is the area of land, on which all the crops $i_1, \ldots, i_p$ were at least once sowed during relevant number of past years. The model is solvable by the stochastic programming methods and enables optimization of the next year crop plan when all the crop rotation restrictions must hold.

2.2 Dynamical programming approach to crop rotation constraints

Alternatively, one can search an optimal sow plan under the crop rotation restrictions not only for the next period but for N upcoming periods. Such a problem can be analyzed using the dynamical programming, when the planning horizon is divided into discrete and finite number of time periods in which the decisions about crop plan are made. Each period n, n, $1 \leq n \leq N+1$, defines one stage of the sequential decision process and is associated with the state space S_n formed by the set of states s_n, which represents all the feasible states of the system in the period n.

The objective in the problem of crop planting is to maximize the total profit from the crops harvested in all N time periods. The only constraint to be considered are the crop rotation restrictions. Suppose that the farmer is planning to sow crops on P particular fields, each of which is to be sowed by one crop each year. The total number of crops is $Q \leq P$ and all of them have to be sowed each year. Denote again r_i the resowing time of the crop i, $1 \leq i \leq Q$ and α the maximum resowing time.

The state of the system can be described by the history of each field up to α periods ago. Hence, the state is defined as

$$s_n = (i_{11} \ldots i_{1\alpha}, i_{21} \ldots i_{2\alpha}, \ldots, i_{P1} \ldots i_{P\alpha}) \in S_n,\ 1 \leq i_{ab} \leq Q, \tag{6}$$

where $1 \leq a \leq P$ and $1 \leq b \leq \alpha$. The symbol $i_{a1} i_{a2} \ldots i_{a\alpha}$ represents the crop history of particular field a, while i_{ab} has the meaning of the particular crop sowed b years ago on the field a. For example, for four fields ($P = 4$) and three crops ($Q = 3$), where crop 1 can be resowed after 2 years ($r_1 = 2$), crop 2 after 1 year ($r_2 = 1$) and crop 3 can be crop each year on the same field ($r_3 = 0$), the feasible state can be

$$s_n = (31, 21, 12, 23). \tag{7}$$

This means that on the first field the crop 3 was sowed one year ago and crop 1 two years ago, on the second field the crop 2 was sowed one year ago and crop 1 two years ago, etc.

The decision maker will determine the state s_{n+1} by invoking management decision $d_n(s_n) \in D(s_n)$ from finite discrete set $D(s_n)$ of decisions associated with s_n. In our problem the decisions are restricted by the crop rotation constraints, number of crops, number of fields and requirement on sowing each of the crops each period at least on one field. Each decision corresponds to one particular crop plan. For example, there are three possible decisions for the state (7) which lead to three distinct states in the stage $n + 1$ (see Tab. 1).

Table 1. Feasible decisions in stage n and related states in stage $n + 1$

$s_n \in S_n$	$d_n \in D(s_n)$	$t(s_n, d_n) \in S_{n+1}$
	$(3, 3, 2, 1)$	$(33, 32, 21, 12)$
$(31, 21, 12, 23)$	$(2, 3, 2, 1)$	$(23, 32, 21, 12)$
	$(2, 3, 3, 1)$	$(23, 32, 31, 12)$

Selecting decision d_n for state s_n earns reward $r(s_n, d_n)$ and cause the transition to state $s_{n+1} \equiv t(s_n, d_n)$. The reward is equal to total profit from harvest when decision d_n is taken:

$$r(s_n, d_n) = \Sigma_{i=1}^{Q} c_i \cdot x_i(d_n), \tag{8}$$

where $x_i(d_n)$ stands for the total area of arable land planted with crop i when decision (i.e. the crop plan) d_n was selected and the parameter c_i is total profit per 1 ha of planted crop i. To enumerate c_i the difference between average income from harvests per hectar and costs per hectar sowed by crop plant i has to be known. By analogy to the procedure showed in Tab. 1 all of the decisions and related consequent states can be generated for any number of periods, fields and crops. The emerging rewards can then be obtained using relation (8).

Let the initial crop design and sow history in the period 1 be given by state s_1 and the aim is to plan the crop design for the next N periods. We do not determine any goal state in stage $N+1$. Such a formulation represents an initial value problem, which can be solved by recursive fixing (see [17]) derived for each state in S_{N+1}. The goal is to identify the sequence of decisions $(d_1^\star, \dots, d_N^\star)$ with highest total reward. General recursive relationship describing the system and determining $(d_1^\star, \dots, d_N^\star)$ is then

$$f(s_n) = \begin{cases} 0, & n = N+1 \\ \underset{d_n \in D(s_n)}{\text{maximum}} \left\{ r(s_n, d_n) + f\left[t(s_n, d_n)\right] \right\}, & n < N+1 \end{cases} \tag{9}$$

where $f(s_n)$ is the maximum total reward obtainable at stages n through $N+1$ if state s_n is occupied.

Note, that restrictions on labor, capital and technological and ecological aspects of farming could be comprised in the model and treated by the dynamical programming unless the number of state variables, states and need of computations made would increase too much. The approach chosen, however, represents to some extent the analogy with the Czech farmer's real-life decision making. They decide about the next period crop plan using their professional experience and the cropping history data. Sometimes, they are forced to break the restriction on crop rotation, because no other possibility is available if the specific crop is to be sowed. Then there is a need of additional fertilizing, which means additional cost and lower profits in respective year. Hence, the short-run planning can decrease the long-run total profit and the dynamical programming approach - optimizing the long run profit of the farm - seems to be a good alternative to expert decisions based on routine and short-run optimizing.

2.3 Planning the forest logging

Typical decision problem for the logging company is planning the machinery use for the forthcoming harvest. The L_{jp} cubic meters of trees of type $j, 1 \le j \le k$ and quality $p, 1 \le p \le l$ are to be harvested and the harvesting shall take d days. The detailed plan must be prepared in advance, because all the harvest equipment is outsourced. Let there be two types of logging machinery to be planned: harvest machinery and transport machinery. The first type is used to cut trees at harvest area and the second type serves for trees transportation to place close to forest roads, accessible by logging trucks. Denote i, $1 \le i \le m$ the types of cutting machinery and i, $m+1 \le i \le n$ the types of transport machinery. The lowest cost solution is sought for if the volumes L_{jp} of trees are to be cut and transported and the total number of machines of type i available is restricted to X_i. Denote c_{ijp} and a_{ijp} the costs per $1\,\mathrm{m}^3$ and daily average volume (in m^3) of wood (type j, quality p) cut/transported by the machinery i, respectively. The respective optimal decision can be found as a solution of the mathematical programming problem:

$$z^\star = \min\ \Sigma_{i=1}^{n} \Sigma_{j=1}^{k} \Sigma_{p=1}^{l} c_{ijp} a_{ijp} x_{ijp}, \tag{10}$$

$$\Sigma_{j=1}^{k} \Sigma_{p=1}^{l} x_{ijp} \le X_i, \tag{11}$$

$$d \cdot \Sigma_{i=1}^{m} x_{ijp} a_{ijp} = L_{jp}, \tag{12}$$

$$d \cdot \Sigma_{i=m+1}^{n} x_{ijp} a_{ijp} = L_{jp}, \tag{13}$$

$$x_i \equiv \Sigma_{j=1}^{k} \Sigma_{p=1}^{l} x_{ijp} : \text{integer}, \quad x_{ijp} \ge 0, \tag{14}$$

where decision variables x_{ijp} stand for the number of the machines of type i cutting/transporting trees of type j and quality p. The desired solution in the form of numbers x_i of outsourced machinery of type i is to be made up using the definition in (14). When planning the machinery need in advance, the cost parameters c_{ijp} have to be treated as random variables, since the harvest costs depend on the weather (e.g. snow conditions, rain or frost introduce additional outsourcing costs). The parameters L_{jp} resp. a_{ijp} are random as well but this can be overcome when an expert estimation based on visiting the area and past experience was made.

No environmental issues were taken into account in the previous optimization model. Using machinery for harvesting in the forest causes damage to the soil depending on the season, soil type, machinery used, weather, etc. These externalities costs should be implemented into the optimization model and affect the optimal set of harvest machinery. Either the cost function c_{ijp} could increase according to environmental impact of particular machinery i or additional constraints could be formulated. We provide the following possible model for local AF decision making. In Czech Republic, mainly harvestors and workers with power saws are employed for harvesting. While the latter cause no damage to the soil during harvesting, the heavy harvestors damage the soil especially in rainy weather. For example, an aditional restriction on total distance travelled by harvestors in the harvest location depending on the soil moisture can be requested. Taking into account that the majority type of harvested trees is pine, the index j can be used for the distance range from the road to place where the machinery x_{ijp} operates. The additional restriction then takes the form:

$$\Sigma_{j=1}^{k}\Sigma_{p=1}^{l}x_{ijp}s_j \leq h_i D_i, \tag{15}$$

where s_j is the average distance to be moved from the road to harvest locality j, D_i is the total distance allowed to be moved by the machinery i during one harvesting period under ideal weather and $h_i < 1$ is the coefficient representing the soil state (soil moisture, type of the soil). The soil moisture is random variable when planning the machinery use in advance, hence h_i is random variable as well. Adding this restriction (15) into the model (10–14) the random variables appears apart from objective function also in the restrictions, which is the problem generally solvable by the methods of stochastic programming (see e.g. [15]).

3 Conclusion

Specific complex restrictions arise when developing optimization models of the AF decision problems. The problem of mathematical design of the respective constraints must be solved when the whole model is to remain solvable and its solution to be well applicable. In this contribution the analysis for the constraints on re-sowing and soil-preserving was performed and the suggested models are prepared to be applied in practical problems.

Acknowledgment The research is supported by the Grant MSM6215648904 of the Ministry of Education, Youth and Sports of Czech Republic.

References

[1] Pochet Y, Wolsey L.A. Production Planning by Mixed Integer Programming, Springer Series in Operations Research and Financial Engineering Springer (2006)

[2] Satyaveer, S. et al, Multi-commodity supply network planning in the forest supply chain Eur. J. Oper. Res. (2009), doi:10.1016/j.ejor.2008.03.024

[3] Bohle C., Maturana S., Vera J.: A robust optimization approach to wine grape harvesting scheduling. Eur. J. Oper. Res. (2009), doi: 10.1016/j.ejor.2008.12.003

[4] Torkamani, J.: Using whole-farm modeling approach to assess prospective technologies under uncertainity. Agric. Syst. 85, 138–154 (2005)

[5] Jones, P.C., Lowe, T.J., Traub, R.: Managing the seed-corn supply chain at Sygenta. Interfaces 33, 80–90 (2003)

[6] Maatman, A., Schweigman, C., Ruijs, A., van der Vlerk, M.H.: Modeling farmers response to uncertain rain fall in Burkina Faso: A stochastic programming approach. Oper. Res. 50, 399–414, (2002)

[7] Bravo M., Gonzales I.: Applying stochastic goal programming: A case study on water use planning. Eur. J. Oper. Res. (2008), doi: 10.1016/j.ejor.2008.04.034

[8] Bin Deris, S.,Ohta, H.: A machine scheduling model for large scale rice production in Malaysia. J. Opl. Res. Soc. 41, 713–723 (1990)

[9] Glen, J.J.: A linear programming model for an integrated crop and intensive beef production enterprise. J. Opl. Res. Soc. 37, 487–494 (1986)

[10] Pla L. M.: Review of mathematical models for sow herd management. Livest. Sci. 106, 107–119 (2007)

[11] Gigler, J.K., Hendrix, E.M.T., Heesen, R.A., et al.: On optimization of agri chains by dynamic programming. Eur. J. Oper. Res. 139, 613–625 (2002)

[12] Zadnik Stirn, L.:Integrating the fuzzy analytic hierarchy process with dynamic programming approach for determining the optimal forest management decisions. Ecol. Model. 194, 296–305(2006)

[13] Duffy, P.A., Taylor, C.R., Long-term planning on a corn-soybean farm: A dynamic programming analysis. Agric. Syst. 42, 57–71 (1993)

[14] Teeter, L., Somers, G.: Optimal forest harvest decission: A stochastic dynamic programming approach. Agric. Syst. 42, 73–84 (1993)

[15] Birge J., Louveaux F., Introduction to Stochastic programming, Springer-Verlag, New York (1997)

[16] Janová, J.: A stochastic programming model of the sowing plan with restrictions on crop resowing. Kybernetika (to appear)

[17] Denardo, E.V.: Dynamic programming, Dover publications, New York (2003)

A REMARK ON EMPIRICAL ESTIMATES VIA ECONOMIC PROBLEMS

Vlasta Kaňková

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
kankova@utia.cas.cz

Abstract. Optimization problems depending on a probability measure correspond to many economic applications. Since the "underlying" measure is usually unknown the decision is mostly determined on the date basis, it means on statistical (mostly empirical) estimates of the probability measure. Properties of the optimal value (and solution) estimates have been investigated many times. There were introduced assumptions under which the asymptotic distribution estimate is normal and the convergence rate is at least exponential. We generalize the assertions concerning the rate convergence. Especially we shall consider distributions with the Pareto tail.
Keywords. Economic problems, stochastic optimization, stochastic estimates, exponential rate convergence, exponential tails, Pareto distribution.

JEL Classification: C44 AMS Classification: 90 C 15

1 Introduction

Economic processes are often influenced by a random factor and a decision parameter. Constructing mathematical models we obtain mostly models depending on a probability measure. They can be static (one stage) or dynamic. A multistage stochastic programming technique can treat an essential class of dynamic cases. Employing a recursive definition (see e.g. [9], [11]), we obtain a system of one–stage (mostly) parametric problems. Consequently, the investigation of one problems is crucial also for multistage cases.

To introduce "classical" one–stage stochastic programming problem let $(\Omega, \mathcal{S}, P)$ be a probability space; $\xi\, (:= \xi(\omega) = [\xi_1(\omega), \ldots, \xi_s(\omega)])$ s–dimensional random vector defined on $(\Omega, \mathcal{S}, P)$; $F(:= F(z),\, z \in R^s)$ and P_F the distribution function and the probability measure corresponding to ξ; $F_i,\, i = 1, \ldots, s$ one–dimensional marginal distribution functions of $\xi_i,\, i = 1, 2, \ldots, s$. Let, moreover, $g_0(:= g_0(x, z))$ be a real–valued (say continuous) function defined on $R^n \times R^s$; $X \subset R^n$ be a nonempty set. If the symbol E_F denotes the operator of mathematical expectation corresponding to F, then a rather general "classical" one–stage stochastic programming problem can be introduced in the form:

Find

$$\varphi(F) = \inf\{\mathsf{E}_F g_0(x, \xi) | x \in X\}. \tag{1}$$

Since in applications very often the problem (1) has to be solved on the basis of empirical data we have to replace the measure P_F by its estimate. An empirical probability measure is a very suitable candidate for it. Consequently, the solution of the problem (1) has to be often sought (in applications) w.r.t. an "empirical problem":

Find

$$\varphi(F^N) = \inf\{\mathsf{E}_{F^N} g_0(x, \xi) | x \in X\}, \tag{2}$$

where F^N denotes an empirical distribution function determined by a random sample $\{\xi^i\}_{i=1}^N$ (not necessary independent) corresponding to the distribution function F. If we denote the optimal solutions sets of (1) and (2) by $\mathcal{X}(F)$, $\mathcal{X}(F^N)$, then (under rather general assumptions) $\varphi(F^N)$, $\mathcal{X}(F^N)$ are "good" stochastic estimates of $\varphi(F)$, $\mathcal{X}(F)$.

The investigation of the empirical (above introduced) estimates started in 1974 by [16]; followed by many papers (see e.g. [3], [15]). The investigation of the convergence rate appeared in [5] and followed e.g. by [1], [4], [12]; for weak dependent random samples e.g. by [6]. The exponential rate convergence has been proven under some relatively strong assumptions on the objective function and on the "underlying" distribution function F. Employing the stability results [8] we can see that the normal distribution is covered by this approach. However, the distribution functions with heavy tails correspond to many random factor in economic problems (see e.g. [13], [14]). The aim of this contribution is to genaralize results [5] to the case of distribution functions with the Pareto tail.

2 Auxiliary Assertion

2.1 Stability

To recall a suitable stability result, let $\mathcal{P}(R^s)$ denote the set of all Borel probability measures on $R^s, s \geq 1$ and let $\mathcal{M}_1(R^s)$ be defined by

$$\mathcal{M}_1(R^s) = \{P \in \mathcal{P}(R^s) : \int_{R^s} \|z\|_s^1 P(dz) < \infty\},$$

$\|\cdot\|_s^1$ denotes the $\mathcal{L}_1$ norm in R^s. We shall introduce a little generalized assertion of [7].

Proposition 1. [7] Let $P_F, P_G \in \mathcal{M}_1(R^s),\ X$ be a compact set. If for every $x \in X$

1. $g_0(x, z)$ is a Lipschitz function of $z \in R^s$ with the Lipschitz constant L,
2. finite $\mathsf{E}_F g_0(x, \xi), \quad \mathsf{E}_G g_0(x, \xi)$ exist,
3. $g_0(x,\, z)$ is a uniformly continuous function on $X \times R^s$,

then

$$|\inf_{x\in X} \mathsf{E}_F g_0(x, \xi) - \inf_{x\in X} \mathsf{E}_G g_0(x, \xi)| \leq L \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| dz_i. \tag{3}$$

Evidently, the right hand side of the relation (3) is in the form of a product of the Lipschitz constant and the Wasserstein metric. Replacing the distribution G by an empirical F^N we can investigate the convergence rate of the empirical estimates $\varphi(F^N)$, $\mathcal{X}(F^N)$.

2.2 Empirical Estimates

Proposition 2. Let $s = 1,\, t > 0$. Let moreover $P_F \in \mathcal{M}_1(R^1)$. If

1. P_F is absolutely continuous with respect to the Lebesgue measure on R^1,
2. there exists $\psi(N,\, t) := \psi(N,\, t,\, R)$ such that the empirical distribution function F^N fulfils for $R > 0$ the relation
$$P\{\omega : |F(z) - F^N(z)| > t\} \leq \psi(N,\, t) \quad \text{for every} \quad z \in (-R,\, R),$$

then for $\frac{t}{4R} < 1$, it holds that

$$P\{\omega : |F(z) - F^N(z)| > t\} \leq (\tfrac{12R}{t} + 1)\psi(N,\, \tfrac{t}{12R},\, R) + P\{\omega : \int_{-\infty}^{-R} F(z)dz > \tfrac{t}{3}\} +$$
$$P\{\omega : \int_{R}^{\infty} (1 - F(z))dz > \tfrac{t}{3}\} + 2NF(-R) + 2N(1 - F(R)).$$

If, moreover,

3. there exists $R := R(N)$ defined on $\mathcal{N}$ such that $R(N) \longrightarrow_{(N\longrightarrow\infty)} \infty$ and simultaneously for $\beta \in (0,\, \frac{1}{2})$

$$N^\beta \int_{-\infty}^{-R(N)} F(z)dz \longrightarrow_{N\longrightarrow\infty} 0, \qquad N^\beta \int_{R(N)}^{\infty} [1 - F(z)]dz \longrightarrow_{(N\longrightarrow\infty)} 0,$$
$$2NF(-R(N)) \longrightarrow_{(N\longrightarrow\infty)} 0, \qquad 2N[1 - F(R(N))] \longrightarrow_{(N\longrightarrow\infty)} 0, \tag{4}$$
$$(\tfrac{12R(N)N^\beta}{t} + 1)\psi(N,\, \tfrac{t}{12R(N)N^\beta},\, R(N)) \longrightarrow_{(N\longrightarrow\infty)} 0,$$

then also

$$P\{\omega : N^\beta \int_{-\infty}^{\infty} |F(z) - F^N(z)| > t\} \longrightarrow_{(N\longrightarrow\infty)} 0 \quad \text{for} \quad \beta \in (0,\, \frac{1}{2}).$$

(the symbol $\mathcal{N}$ denotes the set of natural numbers.)

Proof. The assertion can be proven by the proof technique employed in [8]. □

It is well known (see e.g. [2]) that if $\{\xi^i\}_{i=1}^\infty$ is a sequence of independent random values with a common distribution function F, then we can set

$$\Psi(N,\, t,\, R) = \exp\{-2Nt^2\}, \quad t > 0,\, R > 0,\, N = 1,\, 2,\, \ldots,. \tag{5}$$

Setting $R(N) = N^\gamma$, $\gamma + \beta \in (0,\, \frac{1}{2})$ we can see that the following assertion holds.

Corollary 1. Let $s = 1, t > 0$. Let, moreover, $P_F \in \mathcal{M}_1(R^1),\ \{\xi^i\}_{i=1}^\infty$ be a sequence of independent random values with a common distribution function F. If

1. P_F is absolutely continuous with respect to the Lebesgue measure on R^1 (we denote by f the probability density corresponding to F),
2. there exists constants C_1, C_2 and $T > 0$ such that

$$f(z) \le C_1 \exp\{-C_2|z|\} \quad \text{for} \quad z \in (-\infty, -T) \cup (T, \infty),$$

then

$$P\{\omega : N^{\beta} \int_{-\infty}^{\infty} |F(z) - F^N(z)| > t\} \longrightarrow_{N \longrightarrow \infty} 0 \quad \text{for} \quad \beta \in (0, \frac{1}{2}).$$

Proof. The assertion follows from Proposition 2 and the assumptions. □

The assumption 2 of Corollary 1 covers the normal and empirical distributions. However, many random elements corresponding to economic applications correspond to distributions with heavy tails. We employ the uniform Pareto distribution introduced in [13]. A random value ξ has a Pareto distribution if its probability measure P_F and its probability density f fulfils the relation

$$\begin{aligned} P_F\{\xi > z\} = Cz^{-\alpha}, \quad f(z) &= C\alpha z^{-\alpha-1} \text{ for } z > C^{\frac{1}{\alpha}}, \\ &0 \qquad\qquad\quad z \le C^{\frac{1}{\alpha}}. \end{aligned} \tag{6}$$

Evidently, the Pareto distribution has only one tail. Moreover, we can see that for $\alpha > 1$ it holds $P_F \in \mathcal{M}_1(R^1)$ and for $\beta \in (0, \frac{1}{2})$ and $R := R(N) = N^{\gamma}$, $\gamma > 0$ it holds

$$N^{\beta} \int_{R(N)}^{\infty} [1 - F(z)]dz = N^{\beta}[C(-\alpha+1)z^{-\alpha+1}]_{R(N)}^{\infty} = -C(-\alpha+1)N^{\beta}N^{\gamma(-\alpha+1)},$$

$$N[1 - F(R(N))] = NCN^{-\alpha\gamma} = CN^{1-\alpha\gamma}.$$

Consequently for $\gamma > \max[\frac{\beta}{\alpha-1}, \frac{1}{\alpha}]$ we can obtain that

$$N^{\beta} \int_{R(N)}^{\infty} [1 - F(z)]dz \longrightarrow_{(N \longrightarrow \infty)} 0, \quad \text{and, simultaneously,} \quad 2N[1 - R(N)] \longrightarrow_{(N \longrightarrow \infty)} 0.$$

Employing the relations (4), (5) we can see that the following assertion holds.

Corollary 2. Let $s = 1$, $t > 0$, $\alpha > 1$, and $\beta, \gamma > 0$ fulfil the inequalities $\gamma > \frac{1}{\alpha}$, $\frac{\gamma}{\beta} > \frac{1}{\alpha-1}$, $\gamma + \beta < \frac{1}{2}$. Let, moreover, $\{\xi^i\}_{i=1}^{\infty}$ be an independent random sample corresponding to the distribution function F. If

1. P_F is absolutely continuous with respect to the Lebesgue measure on R^1 (we denote by f the probability density corresponding to F),
2. there exists constants $C > 0$ and $T > 0$ such that

$$f(z) \le C\alpha z^{-\alpha-1} \quad \text{for} \quad z \in (-\infty, -T) \cup (T, \infty),$$

then

$$P\{\omega : N^{\beta} \int_{-\infty}^{\infty} |F(z) - F^N(z)| > t\} \longrightarrow_{(N \longrightarrow \infty)} 0.$$

3 Main Results

In this section we shall try to introduce the results that guarantee the exponential convergence rate of $\varphi(F^N)$ to $\varphi(F)$. First result covers the known case when the tails of the probability density are at least exponential. Evidently, this assertion covers the classical case of normal distribution. The second case will try to cover some new arising economic applications when one dimensional marginal distribution functions have the Pareto tails. It is known that this case appears for example in finance or river flow (for more details see e.g. [13]). The corresponding form of multivariate case can be found e.g. in [10]. However it is over the possibility of this contribution to deal with this case in more detail.

Theorem 1. [8] Let $t > 0$, $\{\xi^i\}_{i=1}^{\infty}$ be a sequence of independent s–dimensional random vectors with a common distribution function F. Let moreover X be a compact set. If

1. F^N is an empirical distribution function determined by $\{\xi^i\}_{i=1}^{N}$, $N = 1, 2, \ldots,$
2. P_{F_i}, $i = 1, \ldots, s$ are absolutely continuous with respect to the Lebesgue measure on R^1 (we denote by f_i, $i = 1, \ldots, s$ the probability densities corresponding to F_i),

3. there exist constants C_1, $C_2 > 0$ and $T > 0$ such that
$$f_i(z_i) \le C_1 \exp\{-C_2|z_i|\} \quad \text{for} \quad z_i \in (-\infty, -T) \cup (T, \infty), \quad i = 1, \ldots, s,$$
4. $g_0(x, z)$ is a uniformly continuous, Lipschitz (with respect to $\mathcal{L}_1$ norm) function of $z \in R^s$, the Lipschitz constant L is not depending on $x \in X$,

then
$$P\{\omega : N^{\beta}|\varphi(F^N) - \varphi(F)| > t\} \longrightarrow_{(N \longrightarrow \infty)} 0 \quad \text{for } \beta \in (0, \frac{1}{2}).$$

Proof. The assertion of Theorem 1 follows from Proposition 1 and Corollary 1. □

Theorem 2. Let $t > 0$, $\alpha > 1$, $\beta, \gamma > 0$ fulfil the inequalities $\gamma > \frac{1}{\alpha}$, $\frac{\gamma}{\beta} > \frac{1}{\alpha - 1}$, $\gamma + \beta < \frac{1}{2}$. Let, moreover, $\{\xi^i\}_{i=1}^{\infty}$ be an s–dimensional independent random sample corresponding to the distribution function F, X be a compact set. If

1. F^N is an empirical distribution function determined by $\{\xi^i\}_{i=1}^{N}$, $N = 1, 2, \ldots,$
2. P_{F_i}, $i = 1, \ldots, s$ are absolutely continuous with respect to the Lebesgue measure on R^1 (we denote by f_i, $i = 1, \ldots, s$ the probability densities corresponding to F_i),
3. there exist constants $C > 0$ and $T > 0$ such that
$$f_i(z) \le C\alpha z_i^{-\alpha-1} \quad \text{for} \quad z \in (-\infty, -T) \cup (T, \infty), \quad i = 1, \ldots, s,$$
4. $g_0(x, z)$ is a uniformly continuous, Lipschitz (with respect to $\mathcal{L}_1$ norm) function of $z \in R^s$, the Lipschitz constant L is not depending on $x \in X$,

then
$$P\{\omega : N^{\beta}|\varphi(F^N) - \varphi(F)| > t\} \longrightarrow_{(N \longrightarrow \infty)} 0.$$

Proof. The assertion of Theorem 2 follows from Proposition 1 and Corollary 2. □

4 Conclusion

The aim of the paper has been to investigate properties of the empirical estimates of the optimal value in the case of one–stage optimization problems depending on a probability measure. The introduced results are based on the stability results corresponding to the Wasserstein metric and $\mathcal{L}_1$ norm in R^s, $s \ge 1$. They do not cover only the normal distribution corresponding to many "classical" approaches in finance, however also the case of Pareto distribution. This result is crucial, namely, it is known that the distributions with "heavy" tails correspond to many new applications (for more see e.g. [10] and [13]). The achieved convergence rate is the best as possible in the case of exponential tails, in the case of Pareto distribution the introduced convergence rate is worse and depends on the parameter α.

Acknowledgement This work was supported by the Grant Agency of the Czech Republic under grants 402/07/1113, 402/08/0107 and the Grant Agency of MŠMT CR under grant LC 06075.

References

1. L. Dai, C. H. Chen, and J. R. Birge: Convergence Properties of Two–Stage Stochastic Programming. J. Optim. Theory Appl. *106* (2000), 489–509.
2. A. Dvoretzky, J. Kiefer, and J. Wolfowitz: Asymptotic minimax character of the sample distribution and the classical multinomial estimate. Ann. Math. Stat. *27* (1956) 642–669.
3. J. Dupačová and R. Wets: Asymptotic Behaviour of Statistical E stimates and Optimal Solutions of Stochastic Optimization Problems. Ann. Statist. *16* (1984), 1517–1549.
4. Y. M. Kaniovski, A King and R. Wets: Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems. Annals of Oper. Res. *56* (1995), 189–208.
5. V. Kaňková: An approximative solution of stochastic optimization problem. In: *Trans. 8th Prague Conf. 1977* Academia, Prague 1978, 349–353.
6. V. Kaňková: A note on estimation in stochastic pProgramming. J Comput. Math. *56*(1994), 97–112.
7. V. Kaňková and M. Houda: Empirical estimates in stochastic programming. In: *Proceedings of Prague Stochastics 2006* (M. Hušková and M. Janžura, eds.), MATFYZPRESS, Prague 2006, 426–436.
8. V. Kaňková: Empirical Estimates via Stability in Stochastic Programming. Research Report UTIA 2007, No 2192, Prague 2007.

9. V. Kaňková: Multistage stochastic programs via autoregressive sequences and individual probability constraints. Kybernetika *44* (2008), 2, 151–170.
10. S. Kotz, N. Balakrishnan, and N. L. Johnson: Continuous Multiviariate Distributions (Volume 1: Models and Applications). Wiley, New York 2000.
11. D. Kuhn: Generalized Bounds for Convex Multistage Programs. Springer, Berlin 2005.
12. Tito Homen de Mello: On Rates of Convergence for Stochastic Optimization Problems under Non–i.i.d. Sampling. Speps 2006.
13. M. M.Meerschaert and H.-P.Scheffler: Portfolio Modeling with Heavy Tailed Random Vectors. In: *Handbook of Heavy Tailed Distributions in Finance* (S. T. Rachev, ae.) Elsevier, Amsterdam 2003, 595–640.
14. S. Rachev and S. Mitting: Stable Paretian Models in Finance (Series in Financial Economics and Quantitative Analysis). Wiley, Chichester 2000.
15. A. Shapiro: Quantitative stability in Stochastic Programming. Math. Program. *67* (1994), 99–108.
16. R. J. B. Wets: A statistical Approach to the Solution of Stochastic Programs with (Convex) Simple Recourse. Research Report, University Kentucky, USA 1974.

DIVIDE AND PRIVATIZE: FIRMS BREAK-UP AND PERFORMANCE[1]

Jan Hanousek, Evzen Kocenda

Charles University and the Academy of Sciences, CERGE-EI, Prague
jan.hanousek@cerge-ei.cz, evzen.kocenda@cerge-ei.cz

Abstract. The literature on corporate divestures in developed countries has provided considerable evidence on their positive effects but our understanding of these operations is still limited; it is not clear whether divestitures are undertaken to correct past development, to affect future course, or as a response to business cycle. Research targeting purely financial aspects documents increases in firms' value and returns after divestiture. Literature analyzing the economic performance following divestment advocates that divestitures improve firms' performance and exhibit various positive distributional effects. However, much of the existing literature does not account for ensuing changes in ownership structure and potential endogeneity of divestiture and ownership with respect to post-divestiture performance. Further, majority of the research centers on developed economies while effects of divestments in emerging markets are largely under-researched as this segment of literature is quite new. We complement the literature by showing at large that effects of divestitures are positive, depend heavily on subsequent ownership structures and diminish with time.
Keywords. Divestiture, breakups, corporate performance, emerging markets, endogeneity.

1. Introduction.

The literature on corporate divestures in developed countries has provided considerable evidence on their positive effects but as survey by [20] shows our understanding of these operations is still limited; it is not clear whether divestitures are undertaken to correct past development, to affect future course, or as a response to business cycle. Research targeting purely financial aspects documents increases in firms' value and returns after divestiture [16, 18, 23, 7, 19, 4]. Literature analyzing the economic performance following divestment advocates that divestitures improve firms' performance and exhibit various positive distributional effects [14, 25, 26, 8, 6, 12, 9]. However, much of the existing literature does not account for ensuing changes in ownership structure and potential endogeneity of divestiture and ownership with respect to post-divestiture performance. Further, majority of the research centers on developed economies while effects of divestments in emerging markets are largely under-researched as this segment of literature is quite new [17, 21, 3, 24, 11]. In this paper we aim at answering the question how divestitures affect short and long term economic performance of firms in emerging market. In doing so we account for ensuing changes in ownership structure and endogeneity issues.

Recent history shows that divestitures in emerging markets occur on larger scale as they are often connected with subsequent privatization, restructuring and changes in ownership structures. [10] show that restructuring and changes in corporate governance are important determinants of post-privatization performance. In this respect divestitures create economy wide phenomena producing very large data sets ideal for empirical research. In this paper, we use a large firm-level data set from an emerging market economy that has became in 2004 a member of the European Union (EU). Wave of divestitures that were followed by privatizations occurred in the Czech Republic as an almost natural experiment as a control group of firms without any divestments existed during the same period. In our analysis we estimate effects of divestitures on corporate performance and contrast the results with such control group. In order to effectively control for endogeneity of divestitures and ensuing changes in ownership we employ instrumental variables to carry our estimations. In terms of the rich data from the new European Union market and robust methodology we differentiate from much of the literature.

Divestures became important in European emerging markets in the early 1990's because they became a way to restructure large state owned conglomerates. In this sense divestitures served as a first step restructuring done by government in the spirit of [22] who discussed restructuring and privatization policies along with their pace

[1] Substantial part of this paper was written while Evzen Kocenda was Visiting Fellow at the Center for Economic Studies (CES), University of Munich, whose hospitality is greatly acknowledged. We are grateful for valuable comments to David Prchal and presentation participants at BESI Conference (Acapulco, 2009), CES (Munich, 2009). Financial support from GAČR grant No. 402/09/1595 is gratefully acknowledged. The usual disclaimer applies.

and sequencing.[2] The second step restructuring, that followed privatization, was left to new private owners. In this context, (i) it may be hypothesized that divestitures improve corporate performance as the new firms strive to establish themselves on the market and to improve corporate governance. Further, as the originally underdeveloped legal and institutional framework improves in emerging market economies, divestitures and certain types of corporate ownership may enhance performance by eliminating diseconomies of scale and by serving as a disciplining device for management.[3] Alternatively, (ii) divestitures may have a negative effect on the performance of the divested units because of weak corporate governance, waning government coordination and regulation, unclear property rights, and underdeveloped legal and institutional framework that exist in emerging market economies.

We also test whether the nature of the effect depends on the type of the new ownership structure. In particular, we are able distinguish the extent to which each firm is owned by an industrial (i.e., non-financial) firm, financial company, individual owner, or state, and we can estimate the effect of different ownership patterns on corporate performance.

2. Estimation Strategy.

To summarize, in analyzing corporate performance after the wave of divestitures and privatization, we use economic and financial indicators for the period 1995-2005. The divestitures occurred in 1991-92, the subsequent privatization in 1992-93, and the distribution of shares of the privatized firms and the major post-privatization swaps of shares in 1993-94. Since for some of the firms the transfer of ownership rights ended in 1994 or very early 1995, we take 1995 to be the first year after divestitures and privatization that truly reflects the new corporate and ownership structure. Moreover, by 1995 the quality of the reported accounting and economic data by and large reflected the international standards. Finally, using data for long period of 1995 - 2005 allows us to test for time-varying effects of divestitures and new ownership structure. Hence, in our estimation we use data on early corporate performance in 1990, firm divestitures in 1991-2 and post-divestiture performance during 1995 - 2005.

Before presenting our formal model, we note that initial conditions, the nature of the divestiture, and the change in ownership may all affect subsequent corporate performance. Moreover, initial conditions are also likely to influence the nature of the divestiture and privatization. Therefore, in analyzing the effects of divestitures, we benefit from the fact that we can identify the parent company and the (to be) divested units within it (i.e., all the operating units of the parent company).

In view of the institutional setting, we model corporate performance as a function of the presence or absence of a divestiture and the type of ownership structure. Since the explanatory variables related to divestitures and ownership structure may be endogenous, we use instrumental variables in estimation

3. Data and Model.

The data originate from the wave of corporate divestitures orchestrated by the Czech government in the early 1990's, an event that satisfies conditions of a natural experiment outlined by [13]. The data were compiled Aspekt, a commercial database, and from the archives of the Ministry of Privatization and the National Property Fund of the Czech Republic. The data allow us to identify unambiguously the parent enterprises and all new units related to a surge of divestitures that occurred in 1991-1992. In this natural experiment 44 large enterprises were broken up into 131 new firms. Along with these 131 newly created firms, as a result of numerous divestitures, we also have data on 780 firms that did not experience any divestment and constitute our control group. The firms in both groups were subsequently privatized in the first wave of the voucher scheme. Our data permit us to use three indicators of corporate performance in our main model: unit labor cost measured by labor costs over sales (labor costs/sales), operating profit over labor costs (profit/labor costs), and operating profit per share (profit/equity). Our main analysis is hence based on a measure of (labor) cost effectiveness and two direct measures of profitability.

Divestiture is considered as a treatment that is absent in our control group of firms without divestment history. Hence, we formulate our panel-data treatment evaluation procedure in the spirit of [2, 15] to model the post-divestiture corporate performance as a function of the presence or absence of a divestiture, initial post-divestment ownership and subsequent changes in the ownership structure. We also control for differences in corporate capital structure and possible endogeneity of ownership. We first outline our model and then expand on relevant methodological issues. Our model of post-divestiture corporate performance is specified as:

[2] [22] voiced his restructuring and privatization arguments in early stage of the economic transformation in Europe but only now we are able to assess the implications.

[3] The theoretical model developed by [5] is relevant in this setting because it shows how divestitures may increase the probability of a takeover by value-improving management that enhances operating performance after the divestiture.

$$\Delta\pi_{it} = \alpha DIV_i + \beta_1 OI(D)_{it} + \beta_2 OI(ND)_{it} + \gamma_1 \Delta OS(D)_{it} + \gamma_2 \Delta OS(ND)_{it} + \phi \Delta CF_{it} + \varepsilon_{it}. \quad (1)$$

In the above specification index *i* denotes firms and *t* time periods. Dependent variable $\Delta\pi_{it}$ is a measure of corporate performance of firm *i* during years 1995-2005; it is defined as a difference in mean performance between two subsequent periods. Further, DIV_i is a dummy variable coded 1 if the enterprise is a divested unit and 0 if it is a firm that did not experience division. Variable $OI(D)_{it}$ measures the initial ownership structure of the divested firms after they were divested (*D*) and privatized; base year refers to 1996 after the post-privatization changes in ownership structures settled. Variable $OI(ND)_{it}$ measures the initial ownership structure in those firms that were not divided (*ND*). Variable $\Delta OS(D)_{it}$ measures the subsequent changes in ownership structure from 1996 until 2005 in those firms that were divested. In a similar spirit variable $\Delta OS(ND)_{it}$ measures the subsequent changes in ownership structure (1996 – 2005) in firms that did not experience divestiture. A control variable ΔCF_{it} measures changes in financial variables of each firm *i* to account for potential changes of a firm's capital structure, firm's development, etc. Finally, ε_{it} is the error term.

Since divestitures and changes in ownership structure may be correlated with firms' unobserved characteristics, we treat the explanatory variables related to divestitures and ownership as endogenous and apply IV estimation. The advantage of the IV procedure over the more efficient maximum likelihood estimation is that it is more robust and does not require numerical integration in the presence of the dummy variable for divestitures and share variable for ownership. Following arguments by [1] we use an instrumental variables technique in a two-stage framework in which we also directly instrument the dummy variable for divestiture. We have also checked that the residuals are free from serial correlation. We employ a two-stage least squares procedure in which we instrument all variables related to ownership. The approach provides consistent estimates that are not affected by potential model misspecification.

4. Empirical Results and Conclusions.

We show that divestitures increase the firms' indicators of profitability and scale of operations (sales), but do not reduce in a significant way their unit labor cost. The performance effects of privatization depend on the resulting ownership structure and on whether or not a firm experienced a divestiture. In particular, smaller state ownership does not result in uniform and widespread improvements in performance. It has a weakly significant positive effect on profitability of firms without divestitures, relative to other types of ownership, but other effects are insignificant or mixed. Industrial (non financial) firms as owners reduce unit labor cost and leave unchanged or increase profitability. Greater ownership by financial companies or individuals reduces profitability in divested firms and increases unit labor cost and reduces sales in firms without divestment.

The overall evidence for divestitures is hence consistent with our first hypothesis, namely that divestitures have a positive effect on performance – presumably by eliminating prior inefficiencies such as diseconomies of scale of large SOE conglomerates, weak managerial incentives and information asymmetries.

We complement the literature by showing at large that effects of divestitures are positive, depend heavily on subsequent ownership structures and diminish with time.

References:

[1] ANGRIST, J., KRUEGER, A. *Instrumental variables and the search for identification: From supply and demand to natural experiments*. Journal of Economic Perspectives 15(4), 2001, pp. 69-85.

[2] ASHENFELTER, O., CARD, D. *Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs*. Review of Economics and Statistics 67, 1985, pp. 648-60.

[3] BARTEL, A., HARRISON, A. *Ownership versus environment: Disentangling the sources of public-sector inefficiency*. Review of Economics and Statistics, 87(1), 2005, pp. 135-147.

[4] BORDE, S., AKHIGBE, A. *Valuation Effects of Foreign Divestitures*. Managerial and Decision Economics, 19(2), 1998, pp. 71-79.

[5] CHEMMANUR, T., YAN, A. *A Theory of Corporate Spin-offs*. Journal of Financial Economics, 72(2), 2004, pp. 259-290.

[6] CHO, M., COHEN, M. *The Economic Causes and Consequences of Corporate Divestiture*. Managerial and decision economics, 18(5), 1997, pp. 367-374.

[7] CUSATIS, P., MILES, J. WOOLRIDGE, R. *Restructuring through Spin-offs*. Journal of Financial Economics 33, 1993, pp. 293-311.

[8] DALEY, L., MEHROTRA, V., SIVAKUMAR, R. *Corporate Focus and Value Creation: Evidence from Spin-offs*. Journal of Financial Economics 45, 1997, pp. 257–281.

[9] DITTMAR, A., SHIVDASANI, A. *Divestitures and Divisional Investment Policies.* Journal of Finance, 58(6), 2003, pp. 2711-43.

[10] D'SOUZA, J., MEGGINSON, W., NASH, R. *The effects of changes in corporate governance and restructurings on operating performance: Evidence from privatizations*. Global Finance Journal, 18(2), 2007, pp. 157–184.

[11] HANOUSEK, J., KOČENDA, E., SVEJNAR, J. *Divestitures, Privatization and Corporate Performance in Emerging Markets*. Economics of Transition, 17(1), 2009, pp. 43-73.

[12] HANSON, R., SONG, M. *Long-term performance of divesting firms and the effect of managerial ownership*. Journal of Economics and Finance, 27(3), 2003, pp. 321-336.

[13] HARRISON, G., LIST, J. *Field Experiments.* Journal of Economic Literature, 42(4), 2004, pp. 1009–1055.

[14] HAUSMAN, J., TARDIFF, T., BELINFANTE, A. *The Effects of the Breakup of AT&T on Telephone Penetration in the United States.* American Economic Review, 83(2), 1993, pp. 178-84.

[15] HECKMAN, J., AND HOTZ, V. *Choosing among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training.* Journal of the American Statistical Association 84, 1989, pp. 862-74.

[16] HITE, G., OWERS, J. *Security Price Reactions around Corporate Spin-off Announcements.* Journal of Financial Economics, 12, 1983, pp.409–436.

[17] MAKHIJA, M. *The Value of Restructuring in Emerging Economies: the Case of the Czech Republic*. Strategic Management Journal, 25(3), 2004, pp. 243-267.

[18] MILES, J., ROSENFELD, J. *An Empirical Analysis of the Effects of Spin-off Announcements on Shareholder Wealth.* Journal of Finance 38, 1983, pp.1597–1606.

[19] MONTGOMERY, C., THOMAS A., KAMATH, R. *Divestiture, Market Valuation, and Strategy*. The Academy of Management Journal, 27(4), 1984, pp. 830-840.

[20] MOSCHIERI, C., MAIR, J. *Research on Corporate Divestitures: A Synthesis*. Journal of Management and Organization, 14(4), 2008, pp. 399-422.

[21] OMRAN, M. *The Performance of State-Owned Enterprises and Newly Privatized Firms: Does Privatization Really Matter*? World Development, 32(6), 2004, pp. 1019-1041

[22] ROLAND, G. *On the Speed and Sequencing of Privatization and Restructuring*. Economic Journal, 104(426), 1994, pp. 1158-68.

[23] SCHIPPER, K., SMITH, A. *Effects of Recontracting on Shareholder Wealth: The case of Voluntary Spin-offs*. Journal of Financial Economics 12, 1983, pp.437–467.

[24] SHIN, G. *The Profitability of Asset Sales as an Explanation of Asset Divestitures*. Pacific-Basin Finance Journal, 16(5), 2008, pp. 555-571.

[25] SLADE, M. *Beer and the Tie: Did Divestiture of Brewer-Owned Public Houses Lead to Higher Beer Prices*? Economic Journal, 108(448), 1998, pp. 565-602.

[26] WOO, C., WILLARD, G., DAELLENBACH, U. *Spin-off performance—A case of over stated expectations.* Strategic Management Journal 13, 1992, pp. 433–447.

RECONSTRUCTION OF TIME DELAY MODEL OF CAPITAL STOCK DYNAMICS BY METHOD OF OPTIMAL TRANSFORMATION.[1]

Jan Kodera, Tran Van Quang

Centre for Dynamic Economics and Econometrics, University of Economics, Prague
kodera@vse.cz, tran@vse.cz

Abstract: Our aim is a revitalization of the traditional version of Kalecki model. We try to introduce a modification of the model where the depreciation of capital is assumed. Similarly as in the Kalecki model we assume investment decision function. The investment delivery is realised with time delay after investment decisions are taken. For that reason the constructed model is described by a delay differential equation. In contrast to scalar ordinary differential equation, the differential equation with delay exhibits more complex dynamics. In empirical analysis we estimate the delay parameter using Renyi's approach of maximal correlation.
Key words: Investment dynamics, delay differential equation, Renyi's concept of maximal correlation, numerical reconstruction

Introduction

The method of optimal transformation has been claimed to be a very effective method of the reconstruction of economics models on the base of observed data recently. In this paper we present an attempt to reconstruct time delay differential equation which describes a capital stock dynamics. It is necessary to emphasize that our aim is to test the method and not to make an estimate of the model of real economy. That is why we have drawn the data from the set generated by a widely theoretically accepted time delay differential equation. The equation used for data generation is a part of famous Kalecki model. A very good description of the model can be found in the original paper of Kalecki (1935) and also in Allen R.D.G.(1956). Using this so called Kalecki investment equation we generate a dataset which serves later as an input for the reconstruction of the dynamics generated by a time delayed differential equation. The paper is structured as follows. In the first section of the paper we will describe the theoretical model. In the second section we will explain the methods which help us to analyze time series generated by a system described in the first section of the paper. It is worthy to mention that in this part our paper will be dealing with the more general time delay differential equation than the classical Kalecki investment equation. When reconstructing the dynamics of the investment equation we use the approach which is an alternative to Takens embedding approach and belongs to non-parametric regression techniques and enables to find the length of the delay and the best transformation at the same time. In the third section the alternating conditional expectation (ACE) algorithm which finds optimal transformation functions will be discussed in details. Also the very important part of the paper is the fourth section where the ACE algorithm is applied to the reconstruction of the Kalecki investment equation.

1 Investment equation

Regarding investment Kalecki (1935) assumes that there is a lag between decision to invest and the delivery of finished capital goods. This lag is equal to θ. We must distinguish these three following notions:

- investment decision (orders of investment) in time t denoted by $I(t)$
- the delivery of finished capital goods in time t denoted by $J(t)$
- capital consumption as a constant portion of capital in time t, denoted by $D(t)$

Assuming a continuous economy, the law of capital stock dynamics is given by the following differential equation

$$\dot{K}(t) = -D(t) + J(t), \qquad (1)$$

[1] The support from Czech Ministry of Education (project LC 06075) and Czech Science Foundation (project 402/09/H045) is gratefully acknowledged.

where $\dot{K}(t)$ denotes the increase of capital stock. Given the assumption on capital consumption and investment made above, we can write:

$D(t) = -\delta K(t),$ where δ is the depreciation rate,

$J(t) = I(t-\theta),$ i.e. there is a time lag between an investment decision and its delivery.

The investment decision depends on capital stock, which justifies us to write:

$I(t) = L(K(t)).$

When analyzing this dependence we must take in consideration that there are investment installation costs which decrease with the increase of capital. Installation costs influence the amount of investment. Relatively small installation costs positively motivate investors to invest more. So investment will grow with capital. But installation costs are not the only process which acts in the field of investing. We must also take into account the diminishing marginal product of capital, which acts against decreasing installation costs, i.e. assuming profit maximizing agents, investment declines as capital increases. These contradictory features justify the use of the increasing function at the beginning of the considered interval when capital is relatively low and thus the effect of installation costs dominates. But if the capital is sufficiently high, the effect of diminishing marginal product of capital prevails over the effect of relatively small installation costs and investment decisions starts to decrease. It explains why investment decision function goes up first and after reaching the maximum it goes down. The function in our contribution is chosen rather arbitrarily (and also speculatively).

Our approach is very similar to the process of calibration of economic models which is frequently used recently. Parameters in economic models are calibrated instead of being estimated which means that economists input numerical values of parameters in (usually linear) economic model. Numerical values are chosen in accordance with economic experience. Their accuracy is definitely verified by the help of model experiments. But here we work with nonlinear model. So we have to choose not only numerical values of parameters but also non-linear functions. Methodological background of our approach is the same – we will input an arbitrary structure (nonlinear function). Then we experiment with the model and the results will tell us if our option was correct or not. Let us try to apply the following investment decision function:

$$I = L(K) = \frac{1}{a + b\, e^{\alpha\,(K-\gamma)} + c\, e^{-\beta\,(K-\gamma)}}. \tag{2}$$

When constructing this function one must take in consideration that there are investment installation costs decreasing with the increasing of capital. So, investment should grow with capital. On the other hand, one must also take into account the diminishing marginal product of capital, i.e. under the assumption of profit maximization agents, the investment declines as the capital increases. These contradictory powers make the use of a function, which is increasing at the beginning of the interval and after reaching the maximum it is declining, reasonable.

Let the values of parameters involved in equation (2) be $a = 2,\ b = 1.2,\ c = 1.2,\ \alpha = 8,\ \beta = 3,\ \gamma = 1$ and we obtain

$$I = L(K) = \frac{1}{2 + 1.2\, e^{8(K-1)} + 1.2\, e^{-3\,(K-1)}}.$$

The graph of this function is shown in Fig.1. The investment decision is placed on vertical axis and capital on horizontal axis. It can be seen in the Fig. 1 that the increase of the function is rather moderate and the decline is relatively sharp.

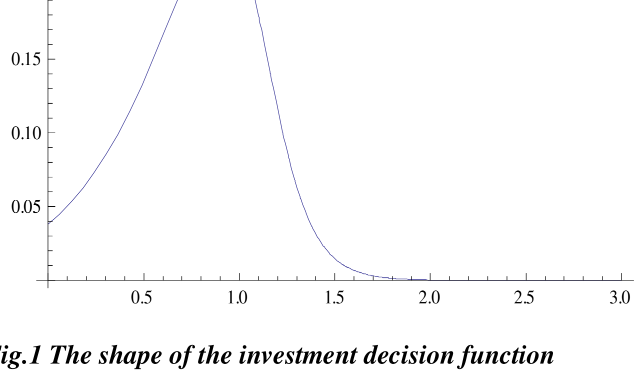


Fig.1 The shape of the investment decision function

Expressing investment as a delayed function of capital stock, the equation for capital dynamics (1) takes the following form:

$\dot{K}(t) = -\delta K(t) + L(K(t-\theta))$. (3)

This equation expresses that the increase of capital at time t equals the investment delivered for the enlargement of capital stock minus depreciated capital. To explain the second term on the right-hand side in equation (3), let consider the following: at time $t-\theta$ a investment decision is taken. During a period θ the investment is being implemented and at the end of the period it is fully realized and that means that at time t it equals the delivered investment. Replacing $L(K(t-\theta))$ in equation (3) with expression (2), we get

$$\dot{K}(t) = -\delta K(t) + \frac{1}{(a + be^{\alpha\,(K(t-\theta)-1)} + be^{-\beta\,(K(t-\theta)-1)}}. \qquad (4)$$

The dynamic of the capital described by differential equation (4) is complex and depends not only on parameters $a = 2,\ b = 1.2,\ c = 1.2,\ \alpha = 8,\ \beta = 3,\ \gamma = 1$, but also on the time delay θ. Let us set the length of lag θ = 20 and the rate of depreciation δ = 0.1 and we get instead of (4)

$$\dot{K}(t) = -0.1K(t) + \frac{1}{(2 + 1.2\,e^{8\,(K(t-20)-1)} + 1.2\,e^{-3\,(K(t-20)-1)}}. (5)$$

As an initial condition we choose $K(t)$ = 0.8 for $t \leq 0$. The solution of equation (5) obtained from Matlab is shown in the following Fig. 2.

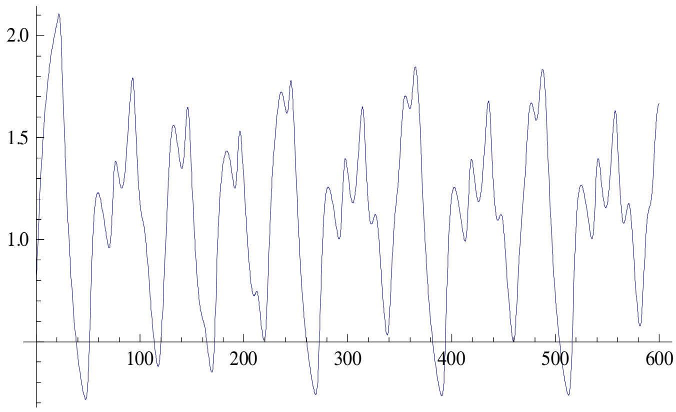


***Fig. 2 Chaotic solution of equation* (5)**

Figure 2 show that the equation (5) with chosen values of parameters is able to generate very chaotic dynamics. More details of possible chaotic behaviour of investment dynamics models can be found for example in Hallegatte et al (2008).

2 Reconstruction of non-linear time delay models from data

Let's consider a nonlinear time delay system which produces chaotic dynamics. This system is represented by a general time delayed differential equation:

$$g(\dot{x}(t)) = f_0(x(t)) + \sum_{j=1}^{k} f_j(x(t-\theta_j)),$$

where g, f_0 and f_j are unknown but continuous functions. For simplicity, let's assume g is only $\dot{x}(t)$. It is well known that the dynamics of this system in a state space could be reconstructed by using Takens theorem (1985) and embedding technique. An alternative approach to do it is to find all the time delays and functions f_0 and f_j. Here it is necessary to mention that the reconstruction is meant to be their numerical estimations, by no means to find their analytical forms. When we reconstruct the dynamics, we have only a series $\{y_t\}_{t=1}^{N}$ at our disposal as a realization of x_t. $\dot{x}(t)$ can be calculated as $\dot{x}(t) = \Delta y_t = \frac{1}{2}(y_{t+1} - y_{t-1})$. Further, the time delays and functions f_0 and f_j would be estimated and the dynamics can be approximated as

$$\dot{x}(t) = \hat{f}_0(x(t)) + \sum_{j=1}^{k} \hat{f}_j(x(t-\theta_j)), .$$

To estimate these functions and the delays, the nonparametric regression technique is used. This technique is based on Renyi's (1959) concept of maximal correlation. The maximal correlation measures the dependence between two random variables $x(t)$ and a group of delayed variables $x(t-\theta_j)$ and is defined as

$$\psi\left(x(t),\ldots,x\left(t-\theta_j\right)\right) = \sup_{\Phi,\ldots,\Phi_k}\left|R\left(\Phi\left(x(t),\sum_{j=1}^{k}\Phi\left(x\left(t-\theta_j\right)\right)\right)\right)\right|$$

where R is the linear correlation coefficient

$$R = \frac{E\left[x(t)x\left(t-\theta_j\right)\right] - E\left[x(t)\right]E\left[x\left(t-\theta_j\right)\right]}{\sqrt{\operatorname{var} x(t)\operatorname{var} x\left(t-\theta_j\right)}}$$

The maximal correlation can be interpreted as following: the functions Φ a Φ_j for which the maximum is achieved maximize the correlation between $x(t)$ and the group of delayed variables $x\left(t-\theta_j\right)$. Therefore, these variables are transformed from a nonlinear relationship to a linear one between Φ a Φ_j. The maximum correlation coefficient is determined by the alternating conditional expectation proposed by Breiman and Friedman (1985). Before going to the computational aspect of this algorithm, let's have a quick look at its principle.

Let $x(t)$, $x(t-\theta_1)$, …, $x(t-\theta_k)$ be random variables and $\Phi\left(x(t)\right),\Phi_1\left(x\left(t-\theta_1\right)\right),\ldots,$ $\Phi_k\left(x\left(t-\theta_k\right)\right)$ be zero-mean functions of the corresponding variables. After regressing $\Phi\left(x(t)\right)$ on $\sum_{j=1}^{k}\Phi_j\left(x\left(t-\theta_j\right)\right)$, the unexplained part of the variability of regressand $\Phi\left(x(t)\right)$ by the regressor $\sum_{j=1}^{k}\Phi_j\left(x\left(t-\theta_j\right)\right)$ is

$$e^2 = \frac{E\left\{\left[\Phi\left(x(t)\right)-\sum_{j=1}^{k}\Phi_j\left(x\left(t-\theta_j\right)\right)\right]^2\right\}}{E\Phi^2\left(x(t)\right)}.$$

The optimal transformations of these variables are those functions $\Phi^*\left(x(t)\right),\Phi_1^*\left(x\left(t-\theta_1\right)\right)$, …, which minimize the sum of residuals of the regression. In the bivariate case, the optimal transformations satisfy

$$R\left(\Phi^*\left(x(t)\right),\Phi_1^*\left(x\left(t-\theta\right)\right)\right) = \max_{\Phi,\Phi_1}\rho\left(\Phi\left(x(t)\right),\Phi_1\left(x\left(t-\theta\right)\right)\right).$$

3 The algorithm

The procedure for finding the optimal transformation functions is iterative. Again for simplicity, let's assume $E\left[\Phi\left(x(t)\right)\right]^2 = 1$ and all functions are zero-mean ones. For a group of functions $\Phi_1\left(x\left(t-\theta_1\right)\right),\ldots,$ $\Phi_k\left(x\left(t-\theta_k\right)\right)$ we minimize

$$e^2 = \frac{E\left\{\left[\Phi\left(x(t)\right)-\sum_{j=1}^{k}\Phi_j\left(x\left(t-\theta_j\right)\right)\right]^2\right\}}{E\Phi^2\left(x(t)\right)}$$ with respect to $\Phi\left(x(t)\right)$ which yields

$$\Phi^1\left(x(t)\right) = \frac{E\left[\sum_{j=1}^{k}\Phi_j\left(x\left(t-\theta_j\right)|x(t)\right)\right]}{\left\|E\left[\sum_{j=1}^{k}\Phi_j\left(x\left(t-\theta_j\right)|x(t)\right)\right]\right\|}.$$

The next step is to minimize e^2 with respect to $\Phi_1\left(x\left(t-\theta_1\right)\right),\ldots,\ \Phi_k\left(x\left(t-\theta_k\right)\right)$. It is achieved through another iterative procedure. At this stage, we minimize e^2 with respect to a single function $\Phi_{j^*}\left(x\left(t-\theta_{j^*}\right)\right)$ for a given set of functions $\Phi\left(x(t)\right), \Phi_1\left(x\left(t-\theta_1\right)\right),\ldots,\ \Phi_k\left(x\left(t-\theta_k\right)\right)$ without $\Phi_{j^*}\left(x\left(t-\theta_{j^*}\right)\right)$. The solution then is

$\Phi^1_{j^*}\left(x\left(t-\theta_{j^*}\right)\right) = E\left[\Phi\left(x(t)\right) - \sum_{j\neq j^*}\Phi\left(x\left(t-\theta_j\right)\right)\middle| x\left(t-\theta_{j^*}\right)\right]$. These steps are repeated until the quantity e^2 fails to decrease. The algorithm therefore is called as alternating conditional expectation (ACE) algorithm. The computational pseudocode for the algorithm would be as follows:

Set $\Phi\left(x(t)\right) = x(t)/\left\|x(t)\right\|$ and $E\left[\Phi_j\left(x\left(t-\theta_j\right)\right)\right] = 0$ for $j = 1, \ldots, k$

Repeat until e^2 fails to decrease

Repeat until e^2 fails to decrease

For j^* = 1 to k

$$\Phi^1_{j^*}\left(x\left(t-\theta_{j^*}\right)\right) = E\left[\Phi\left(x(t)\right) - \sum_{j\neq j^*}\Phi\left(x\left(t-\theta_j\right)\right)\middle| x\left(t-\theta_{j^*}\right)\right]$$

Replace $\Phi_j\left(x\left(t-\theta_j\right)\right)$ with $\Phi_{j^*}\left(x\left(t-\theta_{j^*}\right)\right)$

End for loop

End Inner loop

$$\Phi^1\left(x(t)\right) = \frac{E\left[\sum_{j=1}^{k}\Phi_j\left(x\left(t-\theta_j\right)\middle| x(t)\right)\right]}{\left\|E\left[\sum_{j=1}^{k}\Phi_j\left(x\left(t-\theta_j\right)\middle| x(t)\right)\right]\right\|}$$

Replace $\Phi\left(x(t)\right)$ with $\Phi^1\left(x(t)\right)$

End outer loop

Solutions are $\Phi^*\left(x(t)\right), \Phi^*_1\left(x\left(t-\theta_1\right)\right),\ldots,\ \Phi^*_k\left(x\left(t-\theta_k\right)\right)$

End Algorithm.

4 Application of the ACE algorithm

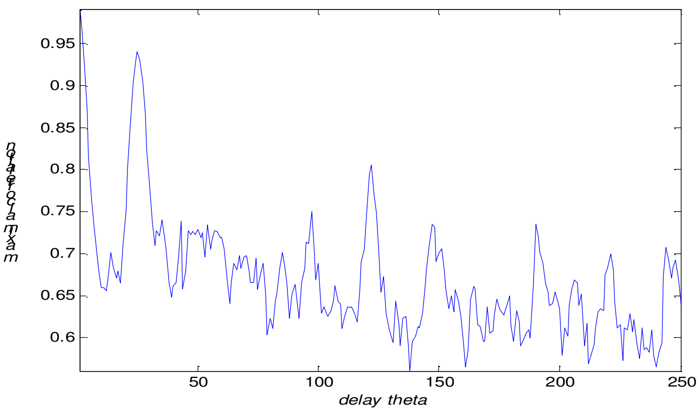


Fig. 3: The relationship between of maximal correlation coefficient and the delay

In this part of our contribution, we apply the ACE algorithm on data generated by Kalecki investment decision model. We use the MATLAB to solve the Kalecki time delayed differential equation setting initial and maximal steps equal 0.1 for a time span t from 0 to 500. The value of the delay is 20. The initial value of $K(t)$ is set to be 0.8. Then we get a series of $K(t)$ with 5001 data points. From this series we choose every tenth data point to form dataset of 500 points. For the delay from 1 to 250 we use the ACE algorithm to find the maximal correlation value between $\Phi\left(K\left(t\right)\right)$ and $\Phi_1\left(K\left(t-\theta\right)\right)$). Fortunately, in our model only one delay is present, so we can use a shorter ACE algorithm just with the outer loop. To estimate expected values of $\Phi\left(K\left(t\right)\right)$ with respect to $\Phi_1\left(K\left(t-\theta\right)\right)$ and vice versa we use the boxcar window method where the conditional expectation value $E\left[Y\middle|X\right]$ at each site *t* is:

$$E\left[Y\middle|X_t\right] = \frac{1}{2W+1}\sum_{i=-W}^{W} Y_{t+i} \text{ ,}$$

where W is the length of the window on one or the other side. Here we choose $W = 5$. The relationship between the values of the maximal correlation and the delay are shown in figure 3. After the fast drop for small delays it reaches the maximum value for $\hat{\theta}$ = 25. This value is higher than the actual delay (θ = 20).

5 Conclusion

As shown by Kalecki, there is a lag between an investment decision and installation of investment goods and the whole investment process can be described by a differential equation with (a) lagged argument(s). Under certain circumstances, this process is able to produce relatively complex chaotic dynamics. Following this theoretical basis, we generate a time series according to the Kalecki investment time delayed differential. Then we use a non-parametric regression technique called the alternating conditional expectation (ACE) algorithm to find the optimal length of delay and transformation functions. Using the ACE algorithm written in MATLAB on our numerically generated data, we find that the optimal delay for out dataset is 25, which is rather higher than the actual value 20. Unlike Voss and Kurths (1997, 2002) who examined the Mackey and Glass model for advertisement and found the length of the delay is exactly the same as the theoretical one, our result show that the ACE algorithm is not always able to determine precisely the actual length of the delay.

6 References

[1] ALLEN R,D.G. (1956) *Mathematical Economics*, Macmillan, London.

[2] BREIMAN L., FRIEDMAN J. H. (1985): *Estimating Optimal Transformations for Multiple Regression and Correlation*, Journal of American Statistical Association, Vol. 80, No. 391, September 1985.

[3] HALLEGATTE S., GHIL M., DUMAS P., HOURCADE J. (2008) *Business cycles, bifurcations and chaos in a neo-classical model with investment dynamics.* Journal of Economic Behavior & Organization, Volume 67, Issue 1, July 2008, Pages 57-77.

[4] KALECKI, M. (1935) *A Macro-Economic Theory of Business Cycles*, Econometrica 3, pp. 327-344.0

[5] RENYI, A (1959): *On Measures of Dependence*, Acta Mathematica Academiae Scientiarum Hungaricae, 10, pp. 441-51.

[6] TAKENS F. (1985): *On the numerical determination of the dimension of and attractor*, in Braaksma B. L. J., Broer H. W., Takens F., Dynamical Systems and Bifurcations. Lecture Notes in Math, Springer, Heidelberg – New York.

[7] VOSS H., KURTHS J (1997): *Reconstruction of non-linear time delay models from data by the use of the optimal transformations*, Physics Letters A 234, pp. 336-344.

[8] VOSS H., KURTHS J (2002): *Analysis of Economic Delayed Feedback Dynamics*, in Modelling and Forecasting Financial Data: Technique of Nonlinear Dynamics, eds. Soofi A. and Cao L., Springer, 2002, ISBN: 978-0-7923-7680-4.

STABILITY OF SSD EFFICIENCY - MONTHLY VERSUS YEARLY RETURNS

Miloš Kopa

Institute of Information Theory and Automation of the ASCR, Department of Econometrics,
kopa@karlin.mff.cuni.cz

Abstract. This paper deals with second-order stochastic dominance portfolio efficiency. In existing portfolio efficiency tests with respect to the second-order stochastic dominance (SSD) criterion, the scenario approach for random returns is assumed. We analyse the stability of SSD portfolio efficient classification with respect to two possible set of historical scenarios: monthly returns and yearly returns. In both cases, 20 years history is considered. For both sets of scenarios we test SSD efficiency of almost one hundred thousand portfolios that can be formed from ten US industry representative portfolios. For each portfolio, we compare the monthly returns results with yearly return results.

Keywords. Second-order stochastic dominance, portfolio inefficiency, scenario approach.

1 Introduction

When solving portfolio selection problem several approaches can be used: mean-risk models, maximising expected utility problems, stochastic dominance criteria, etc. If the information about the risk attitude of a decision maker is not perfectly known one may adopt stochastic dominance approach to test an efficiency of a given portfolio with respect to considered set of utility functions. The second-order stochastic dominance is the most common stochastic dominance relations because of its risk aversion interpretation and relation to Conditional Value-at-Risk, see e.g. Ogryczak and Ruszczynski (2002), Levy (2006) or Kopa and Chovanec (2008).

In the context of portfolio selection problem, Post (2003) and Kuosmanen (2004) develop linear programming tests for testing if a given portfolio is SSD efficient relative to all possible portfolios formed from a set of assets. Using the formulation in terms of concave utility functions and the first-order condition for portfolio optimization, Post derives a computationally efficient LP test. A limitation of this test is that it focuses exclusively on the efficiency classification of the evaluated portfolio and gives minimal information about directions for improved allocation if the portfolio is SSD inefficient. Moreover the Post test is derived for strict SSD portfolio efficiency instead of general SSD portfolio efficiency as it is defined in Kuosmanen (2004) and Ruszczynski and Vanderbei (2003). Using the formulation in terms of second quantile functions, Kuosmanen derives a test that identifies another, SSD efficient portfolio that dominates the evaluated portfolio (if the latter is inefficient). This test involves solving two linear problems; one for a necessary condition and one for a sufficient condition. Unfortunately, the problem for the sufficient condition is large and introduces substantial additional computational burden. And therefore Kopa and Chovanec (2008) derived a new linear programming test. This test is approximately 6-times faster than the Kuosmanen test and identifies another, SSD efficient portfolio that dominates the evaluated portfolio, too.

In all these SSD portfolio efficiency tests, a scenario approach is assumed, that is, asset returns have a discrete probabilistic distribution. To analyse the stability of SSD portfolio efficiency classification with respect to changes in scenarios, Kopa (2009) suggested subsampling methods. Using bootstrap techniques SSD inefficiency of the US market portfolio with high confidence level was shown.

In this paper we compare SSD portfolio efficiency based on monthly returns (240 scenarios) with that based on yearly returns (20 scenarios). We apply the test derived in Kopa and Chovanec (2008) for almost one hundred thousand portfolios that can be formed from ten US industry representative portfolios. For each portfolio, we compare the results.

The reminder of this text is structured as follows. Section 2 introduces notations and basic definitions. In Section 3, a test for testing SSD portfolio efficiency is recalled. Section 4 presents an empirical application to compare the SSD portfolio efficiency classification for both considered sets of scenarios. Finally, Section 5 summaries the results and discuss the ideas for future research.

2 Preliminaries

Consider a random vector $\mathbf{r} = (r_1, r_2, ..., r_N)'$ of returns of N assets and 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

$$\Lambda = \{\boldsymbol{\lambda} \in R^N | \mathbf{1}'\boldsymbol{\lambda} = 1, \ \ \lambda_n \geq 0, \ \ n = 1, 2, \ldots, N\}.$$

The tested portfolio is denoted by $\boldsymbol{\tau} = (\tau_1, \tau_2, ..., \tau_N)'$. Following Ruszczynski and Vanderbei (2003), Kuosmanen (2004), Kopa and Chovanec (2008), we define second-order stochastic dominance relation in a strict form. Let $F_{\mathbf{r}'\boldsymbol{\lambda}}(x)$ 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:

$$F^{(2)}_{\mathbf{r}'\boldsymbol{\lambda}}(t) = \int_{-\infty}^{t} F_{\mathbf{r}'\boldsymbol{\lambda}}(x)dx \tag{1}$$

Definition 1: *Portfolio $\boldsymbol{\lambda} \in \Lambda$ dominates portfolio $\boldsymbol{\tau} \in \Lambda$ by second-order stochastic dominance ($\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$) if and only if*

$$F^{(2)}_{\mathbf{r}'\boldsymbol{\lambda}}(t) \leq F^{(2)}_{\mathbf{r}'\boldsymbol{\tau}}(t) \qquad \forall t \in \mathbb{R}$$

with at least one strict inequality.

The equivalent definition, presented in e.g. Levy (2006) or Kopa and Chovanec (2008) is based on comparison of expected utility of portfolio returns:

$$\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau} \Longleftrightarrow \mathrm{E}u(\mathbf{r}'\boldsymbol{\lambda}) \geq \mathrm{E}u(\mathbf{r}'\boldsymbol{\tau})$$

for all concave utility functions u with strict inequality for at least some concave utility function.

Definition 2: *A given portfolio $\boldsymbol{\tau} \in \Lambda$ is SSD inefficient if and only if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ such that $\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$. Otherwise, portfolio $\boldsymbol{\tau}$ is SSD efficient.*

This definition classifies portfolio $\boldsymbol{\tau} \in \Lambda$ as SSD efficient if and only if no other portfolio is better for all risk averse and risk neutral decision makers.

We follow Pflug (2000) in defining *conditional value-at-risk* (CVaR) for portfolio losses ($-\mathbf{r}'\boldsymbol{\lambda}$).

Definition 3: *Let $\alpha \in \langle 0, 1 \rangle$. Conditional value-at-risk of portfolio $\boldsymbol{\lambda} \in \Lambda$ at level α is the optimal value of objective function of the following optimization problem:*

$$\mathrm{CVaR}_\alpha(\boldsymbol{\lambda}) = \min_{a \in \mathbb{R}} \{a + \frac{1}{1-\alpha} \mathbb{E}[-\mathbf{r}'\boldsymbol{\lambda} - a]^+\}$$

where $[x]^+ = \max(x, 0)$.

It was shown in Uryasev and Rockafellar (2002) that the $\mathrm{CVaR}_\alpha(\boldsymbol{\lambda})$ can be also defined as the conditional expectation of $-\mathbf{r}'\boldsymbol{\lambda}$, given that $-\mathbf{r}'\boldsymbol{\lambda} > F^{(-1)}_{-\mathbf{r}'\boldsymbol{\lambda}}(\alpha)$, i.e.

$$\mathrm{CVaR}_\alpha(-\mathbf{r}'\boldsymbol{\lambda}) = \mathbb{E}(-\mathbf{r}'\boldsymbol{\lambda} | -\mathbf{r}'\boldsymbol{\lambda} > F^{(-1)}_{-\mathbf{r}'\boldsymbol{\lambda}}(\alpha)),$$

where

$$F^{(-1)}_{-\mathbf{r}'\boldsymbol{\lambda}}(\alpha) = \min\{u : F_{-\mathbf{r}'\boldsymbol{\lambda}}(u) \geq \alpha\}.$$

Since we apply scenario approach, following Rockafellar and Uryasev (2002) and Pflug (2000), it can be rewritten as a linear programming problem:

$$
\begin{aligned}
\mathrm{CVaR}_\alpha(\boldsymbol{\lambda}) = \min_{a,w_t} \quad & a + \frac{1}{(1-\alpha)T}\sum_{t=1}^{T} w_t \\
s.t. \quad & w_t \geq -\mathbf{x}^t\boldsymbol{\lambda} - a \\
& w_t \geq 0.
\end{aligned} \tag{2}
$$

There exists a Fenchel duality connection between SSD relation and CVaR proved in Ogryczak and Ruszczynski (2002) and Kopa and Chovanec (2008).

Theorem 1: *Let $\boldsymbol{\lambda}, \boldsymbol{\tau} \in \Lambda$. Then*

$$\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau} \quad \Leftrightarrow \quad \mathrm{CVaR}_\alpha(\boldsymbol{\lambda}) \leq \mathrm{CVaR}_\alpha(\boldsymbol{\tau}) \qquad \forall \alpha \in \{0, \frac{1}{T}, \frac{2}{T}, \ldots, \frac{T-1}{T}\}.$$

with strict inequality for at least some $\alpha \in \{0, \frac{1}{T}, \frac{2}{T}, \ldots, \frac{T-1}{T}\}$.

3 SSD portfolio efficiency test

In this section we present the SSD portfolio efficiency linear programming test in the form of a necessary and sufficient condition derived in Kopa and Chovanec (2008). Let

$$D^*(\boldsymbol{\tau}) = \max_{D_k,\lambda_n,b_k,w_k^t} \sum_{k=1}^{T} D_k \tag{3}$$

$$
\begin{aligned}
\text{s.t.} \quad \mathrm{CVaR}_{\frac{k-1}{T}}(\boldsymbol{\tau}) - b_k - \tfrac{1}{(1-\frac{k-1}{T})T}\textstyle\sum_{t=1}^{T} w_k^t &\geq D_k, & k &= 1,2,\ldots,T \\
w_k^t &\geq -\mathbf{x}^t\boldsymbol{\lambda} - b_k, & t,k &= 1,2,\ldots,T \\
w_k^t &\geq 0, & t,k &= 1,2,\ldots,T \\
D_k &\geq 0, & k &= 1,2,\ldots,T \\
\boldsymbol{\lambda} &\in \Lambda.
\end{aligned}
$$

Theorem 2: *Let $D^*(\boldsymbol{\tau})$ be given by (3). If $D^*(\boldsymbol{\tau}) > 0$ then $\boldsymbol{\tau}$ is SSD inefficient and $\mathbf{r}'\boldsymbol{\lambda}^* \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$. Otherwise, $D^*(\boldsymbol{\tau}) = 0$ and $\boldsymbol{\tau}$ is SSD efficient.*

The alternative SSD portfolio efficiency tests were suggested by Post (2003) and Kuosmanen (2004). The advantages and disadvantages of these three tests were discussed in Kopa and Chovanec (2008).

4 Empirical application

We consider ten US industry portfolios which represents our basic assets. Using a regular grid on set Λ with step size 0.1, we create 92378 portfolios from these assets. In general the number of portfolios from the regular grid with step size s is given by formula:

$$\prod_{i=1}^{N-1}(1 + \frac{1}{si})$$

where N is the number of assets. For each of these portfolios we apply the SSD portfolio efficiency test, solving (3), for 240 monthly excess returns scenarios and then for 20 yearly excess returns scenarios. We consider historical scenarios from September 1987 to August 2007. Our aim is to compare the set of SSD efficient portfolios using monthly excess returns with that using yearly excess returns.

Using monthly excess returns, 94.1% portfolios from the grid were classified as SSD inefficient and 5.9% as SSD efficient. It means that SSD criteria dramatically reduced the set of all possible portfolios and all risk averse investors will chose their optimal allocations from this reduced set.

		Monthly returns	
		SSD efficient portfolios	SSD inefficient portfolios
Yearly returns	SSD efficient portfolios	1.1	1.4
	SSD inefficient portfolios	4.8	92.7

Table 1. Percentage comparison of SSD portfolio efficiency sets - monthly versus yearly excess returns.

	Monthly returns	Yearly returns
SSD efficient portfolios	19	44
SSD inefficient portfolios	99	95

Table 2. Relative percentage levels of results matching.

Using yearly excess returns, 97.5% portfolios from the grid were classified as SSD inefficient and 2.5% as SSD efficient. Again, we can see large portfolio set reduction.

The following table presents the full comparisons of these two cases.

From Table 1, we can conclude that 93.8% portfolios were equally classified in both cases. This high level of coinciding is probably caused by quite large number of SSD inefficient portfolios. If we limit our attention to SSD efficient portfolios then we can see that only 19% of SSD efficient portfolios using monthly scenarios is equally classified in the case of yearly returns. Similarly, 44% of SSD efficient portfolios using yearly scenarios is equally classified in the case of monthly returns. The same analysis can be done for comparing SSD inefficient portfolios. Table 2 summarizes these relative levels of SSD portfolio efficiency results matching. All computations were done in software GAMS 22.8 using solver CPLEX.

5 Summary and concluding remarks

In this study we analyzed the impact of chosen historical data frequency on SSD portfolio efficiency classification. We compared two cases: monthly returns versus yearly returns, both for 20 years history. We constructed almost one hundred thousand portfolios from ten US representative industry portfolios. We applied SSD portfolio efficiency test derived in Kopa and Chovanec (2008) for all portfolios and for both returns cases. Comparing the results, we concluded that 93.8 % portfolios were equally classified in both cases.

To improve the quality of these comparisons more historical scenarios can be used. However, it will increase the computational requests. In addition, including quarterly returns can make this analysis more complex.

Acknowledgement

This work was partially supported by the Grant Agency of the Czech Republic (grant 402/07/1113 and grant 201/07/P107).

References

[1] M. Kopa: An efficient LP test for SSD portfolio efficiency, submitted in Management Science, available at: http://ssrn.com/abstract=1340863.

[2] M. Kopa and P. Chovanec: A second-order stochastic dominance portfolio efficiency measure, Kybernetika, 44, 2 (2008), 243–258.

[3] T. Kuosmanen: Efficient diversification according to stochastic dominance criteria. Management Science, 50, 10 (2004), 1390–1406.

[4] H. Levy: Stochastic dominance: Investment decision making under uncertainty. Second edition. Springer Science, New York 2006.

[5] W. Ogryczak and A. Ruszczyński: Dual stochastic dominance and related mean-risk models. SIAM J. Optim., 13 (2002), 60–78.

[6] G. Ch. Pflug: Some remarks on the value–at–risk and the conditional value-at-risk. In: Probabilistic Constrained Optimization: Methodology and Applications (S. Uryasev, ed.), Kluwer Academic Publishers, Norwell MA 2000, pp. 278–287.

[7] Empirical tests for stochastic dominance efficiency. Journal of Finance, 58 (2003), 1905–1932.

[8] A. Ruszczyński and R. J. Vanderbei: Frontiers of Stochastically Nondominated Portfolios. Econometrica, 71, 4 (2003), 1287–1297.

[9] S. Uryasev and R. T. Rockafellar: Conditional value-at-risk for general loss distributions. Journal of Banking & Finance, 26 (2002), 1443–1471.

THE VALUE OF RISKY PROSPECT RELATIVE TO THE INTERVAL OF REFERENCE POINTS

Donata Kopańska-Bródka, Renata Dudzińska-Baryła

University of Economics in Katowice, Poland
Department of Operations Research,
donata.kopanska-brodka@ae.katowice.pl
renata.dudzinska-baryla@ae.katowice.pl

Abstract. In the Prospect Theory outcomes are understood as gains and losses relative to a reference point. In our approach the reference interval is considered. Handling consequences in this way gains and losses and the value function are evaluated quite differently. The appropriate value function of the changes in relation to the interval of the reference points is constructed. Some interesting properties of the value function will be presented. We will show a particular relationship between the prospect's evaluation in the framework of the Cumulative Prospect Theory and the value in our approach.
Keywords. cumulative prospect theory, reference point, reference interval, utility

1. Introduction.

The Expected Utility and the Rank Dependent Utility are the traditional approaches to analyzing decision making under uncertainty and risk. In these theories the evaluation of the risky alternatives is based on the utility of possible absolute payoffs. The Prospect Theory and the Cumulative Prospect Theory proposed by Kahneman and Tversky [2; 5] are quite different. They are based on the presumption, that peoples are more sensitive to changes than to absolute levels of outcomes. The decision makers valuate the deviations of possible payoffs and the outcomes are interpreted as gains and losses relative to some status quo reference point. The level of the reference point is not specified by prospect theory and is exogenously given. The dependences of the reference points is relatively new concept and raises new problems. The level of this point depends on the base pay, the past and present consumption as well as the expectation of the future consumption. The influence of the location of this point on decision making process is huge and needs to be explore. The reference point is always one of the available options but if it is not available the decision maker can take one of the alternatives as their reference point [1]. Sugden [4] considered the reference point as an uncertain act. The changes of the reference point in context of the risk aversion have been explored by some authors. Levy and Wiener [3] suggested that changes in payoffs also reflect the temporary decision makers' attitude towards risk. This attitude changes and depends on the initial wealth.

In our paper we focus on variations in the reference point within some interval. We suggest that the deviations in the utilities of the reference points are available and influence the choices among risky alternatives. Our research is based on the Cumulative Prospect Theory approach. The risky alternative is expressed as gains and losses relative to some reference point or rather the reference interval and then it is valuated. But earlier the reference point is valuated based on the concept of the utility theory.

2. The Cumulative Prospect Theory.

In the Prospect Theory [2; 5] the value of the risky alternative depends on the value function $v(x)$ and the probability weighting function $g(p)$. The function $v(\cdot)$ assigns to each outcome x a number $v(x)$ which reflects the subjective value of that outcome. The function $g(\cdot)$ assigns to each probability p a decision weight $g(p)$. This decision weight reflects the impact of p on the over-all value of the prospect.

The difference between the Prospect Theory [2] and the Cumulative Prospect Theory [5] is in the formulation of the weights. In the Cumulative Prospect Theory the value of the prospect is multiplied by the difference between transformed adequately cumulated probabilities instead of the simple transformed probabilities.

The main property of the Prospect Theory is that outcomes are understood as gains and losses, rather than the final states. Gains and losses are the changes in the wealth from a reference point. This point can be interpreted as a current wealth of a decision maker or the status quo. The value function in the Prospect Theory is:

- defined for gains and losses relative to some reference point;
- concave for gains (v''(x)<0 for x>0) and convex for losses (v''(x)>0 for x<0). It is consistent with the diminishing marginal value of gains and losses. The concavity of the value function for gains means that decision makers are risk avers and they choose certain decision and do not take risky action connected with even higher gain. Whereas considering losses decision makers are risk-prone (convexity of the value function) and they choose to take risky loss than certain loss;
- steeper for losses and than for gains, because the decision makers feel the loss more than the gain of the same absolute value.

The decision weight measures the impact of event on the desirability of prospect. The weighting function g is an increasing function of p and g(0)=0 and g(1)=1. Furthermore, for small probabilities g(p)>p, it means that the decision makers overweight low probabilities. And the next property of the probability weighting function is that for all $p \in (0,1)$, g(p)+g(1-p)<1. As a consequence of this property we have that g(p)<p for high probabilities.

The risky alternative (or prospect) can be written as gains and losses relative to the reference point in the following form (we assume the ascending order of outcomes):

$$L = \{(x_{-m};p_{-m});\ldots;(0;p_0);\ldots;(x_n;p_n)\}$$

where $x_{-m} < \ldots < 0 < \ldots < x_n$.

The CPT value of such a prospect is defined as:

$$V(L) = V^{+}(L) + V^{-}(L)$$

and

$$V^{+}(L) = v(x_n)g(p_n) + \sum_{k=1}^{n} v(x_{n-k})\left[g\left(\sum_{j=0}^{k} p_{n-j}\right) - g\left(\sum_{j=0}^{k-1} p_{n-j}\right)\right]$$

$$V^{-}(L) = v(x_{-m})g(p_{-m}) + \sum_{k=1}^{m} v(x_{-m+k})\left[g\left(\sum_{j=0}^{k} p_{-m+j}\right) - g\left(\sum_{j=0}^{k-1} p_{-m+j}\right)\right]$$

Kahneman and Tversky have proposed [5] the form of the value function and the probability weighting function and based on the empirical research they estimated the values of the parameters. The value function had the form as follows:

$$v(x) = \begin{cases} -\lambda(-x)^{\alpha}, & x < 0 \\ x^{\alpha}, & x \geq 0 \end{cases}$$

where α=0.88 and λ=2.25, and the probability weighting function was following:

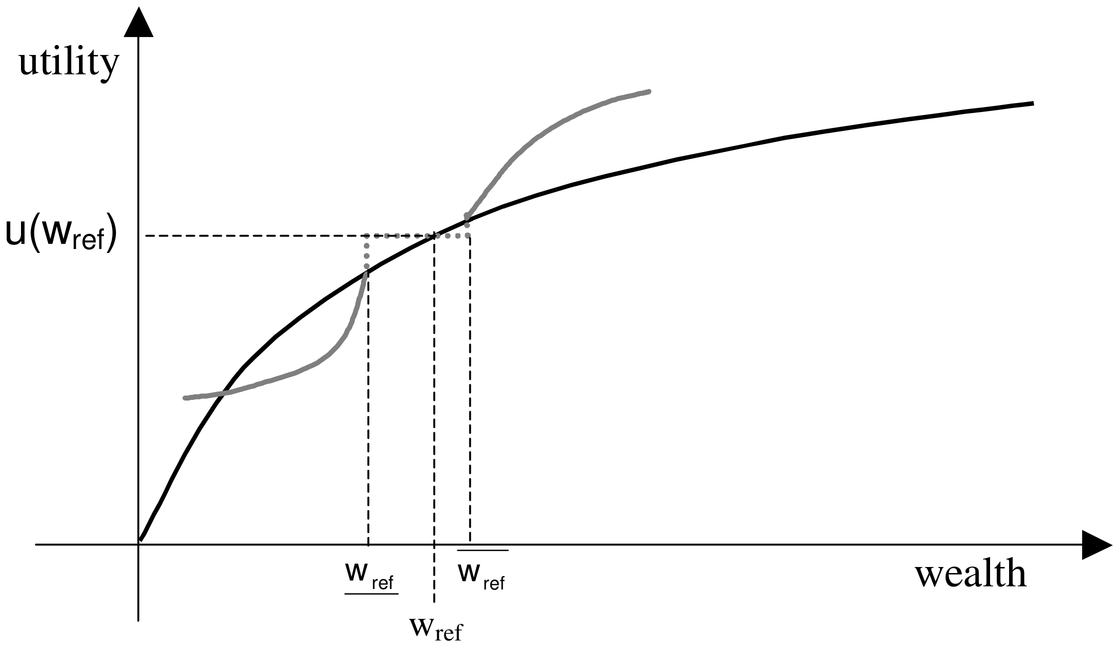
$$g(p) = \frac{p^{\gamma}}{\left[p^{\gamma} + (1-p)^{\gamma}\right]^{1/\gamma}}$$

where γ=0.61 for gains and γ=0.69 for losses.

3. Intervals of the reference points in the Cumulative Prospect Theory.

3.1. The concept.

Often in the real decision situations we are satisfied with a certain interval of outcomes instead of the single reference value, e.g. it doesn't matter if we possess \$100 or \$99 or \$101, because the change in the utilities of these values are slight. So the outcomes of the decision alternative should be valuated relative to the interval of the reference points. The illustration of such a concept presents the graph 1.



Graph 1. The valuation of the risky alternative relative to the interval of the reference points.

The value of the gains and losses relative to the reference interval can be written as follows:

$$V\left(L_{<\underline{w_{ref}},\overline{w_{ref}}>}\right)=V^{-}\left(L_{\underline{w_{ref}}}\right)+V^{+}\left(L_{\overline{w_{ref}}}\right)$$

The next interesting matter is how to determine such a reference interval in which the decision maker is indifferent to the changes in theirs utilities. Saying it in a different way, there exists such an interval $\left\langle \underline{w_{ref}},\overline{w_{ref}}\right\rangle$ that for each $w_{ref}\in\left\langle \underline{w_{ref}},\overline{w_{ref}}\right\rangle$ occurs $\left|u(w)-u(w_{ref})\right|\le\varepsilon$, where ε is small. Any value w belonging to that interval is seen by the decision maker as the reference point or the status quo.

After transformations we get

$$u(w_{ref})-\varepsilon\le u(w)\le u(w_{ref})+\varepsilon$$

where $u\left(\underline{w_{ref}}\right)=u(w_{ref})-\varepsilon$ and $u\left(\overline{w_{ref}}\right)=u(w_{ref})+\varepsilon$

The interval $\left\langle \underline{w_{ref}},\overline{w_{ref}}\right\rangle$ depends on:

- The reference point w_{ref},
- The class of the utility function which reflects the decision maker's profile of risk.

For different utility functions we can obtain intervals of different width, and with the change of the reference point also the width of the interval is going to change. Those matters are shown by the following examples.

3.2. Example 1.

The decision maker has to valuate the following risky alternative:

L={(0;0.5); (5;0.1); (11;0.05); (12;0.1); (13;0.22); (80;0.2); (110;0.1); (120;0.1); (200;0.05); (400;0.03)}

His utility function is $u(w) = \ln w$. Because the values of the utility are not interpreted, and they only rank possibly alternatives, so ε is not given as a certain value, but it is assumed to be 10% of the u(10), then ε=0.2302585.

We consider two reference points:

- $w_{ref} = 10$,
- $w_{ref} = 100$.

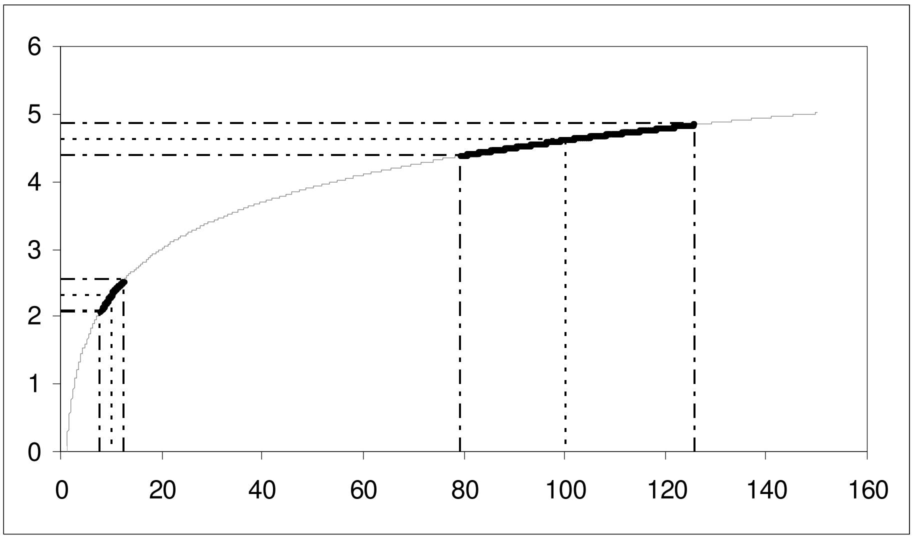
The reference interval for the logarithmic utility function is calculated as follows:

$$\ln(w_{ref}) - \varepsilon \le \ln w \le \ln(w_{ref}) + \varepsilon$$

And after some transformation we get

$$w_{ref} \cdot e^{-\varepsilon} \le w \le w_{ref} \cdot e^{\varepsilon}$$

For the w_{ref}=10 the reference interval is $\langle 7.94; 12.59 \rangle$ and for the w_{ref}=100 the reference interval is $\langle 79.43; 125.89 \rangle$. Graph 2 presents these intervals on the utility function.



Graph 2. Reference intervals on the logarithmic utility function.

To valuate the risky alternative the decision maker has to transform it to get gains and losses relative to the reference point. For the reference point w_{ref}=10 we get prospect:

L(w_{ref}=10) = {(-10,0.5); (-5;0.1); (1;0.05); (2;0.1); (3;0.22); (70;0.2); (100;0.1);
(110;0.1); (190;0.05); (390;0.03)}

And for w_{ref}=100 we obtain the following prospect:

$L(w_{ref}=100)$ = {(-100,0.5); (-95;0.1); (-89;0.05); (-88;0.1); (-87;0.22); (-20;0.2);
(10;0.1); (20;0.1); (100;0.05); (300;0.03)}

The table 1 summarize the CPT values calculated for the chosen reference points and intervals. The value function and probability weighting function are the Kahneman-Tversky functions with theirs parameters estimation.

Table 1. The CPT values.

CPT for	$w_{ref}=10$	$w_{ref}=100$
w_{ref}	36.359	-39.890
$\left\langle \underline{w}_{ref}, \overline{w}_{ref} \right\rangle$	36.105	-28.034
$\underline{w}_{ref}$	37.866	-21.012
$\overline{\overline{w}}_{ref}$	34.296	-65.216

The CPT values for lower bounds of intervals are the highest and for upper bounds are the lowest for both reference points. But the CPT values of reference points and reference intervals don't have fixed order.

3.3. Example 2.

Two decision makers are going to valuate the same risky alternative. The first is characterized by logarithmic utility function and the second by the exponential utility function. We assume that they have the same reference point w_{ref}=10 and that they indifferent to the change of 10% of utility of theirs reference point. The logarithmic utility function is in the form of $u(w) = \ln w$, and the exponential utility function is assumed as $u(w) = -10 \cdot e^{-0.1w} + 10$. For the first decision maker with logarithmic utility function calculations were done in the first example. Now calculations for the exponential utility function. The reference interval for the exponential utility function is calculated as follows:

$$-10 \cdot e^{-0.1w_{ref}} + 10 - \varepsilon \le -10 \cdot e^{-0.1w} + 10 \le -10 \cdot e^{-0.1w_{ref}} + 10 + \varepsilon$$

And after some transformation we get

$$-10\ln\left(e^{-0.1w_{ref}} + 0.1\varepsilon\right) \le w \le -10\ln\left(e^{-0.1w_{ref}} - 0.1\varepsilon\right)$$

The parameter ε equals to 10% of u(10) which is 0.63212. For the w_{ref}=10 the reference interval is $\langle 8.41; 11.89 \rangle$. Because both decision makers are assumed to have the same reference point w_{ref}=10, they evaluate the same prospect expressed by gains and losses relative to the reference point:

$L(w_{ref}=10)$ = {(-10,0.5); (-5;0.1); (1;0.05); (2;0.1); (3;0.22); (70;0.2); (100;0.1);
(110;0.1); (190;0.05); (390;0.03)}

The calculated CPT values are presented in table 2.

Table 2. The CPT values based on both utility functions.

CPT for	DM with logarithmic utility function	DM with exponential utility function
w_{ref}	36.359	36.359
$\left\langle \underline{w}_{ref}, \overline{w}_{ref} \right\rangle$	36.105	36.170
$\underline{w}_{ref}$	37.866	37.520
$\overline{w}_{ref}$	34.296	34.912

For both decision makers prospect relative to the reference point $w_{ref}=10$ is evaluated in the same manner, because they valuated only gains and losses relative to the same reference point, and they do not take into consideration theirs utility functions. The utility functions acquire importance only when the reference interval has to be calculated. Because the reference interval for the logarithmic function is wider than for the exponential function, then the CPT values for the lower and upper bound for the logarithmic function have extreme values.

4. Conclusions.

In this paper we discussed a new way of modelling choice behaviour based on the idea of the Prospect Theory. We proposed a generalization of this theory in which the interval of reference points is available and the probability weighting function is unchanged.

Such an approach seems to be consistent with decision maker's behaviour, who is indifferent to some variations from the status quo. Proposed way of valuation of the risky alternative when the reference interval is considered, shows how the change of the reference interval influences the valuation of the risky alternative.

The proposed concept opens new directions for the research which will widen the Prospect Theory approach. Based on this approach the possible change of the loss aversion and risk aversion need the deeper analysis.

Literature.

[1] Bleichrodt H. *Prospect Theory with Reference Points in the Opportunity Set.* Working Paper. Erasmus University, Rotterdam, 2006

[2] Kahneman D., Tversky A. *Prospect Theory: An Analysis of Decision Under Risk.* Econometrica, 47:2, p.263-291, 1979

[3] Levy H., Wiener Z. *Prospect Theory and Utility Theory: Temporary versus Permanent Attitude Towards Risk.* Working Paper. The Hebrew University, Jerusalem 2002

[4] Sugden R. *Reference-dependent subjective expected utility.* Journal of Economic Theory, 111, p.172-191, 2003

[5] Tversky A., Kahneman D. *Advances in Prospect Theory: Cumulative Representation of Uncertainty.* Journal of Risk and Uncertainty, 5, p.297-323, 1992

COMPETITION AND INNOVATION: IN SUPPORT OF THE INVERTED-U RELATIONSHIP

Ondřej Krčál

Masaryk University
Faculty of Economics and Administration, Economics Department,
krcalo@mail.muni.cz

Abstract. In this article, we show that different profits due to different intensity of competition can influence innovation decisions of otherwise identical firms. We propose a model of innovation decisions in which, under certain conditions, firms with very high and very low profits can be expected to innovate less than firms with moderate profits, which is consistent with the hypothesis of the inverted-U relationship between competition and innovation. Following the "persistence of profits" (POP) literature, we suppose that firms earn profits that are persistently above or below the norm and that the higher are the persistent profits of firms, the lower is the intensity of competition the firms face. It is further assumed that firms are managed by owners and that in each period, the owners earns the profits of their firm. The owner decides whether to adopt or decline an innovation (chooses between innovation and passivity). If the innovation is adopted, the profit of the firm is increased if the innovation is successful (which happens with certain probability), and decreased in the other case. The owner is assumed to choose the alternative with highest value. For valuation of outcomes, the prospect theoretical value function is used. Under certain conditions, the loss aversion and diminishing sensitivity properties of the value function imply the inverted-U relationship between profits and the difference between the value of innovation and the value of passivity. Under the assumption that there is negative value (utility) expected from adopting an innovation, the inverted-U relationship between profits and innovation activity of firms is explained.
Keywords. Innovation, Competition, Prospect Theory, Inverted-U Relationship

1. Introduction

Since Schumpeter´s famous defense of monopolistic practices in *Capitalism, Socialism and Democracy*, there has been a vivid discussion about the relationship between competition and innovation activity. A new approach was proposed by Aghion *et al.* [1]. They presented new empirical evidence supporting a hypotheses dating back to Scherer [4], that the relationship between product market competition and innovation has a form of an inverted-U. In contrast to previous treatments of the problem they used profitability of firms (inverted Lerner index) instead of market concentration for measuring competition. They also developed a theoretical model that is, as far as we know, the only valid theoretical explanation the inverted-U relationship between profitability and competition. In this paper, we propose a different model that offers similar results in terms of the inverted-U relationship between profitability and innovation activity of firms.

The paper is organized as follows: In section 2, we introduce a prospect theoretical model of innovation decision making and establish the inverted-U relationship between profitability and innovation activity of firms. In section 3, we suggest some extensions of the basic model. In section 4, we present the conclusions.

2. Model

Following Schumpeter [5], we see competition as *creative destruction* that acts effectively regardless of the number of firms or industry concentration. If we say that a firm is exposed to low competition, we mean the firm is protected from the forces of dynamic competition, be it entry of new firms, or radical innovations of the incumbents. We assume that profits of firms are persistent because they are persistently protected from the forces of competition.[1] Persistently higher (lower) profits are therefore due to persistently lower (higher) intensity of competition.

[1] This assumption is consistent with the main findings of the persistence of profits (POP) literature. The hypothesis tested by this literature is that "the profits earned in one period, whether from luck or skill, provide the resources to maintain profits into the future. Some companies erect entry barriers through increased product differentiation, others via scarce natural resources or land sites. Some obtain legal protection for their position (e.g. patents, tariffs, licences) by purchasing the services of scientists and technicians, lawyers or lobbyists, or more directly by contributions to politicians and public officials themselves." ([3], p. 370)

In our model, we suppose identical firms that are exposed to dynamic competition of different intensity. We further assume that persistent profits of firms are non-negative *($\pi \geq 0$)*, that firms are managed by owners and that the owners´ compensations are equal to the profits. For the sake of simplicity, we will also limit the decision of the owners to only one innovation project which qualities are determined randomly and are known in advance. Owners have two alternatives: innovation or passivity. There are two possible outcomes of innovation: the innovation is successful, which occurs with probability $p \in (0;1)$, and the firm receives innovation reward $r > 0$ and pays innovation cost $c > 0$, or the innovation is unsuccessful, which occurs with probability $(1- p)$ and the firms pays the innovation cost c and gets no reward.

Following prospect theory ([2], [6]), the innovation decision of an owner is considered as a choice between two prospects: a prospect of innovation ($\pi + r - c, p; \pi - c, 1 - p$) and a prospect of passivity (π). The reference point is the wealth of the owner. If $\pi \geq c$, both outcomes of innovation are perceived as gains and are decomposed into a sure gain of $\pi - c$ and risky prospect of ($\pi + r - c, p$). On the other hand, if $\pi + r < c$, both outcomes are perceived as losses and the prospect of innovating is decomposed into sure loss of $\pi + r - c$ and risky prospect of ($\pi - c, 1 - p$). If $\pi < c$ and $\pi + r \geq c$, a successful innovation would be seen as a gain and failed innovation as a loss, which implies that there is no further editing possible.

In the next step, the owner evaluates all the prospects and chooses the one with the highest value. For it, to every outcome x is assigned a number $v(x)$ that stands for the subjective value of the outcome and every probability p has to be transformed to decision weight $w(p)$. The function that transforms the value of outcomes is called *value function*, the function transforming probabilities is called *weighting function.* The value function reflects two principles: the principle of *diminishing sensitivity* (impact of a change diminishes with the distance from the reference point), which implies that the value function is concave for gains and convex for losses, and the principle of *loss aversion* (losses loom larger than corresponding gains), which implies that the value function is steeper for losses than for gains. The value function is as follows:

$$v(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -\lambda \cdot (-x)^{\beta} & \text{if } x < 0 \end{cases}, \text{ where } \alpha, \beta \in (0;1) \text{ and } \lambda \geq 1. \tag{1}$$

The principle of diminishing sensitivity applies also to the weighting function. In this case, changes are considered less important if they are more distant from certainty ($p = 1$) and impossibility ($p = 0$). This implies that the weighting function is concave near impossibility and convex near certainty. Furthermore, the form of weighting function estimated in [6] overweights small probabilities and underweights medium and high probabilities. The mathematical forms of the weighting function for gains $w^{+}(p)$ and for losses $w^{-}(p)$ are given below:

$$w^{+}(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1-p)^{\gamma})^{\frac{1}{\gamma}}}, \text{ where } \gamma \in (0;1) \tag{2}$$

$$w^{-}(p) = \frac{p^{\delta}}{(p^{\delta} + (1-p)^{\delta})^{\frac{1}{\delta}}}, \text{ where } \delta \in (0;1) \tag{3}$$

The properties of the weighting function and the assumption that $p < 1$ imply that for any potential innovation project must be true that $r > c$. The owner compares the values from the two alternatives, innovation (V_i) and passivity (V_p). We will call the difference between the alternatives *net value* of innovation project $V(\pi) = V_i - V_p$. It follows that

$$V(\pi) = w^{+}(p) \cdot (\pi + r - c)^{\alpha} + (1 - w^{+}(p)) \cdot (\pi - c)^{\alpha} - \pi^{\alpha}, \text{ if } \pi \geq c \text{ and} \tag{4}$$

$$V(\pi) = w^{+}(p) \cdot (\pi + r - c)^{\alpha} + \lambda \cdot w^{-}(1-p) \cdot (c - \pi)^{\beta} - \pi^{\alpha}, \text{ if } \pi < c.$$

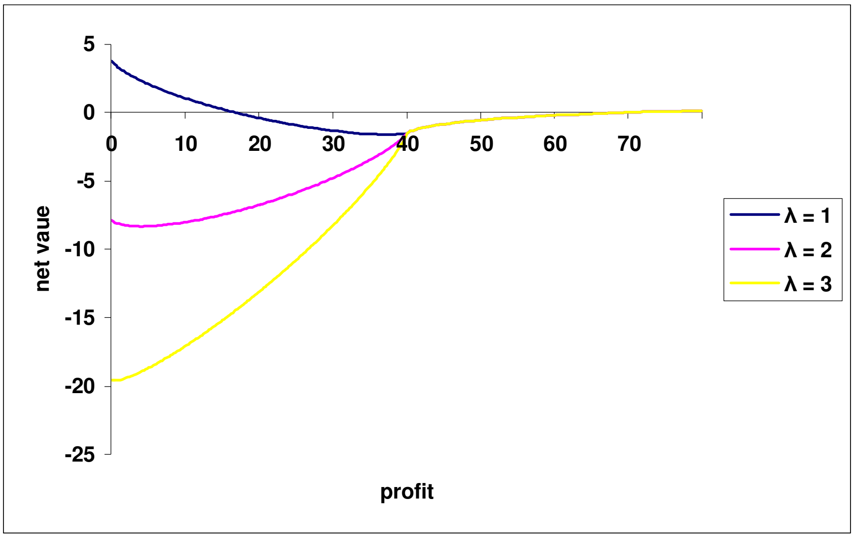


Fig. 1: The relationship between $V(\pi)$ and π for r = 100, c = 40, p = 0.5, $\alpha = \beta = 0.88$, $\gamma = 0.61$, $\delta = 0.69$.[2]

The inverted-U form of $V(\pi)$ is due to the principles of diminishing sensitivity and loss aversion. The form of $V(\pi)$ depends on how fast is the value of each possible outcome growing with growing π. From the principle of diminishing sensitivity follows that the closer is the change to the reference point, the more strongly is it perceived. If loss aversion is minimal ($\lambda = 1$), for small π, $V(\pi)$ *is* decreasing and convex because the value of passivity grows faster relatively to the value innovation, but with π coming close to c the value of unsuccessful innovation starts growing faster than the value of passivity minus successful innovation and the function $V(\pi)$ starts growing (see Fig. 1, $\lambda = 1$). The profit for which $V(\pi)$ is minimal will be called m. With growing λ, the value of m is decreasing (in Fig. 1, for $\lambda = 2$ is it around 4). If λ is high enough, the function $V(\pi)$ for $\pi \in \langle 0; c \rangle$ is growing (see Fig. 1, $\lambda = 3$[3]).

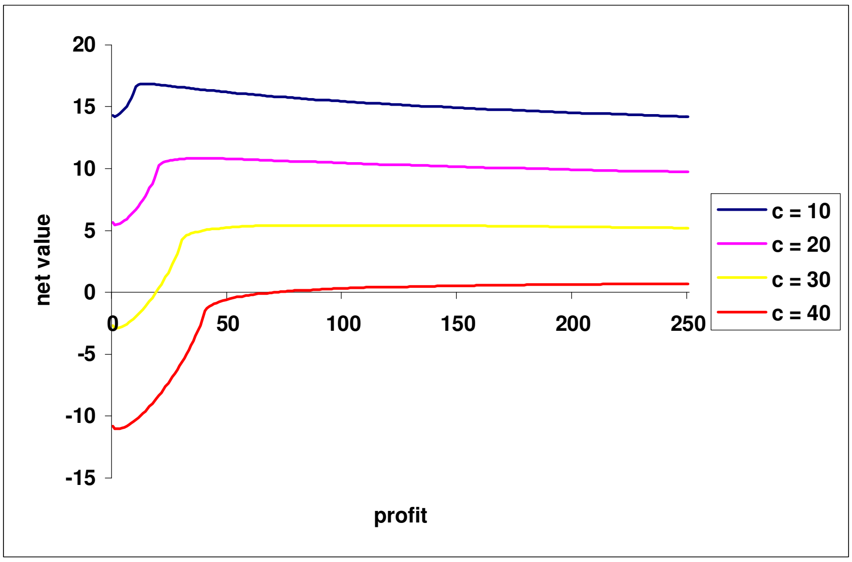


Fig. 2: The relationship between $V(\pi)$ and π for r = 100, p = 0.5, $\alpha = \beta = 0.88$, $\lambda = 2.25$, $\gamma = 0.61$, $\delta = 0.69$.

If for any $\pi > m$, $V(\pi) > 0$, then for π close to c the value of unsuccessful innovation grows faster then the value of passivity minus the value of successful innovation, which implies growth, but the difference between the velocities of the growth is decreasing, which implies concave shape of $V(\pi)$ and maximum of the function $V(\pi)$ for profit n. Moreover, the total value of all three outcomes is decreasing due to the principle of diminishing

[2] The values of the parameters of the value function and weighting function are taken from [4]. λ will be 2.25.
[3] In fact, m for $\lambda = 3$ is around 0.3535.

sensitivity, which implies that for high π, the function is again decreasing and convex (see in Fig. 2, $c = 10$) and finally that $\lim_{\pi \to \infty} V(\pi) = 0$. If for any $\pi > m$, $V(\pi) > 0$, then the function $V(\pi)$ takes form of the inverted-U.[4]

However, the inverted-U form of $V(\pi)$ doesn´t automatically imply that the *innovation activity* will be also inverted-U shaped. As we can see in Fig. 2, for high levels of c and for low levels of π, $V(\pi)$ is negative; which means that the owner of the firm prefers passivity to innovation. But even in this case, if the owner innovates for moderate profits, she will also innovate for high profits (see Fig. 2, $c = 40$). We can therefore explain just the increasing part of the inverted-U relationship between profits and innovation activity. For explaining the entire relationship, we have to introduce another property of innovation, *innovation disutility*, to the model. Innovation disutility d corresponds to additional effort and attention and to unpleasant decisions and changes connected to the innovation project. The net value of innovation project is now

$$V(\pi) = w^{+}(p)\cdot(\pi + r - c)^{\alpha} + (1 - w^{+}(p))\cdot(\pi - c)^{\alpha} - \pi^{\alpha} - d \text{, if } \pi \geq c \text{ and} \qquad (5)$$

$$V(\pi) = w^{+}(p)\cdot(\pi + r - c)^{\alpha} + \lambda\cdot w^{-}(1-p)\cdot(c - \pi)^{\beta} - \pi^{\alpha} - d \text{, if } \pi < c\text{.}$$

From this new equation follows that if $d > 0$, $\lim_{\pi \to \infty} V(\pi) < 0$, which means that not only for low profits but also for high profits, firms will prefer passivity to innovation even if firms with moderate profits innovate. The inverted-U relationship is therefore established (see Fig. 3).

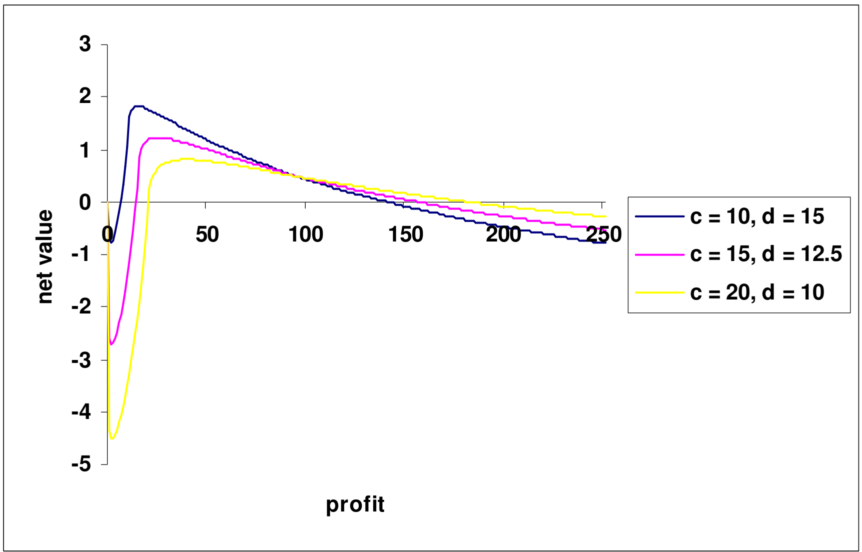


Fig. 3: The relationship between $V(\pi)$ and π for $r = 100$, $p = 0.5$, $\alpha = \beta = 0.88$, $\lambda = 2.25$, $\gamma = 0.61$, $\delta = 0.69$.

3. Discussion

In this part, we will present some extensions of the model. Until now, we assumed that qualities of the innovation projects have been set independently from the previous innovations. In reality, we observe that there are innovations (radical innovations - blue line in Fig. 4) that set stage for other innovations (incremental innovations - pink line in Fig. 4). It seems reasonable to assume that radical innovations have typically higher innovation disutility that incremental innovations.[5] We further assume that if firms don´t adopt radical innovations, the source of profitable incremental innovations dries up. It can be expected that high profits firms will be less willing to adopt radical innovations and therefore, will have access to less profitable incremental innovations.

[4] If for any $\pi > m$, $V(\pi) \leq 0$, the function $V(\pi)$ is non-decreasing for $\pi > m$.

[5] The rationale for the assumption can be easily provided: Firms implementing radical innovation are often forced to redefine the scope of their business. The qualification of the workers, and often also of the management (owner) might become obsolete. Radical innovation might involve important organizational changes, including downsizing or restructuring of the business model, which are typically very unpleasant decisions to make. On the other hand, in some cases radical innovations can be perceived as more exciting than incremental innovations, which would decrease the innovation disutility attached to the project.

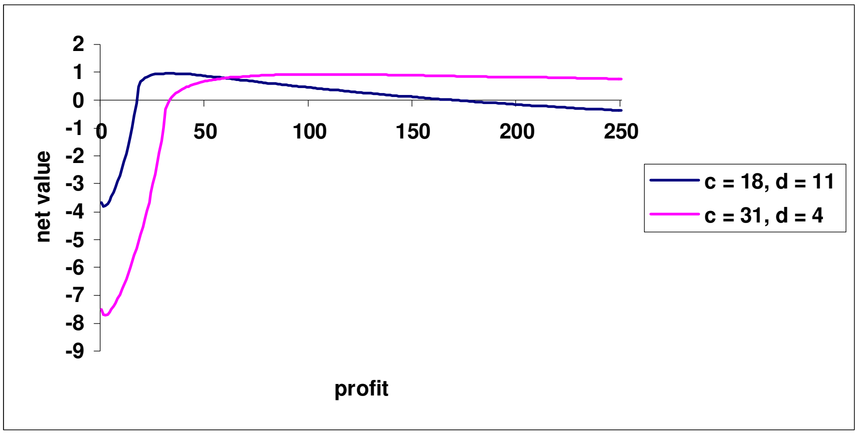


Fig. 4: The relationship between $V(\pi)$ and π for $r = 100$, $p = 0.5$, $\alpha = \beta = 0.88$, $\lambda = 2.25$, $\gamma = 0.61$, $\delta = 0.69$.

We achieve similar effect if we substitute manager for owner. We assume that like in case of the owner, the compensation of the manager reflects the size of the persistent profits of the firm, meaning that firms with persistently low profits pay lower *fixed salary* to the managers than the firms with persistently high profits, and that the value of the fixed salary as perceived by the managers is the same like the value of the profits perceived by the owners (we use the same numbers on the profit scale for the owner and for the manager in Fig. 5 and 6).

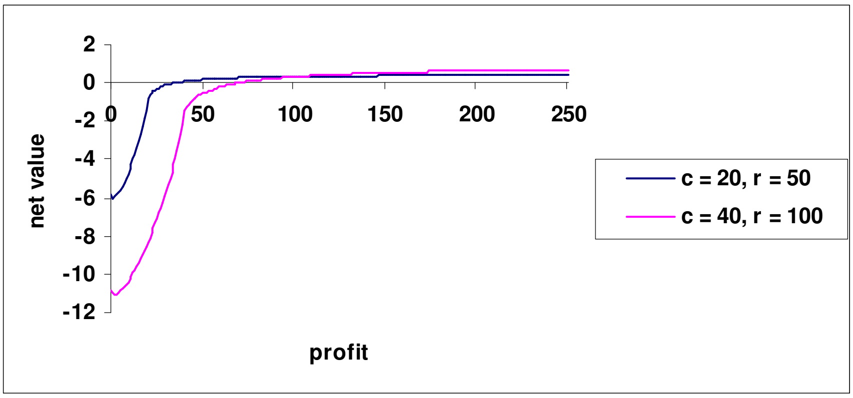


Fig. 5: The relationship between $V(\pi)$ and π for $d = 0$, $\alpha = \beta = 0.88$, $\lambda = 2.25$, $\gamma = 0.61$, $\delta = 0.69$

Unlike the previous case, in which the proportion between expected profits and reward and innovation cost has been clearly defined, in case of manager´s compensation, the proportion between fixed salary and the reward or punishment from the innovation is settled in manager´s contract. If the proportion is smaller and if innovation disutility decreases in proportion or is equal to zero, managers of low profit firms will be more willing to innovate and managers of high profit firms will be less willing to innovate than the owner of the same firm (see Fig. 5; managers – blue line, owners – pink line). If all the previous conditions hold and innovation disutility is negative and the same for managers like for owners, than of course, managers (pink line in Fig. 6) are less willing to innovate than owners (yellow line). In order to motivate managers to innovation, the owners may change the proportion between the effect of successful and unsuccessful innovation on manager´s compensation. In this case again, managers of firms with lower profits will be more willing to innovate than owners (blue line). Even if the fixed salary of managers is so low, that the outcomes of the innovation for their compensation is higher than proportional, the owners may still be interested in changing the proportion between reward and cost for managers in favor of the reward to increase their motivation to innovate.[6]

---

[6] There is another possibility. The owner may leave the proportion between reward and cost from innovation for the manager intact, but he may cover some of the potential losses. This way can be reduced the loss aversion effect of managers.

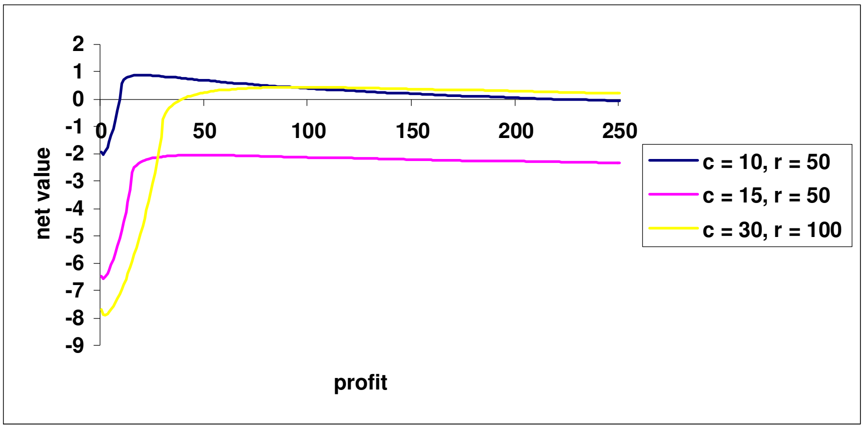


Fig. 6: The relationship between $V(\pi)$ and π for $d = 5$, $\alpha = \beta = 0.88$, $\lambda = 2.25$, $\gamma = 0.61$, $\delta = 0.69$.

Similar effects are in place if we allow firms to implement more innovation projects with more or less independent probabilities at the same time. The more independent projects are implemented, the less probable are the extreme outcomes of the innovation activity (like all innovations successful or unsuccessful), and the less extreme are the effects of innovation activity on the compensation of the owners (managers). If we assume that there are extra costs connected to having more (smaller) innovation projects running at the same time, we may conclude that owners (managers) of firms with rather low profits are more willing to cover the extra costs to limit their exposure to risk that owners (managers) of high profit firms. And consequently, that this ability to combine more innovation projects again, *ceteris paribus*, increases the innovation activity of less profitable firms.

4. Conclusions

In this paper, we propose a model that explains the inverted-U relationship between competition and innovation. One of the problems of the inverted-U shape generated by our model is that it declines very slowly, i.e. that contrary to empirical evidence, the innovation activity of firms under less intense competition is similar to innovation activity of firms exposed to more intense competition. In section 3, we present some modifications of the basic model that might reduce the innovation activity of high profit firms, and consequently make the model more consistent with the data.

Bibliography

[1] AGHION, P., BLOOM, N., BLUNDELL, R., GRIFFITH, R., HOWITT, P. (2005): Competition and Innovation: An Inverted-U Relationship. *The Quarterly Journal of Economics*, Vol. 120, No. 2, pp. 701–728

[2] KAHNENAN, D., TVERSKY, A. (1979): Prospect Theory: An Analysis of Decision under Risk. *Econometrica,* Vol. 47, No. 2, pp. 263–291

[3] MUELLER, D. C. (1977): Persistence of Profits above the Norm. *Economica,* Vol. 44, pp. 369 - 380

[4] SCHERER, F. M. (1967): Market Structure and the Employment of Scientists and Engineers. *The American Economic Review*, Vol. 57, No. 3, pp. 524–531

[5] SCHUMPETER, J. A. (1994): *Capitalism, Socialism and Democracy*. Routledge, 437 pages, ISBN 0415107628

[6] TVERSKY, A., KAHNENAN, D. (1992): Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty,* Vol. 5, pp. 297–323

QUALITY OF THE SOLUTIONS OF THE TIME LIMITED VEHICLE ROUTING PROBLEM OBTAINED BY SELECTED METHODS

Petr Kučera

Czech University of Life Sciences in Prague
Faculty of Economics and Management, Department of Operational and Systems Analysis
kucera@pef.czu.cz

Abstract. The time limited vehicle routing problem (TLVRP) is related to the vehicle routing problem (VRP). The main difference is that the routes are paths (not cycles), i.e. vehicles do not return to the central city. Costs are given for straight routes between each pair of cities and represent time necessary for going through. Each path must not exceed a given time limit. The sum of time for all routes is to be minimized.
Several tens of instances of the TLVRP were solved by the nearest neighbour method and by the tree approach which is the combination of the Mayer method for the VRP and the Christofides method for the traveling salesman problem. The results were analyzed and an interesting dependency of the solution quality on the location of the central city was found out.
Keywords. Time limited vehicle routing problem, vehicle routing problem, traveling salesman problem, Mayer method, Christofides method, tree approach, nearest neighbour method.

1. Introduction

The time limited vehicle routing problem (TLVRP) is defined as follows: One central city and other n (ordinary) cities are given and for each pair of cities a cost is given, representing time necessary for going through the straight route between them. The cost matrix is supposed to be symmetric. The goal is to find a set of paths so that each of them has one of its endpoints in the central city, its length does not exceed a given time limit and each city except for the central one lies on exactly one of the paths.

This problem has many practical instances, e.g. transportation newspapers from the printers' to shops, grocery products (dumplings etc.) from the manufactory to restaurants, daily reports from affiliated branches to the headquarters etc. Each vehicle is required to visit all the cities on its route until a given time, but we do not mind how it gets from the end back to the start of its route to realize it next time.

Nevertheless, TLVRP has been studied relatively little. It belongs to the NP-hard problems, for which there is no efficient algorithm finding their theoretical optimum. It is related to the vehicle routing problem (VRP), where routes are cycles instead of paths, and so to the traveling salesman problem (TSP), i.e. the task to construct one cyclic route containing all the cities, too. Thus heuristics (approximation methods) for the TLVRP can be derived from the methods for the VRP and the TSP.

Let us introduce some notation. The central city will be indexed by 0 and the other cities by numbers from 1 to n. The cost matrix will be denoted by $\mathbf{C}$ (and so single costs c_{ij}, $i, j = 0, \ldots, n$).

The aim of this contribution is to test two methods for the TLVRP on several test cases and find out some interesting dependencies of the obtained solution quality on properties of the task.

2. Tested Methods

3. Tree approach

The tree approach is a combination of the well-known Christofides method for the TSP [1] and the Mayer method for VRP, more preciously for its special case where capacities (demands) of cities and vehicles are specified (so called multipletours traveling salesman problem, cf. e.g. [5])

The Mayer method makes the minimum spanning trees starting in the remotest cities from the central one by the same way as the Prim (Jarník) algorithm ([7], [2]), as large as the capacities of the cities and of the vehicle allow. It only separates cities into groups so that each group contains cities on a route for one vehicle and then some method for the TSP must be used to determine the order of the cities on the single cycles. The detailed description of the Mayer method is available in [3], where more information about the tree approach for the TLVRP is given.

The algorithm of the tree approach is here:

1. Choose from the cities, which have not been put onto any route, such a city i so that c_{0i} is maximum possible (i.e. the remotest city from the central one) as the first city of the currently created route. Set T to the graph consisting of the only vertex i.
2. Let P denote the set of non-central cities in the currently created route and Q the set of non-central cities, which have not been put onto any route.
   If $Q = \emptyset$
   then stop
   else find the minimum c_{ij} so that $i \in P$ and $j \in Q$.
3. Add the vertex j and the edge $\{i, j\}$ to T. Derive T_0 from T by adding the vertex 0 and the edge $\{0, k\}$ such that $k \in P \cup \{j\}$ and c_{0k} is minimum possible. (Note that both T and T_0 are trees).
4. Find the minimum cost "nearly perfect" matching M of all such cities of T that have odd degree in T_0. Obviously, there is odd number of such cities, so exactly one of them, say l, is to be isolated in this matching.
5. Now there are exactly two vertices of an odd degree in $T \cup M$: l and 0. Find a walk from l to 0 containing all the edges of $T \cup M$.
6. Create a path by preserving the first occurrence of each city in the walk and deleting all other occurrences.
7. If the total cost (the time for going through) of this path does not exceed the time limit given in the input of this task
   then declare the path from the last executing of the step 6) to be the currently created route and go to 2) to search for other cities for this route
   else declare the path obtained *before* the last executing of the steps 2) to 6) to be a finished route and go to 1) to start creating a new path.

Let us remark that all paths are supposed to contain at least two non-central cities in this formulation of the algorithm.

4. Nearest Neighbour Method

This method is perhaps the simplest possible method for solving the TSP. Its name describes appositely what it is based on. Its algorithm is taken from [6]:

1. Let Q denote the set of non-central cities, which have not been put onto any route. Join the closest city of Q (with the minimum c_{0i}) to the central one (i.e. start the construction of a new path with the edge $\{0, i\}$).
2. If $Q = \emptyset$
   then stop
   else denote i the last city added to the route and find the minimum c_{ij} so that $j \in Q$. If adding the edge $\{i, j\}$ the total cost (the time for going through) of this path does not exceed the time limit given in the input of this task
   then add the edge $\{i, j\}$ to the solution and continue constructing this route repeating the step 2)
   else go to 1) to start creating a new path.

5. Test Computations and Their Results

For testing three types of randomly generated cases were taken. In all the types all the cities were located in a circle with 100 time unit radius (the time necessary for going through a given route is supposed to be directly proportional to the distance). The time limit for routes was set to 250. In the first two types the central city was in the middle of this circle.

Type 1: First let us consider the area between a circle given above and another circle with the radius of 20 time units with the same center. In this area 20 cities are originally randomly generated with the uniform distribution. Then the closest pairs of cities are joined into "regions" so that at most four of them may form one "region" and the final number of non-central cities ("regions") is 12.

Type 2: This type of test cases contains 24 (non-central) cities randomly generated with the uniform distribution with no additional conditions or modifications.

Type 3: This type differs from the type 1 only that the city taken as the central one is the remotest one from the centre of the circle (the closest one to its perimeter).

For each type ten test cases were computed using both the tree approach and the nearest neighbour method and for better evaluation of the obtained solutions also using the Habr frequencies approach which has shown out to be generally one of the best methods for the TLVRP (for the definition and more information see [4]). The results are summarized in the tables 1, 2 and 3 in the percentage form, where 100 p.c. is the best of the results.

However, it is interesting to notice the difference between the results by these methods at single cases of the same type. As far as the type 1 is concerned, in the cases 2 and 10 the tree approach gave relatively good results while in the cases 5 and 7 it completely failed. Thus some properties, which would have influenced the quality of the tree approach solution, were searched for. It was found that there is an important dependence of the results on the ratio between the farthest and the nearest city from the central one among cities lying on the convex hull of the set of all cities. Big values of this ratio properly indicate that the central city does not lie near the middle of the actually attended region. This property will be called eccentricity and it is added to the tables 1, 2 and 3 (NNM denotes the nearest neighbour method, TA the tree approach, HFA the Habr frequencies approach and Ecc. the eccentricity).

Thus for the type 1 the tree approach achieves good results for the cases with relatively high eccentricity. On the contrary, for the type 3 the tree approach gives good results for cases with the small eccentricity (e.g. cases 1, 7 and 9) and bad results in the opposite cases (3 and 10). Thus the conclusion is that the tree approach usually gives good results for cases with the eccentricity of the medium size.

A similar situation occurs at the nearest neighbour method for the types 2 and 3. For the type 2 the results are rather worse for the high eccentricity (cases 3 and 6) while the best result is in the case 10 with low eccentricity. For the type 3 there usually are good results for the high eccentricity (cases 5, 6 and 10) and bad results for the low eccentricity (cases 1, 4, 7 and 9). So the nearest neighbour method is suitable for the cases with extreme values of the eccentricity.

	NNM	TA	HFA	Ecc.
Case 1	108,8%	114,6%	100,0%	1,68
Case 2	112,6%	109,4%	100,0%	1,71
Case 3	107,1%	113,2%	100,0%	1,76
Case 4	110,9%	117,5%	100,0%	2,10
Case 5	104,1%	126,6%	100,0%	1,26
Case 6	100,0%	108,4%	100,0%	1,30
Case 7	106,0%	125,9%	100,0%	1,32
Case 8	108,6%	106,6%	100,0%	1,52
Case 9	107,1%	114,5%	100,0%	1,24
Case 10	104,2%	104,8%	100,0%	1,66

Table 1: Test Results – Type 1

	NNM	TA	HFA	Ecc.
Case 1	112,2%	130,9%	100,0%	1,24
Case 2	108,2%	112,8%	100,0%	1,31
Case 3	131,4%	128,9%	100,0%	1,44
Case 4	115,0%	108,9%	100,0%	1,14
Case 5	102,3%	101,6%	100,0%	1,33
Case 6	119,0%	100,2%	100,0%	1,39
Case 7	120,8%	109,4%	100,0%	1,22
Case 8	101,0%	100,0%	105,2%	1,31
Case 9	120,7%	119,9%	100,0%	1,41
Case 10	100,0%	106,1%	114,5%	1,21

Table 2: Test Results – Type 2

	NNM	TA	HFA	Ecc.
Case 1	108,4%	100,0%	111,0%	2,49
Case 2	115,1%	100,1%	100,0%	3,61
Case 3	115,5%	120,7%	100,0%	3,97
Case 4	115,4%	100,0%	105,3%	2,89
Case 5	100,0%	100,9%	117,7%	3,36
Case 6	103,0%	111,8%	100,0%	3,66
Case 7	104,9%	100,0%	100,9%	2,51
Case 8	103,7%	106,9%	100,0%	3,19
Case 9	119,0%	100,6%	100,0%	1,80
Case 10	103,8%	112,0%	100,0%	5,80

Table 3: Test Results – Type 3

All these four dependences:

- for the tree approach for the type 1,
- for the tree approach for the type 3,
- for the nearest neighbour method for the type 2,
- for the nearest neighbour method for the type 3,

have been confirmed using the linear regression analysis (with p-values less than 0.15).

6. Conclusion

Both the tree approach and the nearest neighbour method do not be the best ones for the TLVRP. The Habr approach which serves for comparing the results achieved by single methods is much better. Thus it is important to know when to use which of the relatively worse methods according to the properties of a particular task. The eccentricity which describes the location of the central city in the task has shown out to be useful for these purposes. In the test the tree approach gave good solution for the cases with the medium size eccentricity and the nearest neighbour method for the extremely low or high eccentricity. So both these methods complement each another. There is a question how the methods will behave at some different instances, namely larger ones with a greater number of cities, but it can be expected that they will achieve their good results at the similar types of instances as during the observations here.

References to literature:

[1] CHRISTOFIDES, N.: *Worst-case Analysis of a New Heuristic for the Traveling Salesman Problem*, Technical Report CS-93-13, Carnegie Mellon University, 1976

[2] JARNÍK, V.: *O jistém problému minimálním*, Práce moravské přirodovědeské společnosti, 4, 1930

[3] KUČERA, P. *Tree Approach to the Time Bounded Transportation Problem*, Mathematical Methods in Economics, Hradec Králové, 2005. ISBN 80-7041-535-5

[4] KUČERA, P., DÖMEOVÁ, L. *Comparing Various Heuristics for Special Types of the Vehicle Routing Problem.* Increasing Competitiveness or Regional, National and International Markets, Ostrava, VŠB-TU, 2007. ISBN 978-80-248-1458-2

[5] KUČERA, P, HOUŠKA, M., BERÁNKOVÁ, M. *Methods for the Multipletours Traveling Salesman Problem Making the Final Solution in One Time*, Mathematical Methods in Economics, Plzeň, 2006. ISBN 80-7043-480-5

[6] PELIKÁN, J., *Časově omezený rozvozní problém.* E+M. Ekonomie a Management, 2005, Vol. VIII, No. 3. ISSN 1212-3609

[7] PRIM, R.C.: *Shortest Connection Networks and Some Generalizations*, Bell Syst. Tech. J., 36, 1957

SIMPLE AGENT-BASED COMPUTATIONAL MODEL OF MARKET WITHOUT INTERMEDIATION

Michal Kvasnička

Masaryk University, Faculty of Economics
michal.kvasnicka@econ.muni.cz

Abstract. This paper describes an agent-based computational model of a market with reservation prices without intermediation. The market institutions are parameterized, and cover Smith's CDA market, Chamberlin's unorganized market, and many other types of markets. The model allows us to estimate the relative contribution of the individual institutional elements to the market efficiency.
Keywords. Simulation, agent-based economics, market, market institutions, market efficiency, mechanism design.
JEL C150 D020 D400 D440 D490 D800

1 Introduction

The goal of this study is to explore the impact of institutions on the efficiency of a market with reservation prices without intermediation. Specifically, we try to explore how much the individual institutional elements contribute to the experimentally known efficiency of the continuous double auction. For the reasons explained bellow, we use an agent-based computational model for this. The structure of the paper is as follows: 1) we review the relevant literature, 2) we describe our model, and 3) we present the results of the simulations.

2 Literature Review

Hayek [5] claimed that knowledge is dispersed in the society and that "[w]e must look at the price system as such a mechanism for communicating information..." (p. 526). However, there had been no way to test the hypothesis that markets are able to gather, process, and spread information in the way described by the competitive price theory until the first experiments were conducted.

The first market experiment was conducted by Chamberlin [3]. He divided students into buyers and sellers, gave each of them one unit to trade, a secret reservation price (the lowest price for which a seller can sell, or the highest price for which a buyer can buy), and asked them to trade. Neither speculative purchases, nor short sales were allowed. Beside these, he imposed no market institutions. The students were going around the class, meeting in pairs or small groups, and haggling. As soon as they traded, they dropped from the market, and the price was recorded. The market was closed when no transaction took place for some time. The outcomes did not support the Hayek's hypothesis: 1) the quantity traded was bigger than equilibrium, 2) the prices were in average below equilibrium, and 3) there was no tendency toward equilibrium. Chamberlin claimed: "My own skepticism as to why actual prices should in any literal sense tend toward equilibrium during the course of a market has been increased not so much by the actual data of the experiment ... as by failure ... to find any reason why it should be so" (p. 102).

Later, Smith [7] conducted other experiments. They were similar to the Chamberlin's one with two important differences: 1) there were many trading days, i.e. when the market was closed, the traders were given new units to trade and the market was reopened, which allowed the traders to learn from the past, and 2) he imposed a peculiar market organization: the continuous double auction (CDA). In this kind of auction all traders stand together, can ask or bid at any time, accept the last proposal, and hear what is going on on the market. Smith reported radically different outcomes than Chamberlin: he claimed that "[e]ven where numbers are 'small', there are strong tendencies for a supply and demand competitive equilibrium to be attained..." (p. 134).

Thus from Smith [7] and subsequent experiments we know that markets are indeed able to communicate private knowledge (reservation prices) efficiently, and reach the competitive equilibrium. However, for many decades we lacked a formal model explaining how the equilibrium is attained. The neoclassical models cannot explain this because they *presuppose* the equilibrium. Another modeling technique was needed that would be able to model the *adjusting processed* explicitly. One such technique is the agent-based computational economics (ACE). In ACE we proceed in two steps: 1) we describe the behavior of the economic agents. The agents are not optimizing units, but software robots pursuing their goals based on their reaction functions which need to be neither optimal, nor analytically tractable. 2) We simulate interactions among the robots and observe both their behavior and the behavior of the overall system consisting of them. The interactions may, or may not lead to an equilibrium. The model provides no analytical solution, only simulation data to be statistically analyzed. For more information on ACE see [8].

The first ACE model of the CDA market with reservation prices was created by Gode and Sunder [4]. Their software robots submitted random bids and asks within the limits of their reservation prices, and gained surprisingly high market efficiency measured by total surplus gained. They claimed that "[a]llocative efficiency of a double auction derives largely from its structure, independent of traders' motivation, intelligence, or learning" (p. 119). Later Cliff and Bruten [1, 2] criticized this approach, and proposed software robots able to learn. They claimed that "[m]ore than zero intelligence [is] needed for continuous double-auction markets" [2].

So far, the literature can be summarized in this way: it is the market institution (e.g. CDA) what makes markets efficient. However, the CDA is a complex institution consisting of many elements. The goal of this paper is to test the hypothesis that the market efficiency depends on the market institutions (it decreases as we depart from the CDA market) and explore how much the individual elements of the CDA contribute to its efficiency.

3 Description of the Model

Our model consists of a rectangular toroid world on which traders (Cliff and Bruten's software robots) are randomly positioned. The space allows us to define a simple metric of distance between the traders. The traders have a private reservation price, and one unit to trade in each trading day. They trade directly (without intermediation). Each trader has his private haggling price—if he quotes, he quotes this price; if he is asked or bid, he accepts it if the ask or bid price is better or equal to his haggling price. The initial value of the haggling price is the seller's (buyer's) reservation price plus (minus) a random margin in percents. The trader is said to be active if he has not yet traded his unit.

The market institution is described by four parameters: 1) the `vision%` parameter (0–100 %) describes how much of the world can each trader see. This describes the market integration. If the `vision%` is 100 %, each trader can see all other traders, and the market if fully integrated. If the `vision%` is lower (e.g. 30 %), then each trader can see only a circle area in which center he stands and which covers 30 % of the world—thus he can see in average only 30 % of other traders. Then the total market consists of many overlapping sub-markets. 2) The `moving-type` parameter (moving / not-moving) describes whether the traders can move; if they can move, the market is integrated in this way too. 3) The `public-offers?` parameter (true / false) determines whether a trader can ask or bid publicly (i.e. every trader in his vision range can possibly accept it, true value), or privately (i.e. he must select one other agent of the opposite type to trade with, false value). And 4) the `public-hearing?` parameter (true / false) determines whether all traders within the vision range of the asking or bidding trader can hear whether the proposal was accepted, and at what price (true value). Thus this parameter determines whether the trader can learn from the experience of the whole market (true value), or from his own experience only (false value).

These parameters allow us to mix many types of markets. For instance,the CDA means `vision%` = 100 %, `public-offers?` = `public-hearing?` = true. The Chamberlin's market means lower `vision%`, uncertain `public-offers?` and `public-hearing?`, and `moving-type` = moving.

Each simulation consists of many trading days. At the beginning of each simulation, traders are created and randomly positioned in the world. They are assigned their reservation price derived from a given supply or demand curve. Their haggling prices and behavioral parameters are assigned their initial values. Other traders' state variables are reset to zero. At the beginning of each trading day, each trader is given a new quantity to trade (one unit), and traders' state variables other than quantity to trade, reservation, haggling price, and behavioral parameters are reset to zero.

A trading day proceeds like this: 1) if `moving-type` = moving, all traders move. 2) One active trader is randomly selected. If `public-offers?` = false, he randomly selects one active trader of the opposite type within his vision-range as his partner; otherwise he tries all active traders within his vision-range. 3) If there is a partner, and their haggling prices allow them to trade, they trade and drop out of the market. The price is recorded. 4) If `public-hearing?` = true, then the selected trader and all traders within his vision-range update their haggling prices; otherwise only the selected trader and his partner adjust them. The steps 1–4 are repeated unless there is no room for trade or no deal takes place for some time; then the day is over, the market is closed, and reopened for the next trading day.

The haggling price adjustment proceeds in two steps (see [1]): 1) the trader determines whether to rise, or lower his haggling price, and 2) he determines how much to change it. We will describe the algorithm for a seller; the buyer's algorithm is similar. Let us suppose that the seller's haggling price is p_t and the last quotation price was q_t. If the last quotation was accepted, then if $p_t \leq q_t$, the seller raises his price because he can ask more; if $p_t \geq q_t$ and he is still active, he lowers his price because otherwise he would be undercut by a competitive seller (if he has already sold, he does not change the price since he can hope he would sell at the same price the next trading day). On the other hand, if the last quotation was not accepted, it was an offer, the seller is still active, and $p_t \geq q_t$, then the seller lowers his price because he would not get the deal either. Otherwise, he does not change the price.

The new haggling price p_{t+1} is then determined by equations 1–3:

$$\tau_{t+1} = R_{t+1} \cdot q_t + A_{t+1}, \tag{1}$$

$$\Gamma_{t+1} = (1-\gamma) \cdot \beta \cdot (\tau_{t+1} - p_t) + \gamma \cdot \Gamma_t, \tag{2}$$

$$p_{t+1} = p_t + \Gamma_{t+1}, \tag{3}$$

First, the seller sets his target price τ_{t+1} to "test" the market (the R_{t+1} and A_{t+1} are random numbers generated each time; if he is raising the price, they are positive, otherwise they are negative). The Γ_{t+1} is the motion of the haggling price. It has two features: 1) the seller does not want to move to his target price τ_{t+1} at once, but only the proportion $\beta \in \langle 0, 1\rangle$, and 2) there is inertia in price changes governed by the behavioral parameter $\gamma \in \langle 0, 1\rangle$. The initial momentum $\Gamma_0 = 0$. The p_{t+1} is rounded. If the p_{t+1} is not allowed by the reservation price, it is reset to its old value p_t.

The overall market efficiency is measured by Smith's coefficient α [7]. He defined it in the manner similar to the coefficient of variation as

$$\alpha = \frac{\sqrt{\frac{\sum_{i=1}^{n}(P_i - P^*)^2}{n-1}}}{P^*} \cdot 100\ \%, \tag{4}$$

where P_i is the i-th market price within the trading day, P^* is the equilibrium price predicted by the competitive market theory, and n is the number of transactions within the trading day.

The model was implemented in NetLogo (see [6]), the data was analyzed in MatLab. The interactive web version of the model and its source code are available at http://www.econ.muni.cz/~qasar/marketmodel/.

4 Results of Simulations

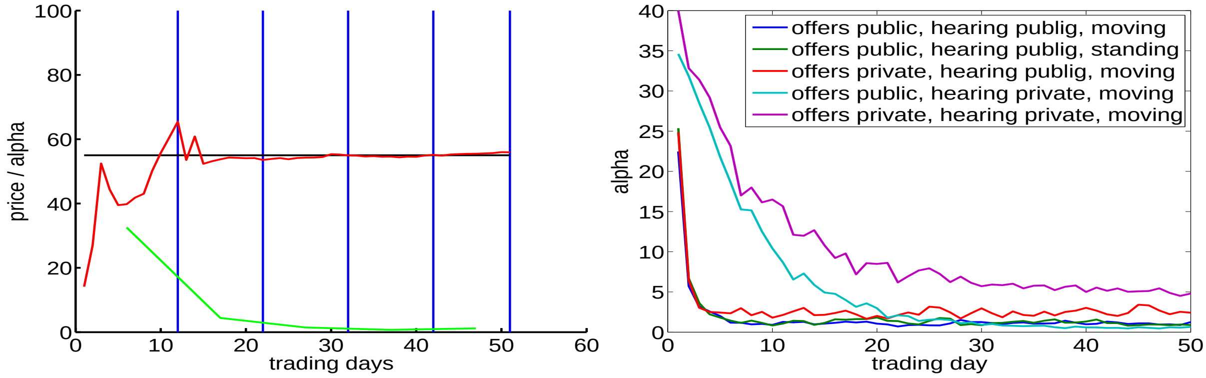
We have simulated the model with 19 buyers and 19 sellers. Their reservation prices were generated from symmetric linear demand and supply curves such that their reservation prices were 5, 10, 15, ..., 100. There were 10 intra-marginal pairs of traders, the equilibrium price was 55. Traders' behavioral parameters were randomly generated for each trader; their distributions and distributions of R_t and A_t in eq. 1 were taken from [1]. We simulated 50 trading days, and ran the simulation ten times for each combination of parameters (`vision%` being 10 %, 20 %, ..., 100 %).

We explored the market efficiency in two ways. 1) We measured with the α how close the market prices converge to the equilibrium price P^*; for the average eventual levels of the α see tab. 1. This should be

Table 1. The average eventual levels of the α. "T" in the first column stands for true, "F" for false. The order of parameters is `public-offers?`, `public-hearing?`, and `moving-type`. Thus "T/T/F" means `public-offers?` = `public-hearing?` = true and `moving-type` = not-moving. The minimal value in the row is typeset in boldface.

`vision%`	10 %	20 %	30 %	40 %	50 %	60 %	70 %	80 %	90 %	100 %
T/T/T	4.43	2.19	1.24	0.87	**0.78**	0.88	0.92	0.96	0.94	1.08
T/F/T	6.57	2.88	1.58	1.15	1.06	0.76	0.67	0.60	0.63	**0.55**
F/T/T	6.50	4.28	3.70	2.96	3.47	2.86	2.64	2.60	2.60	**2.59**
F/F/T	7.94	6.63	5.27	5.22	5.69	4.89	4.82	5.03	**4.70**	5.10
T/T/F	35.19	24.54	11.71	6.42	5.13	3.52	2.16	1.80	1.24	**1.01**
T/F/F	35.43	30.06	20.88	13.25	9.84	7.32	4.93	1.88	1.23	**0.53**
F/T/F	37.26	20.84	18.37	10.64	8.11	5.69	4.59	3.13	2.95	**2.84**
F/F/F	35.45	29.52	24.21	20.29	13.85	9.83	9.63	6.23	5.33	**4.97**

Fig. 1. The left panel shows a typical evolution of prices (red line) and the α (green line) on the CDA market. The trading days are separated with blue lines. The right panel shows the time evolution of the α under various institutions with `vision%` = 100 %.



compared with results of Smith's experiments 1–7 [7, p. 117] where the eventual α lied between 0.6 % and 9.4 % with median 3.5 % on the CDA markets. 2) We examined the speed of convergence of prices to their stationary values; for this see the right panel of fig. 1.

Generally, the integrated markets provided relatively high efficiency, regardless whether they were integrated through traders' "walking", or through formal market integration (`vision%`). However, even when traders could not walk and the vision was limited, the markets were still quite efficient for most combinations of the other institutional elements.

We expected that `public-offers` = false would slow the price convergence down and not affect the eventual α. However, the opposite happened. With `public-hearing` = false, we expected both a slower convergence and a lower eventual α. The price convergence was indeed slower, but the eventual α was surprisingly *lower* than on the CDA market. The combination `public-offers` = `public-hearing` = false led both to slow convergence and to lower efficiency.

Surprisingly, the CDA was not the most efficient market institution: the limited hearing and limited market integration, which both limited the traders' ability to learn from the experience of the others, improved the eventual α significantly. A probable explanation is that the traders' ability to learn from each other leads to "bubbles" in prices. This is confirmed by fact, that the CDA prices were much more strongly autocorrelated.

The prices in the first trading day were in average below equilibrium (e.g. on the CDA market 76 % of simulations had the average price below equilibrium in the first day), which agrees with the experimental findings [3]. It could be easily explained from the initial setting of the haggling prices: if the margin is given in percents, the intra-marginal sellers (those with the lowest reservation prices) add low margins (e.g. 30 % × \$5) while the intra-marginal buyers (those with the highest reservation prices) subtract high margins (e.g. 30 % × \$100). Thus the equilibrium given by the haggling-price supply and demand is downward-biased. Moreover, if a transaction takes place below the competitive equilibrium, the next one is likely to be there as well because of the learning.

References

1. Cliff, D.: *Minimal-Intelligence Agents for Bargaining Behaviors in Market-Based Environments*, HP Laboratories Bristol, 1997.
2. Cliff, D., Bruten, J.: "More than Zero Intelligence Needed for Continuous Double-Auction Trading", HP Laboratories Bristol, 1997.
3. Chamberlin, E. H.: "An Experimental Imperfect Market", *The Journal of Political Economy*, 1948, vol. 56, no. 2, pp. 95–108.
4. Gode, D. K., Sunder, S.: "Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality", *The Journal of Political Economy*, 1993, vol. 101, no. 1, pp. 119–137.
5. Hayek, F. A.: "The Use of Knowledge in Society The Use of Knowledge in Society", *The American Economic Review*, 1945, vol. 35, no. 4, pp. 519–530.
6. NetLogo Homepage, http://ccl.northwestern.edu/netlogo/.
7. Smith, V. L.: "An Experimental Study of Competitive Market Behavior", *The Journal of Political Economy*, 1962, vol. 70, no. 3, pp. 322–323.
8. Tesfatsion, L.: "Agent-Based Computational Economics: Growing Economies from the Bottom Up", http://www.econ.iastate.edu/tesfatsi/ace.htm.

VISUAL RECURRENCE ANALYSIS OF SIMULATED FX RATE – GRAPHICAL AND ENCRYPTION ALGORITHMS

Ladislav Lukáš

University of West Bohemia in Pilsen, Faculty of Economics,
Department of Statistics and Operations Research,
lukasl@kso.zcu.cz

Abstract. Paper presents graphical and encryption algorithms implemented to visual recurrence analysis of simulated FX rate. The simulation models constructed by CA technique provide versatile generators of time series, which are further processed by VRA algorithms. The proposed procedure implements various measures of time series variations to build VRA specific color images. Public available encryption algorithm is applied on VRA images in order to enable their safe transfer via electronic networks. The paper is focused on algorithmic and numerical aspects of elaborated VRA of simulated FX rate time series.
Keywords. Foreign exchange rate, visual recurrence analysis, numerical simulation, graphical algotithms.

JEL: C15, C51, C63, G15.

Introduction

The foreign exchange (FX) money markets attract lasting interest of both academic and applied finance families. Standard analysis of FX rate examines economic fundamentals to explain FX rate movements, but, in many cases, the fundamentals-based models fail to explain the past adequately, or predict the future reliably. On the contrary, a lot of FX money market practitioners apply various forms of technical analyses either in short or longer period quite successfully. There is well-known that technical trend signals can affect traders and price behaviour, generating excess market reactions without any fundamental reason. Hence, the instant processing of market messages plays specific role, thus leading to permanent interactions among traders. Since an empirical evidence of trading volumes in spot FX money markets shows that the overwhelming part of the turnover is merely due to short-term speculative trading transactions, one usually accepts neglecting of volumes due to an international trade transactions.

Following the classical ideas, there are typical two kinds of traders applying different trading rules based upon either technical analyses or fundamental ones. The technical analysis concerns with identifications of both trends and trend reverses using more or less sofisticated procedures to predict future price movements from those of the recent past. The fundamental analysis searches and looks for various reasons and/or events thus explaining market actions in order to set up future expectations.

Basically, these two different concepts are used for building various problem-oriented computer-based agent interaction models. The basic idea of such models is rather simple - in mathematical form to allow a self-development of FX money market under driving forces caused by interactions among computer agents which are designed to model traders', i.e. chartists' and fundamentalists', behaviour. Since these models are formulated in abstract forms, their main issues are usually, rather than prediction of actual foreign exchange market movements, studying complex dynamic effects on the base of non-linear simulation models, and investigation of possible quantitative and/or qualitative responses upon various modes of control.

For more details, the interested readers are referred to the book [1] as regards the financial market theory, and to the papers [2]-[5],[9],[10] as to both various aspects of technical and fundamental analyses and influences of central bank, in particular. The paper [6] gives the motivation and details of color image encryption. Finally, the papers [7],[8] present the author's contributions to the topic.

The paper is composed in following way. In Section 2 we briefly summarize dynamic equilibrium FX rate simulation models with special attention to modelling CB's influences. In Section 3 we provide details of generalization VRA (Visual Recurrence Analysis) procedure, which will enable more detailed and flexible detection and investigation of patters hidden in FX rate time series. In Section 4, there are given some details regarding encryption schemes of VRA images. The final Section gives conclusions and topics of further research.

Dynamic equilibrium models

Based upon general concept of market equilibrium conditions we are able to formulate FX money market clearing conditions taking into account the CB's influence in form following [7],[8]

$$m(t)\, d^{C}(t) + (1-m(t))\, d^{F}(t) + d^{B}(t) = 0, \tag{1}$$

where $d^{C}(t)$, $d^{F}(t)$, $d^{B}(t)$ denote temporal money demand of chartists, fundamentalists and the central bank, respectively, at the trading period t. The money market shares of chartists and fundamentalists are denoted $m(t)$ and $(1-m(t))$, respectively.

The FX rate $S(t)$ represents a temporal equilibrium between two currencies available during the trading period t, as usual $m_{I}(t) = S(t)\, m_{II}(t)$, or in an additive form $log(m_{I}(t)) = log(S(t)) + log(m_{II}(t))$. We assume $t \rightarrow S(t)$ is a real positive discrete function defined on $\{..., t\text{-}2, t\text{-}1, t, t\text{+}1, ...\}$, which represents a trajectory of a discrete stochastic process with continuous state space.

The demands $d^{C}(t)$, $d^{F}(t)$, $d^{B}(t)$ are formulated in particular as a mix of systematic and unsystematic components as follows

$$d^{C}(t) = a^{C,1}\varphi(.) + a^{C,2}\omega(.), \tag{2}$$

where $a^{C,1}$, $a^{C,2} > 0$ represent corresponding reaction coefficients levering both modes. The general form is

$$d^{C}(t) = a^{C,1}\left(\textstyle\sum_{(i)} \alpha_{i}\, log(S(t-i)/S(t-i-1)) \right) + a^{C,2}\delta(t-1), \tag{3}$$

where $\delta(t-1) \sim N(0,\sigma_{C}^{2})$ is a typical noise with zero mean and time invariant finite variance σ_{C}^{2}, index i identifies truncated history of past rates $\{S(t-i),\ i = 1,...,n_{C}+1\}$, with n_{C} given.
The demand excess of fundamentalists $d^{F}(t)$ within the period t is

$$d^{F}(t) = a^{F,1}(\ \mathrm{E}^{F}_{t}\,[S(t+1)] - S(t))\ /\ S(t), \tag{4}$$

where $\mathrm{E}^{F}_{t}\,[S(t+1)]$ expresses the expected future FX rate made by fundamentalist agent at t, and $a^{F,1} > 0$ is reaction coefficient coping with relative distance between expected future FX rate and the spot rate. Rather general way of forming that expectation, sometimes called an anchoring heuristics, is following

$$\mathrm{E}^{F}_{t}\,[S(t+1)] = \gamma S^{F}(t-1) + (1-\gamma)\ (\textstyle\sum_{(j)} \beta_{j} S(t-j)), \tag{5}$$

where $\beta_{j} \geq 0$ are weights of $S(t-j)$, $j = 1,...,n_{F}$, such that $\sum_{(j)} \beta_{j} = 1$, with n_{F} given, and $S^{F}(t-1)$ denotes the last fundamental value. Development of $S^{F}(t)$ is quite natural to assume that due to rate-related news it behaves like a discrete random walk

$$logS^{F}(t) = logS^{F}(t-1) + a^{F,2}\varepsilon(t), \tag{6}$$

where $a^{F,2}$ is coefficient representing probability of news shock hitting FX money market, $0 < a^{F,2} < 1$, and $\varepsilon(t) \sim N(0,\sigma_{F}^{2})$ is a typical noise with zero mean and time invariant finite variance σ_{F}^{2}, again.

The market share $m(t)$ is formulated generally by following mappings

$$m(t) = (1 + w(t))^{-1}\ ,\ \text{or equivalently } w(t) = (1-m(t))/m(t) \tag{7}$$

with $w(t)$ ranging $[0,+\infty[$ being defined in form of a compound function

$$w(r(t)) = \chi_{0} + (\textstyle\sum_{(k)} \chi_{k}\, r(t)^{2k})^{1/p}\ , \tag{8}$$

where $\chi_{0} \geq 0$, $\chi_{k} > 0$, $k = 1,...,n_{S}$ are real constants given, and $n_{S} \geq 1$, $p \geq 1$, too. The function value $r(t)$ expresses a relative distance between $S^{F}(t-1)$ and $S(t-1)$, in a way similar to (9) in some sense

$$r(t) = (S^{F}(t-1) - S(t-1))/S(t-1). \tag{9}$$

Hence, $m(t)$ can be calculated form set of past observations $\{S(t-i),\ S^{F}(t-j),\ i, j = 1,...\}$ known.

Finally, the $d^{B}(t)$ may be expressed as convex combination of both basic CB strategies, i.e reversing and targeting one

$$d^{B}(t) = \eta\ d^{B,1}(t) + (1-\eta)\ d^{B,2}(t), \tag{10}$$

where η gives a leverage between both intervention strategies, $\eta \in [0,1]$ given. It holds for reversing strategy

$$d^{B,1}(t) = a^{B,1}\ (\log S(t-2) - \log S(t-1)),$$

whereas for targeting strategy

$$d^{B,2}(t) = a^{B,2}\ (S^{F}(t-1) - S(t-1))/S(t-1) = a^{B,2}\ r(t),$$

with $a^{B,1}$, $a^{B,2} > 0$, given reaction coefficients of the corresponding strategies.

Substituting all given expressions $d^{C}(t)$, $d^{F}(t)$, $d^{B}(t)$ into (1) and solving for $S(t)$ we get

$$S(t) = \mathrm{E}^{F}_{t}\,[S(t+1)]/\{1-(m(t)\ d^{C}(t) + d^{B}(t))/(a^{F,1}(1 - m(t)))\}, \tag{11}$$

which is the desired dynamic stochastic simulation model of FX rate respecting various CB's influence effects, and also a time series generator for next VRA procedure and numerical experiments.

Visual Recurrence Analysis

Visual recurrence analysis (VRA) is a tool for time series analysis based on computer graphics approach profiting from various image processing algorithms. It has become a useful and popular technique recently, and the main reasons are following:

1. it is well adapted for analysis of non-linear systems,
2. the output of VRA is an image which is well suited for human perception, thus enabling accumulation of experience,
3. it provides an effective way for detection of various patterns hidden in time series.

The basic steps of VRA procedure are summarized as follows:

1. Input data – the given time series at equidistant time steps δ

$$\{S_t\},\ t = 0,...,T$$

2. Transform data – find a suitable affine transformation $\mathtt{A}$, i.e. moving and scaling of time series state space, in order to get a new series with non-negative values: $\xi_t \geq 0,\ t = 0,...,T$

$$\mathtt{A} : \{S_t\} \rightarrow \{\xi_t\}$$

– calculate their global thresholds

$$d = \min(\xi_t),\ \ h = \max(\xi_t),\ \ t = 0,...,T$$

– extract a sample of time series by moving time window $\mathtt{W}$ defined by given local start at $t = w$, and a span K:

$$\mathtt{W} : \{\xi_t\} \rightarrow \{x_i\},\ i = 0,...,K$$

i.e. by assignment

$$x_i = \xi_t,\ t = i + w,\ \ i = 0,...,K\ ,$$

where an integer $w \geq 0$ defines a beginning of the sample, whereas an integer K defines its local end. The value of K is usually determined by pixel resolution of a screen,

$$K \sim \min(\text{horizontal axis, vertical axis}),\ \text{e.g. } K = 700.$$

Hence, the last sample possible to extract is given by indices

$$t = i + w\ ,\ \ i = 0,...,K\ ,\ \text{with } w = T - K.$$

3. Expand an 1-d time series sample $\{x_i\},\ i = 0,...,K$ into a 2-d image I(RGB).
Well known motivating idea of VRA is to convert numerical values into some graphical information. In particular, it creates a color of each image pixel $(i,j),\ i,j = 0,\ldots,\ K$ in correspondence with numerical values of a couple (x_i,x_j).
Let $\mathtt{M}$ denote a mapping generating the RGB image I(RGB)

$$\mathtt{M} : (x_i,x_j) \rightarrow c(i,j),\ i,j = 0,...,K \tag{12}$$

where $c(i,j)$ denotes RGB color of pixel (i,j).

Traditional approach – a color $c(i,j)$ is represented by 24 bits, with 8 bits for each RGB color, and it is set $c(i,j) \sim |x_i - x_j|$ using selected color palette (with 256^3 = 16 777 216 different colors, in general) and a proper time series state space discretization.

<u>Proposed approach</u> – a color $c(i,j)$ is tackled as 3-d vector $\boldsymbol{c}(i,j) = (r(i,j),\ g(i,j),\ b(i,j))^{\mathrm{T}}$, with RGB color components denoted $r(i,j)$, $g(i,j)$, $b(i,j)$ ranging [0,255], respectively, i.e. with 256 different levels each. Such individual acceptance of basic color components gives us larger flexibility to express and investigate various relations between selected cue values x_i, x_j. However in such case, the finest discretization is restricted to 256 levels, which is much coarser than about 16 million of ones provided by traditional approach, but we hope it is still acceptable for various VRA applications in finance.
Let us introduce three color mapping functions

$$r(i,j) = \varphi(x_i, x_j)\ ,\quad g(i,j) = \chi(x_i, x_j),\quad b(i,j) = \psi(x_i, x_j), \tag{13}$$

hence, we able to define a 3-d mapping $\mathbf{M}$ as follows

$$\mathbf{M} : (x_i,x_j) \rightarrow \boldsymbol{c}(i,j) = (\varphi(x_i,x_j),\ \chi(x_i, x_j),\ \psi(x_i, x_j))^{\mathrm{T}},\ i,j = 0,...,K. \tag{14}$$

Some possible choices of color mapping functions:
a) $\varphi(x_i, x_j) = x_i\ ,\quad \chi(x_i, x_j) = x_j\ ,\quad \psi(x_i, x_j) = |x_i - x_j|,$

which enables construction of mapping corresponding with the traditional approach
$\varphi(x_i, x_j), \chi(x_i, x_j) = 0, \psi(x_i, x_j) = |x_i - x_j|$ provided the distance $|x_i - x_j|$ is converted into blue colors only,

b) $\varphi(x_i, x_j) = x_i$, $\chi(x_i, x_j) = x_j$, $\psi(x_i, x_j) = (x_i + x_j)/2$,

c) $\varphi(x_i, x_j) = x_i$, $\chi(x_i, x_j) = x_j$, $\psi(x_i, x_j) = ((x_i)^2+(x_j)^2)^{1/2}$,

There is evident we may either build any permutation of RGB color components, and/or to combine any different distances in order to construct a new suitable mapping $\mathbf{M}$.

Note, another interesting possibility will occur if we substitute local values x_i and x_j by some local averages at x_i and x_j, but that topic will not be discuss in this paper.

Concluding, an advantage of the proposed approach is evident – it provides a lot of various ways how to define a color $c(i,j)$ at the pixel (i,j), and how to produce an image I(RGB), since any RGB color component may represent either the first or the second cue value of the time series, i.e. x_i, x_j, their average $(x_i+x_j)/2$, an absolute distance $|x_i - x_j|$, an Euclidean distance $((x_i)^2+(x_j)^2)^{1/2}$, etc. Such flexibility gives a new opportunity to use VRA for deeper understanding of time series.

Encryption scheme

Motivation for image encryption stems from fact that generated VRA images, in particular, such ones containing CB's influences upon FX rate, may be declared confidential. Hence, cryptographic solutions should be used to ensure privacy and confidentiality when such VRA images are to be transmitted over un-trusted or public networks.

We adopt the encryption scheme proposed in the paper [6] to produce VRA image encryption, which is well-adapted for implementation and cost-effective. The basic idea of this encryption scheme stems from well-known principles of human perception of color images.
Any pixel color image I(RGB), e.g. VRA image, is given by pixel mapping π: $(Z_+)^2 \rightarrow (Z_+)^3$

$$\pi: (i,j) \rightarrow \underline{\boldsymbol{c}}(i,j) = (r(i,j), g(i,j), b(i,j))^T, \; i,j = 0,..,K.$$

Hence, we may summarize the encryption/decryption procedure as follows:

1. Extract magnitude and orientation of color vector $\boldsymbol{c}(i,j)$ at the pixel (i,j), which correspond to luminance and chromaticity in human perception of colors, respectively –

Luminance ~ color vector magnitude is given by

$$G(i,j) = \|\boldsymbol{c}(i,j)\| = ((r(i,j))^2 + (g(i,j))^2 + (b(i,j))^2)^{1/2}, \tag{15}$$

Chromaticity ~ color vector orientation is given by

$$\boldsymbol{D}(i,j) = \boldsymbol{c}(i,j)\,(\|\boldsymbol{c}(i,j)\|)^{-1} = \boldsymbol{c}(i,j)/G(i,j), \text{ where } \|\boldsymbol{D}(i,j)\| = 1. \tag{16}$$

2. Extract bit-level components in order to alter both luminance and chromaticity simultaneously –
Let redefine the color vector $\underline{\boldsymbol{c}}(i,j) = (r(i,j), g(i,j), b(i,j))$ component-wise

$$\underline{\boldsymbol{c}}(i,j) = (r(i,j), g(i,j), b(i,j)) = (c_1(i,j), c_2(i,j), c_3(i,j)), \text{ each } c_n(i,j),\ n = 1,2,3 \text{ ranges } [0, 255],$$

we may express the bit-level representation of pixel color as follows

$$\underline{\boldsymbol{c}}(i,j) \sim \sum_{b=1}^{B} {}^{b}\underline{\boldsymbol{c}}(i,j)\,(2)^{B-b} \tag{17}$$

where

$${}^{b}\underline{\boldsymbol{c}}(i,j) = ({}^{b}c_1(i,j), {}^{b}c_2(i,j), {}^{b}c_3(i,j)) = ({}^{b}r(i,j), {}^{b}g(i,j), {}^{b}b(i,j)) \in \{0,1\}^3,$$

is a b-th bit-level binary vector composed from corresponding RGB bit-level components ${}^{b}c_n(i,j)$, $n = 1,2,3$, $b = 1,..,B$, taking $B = 8$, as usual in 24-bit color representation.

3. Encryption – transfer – decryption steps

- encryption – use bit-level binary vectors ${}^{b}\underline{\boldsymbol{c}}(i,j)$ to generate two binary auxiliary/share vectors ${}^{b}\underline{\boldsymbol{u}}(i,j)$, ${}^{b}\underline{\boldsymbol{v}}(i,j) \in \{0,1\}^3$, which are further used for creation VRA encrypted image,
  – apply component-wise choice-dependent assignment

$$({}^{b}u_n(i,j), {}^{b}v_n(i,j)) \in \{(0,1),(1,0)\} \text{ iff } {}^{b}c_n(i,j) = 1, \; n = 1,2,3$$

$\{(0,0),(1,1)\}$ iff ${}^{b}c_n(i,j) = 0$,

– create two share color vectors

$$\underline{\boldsymbol{u}}(i,j) \sim \sum_{b=1}^{B} {}^{b}\underline{\boldsymbol{u}}(i,j)\,(2)^{B-b}, \qquad \underline{\boldsymbol{v}}(i,j) \sim \sum_{b=1}^{B} {}^{b}\underline{\boldsymbol{v}}(i,j)\,(2)^{B-b} \qquad (18)$$

– create pixel-wise two b-th bit-level share images ${}^{b}U(\mathrm{RGB})$, ${}^{b}V(\mathrm{RGB})$

$${}^{b}U := \{\underline{\boldsymbol{u}}(i,j)\}, \quad {}^{b}V := \{\underline{\boldsymbol{v}}(i,j)\}, \quad i,j = 0,..,K,$$

which differ in luminance and chromaticity pixel-wise each other and the original VRA image, too.

- transfer VRA encrypted images, $\{ {}^{b}U, {}^{b}V \}$, $b = 1,..,B$, over public networks,
- decryption – compose the original VRA image $I(\mathrm{RGB})$ from the VRA encrypted images by decryption procedure, which is 'inverse' to encryption one.

For more technical details we refer the paper [6].

Conclusions

1. Various FX rate stochastic simulation models were presented. The dynamic money market clearing conditions were used for derivation of such models and special focused upon modelling of CB's influence.
2. Generalization of VRA procedure was proposed. There were discussed various possibilities how to construct color mapping functions more suitable for subtle investigation of a time series sample. We hope the proposed variety of functions gives more flexibility for deeper understanding of correlations patterns hidden in time series.

3. Encryption scheme applied to VRA images in particular was given in details. Such procedure is to provide safe transmit of case sensitive VRA images over public network.

Work on progress and field of further research

a. Development of sw package implementing VRA with encryption on the Java 2D platform implementing the java.ai purpose-oriented package for advanced image processing.
b. Identification and calibration of model constitutive parameters included in presented FX rate models using various estimation procedures. Accumulate and evaluate empirical evidence of real FX rate time series.
c. Further development of VRA procedure, which represents an interesting and challenging instrument for both detection and investigation of various patterns caused by endogenous nonlinear effects already observed and reported in the field of time-dependent FX rate development.

Acknowledgement

This research work was supported both by the grant No. 402/09/1536 of the Grant Agency of Czech Republic and by funds of the project LC06075 of the Ministry of Education of Czech Republic.

References

[1] Blake, D.: *Financial Market Analysis.* (*Analýza finančních trhů*), Grada Publishing, Praha, 1995

[2] Fiess, N.M. – MacDonald, R.: *Towards the fundamentals of technical analysis*: *analysing the information content of High, Low and Close prices*. Economic Modelling, 19 (2002), 353-374

[3] LeBaron B.: *Technical trading rule profitability and foreign exchange intervention.* J.Int. Economics, 49 (1999), 125-143

[4] LeBaron B.: *Agent-based computational finance*: *Suggested readings and early research*. J.Econ. Dynamics and Control, 24 (2000), 679-7002

[5] Lui, Y.H. – Mole, D.: *The use of fundamental and technical analyses by foreign exchange dealers*: *Hong Kong evidence*. J.Int. Money and Finance, 17 (1998), 535-545

[6] Lukac, R. – Plataniotis, K.N.: *A cost-effectie encryption scheme for color images*. Real-Time Imaging 11 (2005) 454-464, Elsevier

[7] Lukáš, L.: *Market clearing conditions used in FX rate simulation models*. Proceedings 22-th Int.Conf. Mathematical Methods in Economics, Masaryk University Brno, 2004, ISBN 80-210-3496-3, 184-188

[8] Lukáš, L.: *Variety of Agent-based Models for Computer Simulation of FX Rate*. 9-th Annual Conf. Computing in Economics and Finance, July 11-13, 2003, Seattle, Washington, USA

[9] Saacke, P.: *Technical analysis and the effectiveness of central bank intervention*. J.Int. Money and Finance, 21 (2002), 459-479

[10] Westerhoff, F.H.: *Central bank intervention and feedback traders.* Int. Fin. Markets, Inst. and Money, 13 (2003), 419-427

DISEQUILIBRIUM OPEN ECONOMY WITH PAPER MONEY (A MODEL ACCORDING TO A. RAŠÍN)

David Martinčík

University of West Bohemia in Pilsen
Faculty of economics, Department of economics and finance,
martinci@kef.zcu.cz

Abstract. The article contains a model built according to thinking of Alois Rasin, the famous Czech economist and politician at the beginning of 20-th century. Author of this article used a selected parts of Rašín's textbook „National economy" which treat of the international adjustment under paper (uncovered) money. Only a few pages of original Rašín's text deal with a lot of economics variables and their relations, often without detailed explanation. Author selected only 12 variables and their relations to construct a sufficiently simple model that transforms the Rašín's ideas to formal mathematical language. The model describes the behavior of domestic economy after the monetary expansion in relation to the rest of the world with stable currency. The model consist of one exogenous variable (money stock in domestic country), 11 endogenous variables (exchange rate, real export and import, domestic and foreign price level, domestic and foreign nominal wage rate, domestic and foreign flow of nominal savings, domestic and foreign interest rate) and 11 equations. After rearrangement the model contains 6 endogenous variables and 6 ordinary linear differential equations. A numerical solution under selected parameters gives the same result as Rašín's claim and it could be considered as a proof of logical consistency of his thinking about the disequilibrium of this economic system. The transformation of historic text to the formalized mathematical language is difficult job and author is familiar with the danger of distortion through misunderstanding of given historic economic paradigm. Nevertheless author maintains that this approach is useful – many times forgotten old ideas can upgrade the contemporary economic thought.
Keywords. Rašín Alois, monetary expansion, disequilibrium open economy, continuous model

1. Introduction, methods and objectives

The problem of open economies behavior is more and more important. The cohesion of national economies raise through the continuing globalization and economic affairs in one part of the world influence the economics variables in another part of world. One of the feasible inspirations for description of present economies could be the "old" texts.

By default the abroad influence on domestic economy is explained by the help of automatic adjustment mechanism of the balance of payment or of the trade balance – price, income, monetary and exchange rate mechanism. However in some historic texts we can find mechanisms that are marginalized in contemporary economics.

The aim of the article is to construct an open economy model based on ideas of Alois Rašín expressed in his textbook *National Economy* published in 1922. Mr. Rašín wasn't an excellent theoretician, more likely he was practical economist and politician. His opinions were influenced by economic paradigms overbearing in central Europe of his days – German historic school, Austrian subjective-psychological school and universal ideas of liberalism. Rašín is known as unyielding advocate of "gold currency" however at the same time he refused the quantity theory of money and often used endogenous money concept. He put emphasis on active interest rate policy performed by central bank which is typical postkeynesian idea in contemporary economics.

From the reason, using different points of view, and further from the reason non-use the formalization for economic relations (the neoclassical economics wasn't wide-spread in central Europe in Rašín's time furthermore), is the transformation of Rašín's text to the formalized model controversial. Author of this article knows all the dangers arisen from this method but is confident that some of this way detected relation may be useful for contemporary economics.

This article connects to the author's previous text [1] where the discrete model of open economy with paper money is shown among others. Now the continuous model is constructed and the numerical solution is demonstrated.

To keep the text well-arranged and transparent, the original Rašín's text (strictly speaking the text parts which the author considers as important for given topic – the resource is the chapter *Money in international trade* from the textbook *National Economy*, p. 185-190, [2]) is italic and the formalization and explanation of used variables follow. After the equations reconstruction from the original text is performed the rearrangement to the system of ordinary differential equations and numerical solution.

To keep the model simple, all economic relations are described as linear with eliminated constant terms. Further all equations are constructed, to be all applied parameters positive. The Latin capital letters denote the variables, the small letters (Latin or Greek alphabet) denote the parameters in case of need constants.

It will be seen that Rašín's approach is very unusual, different from "standard economics". The present economic theory distinguishes four ways of automatic adjustment mechanism of trade balance: price, income, monetary and exchange rate. The price adjustment mechanism (price-specie flow) was originally published by David Hume in 1752, for interpretation of this see another author's text [3] and [4]. The income adjustment mechanism was developed by Roy Forbes Harrod in 1932 at the beginning of period of Keynesianism, for a comparison of the price and income adjustment mechanism see [5] and [6]. The monetary adjustment mechanism appeared the first time in the *Wealth of Nations* by Adam Smith in 1776 but the classical political economy didn't understand this approach and it reappeared again in 70's of 20th century. The rigorous exposition of exchange rate adjustment mechanism was developed in first half of 20th century by Alfred Marshall (1924), Joan Robinson (1937) and Abba Lerner (1944).

In the period of Rašín's live is remarkable also another Czech economist Karel Enlgliš. His work contains a combination of price and exchange rate adjustment. In the concrete it is a macroeconomic partial equilibrium approach with following variables: exchange rate, domestic price level and given world price level. The Engliš's approach is able to be used for both exchange rate regimes – fixed and floating. Under the floating exchange rate regime the adjustment goes through depreciation or appreciation and under the fixed one the all adjustment goes through domestic price level changes i.e. through inflation or deflation, for detailed description of this see [7]. The Rašín's contribution is very unconventional in comparison with the one of Engliš's. It leads up to the absence of automatic adjustment and to the deepening disequilibrium of balance of payment.

2. "Model extraction" from original text

... in the countries with paper currency, if circulation of money overflows its actual needs which happen as a rule, because the state uses the printing of money as a source of revenue and reproduces it not for the needs of national economy, but for the help to government budget (e.g. during the war). It is not possible to reduce the circulation of this money if its need is depressed, specie flows out through the payments to abroad, paper money becomes cheaper and gets disagio in relation to species, this means that species become expensive and get agio because of their scarcity and surplus of paper money. Indeed all this influences the exchange rate and commercial papers (bills) in foreign currency rise a lot.

The exchange rate is positively affected by the money stock which could be expressed by simple equation:

$$E = \omega M \tag{1}$$

where is E nominal exchange rate (outright quote)
M nominal money stock in domestic country
ω parameter of sensitiveness

Unfortunately Rašín didn't explain how he came to this relation – evidently he had on mind the determination of exchange rate as a relative price between domestic (paper) currency and foreign (gold) currency. He didn't mention other economic variables through them the money stock is connected with the exchange rate and therefore author doesn't want to speculate and accepts this relation as given and the equation as written.

This paper money disagio (agio of species) is profitable for export at the beginning because the exporter can bid the goods very cheap in abroad in foreign currency whereas he become a lot of domestic currency for one unit of foreign currency and purchases the same quantity of goods. …However the exporter doesn't become the whole agio profit because the wants to export as much as possible and preferably he cuts the prices for the abroad.

Rašín had in mind the physical size of export (real export) which is increasing while domestic currency is depreciating. We will use the linear form of this relation:

$$X = bE \tag{2a}$$

where is X physical size of export (real export)
b parameter of sensitiveness

However the agio takes effect as a protective tariff, inhibits the import because the importer have to ask to much in domestic paper currency, to obtain a adequate sum in his own currency.

The relation of physical size of import and exchange rate is naturally opposite than by export:

$$Z = -dE \tag{3a}$$

where is Z physical size of import of domestic country (real import)

d parameter of sensitiveness

This implies that domestic producers prosper from less foreign competition and increase the goods prices. Besides the things which have to be imported (e.g. cotton, wool, leather, cooper ore) are sold – calculate in domestic money – very expensive to provide the importer adequate amount in his money and he could buy it in abroad in expensive currency.

The fall of import reduces the goods supply on domestic market and causes the increase of domestic price level:

$$\dot{P}_d = -\alpha \dot{Z} \tag{4a}$$

where is $\dot{P}_d$ time derivative of domestic price level

$\dot{Z}$ time derivative of real import

α parameter of sensitiveness

All the goods made from this raw materials become expensive which causes the increasing of wages, salaries, etc. and all production costs.

Rašín thought the situation when the wages (wage rates) aren't rigid on the same level but the wages development follows the development of price level:

$$\dot{W}_d = \dot{P}_d \tag{5}$$

where is $\dot{W}_d$ time derivative of domestic nominal wage rate

to keep simplicity the sensitiveness parameter is missing (thereby is reduced the final number of equations and variables)

…the state with depreciating currency begins to feel the lack of capital because the increasing prices (calculate in depreciate paper money) cause that everybody consumes the greater part of his income then earlier, less saves, …

The nominal savings fall as a result of increasing price level[1]:

$$\dot{S}_d = -g\dot{P}_d \tag{6}$$

where is $\dot{S}_d$ time derivative of nominal domestic savings

g sum of instantaneous intensities of domestic real production and real consumption

…but the production requires more operation capital (wages and interest rates increase), capital isn't accumulated in domestic economy, it is necessary to borrow form abroad. The abroad lends unwillingly … the uncertain situation must be paid through higher interest rate. …By this the production margin becomes wrong, interest costs increase in country with depreciating currency.

The interest rate is determined by the capital market where the supply is constituted by savings and demand is constituted by the need of operation capital (wages and interests[2]). The simplest way is to define the time derivative of interest rate as a linear function of exceed capital demand:

$$\dot{R}_d = \rho(W_d l_d + R_d k_d - S_d) \tag{7}$$

where is $\dot{R}_d$ time derivative of domestic interest rate

ρ adjustment speed of domestic interest rate to the exceed capital demand

[1] This relation we could obtain by limiting the discrete equation with an analogy of Robertson's time lag (lag of expenditures behind income): $S_{t+1} = YP_t - CP_{t+1}$, where S denotes the nominal savings, Y is constant real production, C is constant real consumption and the index t denotes discrete time. By equation redating one time period back we obtain $S_t = YP_{t-1} - CP_t$ and by subtraction of both equations we obtain $S_{t+1} - S_t = Y(P_t - P_{t-1}) - C(P_{t+1} - P_t) = Y\nabla P - C\Delta P$ where ∇ denotes the backward difference and Δ denotes the forward difference. By limit process we obtain $\dot{S} = -y\dot{P} - c\dot{P} = -(y+c)\dot{P}$ where y and c are the instantaneous intensities of domestic real production and real consumption (rem.: the term with backward difference changes the sing of derivative).

[2] Rašín didn't mention the expected inflation so it is not possible to distinguish the real and nominal interest rate.

	l_d	labor size in domestic country (constant)
	k_d	size of physical capital in domestic country (constant)

The purchasing power of money increases in the state with appreciating currency, the goods prices and wages decline and the export becomes higher. …in the state with appreciating currency the more of capital remains at home, interest rate declines a through that the country becomes more-competitive in despite of the agio the depreciating currency.

From the last sentence we can deduce (although it describes the abroad) that interest rate determines real export and according to Rašín even more then exchange rate. The increasing interest rate declines the competitive advantage of the country. Therefore we reformulate the equation (2a) this way:

$$X = -aR_d + bE \tag{2}$$

where is	a	parameter of sensitiveness

By the reason that import to domestic country is the same as export from abroad, we have to reformulate the equation (3a) by the same way:

$$Z = -cR_f - dE \tag{3}$$

where is	R_f	foreign interest rate
	c	parameter of sensitiveness

The competitive advantage mentioned by Rašín is to understand as price competitiveness. It is suitable to complete the equation (4a) by the changes of interest rate:

$$\dot{P}_d = -\alpha\dot{Z} + \beta\dot{R}_d \tag{4}$$

where is	β	parameter of sensitiveness

From the Rašín's text is noticeable that the process in domestic economy and abroad are "complementary" and therefore we can define for abroad the phenomena described for domestic country by equations (4), (5), (6) a (7) by analogy:

$$\dot{P}_f = -\gamma\dot{X} + \delta\dot{R}_f \tag{8}$$

where is	$\dot{P}_f$	time derivative of abroad price level
	$\dot{X}$	time derivative of real export from domestic country (real import to abroad)
	$\gamma;\delta$	parameters of sensitiveness

$$\dot{W}_f = \dot{P}_f \tag{9}$$

where is	$\dot{W}_f$	time derivative of abroad nominal wage rate

to keep simplicity the sensitiveness parameter is missing

$$\dot{S}_f = -h\dot{P}_f \tag{10}$$

where is	$\dot{S}_f$	time derivative of abroad nominal savings
	h	sum of instantaneous intensities of abroad real production and real consumption

$$\dot{R}_f = \sigma(W_f l_f + R_f k_f - S_f) \tag{11}$$

where is	$\dot{R}_f$	time derivative of abroad interest rate
	σ	adjustment speed of abroad interest rate to the exceed capital demand
	l_f	labor size in abroad (constant)
	k_f	size of physical capital in abroad (constant)

The model contains 11 equations which are written again for lucidity and transparency:

$$E = \omega M \tag{1}$$

$$X = -aR_d + bE \tag{2}$$

$$Z = -cR_f - dE \tag{3}$$

$$\dot{P}_d = -\alpha\dot{Z} + \beta\dot{R}_d \tag{4}$$

$$\dot{W}_d = \dot{P}_d \quad (5)$$
$$\dot{S}_d = -g\dot{P}_d \quad (6)$$
$$\dot{R}_d = \rho(W_d l_d + R_d k_d - S_d) \quad (7)$$
$$\dot{P}_f = -\gamma\dot{X} + \delta\dot{R}_f \quad (8)$$
$$\dot{W}_f = \dot{P}_f \quad (9)$$
$$\dot{S}_f = -h\dot{P}_f \quad (10)$$
$$\dot{R}_f = \sigma(W_f l_f + R_f k_f - S_f) \quad (11)$$

The money stock in domestic country M is exogenous variable and the number of endogenous variables is 11 ($E, X, Z, P_d, W_d, S_d, R_d, P_f, W_f, S_f, R_f$) so the model allows the solution.

First we rearrange the model: substituting equation (1) to the equation (2) and (3), their time differentiation and substitution to equation (4) and (8). Further we rearrange the whole structure to the system of linear ODE:

$$\dot{W}_d = -\alpha[-c\sigma(W_f l_f + R_f k_f - S_f) + d\omega\dot{M}] + \beta\rho(W_d l_d + R_d k_d - S_d) \quad (12)$$
$$\dot{S}_d = g\alpha[-c\sigma(W_f l_f + R_f k_f - S_f) + d\omega\dot{M}] - g\beta\rho(W_d l_d + R_d k_d - S_d) \quad (13)$$
$$\dot{R}_d = \rho(W_d l_d + R_d k_d - S_d) \quad (14)$$
$$\dot{W}_f = -\gamma[-a\rho(W_d l_d + R_d k_d - S_d) - b\omega\dot{M}] + \delta\sigma(W_f l_f + R_f k_f - S_f) \quad (15)$$
$$\dot{S}_f = h\gamma[-a\rho(W_d l_d + R_d k_d - S_d) - b\omega\dot{M}] - h\delta\sigma(W_f l_f + R_f k_f - S_f) \quad (16)$$
$$\dot{R}_f = \sigma(W_f l_f + R_f k_f - S_f) \quad (17)$$

It remains 6 endogenous variables and 6 equations.

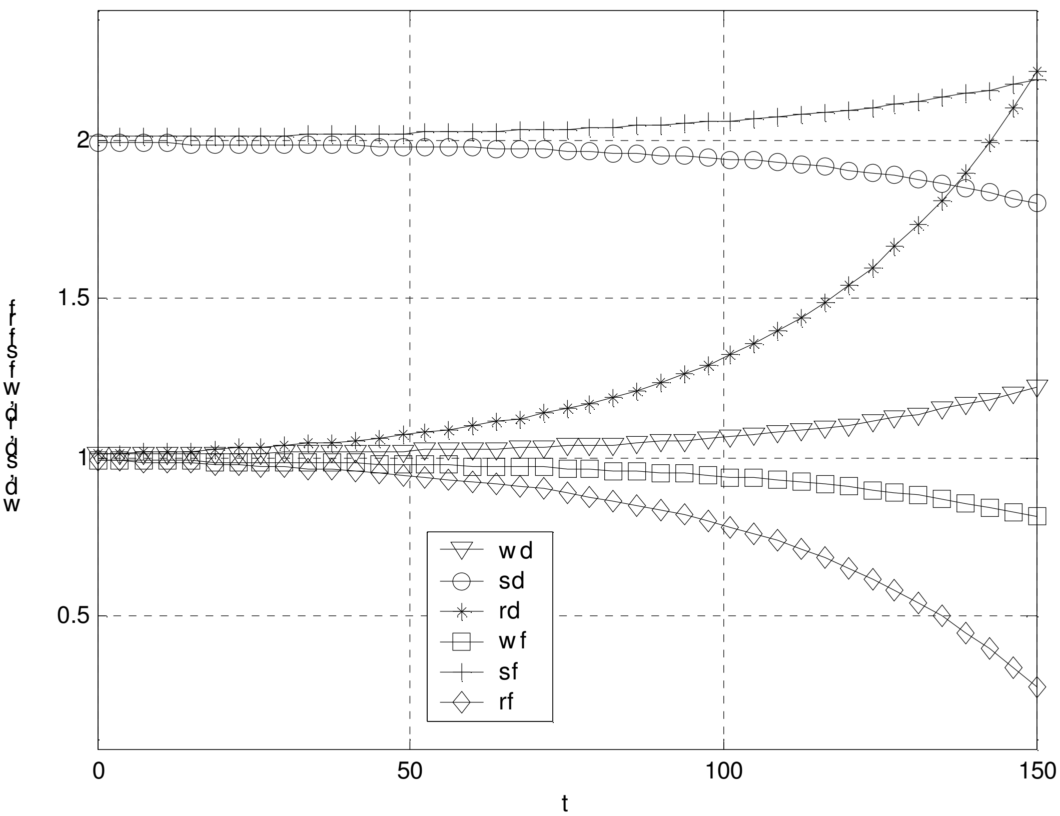
3. Numerical solution

Putting all the left sides of equations equal to zero we solve the steady state. First it is necessary to choose constants suitable way: $l_d = l_f = k_d = k_f = 1$. For constant exogenous variable ($\dot{M} = 0$) this choice gives as one of possible steady states e.g.: $W_d = R_d = W_f = R_f = 1$; $S_d = S_f = 2$ (the equations are linear dependent) – this steady state or deviation from it we use as initial conditions for numerical solution of ODE system.

We choose the parameters as follows: $\omega = 1$; $a = 0.12$; $b = 0.08$; $c = 0{,}08$; $d = 0.18$; $g = 0.9$; $h = 0.8$; $\alpha = 0.08$; $\beta = 0.18$; $\gamma = 0.12$; $\delta = 0.3$; $\rho = 0.02$; $\sigma = 0.016$ (the choice is determined by non-overlaying of the trajectories on graph).

Further we have to choose the initial conditions. The system behaves chaotically (different deviation of different variables from steady state leads to totally different directions of variables progression) and therefore we must deviate the variables in the "correct" direction – according to Rašín:

As a result of increase of money stock in domestic country the immediate depreciation of domestic currency passes over, next follows the physical size of export increase and physical size of import decrease. This reduces the supply of goods on the domestic market a leads to rise of domestic price level and nominal ware rate. Further follows decrease of domestic nominal savings (as a consequence of lag of spending behind income). This and the higher demand for operation capital (wages changes proportionally to price level) induce the increasing of domestic interest rate. The reverse process occurs abroad. Hence we choose the small deviation from steady state this way: $W_d = R_d = 1.01$; $W_f = R_f = 0.99$; $S_d = 1.99$; $S_f = 2.01$. The exogenous variable we left stable of course, i.e. $\dot{M} = 0$.



Graph 1: Variables trajectory

4. Conclusion

The theoretically weakest element of the model is the equation (1) which says: the money stock increase causes the lower price of this money in comparison with foreign currency (exchange rate E means the outright quote) – in general: the supply increase decline the relative price whereas the type of commodity doesn't matter, it could be what you want. But a more sophisticated exchange rate understanding is required. Rašín mentioned the "money need in national economy" – money demand – but he didn't specify it. It is possible to define the exchange rate changes as a function of excess money demand, for Cambridge money demand by this simple way: $\dot{E} = \omega(M - (1/v)PY)$, thereby we get close to monetary concept of exchange rate. Otherwise Rašín was an advocate of quantitative theory of money in the country with paper currency (more money increases the prices) but the higher prices in open economy he explained using the decline of import, less competition on domestic market and increase of raw materials prices. Therefore the author considers the formulation of equation (1) as adequate and applicable.

Rašín used on a few text pages a lot of economic variables and relations which aren't quite rigorous formulated and explained. The author of this article chose 12 variables and above described relations of them so that the formulation of the model kept simply and could show the consistency of Rašín's ideas: the deviation from steady state (the monetary expansion) leads to explosive trajectory of variables. The permanent increasing domestic interest rate and permanent decreasing abroad interest rate worsen the trade balance of domestic country (the physical size of export is going down and the physical size of import is going up) which attaches the borrowing from abroad. The foreign subjects require a high pay interest on their assets (e.g. as effect of raising risk premium) which causes more and more higher increase of domestic interest rate that is not incorporated in presented model. (*Therefore the foreign capital wants to lend in its own currency only to move the risk to debtor and the lack of appetite to undergo the uncertainty have to be compensate by higher interest payments.*)

Rašín saw in raising foreign debt the another risk: *A state with depreciate currency has a higher spending: the salaries of official and bookkeepers have to be increased, if the state owns a monopoly, the raw materials prices and wages become higher and the profit becomes lower, if the state runs the railways and mines the costs and investments raise, the taxes, monopoly prices and transport ticket prices have to increase and at last the borrowings are necessary but the domestic capital poverty forces to borrowing from abroad in foreign currency or with condition pay-off in species. …The pay interest on this state debt as well as the private debt causes the* *high demand for foreign currency, for foreign bills and a great exchange shocks appear during the time of balance of payment disequilibrium.*

On the score of this problem Rašín extrapolated the urgency of “robust anchor” for domestic currency – the gold. The tragic Rašín’s death at the beginning of 1923 causes the initiation of the end of effort to restore the gold cover of Czechoslovak currency.

References:

[1] MARTINČÍK, D. Náprava vnější ekonomické nerovnováhy v díle Aloise Rašína. In *Ekonomické znalosti pro tržní praxi*. Olomouc: Filosofická fakulta Univerzity Palackého, 2007, s.282-289. ISBN 978-80-903808-8-2. [on-line]: http://home.zcu.cz/~martinci/publikace_pdf/UP07.pdf

[2] RAŠÍN, A. *Národní hospodářství*. Praha: Český čtenář, 1922. 371 s. [on-line]: Zlatý fond českého ekonomického myšlení, http://www.econlib.cz/zlatyfond/html/zf0011.htm

[3] MARTINČÍK, D. Humeův vyrovnávací mechanismus obchodní bilance. In *Firma a konkurenční prostředí 2008*. Brno: MSD, 2008, s.96-102. ISBN 978-80-7392-020-3.Národní hospodářství. Praha: Český čtenář, 1922. 371 s. [on-line]: http://home.zcu.cz/~martinci/publikace_pdf/FKP08.pdf

[4] MARTINČÍK, D., ŠEDIVÁ, B. Dynamics of Hume's Law . In *Acta Oeconomica Pragensia*. 2007, roč.15. č.4, s.69-78, ISBN 0572-3043. [on-line]: http://home.zcu.cz/~martinci/publikace_pdf/AOP07.pdf

[5] MARTINČÍK, D. Cenový a důchodový vyrovnávací mechanismus platební bilance - srovnání. In *Hradecké ekonomické dny 2007/II*. Hradec Králové : Gaudeamus, 2007, s.22-27. ISBN 978-80-7041-824-6.

[6] MARTINČÍK, D. Cenový a důchodový vyrovnávací mechanismus platební bilance –podobnost v dynamice a jejich integrování do jednoho obecného modelu. In *Sborník 4. semináře Výpočtová ekonomie (18. 12. 2008).* Plzeň : Západočeská univerzita, (presently printing)

[7] MARTINČÍK, D. Vyrovnávací procesy obchodní bilance v díle Karla Engliše. In *Filosoficko-ekonomické myšlení brněnské normativní školy*. Brno : Masarykova univerzita, 2008, s.71-80. ISBN 978-80-210-4668-9. [on-line]: http://home.zcu.cz/~martinci/publikace_pdf/MUNI07.pdf

Acknowledgments:

The support from the Ministry of Education, Youth and Sports of the Czech Republic under the grant LC06075 is gratefully acknowledged.

MODELING THE CRISIS IN AUTOMOBILE INDUSTRY THE APPLIED CGE STUDY

Veronika Miťková

Comenius University in Bratislava
Faculty of Social and Economic Sciences, Institute of Economics,
veronika.mitkova@fses.uniba.sk

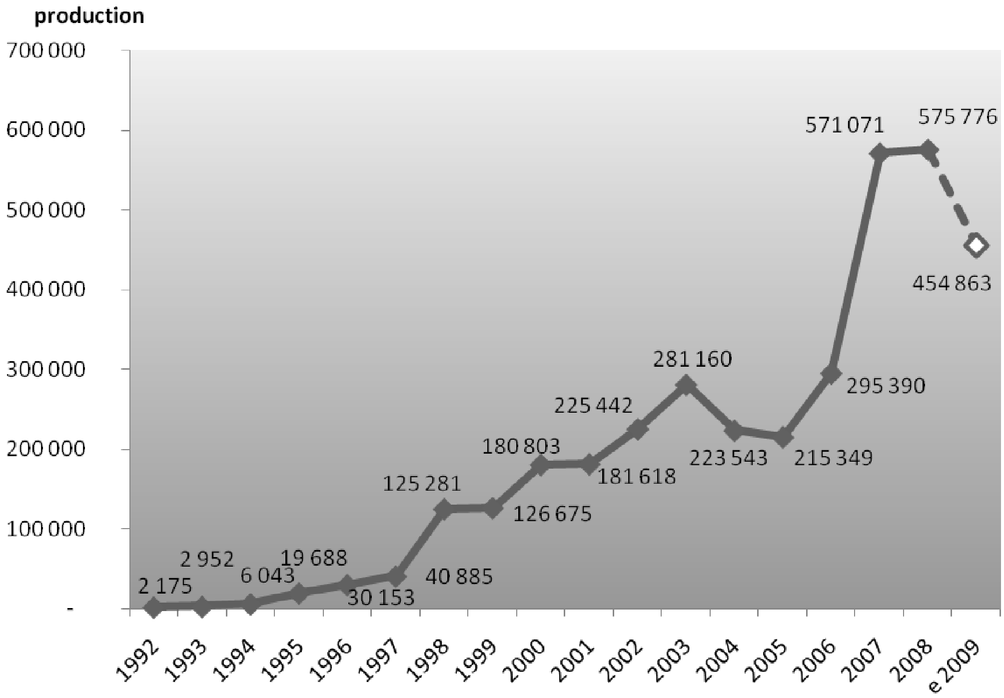
Abstract. The aim of the paper is to analyze the impact of the current crisis on the automobile industry in Slovakia. A computable general equilibrium model is used to model such shocks as decline of domestic production of automotives and production drop in all other industries. Also the scrapping subsidy program in which car owners who scrap cars manufactured before December 31, 1999 can receive a subsidy of up to €2,000 when buying a new car will be analyzed. The total welfare impact on Slovakian households will be evaluated.
Keywords. CGE model, automotive industry, crisis, scrapping subsidy program

1. Preliminary

Automobile industry, which is in the limelight of this paper, meets the controversial function of former armaments industry in Slovak mechanical engineering industry. As a one of most important sector of Slovak industries experienced a dynamic growth and it was a driving force of Slovak economy development. This industry started in the nineties of 20. century after the period of restructuring and finding of new production programs and markets. Partly thanks to the "Program of development of automotive industry" noted by Slovak government in 1998, which defined the vision, strategy and development goals, has Slovakia in 2007 according to Slovak Automotive Industry Association the first place in world in car production per 1000 inhabitants, which counts 105,7 cars [12]. Currently in Slovakia operate three automobile companies: Volkswagen, PSA Peugeot Citroën and Kia-Hyundai with total production capacity up to 900 000 cars per year. Graph 1 shows the production trend from 1992 through 1994 when Volkswagen started the production, 2006 when PSA Peugeot Citroën and Kia-Hyundai produced their first cars in Slovakia, to 2008. There was more than 94% growth in 2007 due to start of PSA Peugeot Citroën and Kia-Hyundai production. The impact of financial crisis is noticeable on 2008 data, when the production stagnated and according to estimates for 2009 the production will decline by 21%. Some European countries introduced the scrapping subsidy program which goal is to reduce the negative impact of crisis.

Since one quarter of Slovak GNP was produced by three automobile companies and a number of producers of car components and more than 40% of Slovakian export makes the automobile industry the crisis strongly hurts its economy. Slovakia joined the scrapping scheme program and the government allocated €33.2 million in the first wave [9] and €22.3 million in the second wave [10] for those scrap their old car and buy a new one. The main qualifying characteristics of cars eligible for the bonus were the same for both waves: they had to have been in use for at least 10 years and been registered in Slovakia before December 31, 2008. The price limit for new cars which can be purchased under the scheme is €25,000 and the maximum scrapping bonus was €2,000. On the part of this amount contributed the state, the rest was provided by the car seller as discounted normal price of the car.

This paper deals with both above mentioned problems: the economic crisis represented by automotives production decline and drop in production in all sectors of economy and with the scrapping subsidy program. The computable general equilibrium (CGE) model was used to simulate these situations and evaluate the impact to Slovak economy. The paper extends the current state of the CGE models in Slovakia (see [1], [2], [3], [7]) by applied model of Slovakia with special emphasis to the sector of automotives.



Graph 1: The automobile production in Slovakia from 1992 to 2008, estimation for 2009
Source: ZAP SR

Computable general equilibrium models are a class of economy-wide models used in policy analysis because they explicitly recognize that an exogenous change that affects any one part of the economy can produce repercussions throughout the system. CGE models describe by a system of equations the main idea that in economy resources equal to their uses. They are based on the microeconomic assumptions according to producers maximize profits subject to production functions, with primary factors as arguments, while households maximize utility subject to budget constraints. The model presented in this work is based on the model presented in [4] and is adapted to Slovak conditions. Changes were made in order to apply the model to the data base of System of national account according to ESA95 [5].

Benčík [1] elaborated the CGE problems in his work in which he constructs an experimental computable general equilibrium model of Slovakia based on Branch-and-Commodity Tables for the Supplies and Usage for year 1996. The model is used for import price increase simulation in primary industry by 8%, for import price increase of agricultural commodities, raw materials and products of primary industry. Benčík concludes that the model is able to provide rational results in spite of its considerable instability. Kotov and Páleník [3] quantify the positives and the negatives of Slovakia's joining the European Union by the CGE model based on social accounting matrix for year 1998. The overall conclusion of this work is that the entry of Slovakia to the European Union does not tend to the significant increase of the standard of living but that the total benefit prevail the total cost. Papers [11] and [8] extend the results of this work.

2. The CGE Model

The computable general equilibrium model from [6] was used to simulate the production drop and its modified version was applied to the government subsidy program evaluation.

Notations sets are:

activities α, $\alpha \in A = \{S-A, O-A\}$,

commodities c, $c \in C = \{S-C, O-C\}$,

factors f, $f \in F = \{LAB, CAP\}$,

institutions i, $i \in I = \{U-HHD, R-HHD, GOV, ROV\}$,

households h, $h \in H(\subset I) = \{U-HHD, R-HHD\}$,

where: $S-A$ analyzed activity,

$O-A$ aggregated other activities,

$S-C$	analyzed commodity,
$O-C$	aggregated other commodities,
LAB	labor,
CAP	capital,
GOV	government,
ROW	rest of the world,
U - HHD	urban household,
R - HHD	rural household.

The model is based on two basic microeconomic foundations: consumer utility maximization problem and firm's cost minimization problem. The problem of consumer utility maximization subject to household's income from ownership of production factors and household's expenditures less income tax and savings equality is formulated as follows:

$$\begin{aligned} &\max_{QH_{ch}} U(QH_{ch}) = \prod_{c\in C}\prod_{h\in H} QH_{ch}^{\beta_{ch}} \\ &s.t.\quad (1-mps_h)\cdot(1-ty_h)\cdot YH_h - \sum_{c\in C} PQ_c \cdot QH_{ch} = 0 \end{aligned} \tag{1}$$

and the household's demand function derived from this problem is

$$QH_{ch} = \frac{\beta_{ch}\cdot(1-mps_h)\cdot(1-ty_h)\cdot YH_h}{PQ_c}, \qquad c\in C,\ h\in H \tag{2}$$

where:

mps_h	ratio of disposable income of household h to savings,
ty_h	rate of household h income tax,
β_{ch}	share of commodity c in the consumption of household h,
PQ_c	composite commodity c price,
QH_{ch}	quantity of consumption of commodity c by household h,
YH_h	household h income.

The production side is described by production function:

$$QA_\alpha = \alpha d_\alpha \cdot \prod_{f\in F} QF_{f\alpha}^{\alpha_{f\alpha}} \qquad \alpha\in A \tag{3}$$

and demand on production factors is derived from the production function such that the wage of factor equals the value of marginal product of that factor:

$$WF_f \cdot WFDIST_{f\alpha} = \frac{\alpha_{f\alpha}\cdot PVA_\alpha \cdot QA_\alpha}{QF_{f\alpha}} \qquad f\in F,\ \alpha\in A \tag{4}$$

where:

αd_α	production function efficiency parameter in activity α,
$\alpha_{f\alpha}$	value-added share for factor f in activity α,
PVA_α	value-added price of activity α,
QA_α	level of activity α,
$QF_{f\alpha}$	quantity demanded of factor f by activity α,
WF_f	average wage of factor f,
$WFDIST_{f\alpha}$	wage distortion factor for factor f in activity α.

The model embodies of two types of representative households which own two types of factors: labor and capital which are used in production of commodities. Two types of commodities are demanded by domestic consumers – households and foreign consumers – export. Besides that domestic consumers demand foreign commodity – import which are, with domestic commodity, also used as intermediate goods in production. Households receive income from the ownership of factors and the government and foreign transfers and they have to pay to the government sales tax and income tax.

The equations of the model are grouped in four blocks: price block, production and commodity block, institution block and system constraint block. In the price block, the exogeneity of import and export prices indicates that the country is small relative to the rest of the world. The export and import prices in domestic currency are simply expressed as multiples of export/import price in foreign currency and the exchange rate. The absorption, total domestic expenditures on commodities in domestic prices, is expressed as sum of the spending on domestic output and import. The domestic price together with the import price enters the composite commodity price. All domestic agents as households, government, producers and investors pay a composite price. The value of domestic output is set in producers prices as sum of domestic output sold in domestic country and the value of export for each commodity. It is distinguished among commodity prices, activity prices and value-added prices. The activity price is derived from suppliers' price and it depends on yield of commodity used per unit of particular activity. Value added per activity is given as the difference between the price of the activity and the price of the commodity used as intermediate good. Prices of factors tend to be distorted in the real world in the broad sense that they differ across activities.

In the production and commodity block it is assumed that all domestic consumers use composite commodities. Constant elasticity of substitution (CES) aggregation function captures the imperfect substitutability of domestic output sold in domestic country and import. Composite commodity is "produced" by domestic commodities and imported commodities and enters the production function as input. It means that demander preferences over import and domestic outputs are expressed as CES function which is called Armington function. Constant elasticity of transformation (CET) function is identical to the CES function, expect to negative elasticities of substitution. The difference between CES and CET functions is that arguments of CES function are inputs and arguments of CET function are outputs. In this sector equations defining optimal mix between imports and domestic output and export and domestic sales, respectively, are also included. The difference between import demand equation and export supply equation is that demanded quantity of import is in an inverse relationship to the price of import and supplied quantity of export is directly related to the price of export.

The institution block states that factors are mobile across activities, available in fixed supplies and demanded by producers at market-clearing prices. On the basis of fixed shares, factor incomes are passed on to the households, providing them with their only income besides the transfers from the sector of government. The composite commodity is therefore demanded by the households at market-clearing prices. Thus it is assumed for the factor market that each activity pays a wage expressed as the product of a wage variable and a distortion factor. In each factor market, adjustments in the average wage clear the market. The investment demand is simply given as base-year investment adjusted by investment adjustments factor. Two types of taxes are implemented to the model: income tax and commodity tax, which make up the revenue of the sector of the government. On the other side the expenditures of the government consist of government demand on commodities and transfers to the households. The corresponding equations are for the income of the government:

$$YG = \sum_{h \in H} ty_h \cdot YH_h + EXR \cdot tr_{gov,row} + \sum_{c \in C} tq_c \cdot \left[PD_c \cdot QD_c + \left(PM_c \cdot QM_c \right) \right] \tag{5}$$

and for the expenditures of the government:

$$EG = \sum_{h \in H} tr_{h,gov} + \sum_{c \in C} PQ_c \cdot qg_c \tag{6}$$

where:
- EG government expenditure,
- EXR foreign exchange rate,
- PD_c domestic price of domestic output,
- PM_c import price,
- PQ_c composite commodity price,
- QD_c quantity of domestic output sold domestically,
- qg_c government commodity demand,
- QM_c quantity of imports,
- tq_c sales tax rate,
- $tr_{ii'}$ transfer from institution i' to institution i,
- ty_h rate of household income tax,
- YG government revenue,
- YH_h household income.

The system constraints block contains the set of equilibrium conditions that is functionally dependent for the economic system without being considered by its individual agents. The "micro constraints" consider individual factor markets and commodity markets and it is assumed (with few exceptions for labor, export and import) that flexible prices clear the markets. The “macro constraints” apply to the government, savings-investments balance and the rest of the world. The equilibrium for the government (government closure) is satisfied by its savings and the investment value adjusts to changes in the value of total savings. The model enables two types of closures for the rest of the world sector: the flexible exchange rate or flexible rest of the world savings. In this model the flexible exchange rate ensures the rest of the world closure. The equilibrium condition for factors market assumes that unemployment is fixed with activity specific real wage for labor and fixed use of capital for each activity. The equilibrium condition on the composite commodity market imposes equality of the demand side represented by all types of domestic commodity use with the supply coming from the Armington function that aggregates import and domestic output sold domestically.

2.1. Modified model with scrapping subsidy

Implementing the scrapping subsidy to the model, equations applied to households demand on automotive commodity (2) and government expenditures (6) were modified, the producer price was lowered and two new parameters related to the scrapping subsidy were added to the model:

$$QH_{ch} = \frac{\beta_{ch} \cdot (1 - mps_h) \cdot (1 - ty_h) \cdot YH_h}{PQ_c \cdot (1 - s)}, \qquad c \in S - C,\ h \in H \tag{7}$$

$$EG = \sum_{h \in H} tr_{h,gov} + \sum_{c \in C} PQ_c \cdot qg_c + S \tag{8}$$

where: s parameter lowering the price of automotive commodity faced by households,
S the total value of subsidy.

3. Simulations

There were made two simulation runs on the model. In the first run the impact of the economic crisis was modelled and in the second run the scrapping subsidy program was evaluated.

3.1. The Economic Crisis

As it was noted above, the decline in automotive commodity production is estimated on 21%. The drop in production in other sectors may be approximated on level of 15%. Implementing the shocks to the mode leads to the following results: the initial shock results in more than 28% decrease in the value production of automotives and more than 23% fall in the production of other sectors. This massive dropdown was caused partly by the initial shock, partly by multiplicative effect, and partly fall of the domes-tic price of automotives by 11.55% and other commodities by 9.9%. All prices in economy (domestic, export, import) fall resulting in CPI drop by more than 10%. The GDP falls by 6.3%. The welfare analysis made by Hicksian equivalence states that the economic situation of households measured by changes in their utility level and income is better since their income rose by more than 7.5% and the prices they face fell. The impact of the crisis represented by dropdown in production is hence ambiguous.

3.2. The Scrapping Subsidy Program

This program was introduced to the model by lowering the price of automotive commodity the households face by 12.5%. This number is an average percentage subsidy for households fulfilling the conditions when buy a new car. Another change was lowering the producer price of automotives by 4% and raising the government expenditures by the total amount of subsidy, which is €55.3 million. The new equilibrium condition implemented to the model is such that the total supply of automotive commodity equals the total demand plus scrapping subsidy. As a result the total production in automotive sector increased by 0.58%, while the production in other sectors drop by 1.75%. The value consumption of cars by households increased by 4.44%, the value of consumption of other commodities decreased. Domestic price of automotives rose by 4.44%, considering the subsidy the price in fact decreased by 3.91%. Also the price of other aggregated commodity decreased by 2.13%. The average price level measured by CPI decreased by 1.53% what is the reason for decreased value of GDP. The government deepens its debt, partly because the small amount of government income in form of duty stamps and registration fee of new cars was not considered in the model. The total export value in automotives decreased although the export price decreased by 0.79%. The import value in automotives increased while value of import of other commodities dropped. This fact is induced by high domestic consumption in automotives.

The income of both groups of households decreased due to factor prices distortion. Using Hicksian equivalent variation it must be taken off 109 640 monetary units from households to reach their pre-change utility level. It concludes that households as representative agents in economy are better off with the scrapping subsidy program.

4. The Conclusions

The orientation of Slovak industry toward automotives is its strong and at the same time weak feature. In one side Slovakia as the leading world producer of automobiles per habitant takes advantage of the stabile engineering background developed by decades using its capital and human resources, but in the other side as small open economy exporting more than 90% of its automotive production sensitively responses to the economic situation in other countries. The estimation for 2009 is that production in automotive sector drop by 21%. This shock hurts all sector of Slovak economy and leads to decreasing GDP by more than 6%. The government intervenes against this inconvenient situation by joining the car scrapping subsidy program. It was shown that the subsidy of €55.3 million leads to higher domestic consumption constituted by household demand on domestic and imported cars but in other side decreased consumption in other sectors. It is necessary to keep in mind that the scrapping program incurs only short run effects and distorts the market relations.

References

[1] BENČÍK, M. *Konštrukcia experimentálneho modelu všeobecnej ekonomickej rovnováhy a jeho vlastnosti*. 2001, Národná banka Slovenska: Inštitút menových a finančných štúdií. Bratislava.

[2] KOTOV, M. *Modely všeobecnej ekonomickej rovnováhy*. Diplomová práca. 2002, Fakulta matematiky, fyziky a informatiky, Univerzita Komenského. Bratislava.

[3] KOTOV, M., PÁLENÍK, V. *Aplikácia CGE modelu na kvantifikáciu prínosov a nákladov vstupu SR do EÚ.* 2002, Ekonomický časopis 50.

[4] LÖFGREN, H. *Exercises in General Equilibrium Modeling Using GAMS.* Microcomputers in Policy Research 4a. 2003, International Food Policy Research Institute.

[5] MIŤKOVÁ, V. *General Equilibrium Model of Closed Economy: Case of Slovakia.* In: V. Miťková, N. Mikušová, A. Kavkler, D. Boršič (eds.) *Micro and Macro Models in Economics.* 2006, Bratislava.

[6] MIŤKOVÁ, V. *The Influence of Expansion in Automotives in Slovakia - An Applied CGE Study*. Proceedings of international scientific conference Firma a konkurenční prostředí 2009. Brno: MSD, 2009, pp. 355-363. ISBN 978-80-7392-084-5.

[7] PÁLENÍK, V., ĎURAŠ, J., HRIVNÁKOVÁ, J., KVETAN, V. *Prognóza efektov hospodárskej stra-tégie CGE modelom*. Ústav slovenskej a svetovej ekonomiky SAV, 2004. EŠ 58.

[8] PÁLENÍK, V., KOTOV, M., KOTULIČ BUNTA, J. *Review of the Recent CGE Applications in the Slovak republic's Economy.* Input – Output and General Equilibrium Data, Modeling and Policy Analysis. 2004, Brussels.

[9] Regulation of Ministry of Economy of Slovak Republic No. 2/2009

[10] Regulation of Ministry of Economy of Slovak Republic No. 3/2009

[11] ŠIKULA, M. ET AL. *Ekonomické a sociálne dôsledky integrácie Slovenskej republiky do Európskej únie.* 2003, Inštitút slovenskej a svetovej ekonomiky SAV. Bratislava.

[12] UHRÍK, J. *Správa prezidenta o výsledkoch a smerovaní ZAP SR*. Presented on Slovak Automotive Industry Association general assembly, Bratislava, April 2008, available on internet: http://www.zapsr.sk/global/files/file/Sprava_prezidenta_VH2008.pdf

TOOLS FOR A PENSION FUND INVESTMENT STRATEGY OPTIMIZATION

Vladimír Mlynarovič

Comenius University in Bratislava
Faculty of Social and Economic Sciences, Institutute of Economics,
mlynarovic@fses.uniba.sk

Abstract. The paper starting from a Slovak pension funds competition environment analysis, an econometric modeling of a key security index developing in dependence with USA macroeconomic development and technical analysis of historical development of key security indices presents an optimization of pension investment fund strategy. The strategy building process uses short term historical data that describe the development of the average income unit of the market competition on Slovak Pension funds market, the long term historical data on selected regional and sectoral indices, bonds and money market tools and long – term historical data on representative subsets of individual assets for selected indices. At the first phase the methodology applies replication procedures to construct the short – term benchmark of average pension unit, at the second phase the same procedures are used to optimize the investment strategy that outperform the benchmark using optimized individual investment strategies within stated risk budget. An optimization of individual strategies is based on a momentum dynamization of Markowitz approach, omega function and outranking procedures. At the third phase the indices are replicated through individual assets.
Keywords. Benchmark, replication procedures, CVaR, lower semi absolute deviation, momentum, omega function

1. Introduction

The subject of the paper is an investment strategy of a pension fund that is running in the so called second pillar of the Slovak pension system. The process of a pension fund strategy optimization uses supporting tools that provide views on the market competition and investment environment and its assumed development and main tools for the optimization of the strategy with a goal to outperform a benchmark.

At the first the applied approach aims to analyze the market of the corresponding pension funds as a whole, to rank funds in the multiple criteria space, and to derive conclusions about investment efficiency strategies in the selected return-risk spaces. Initially the problem is formulated and solved as a multiple criteria decision one, and Promethee methodology is used for outranking the pension funds in a dynamic context. Applications of modern portfolio theory are then used to approximate efficient frontiers and allocate individual funds. The Markowitz [4] portfolio selection model and its modification suggested by Black and Litterman [1] are used to model efficient frontiers, where the risk is measured by standard deviations, lover semi standard deviations, lower semi absolute deviations or conditional value at risk. The result of the second approach is illustrated in Figure 1, where together with the efficient frontier and individual positions of the Slovak growth funds one can see positions of the market competitions that are computed as simple or weighted averages of individual funds.

For the investment environment modeling and analysis an econometric approach and technical analysis are applied. Technical analysis uses the developed decision support system in Excel environment [9] that applies such indicators as moving averages, relative strength index, stochastics, Bollinger bands, directional movement indicators, moving averages convergence-divergence analysis and their combinations to describe trends, identify divergences and optimize buy – sell signals for a selected security index, e.g. S&P 500. Optimization is realized from a viewpoint of a pension fund investment with a starting endowment and a possibility invest additional money in following time periods. System looks for such time periods for individual indicators that for a selected historical period generate buy – sell signals with maximal current value of the investments and expected buy - sell signal for simulated short term index scenarios.

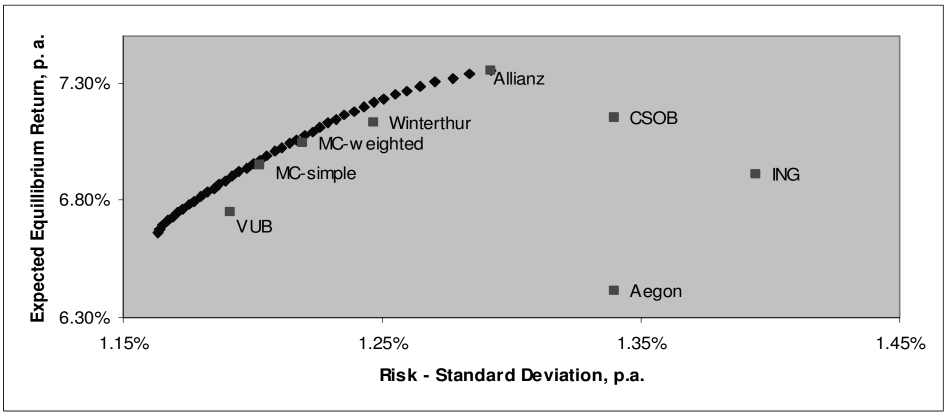


Figure 1: The Efficient Frontier of the Slovak Growth Pension Funds

As a supporting tool the single equation econometric models for key world security indices that explain index volatility through macroeconomic indicators development is used as well. A short term forecasting is then based on macroeconomic expectations of reputable financial institutions as it is illustrated in Figure 2.

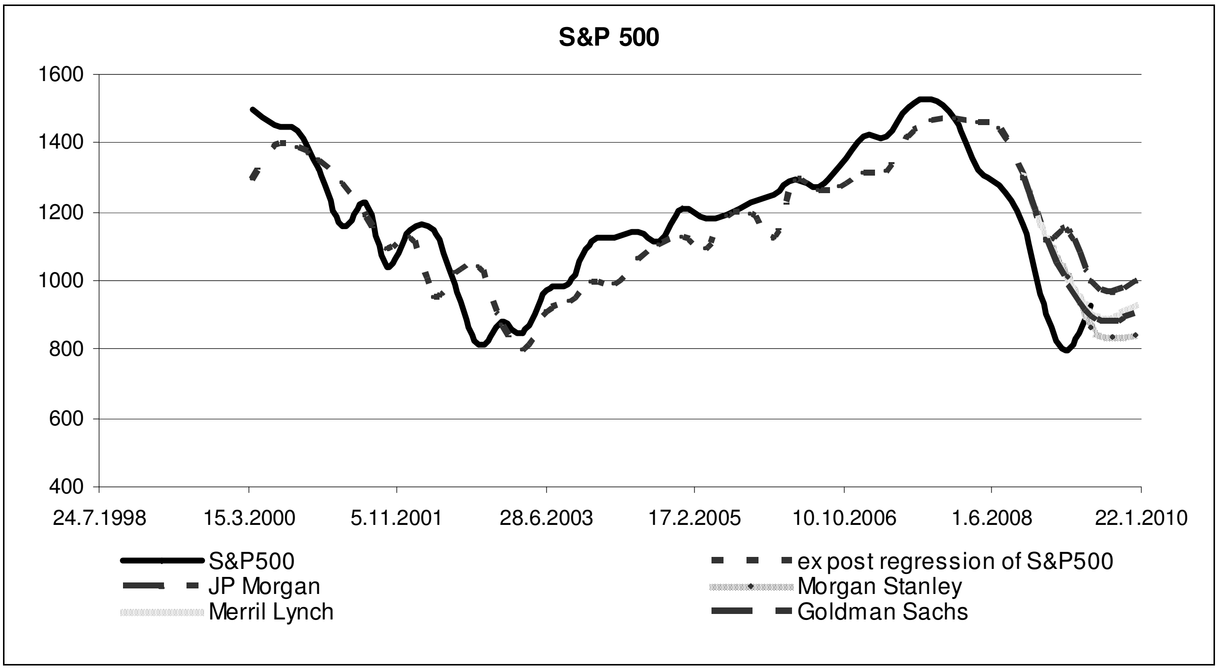


Figure 2: S&P 500 via USA macroeconomics

Views resulted from the market competition analysis and investment environment analysis enter into the process of a fund strategy optimization in the form of basic asset classes limits and decisions concerning positions (neutral, over weighted, under weighted) in security assets in comparison with the market competition. The pension fund strategy building process then uses short term historical data that describe the development of the average income unit of the market competition on Slovak Pension funds market, the long term historical data on selected regional and sectoral indices, bonds and money market tools and long – term historical data on representative subsets of individual assets for selected indices. At the first phase the methodology applies replication procedures to construct the short – term benchmark of average pension unit, at the second phase the same procedures are used to optimize the investment strategy that outperform the benchmark using optimized individual investment strategies within stated risk budget. An optimization of individual strategies is based on a momentum dynamization of Markowitz approach, omega function and outranking procedures. At the third phase the indices are replicated through individual assets. The basic tools are presented in two following parts of the paper.

2. Replication procedures

Let us assume that $\mathbf{R} = (R_1, R_2, \ldots, R_n)$ is distributed over finite set of points $\mathbf{r}_t = (r_{1t}, r_{2t}, \ldots, r_{nt})$, $t = 1, 2, \ldots, T$. These are obtained for example by historical data. Let p_t be the probability that $\mathbf{R}$ attains $\mathbf{r}_t$, $t = 1, 2, \ldots, T$, and W is the set of feasible portfolios, where

$$W = \left\{ \mathbf{w} \middle| \sum_{j=1}^{n} E_j w_j \geq \tau, \sum_{j=1}^{n} w_j = 1, \quad l \leq w_j \leq u, \quad j = 1, 2, \ldots, n \right\}$$

and τ, l and u are some appropriate constants.

Two of the replication procedures are based on lower partial risk measures, such as Conditional Value at Risk (CVaR) and Lower Semi Absolute deviation (LSAD) [3] and in the last approach the tracking error is minimized [5]. Let q_t, $t = 1, 2, ..., T$, is return of index, then Mean - CVaR model can be written in the form

$$\min \quad \alpha + \sum_{t=1}^{T} \frac{z_t}{(1-\beta)T}$$

subject to

$$z_t \geq -\sum_{j=1}^{n} r_{jt} w_j + q_t - \alpha, \; t = 1,2, \ldots T \,,$$

$$z_t \geq 0, \; t = 1, 2, \ldots T, \; \mathbf{w} \in W$$

and Mean – LSAD model in the form

$$\min \quad \alpha + \sum_{t=1}^{T} \frac{z_t}{T}$$

subject to

$$z_t \geq -\sum_{j=1}^{n} r_{jt} w_j + E_P, \; t = 1,2, \ldots T \,,$$

$$z_t \geq 0, \; t = 1, 2, \ldots T, \; \mathbf{w} \in W \,.$$

Let $\mathbf{d}$ is the vector of elements d_t $t = 1, 2, \ldots, T$, where

$$d_t = \sum_{j=1}^{n} r_{jt} w_j - q_t \,,$$

then in the last approach the following model is used:

$$\min \quad \mathrm{var}(\mathbf{d})$$

subject to

$$\mathbf{w} \in W \,.$$

3. Individual strategies optimization principles

Although the Markowitz`s portfolio selection model is usually used for long term strategic asset allocation, but no for short term allocations. Riberio a Loyes [10] just have shown that Markowitz optimization in a combination with so called momentum based tactical asset allocation can provide significant value added in the process of tactical investment strategy construction.

Most investor base tactical asset allocation decisions on consideration of value but lately also momentum, speed, in relative return across a wide set of asset classes is frequently used. In the most basic form, the strategy invests each 6 months in 5 asset classes, out of a possible 10, choosing those asset classes with the highest returns over the past 6 months. Momentum is an empirical phenomenon that contradicts market efficiency. In an efficient market, it should not be possible to build a profitable trading strategy without moving into riskier assets. But at now even well-known guardians of the efficient market hypothesis, such as Eugen Fama [6], recognize the possibility that momentum profits could be due to a market inefficiency. There could be other explanation, even base on a rational framework, and there are a lot of theories that try to explain momentum. Arguments of behavioral finance are considered as the most convincing and suggest as reason for momentum underreaction and overreaction to information.

In the under reaction case, investors are unable to process available information in a timely fashion. Thus, security prices undereact to new information. In that case, prices will slowly adjust in the direction on intrinsic value, producing short-term trends. For example, new information arrives that tell investors that fundamentals are better than expected. But investors are suspicious and do not adjust their expectations fully. Prices move less than fundamentals would imply. Consequently, prices will only slowly absorb the positive past information. The overreaction story is also based on other investors` cognitive biases. Investors learn that fundamentals are better than expected. Most of the investors adjust their expectations fully, but some of the investors extrapolate this positive news into the future. Prices increase as fundamentals would predict, but they continue the upward move beyond fundamentals because of extrapolation or even trend following behavior. Both behavioral biases will make prices deviate persistently from intrinsic value.

The performance of the dynamic Markowitz strategy, which is the result of Markowitz optimization and momentum based tactical allocation combination, e.g. in comparison with a naïve strategy based on the momentum strategy, comes from the following sources:

- momentum in asset class returns – as best performing asset classes in the recent past are also more likely to outperform in the near future,
- persistence (clustering) in asset class volatility and correlation,
- stability in total risk exposure – as we can combine asset classes to maintain a reasonably constant total volatility, thus introducing a volatility timing feature in the strategy.

Many of the difficulties one can encounter in performance measurement and attribution of financial assets or their portfolios are rooted in the over – simplification that mean and variance fully describe distribution of returns. It is a generally accepted stylized fact of empirical finance that few, if any, would now challenge that returns from investment are not distributed normally. Thus in addition to mean and variance higher moments are required for a complete description.

In the last decade one can see many approaches to the portfolio and securities return analyzes and explanations that try to capture an asymmetry in investments returns. Cascon, Keating and Shadwick [8], Keating, Shadwick [2] and Cascon, Keating, Shadwick [7] suggested so called *Omega* measure (function), which employs all the information contained within the return series and can be used for evaluation and ranking of portfolio assets.

The approach based on the mathematical technique, which facilitate the analysis of returns distributions. This approach enables partitioning returns into loss and gain above and below a return threshold and then considering the probability weighted ratio of returns above and below the partitioning. The formally the measure is defined in the form

$$\Omega(r) = \frac{\int_r^b (1 - F(x))dx}{\int_a^r F(x)dx},$$

where (a, b) is the interval of returns and F is the cumulative distribution of returns. It is in other words ratio of the two areas I_2 and I_1 shown in the Picture 1a for defined loss threshold r, where I_2 is the area above the function F on the right from r and I_1 is the area below the function F on the left from r. If we assume this ratio for all possible return thresholds, we will have the function, illustrated in the Fig. 3a, which is a characteristic one for assumed asset or portfolio.

The *Omega* function possesses many pleasing features [7] that can be intuitively and directly interpreted in financial terms. As it is illustrates in the Figure 3b, *Omega* takes the value 1 when r is the mean return. An important feature of the measure is that it is not plagued by sampling uncertainty, unlike standard statistical estimators, as it is calculate directly from the observed distribution and requires no estimates. This function is, in a rigorous mathematical sense, equivalent to the returns distribution itself, rather than simply being an approximation to it. It therefore omits none of the information in the distribution and it is as statistically significant as the return series itself.

As a result, *Omega* is a perspective tool for financial performance measurement where what is of interest to the practitioner is the risk and reward characteristics of the return series. This is the combined effect of all its moments, rather then the individual effects of any of them, which is precisely what *Omega* provides

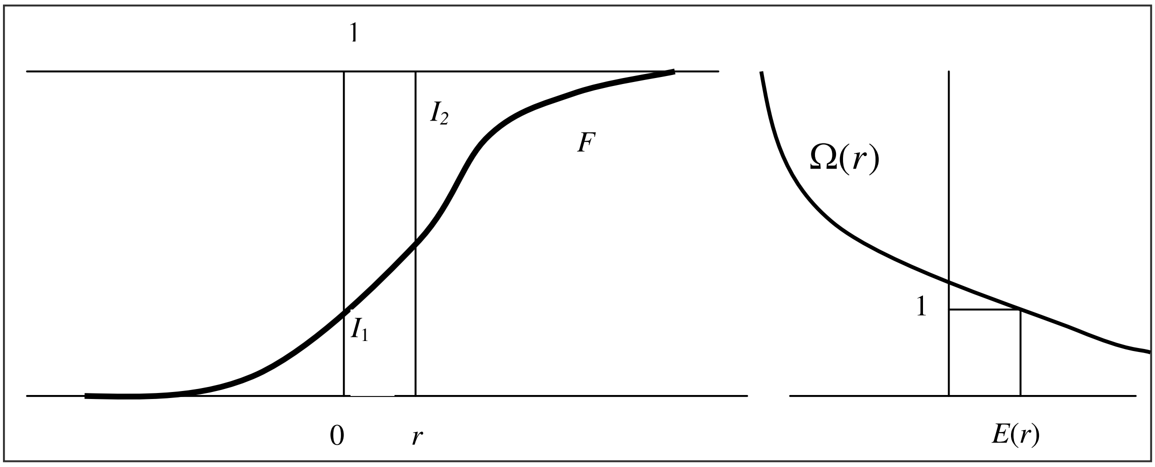


Figure 3a -b: Omega ratio and Omega function

If E_P is the required portfolio return, then as a further approach to portfolio selection can be used just the one, which for the stated performance threshold E_P looks for such portfolio that maximizes *Omega* function value. The problem can be formally written in the form

$$\max \quad \Omega(\mathbf{w}, E_P)$$

subject to

$$\mathbf{e}^T \mathbf{w} = 1,$$

$$\mathbf{w}^l \le \mathbf{w} \le \mathbf{w}^u,$$

$$\Omega(\mathbf{w}, E_P) = \frac{\int_{E_P}^{b} (1 - F(\mathbf{w}, x))dx}{\int_{a}^{E_P} F(\mathbf{w}, x)dx}.$$

4. Conclusions

The paper presents methodology for pension fund investment strategy modelling based on a synthesis such methodological tools as econometrics, technical analysis and optimisations. It shows how such tools can be combined to create a complete decision support system for investment decisions in Excel environment.

References

[1] BLACK, F. – R. LITTERMAN. *Global Portfolio Optimization.* Financial Analysts Journal, vol. 48, no. 5, 1990, 28 - 43

[2] KEATING, C. and W. F. SHADWICK. *An Introduction to Omega.* The Finance Development centre, 2002.

[3] KONNO, H. - WAKI , H.–A.YUUKI. *Portfolio Optimization under Lower Partial Risk Measures.* Asia/Pacific Financial Markets **9**, 2002, 127-140.

[4] MARKOWITZ, H. *Portfolio Selection*, Journal of Finance, Vol. 7, 1952, pp.77-91.

[5] SHARPE, W.F. *Asset Allocation: Management Style and Performance Measurement.* Journal of Portfolio Management, Winter 7 – 9, 1992.

[6] FAMA E. and K. FRENCH. *Multifactor Explanations of asset Pricing Anomalies*, Journal of Finance, January 1996

[7] CASCON, A., C., KEATING, and W. F. SHADWICK. *The Mathematics of Omega Measure*. The Finance Development centre, 2002

[8] KEATING, C. and W. F. SHADWICK: *A Universal Performance measure.* The Finance Development centre, 2002

[9] MLYNAROVIČ, V. *A spreadsheet decision support system for technical analysis.* Proceedings of the international academic conference on increasing competitiveness or regional, national and international markets, Ostrava, Ekonomická fakulta VŠB TU, 2007 s. 1-8 , Proceedings of the 25th International Conference.

[10] RIBERIO, R. and J. LOYES. *Markowitz in Tactical Asset Allocation.* Global Market Strategy, J.P. Morgan Securities Ltd. August 7, 2007

SMALL DSGE MODEL FOR THE CZECH ECONOMY*

Karel Musil**

Faculty of Economics and Administration, Masaryk University
Monetary and Statistics Department, Czech National Bank
karel.musil@cnb.cz

Abstract. The paper introduces a New Keynesian Dynamic Stochastic General Equilibrium (NK DSGE) model of the Czech economy. It is a closed economy model strictly based on microeconomic foundations. The model consists of three representative agents. Behavior of a representative household contains rigidities in consumption in a form of a habit formation. A production sector is divided between tradables and non-tradables. Price setting behavior of a representative firms in both sectors is described by the Calvo style resulting into a New Keynesian Phillips Curves (NKPCs). Monetary authority implements monetary policy according to a modified Taylor rule in the inflation targeting regime. The model contains extra rigidities connected to labor market (wages NKPC and labor adjustment costs). A Bayesian method is used for the estimation of the linearized model equations. A basic analysis of the estimated model is carried out through parameters description and impulse response functions. The model seems to offer a suitable approximation of the Czech economy behavior with respect to its structure.
Keywords. NK DSGE model, modified Taylor rule, inflation targeting, Bayesian estimation, wage and labor rigdity.

1 Introduction

The basic aim of the paper is to introduce a suitable conceptual model of the Czech economy with a simple structure, which is estimated. The structure and equations, describing behavior of agents in the model, are required to capture basic characteristics of the Czech economy. Moreover, the model tries to describe fundamental features connected to the macroeconomic behavior of labor market by using a nominal wage rigidity together with a simple real rigidity connected to labor.[1]

A structure of the employed model proceed from [3][2] and it is extended by [8]. The model is adjusted and adapted to the conditions of the Czech condition. A closed economy (CE) model is introduced to meet the requirement of a simple structure.

2 Conceptual Model

The employed model is a New Keynesian Stochastic General Equilibrium (NK DSGE) model strictly based on microeconomic foundations. It is a closed economy model; this could create a limitation in case of using it for the Czech economy, however, its simplicity helps to understand some basic features of a model framework and also structural characteristics of the economy. The model is composed of domestic representative agents - households, firms, and central monetary authority. There are two production sectors: tradables and non–tradables. The optimizing behavior of agents contains some rigidities. The model is an extension of [6] and [7].

2.1 Representative Households

A representative household solves an optimizing problem. It tries to maximize its utility function

$$E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[\frac{(CC_t \; A_{W,t})^{1-\sigma}}{1-\sigma} - \frac{NN_t^{1+\varphi}}{1+\varphi} \right] \right\} \tag{1}$$

* This paper is supported by MŠMT project Research centers 1MO524, funding of specific research at ESF MU, and by CNB.

** The opinions expressed are those of the author and do not necessarily reflect the views of the Czech National Bank.

1 For a more detailed analysis of real wage friction and labor market rigidities see e.g. [1], or [2].

2 A simpler structure of the model without tradables and non–tradables is introduced e.g. in [5].

subject to its budget constraint

$$P_tC_t + E_t\left(\frac{D_{t+1}}{R_t}\right) \leq D_t + W_tN_t \tag{2}$$

for $t = 0, 1, 2, \ldots$, where where β is a discount factor, σ is an inverse elasticity of intertemporal substitution, and φ is an inverse elasticity of labor supply. N_t denotes hours of labor, W_t is a nominal wage for labor supply, R_t is a gross nominal interest rate (for payments from a portfolio of assets), P_t is an overall consumption price index (CPI), D_t is a nominal pay–off on a portfolio held at t, and the whole term $E_t\left(\frac{D_{t+1}}{R_t}\right)$ expresses a price of a portfolio purchased at time t. Some rigidities connected to behavior of the representative household are connected to the effective consumption (CC_t) and the adjusted hours of labor (NN_t).

The effective consumption depends on a current and last period consumption (C_t and C_{t-1}), on a habit formation parameter h, and on parameter γ, which is a steady state growth rate of $A_{W,t}$ (for all t):

$$CC_t = C_t - h\gamma C_{t-1} \tag{3}$$

A long–run technology shock A_t is non–stationary and $z_t = \frac{A_t}{A_{t-1}}$. The evolution of z_t follows a stationary autoregressive process (in a log–linearization form)[3]:

$$\tilde{z}_t = \rho_z\tilde{z}_{t-1} + \epsilon_{z,t}, \tag{4}$$

for all t, with parameter ρ_z, where $\tilde{z}_t$ is a deviation from its steady state and the serially uncorrelated innovation $\epsilon_{z,t}$ has a standard normal distribution.

The adjusted hours of work NN_t evolves according to the following relationship:

$$NN_t = N_t - gN_{t-1} \tag{5}$$

for all t, where g is a parametr of adjustment.

The first order conditions (FOCs) derived from the optimizing behavior of the representative household are:

$$A_{W,t}\Lambda_t = CC_t^{-\tau} - h\gamma\beta E_t\left(\frac{A_{W,t}}{A_{W,t+1}}CC_{t+1}^{-\tau}\right) \tag{6}$$

$$\frac{1}{R_t} = \beta E_t\left(\frac{\Lambda_{t+1}}{\Lambda_t}\frac{P_t}{P_{t+1}}\right) \tag{7}$$

for $t = 0, 1, 2, \ldots$, where Λ_t is a Lagrangian multiplier at time t.

The total consumption C_t is composed of two groups of goods - tradable consumption $C_{T,t}$ and non–tradable consumption $C_{N,t}$ (for all t):

$$C_t \equiv \{(1-\lambda)^{\frac{1}{\nu}}C_{T,t}^{\frac{\nu-1}{\nu}} + \lambda^{\frac{1}{\nu}}C_{N,t}^{\frac{\nu-1}{\nu}}\}^{\frac{\nu}{\nu-1}}, \tag{8}$$

where λ is a share of non–tradable goods in total consumption and ν is an interteporal elasticity of substitution between tradable and non–tradable consumption.

Then it is possible to declare an overall CPI:

$$P_t \equiv \{(1-\lambda)P_{T,t}^{1-\nu} + \lambda P_{N,t}^{1-\nu}\}^{\frac{1}{1-\nu}}, \tag{9}$$

for $t = 0, 1, 2, \ldots$, where $P_{T,t}$ is a price of tradable and $P_{N,t}$ is a price of non–tradable consumption bundle.

The labor employed by tradable and non–tradable goods producers is supplied by a representative firm which hires the labor supplied by each household. The wage setting policy of households is subject to the Calvo style. Under the Calvo pricing only a fraction $(1-\theta_W)$ of households sets wages optimally. The rest fraction θ_W adjusts their prices with respect to the steady state wage growth rate $\tilde{\pi}_w$. A result of their behavior is a wage inflation New Keynesian Phillips Curve (NKPC):

$$\tilde{\pi}_{w,t} = \beta E_t\tilde{\pi}_{w,t+1} + \frac{(1-\beta\theta_W)(1-\theta_W)}{\theta_W}\tilde{mc}_{w,t} \tag{10}$$

for all t, where real marginal costs of labor follows the following rule:

$$MC_{w,t} = \frac{\beta g E_t NN_{t+1}^{\varphi} + NN_t^{\varphi}}{\Lambda_t\, W_t}. \tag{11}$$

[3] Henceforth all variables in small letters with tildes express a deviation of the original variables (in capital letters) from their steady states.

2.2 Representative Producers

A continuum of monopolistic competitive producers; each representative producer uses the following production function:

$$Y_{j,t}(i) = A_t A_{j,t} N_{j,t}(i), \tag{12}$$

for $t = 0, 1, 2, \ldots$, $j = T, N$ (tradable and non–tradable goods), and for the i–the producer, where $A_{j,t}$ is a stationary sector specific technology shock described by an AR(1) process for $j = T, N$ (in a log–linearized form):

$$\tilde{a}_{j,t} = \rho_j \tilde{a}_{j,t-1} + \epsilon_{j,t}, \tag{13}$$

for all t, where ρ_j is a parameter, $\tilde{a}_{j,t}$ is a deviation of $A_{j,t}$ from its steady state, and $\epsilon_{j,t}$ is a serially uncorrelated innovation with a standard normal distribution.

A solution of a cost minimizing problem of representative producers in both sectors with respect to their demand constraints follows a development of real marginal costs for $j = T, N$ and for $t = 0, 1, 2, \ldots$ according to:

$$MC_{j,t} = \frac{W_t}{RP_{j,t}\, A_{j,t}}, \tag{14}$$

where $RP_{T,t}$ and $RP_{N,t}$ is a relative price of tradables and non–tradables respectively.

Producers set their prices by following a Calvo's setting. In each period, only a fraction $(1 - \theta_j)$ of firms reoptimize (sets their their prices optimally). The rest fraction θ_j of firms in the j–th sector adjusts their prices with respect to the steady state inflation rate π.

The i–th representative firm in the j–th production sector optimizing its price of production at period t maximizes its revenue streams subject to the demand function in the proper sector.A result of this kind of behavior implies that prices are set as a mark–up over the marginal costs. The relationship between inflation and marginal costs after aggregation over all individual firms is a NKPC for $j = T, N$ and for all t:

$$\tilde{\pi}_{j,t} = \beta E_t \tilde{\pi}_{j,t+1} + \frac{(1 - \beta\theta_j)(1 - \theta_j)}{\theta_j} \tilde{mc}_{j,t}. \tag{15}$$

2.3 Central Monetary Authority

An implementation of monetary policy is approximated by a modified Taylor rule (in a gap form). The rule describes behavior of a central bank in an inflation targeting regime and can be understood as a growth rule for output. The bank is interested in a development of inflation and output gap. The output gap is used in a form of the adjusted output gap by a true value (not the steady state) growth rate of the output. The rule has the following form:

$$\tilde{r}_t = \rho_r \tilde{r}_{t-1} + (1 - \rho_r)\left[\phi_1 \tilde{\pi}_t + \phi_2(\Delta\tilde{y}_t + \tilde{z}_t)\right] + \epsilon_{R,t} \tag{16}$$

for all t, where $\tilde{\pi}_t$ is a deviation of inflation from an inflation target and the output gap $(\Delta\tilde{y}_t + \tilde{z}_t)$ is a deviation of adjusted output gap; ρ_r, ϕ_1, and ϕ_2 are parameters of the rule

2.4 Goods Market Clearing Condition

The condition expresses a fact that the total production is divided between domestic tradable and non–tradable consumption (for all t):

$$Y_t = C_{H,t} + C_{N,t}. \tag{17}$$

Consumption in the non–tradables (optimal allocation function for non–tradabel consumption) is a result of the optimizing behavior of a representative household and has the following form (for all t):

$$C_{N,t} = \lambda\, RP_{N,t}^{-\nu}\, C_t. \tag{18}$$

Similar equation can be derived for the tradables:

$$C_{T,t} = (1 - \lambda)\, RP_{T,t}^{-\nu}\, C_t. \tag{19}$$

A combination of three previous equations together yields the goods market clearing condition:

$$Y_t = \{(1 - \lambda)\, RP_{T,t}^{-\nu} + \lambda\, RP_{N,t}^{-\nu}\}\, C_t. \tag{20}$$

3 Model Equations and Solving of the Model

Log–linearizing and stationarizing of the model forms a linearized system. The system is composed of linearized and recalculated equations (3), (4), (5), (6), (7), (9), (10), (11), (13), (14), (15), (16), and (20), together with the following price and wage inflation identities: $\tilde{\pi}_t = \tilde{p}_t - \tilde{p}_{t-1}$, $\tilde{\pi}_{j,t} = \tilde{p}_{j,t} - \tilde{p}_{j,t-1}$, and $\tilde{\pi}_{w,t} = \tilde{w}_t - \tilde{w}_{t-1}$. The parameters of the model for an estimation are θ_T, θ_N, θ_W, λ, σ, h, g, ν, φ, ρ_r, ϕ_1, ϕ_2, ρ_z, ρ_T, ρ_N, and steady state value for $\tilde{r_t}$.

The log–linearized model can be rewritten into a linear rational expectations (LRE) model, which can be solved and expressed as a general equilibrium (GE) in a form of the following state space model:

$$s_t = A(\theta)s_{t-1} + B(\theta)\epsilon_t \tag{21}$$

$$y_t = C(\theta) + Ds_t \tag{22}$$

for all t, where s_t is a vector of states, ϵ_t is a vector of innovations of the exogenous processes, $A(\theta)$ and $B(\theta)$ are matrices of model's deep parameters, which are collected in a vector θ, representing the dynamic core of the model, y_t expresses a vector of observables and matrices $C(\theta)$ and D.

The Bayesian approach is employed for an estimation of the linearized model. According to the Bayesian formula all inference about the parameter vector is contained in the posterior density. The density is calculated with using a likelihood function and a prior density. To generate draws from the posterior distribution of the model parameters, it is used Markov Chain Monte Carlo (MCMC) method.[4]

4 Estimation Results

A data set for the Czech Republic from I.Q 1999 to IV.Q 2008 is used for the estimation. The quarterly data has the following structure:

- a deviation of the real GDP growth from a development of its balanced growth,
- a deviation of the CPI from a development of its dynamic equilibrium and from I.Q 2006 a deviation from a CPI level corresponding to the inflation target,
- a deviation of the level of tradable prices from a development of its dynamic equilibrium,
- domestic nominal interest rate – 3M Pribor.

With using the data and information included in prior densities, 250 000 draws of Markov Chain were generated. Parameter β was set in the following way: $\beta = \frac{1}{1+rr^{ss}/100}$, where rr^{ss} is a steady state interest rate. Table 1 presents the posterior estimates of parameters and 90 % probability intervals.

The estimation of parameter θ_T is 0.41, or in other words an average duration of price contracts of a representative firm in the tradable sector is more than 1.7 quarter. An average duration of the price contracts in the non–tradable sector is slightly shorter – about 1.5 quarter, and for the wage contracts is about 1.3 quarter. A share of the non–tradable goods in consumption (λ) corresponds to the prior information. An assumption of a lower possibility of substitution between tradables and non–tradables is reflected in the estimation by 0.60.

Behavior of a representative household is characterized by an elasticity of substitution – 0.81 and by a relatively high habit parameter in consumption (h) – 0.90. There is also a relatively high elasticity of labor supply about 6.3 which is connected to the relatively high degree of adjustment in hours of work (parameter g) about 0.45. The estimated value of the discount factor is 0.99.

A monetary policy implementation according to the adjusted Taylor rule is estimated as (for all t):

$$\tilde{R}_t = 0.92\, \tilde{R}_{t-1} + (1 - 0.92)\, [1.17\, \tilde{\pi}_t + 0.38\, (\varDelta\tilde{y}_t + \tilde{z}_t)] + \epsilon_{R,t}, \tag{23}$$

which indicates a relatively high stress on the interest rate smoothing. A standard deviation of the monetary policy shock (σ_{ϵ_R}) is low.

Exogenous processes are characterized with a relatively high persistence in a development of all technologies. This is in a combination with a high volatility of the tradable and non–tradable technology shock (σ_{ϵ_T} and σ_{ϵ_N}). On the other hand, the standard deviation of the growth rate of the long–run technology shock is relatively small ($\sigma_{\epsilon_z} = 1.43$).

[4] The extension of the paper forbids any detailed description of the Bayesian method. For general description of the Bayesian approach see e.g. [4].

Parameter	Prior Mean	Posterior Median	90 % Posterior Interval
θ_T	0.50	0.4117	$\langle 0.2520;\ 0.5559\rangle$
θ_N	0.50	0.3112	$\langle 0.1850;\ 0.4523\rangle$
θ_W	0.50	0.2407	$\langle 0.0888;\ 0.3920\rangle$
λ	0.48	0.5061	$\langle 0.4720;\ 0.5368\rangle$
σ	1.00	1.2320	$\langle 0.4173;\ 2.1029\rangle$
φ	1.00	0.1587	$\langle 0.0278;\ 0.2884\rangle$
h	0.70	0.9015	$\langle 0.8172;\ 0.9962\rangle$
g	0.50	0.4455	$\langle 0.1429;\ 0.7493\rangle$
ν	1.00	0.6046	$\langle 0.3644;\ 0.8500\rangle$
ρ_R	0.75	0.9216	$\langle 0.8987;\ 0.9455\rangle$
ϕ_1	1.30	1.1740	$\langle 1.0000;\ 1.3226\rangle$
ϕ_2	0.40	0.3839	$\langle 0.1756;\ 0.6028\rangle$
ρ_z	0.75	0.5581	$\langle 0.4273;\ 0.6896\rangle$
ρ_T	0.60	0.6636	$\langle 0.5221;\ 0.8010\rangle$
ρ_N	0.60	0.5721	$\langle 0.4210;\ 0.7136\rangle$
rr^{ss}	0.25	0.2512	$\langle 0.0931;\ 0.3958\rangle$
σ_{ϵ_T}	2.00	3.6979	$\langle 2.1569;\ 5.1619\rangle$
σ_{ϵ_N}	2.00	3.1330	$\langle 2.0063;\ 4.5332\rangle$
σ_{ϵ_R}	0.25	0.1098	$\langle 0.0848;\ 0.1327\rangle$
σ_{ϵ_z}	1.20	1.4326	$\langle 1.1368;\ 1.7091\rangle$

Table 1. Posterior Estimates of Parameters and Innovations

5 Analysis of Behavior

An analysis of behavior is conducted by impulse response functions as a reaction to a unit change of the exogenous variables of the model – an increase in technology in tradable and non–tradable sector, monetary policy shock, and growth rate of the world wide technology shock. A tradable technology shock is introduced in Figure 1.

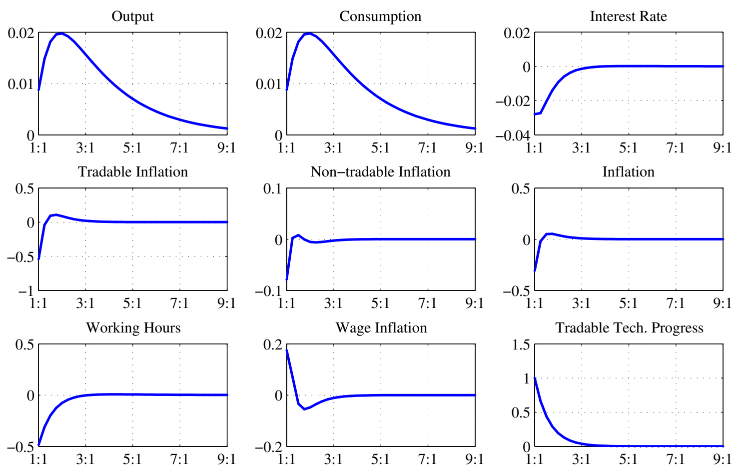


Fig. 1. Impulse Response Function from One Unit of Monetary Shock

A unit tradable technology shock increases domestic output and consumption as well. It also reduces tradable inflation. Because production sectors are interconnected, there is a drop in non–tradable inflation. This corresponds to the estimated values of the Calvo parameter in both sectors. Lower inflation causes a reduction in interest rate according to the modified Taylor rule. Lower domestic inflation creates an increase in real wages. On the other hand, hours of work are decreased, but an existence of a rigidity ensures a soften reaction.

6 Conclusion

The theoretical background, estimation of the model, and results from the impulse analysis offer an appropriate characteristics of behavior of the Czech economy with respect to the model structure. An extended structure of tradable and non–tradable sectors and an introduction of extra frictions in the modeled labor market can improve understanding of behavior of the economy. The analysis of behavior aimed at the tradable technology shock offers a valuable look at the model structure. The employed model seems to be a suitable starting point for a model with a more complex structure.

References

[1] Christoffel, K., Linzert, T.: *The Role of Real Wage Rigidity and Labor Market Frictions for Unemployment and Inflation Dynamics*. European Central Bank, WP 556, 2005.

[2] Gertler, M., Sala, L., Trigari, A.: An Estimated Monetary DSGE Model with Unemployment and Staggered Nominal Wage Bargaining. *Journal of Money, Credit and Banking*, Vol. 40, No. 8 (Dec.), pages 1713–1764, 2008.

[3] Lubik, T. A., Schorfheide, F.: *A Bayesian Look at New Open Economy Macroeconomics*. University of Pennsylvania, F. Schorfheide Web Site, 2005, access from internet <www.econ.upenn.edu/ ˜schorf/papers/nber-final.pdf> (cited to date 12. 5. 2007).

[4] Lubik, T. A., Schorfheide, F.: Testing for Indeterminacy: A Reply to Comments by A. Beyer and R. Farmer. *The American Economic Review*, Vol. 94, No. 1 (Mar.), pages 190 – 217, 2004.

[5] Monacelli, T.: *Monetary Policy in a Low Pass-through Environment*. European Central Bank, WP 227, 2003.

[6] Musil, K., Vašíček, O.: A Basic DSGE Model of the Czech Economy. *Mathematical Methods in Economics 2008*. pages 369–377, Liberec, Technická univerzita v Liberci, ISBN 978-80-7372-387-3, 2008.

[7] Musil, K. and Vašíček, O.: A Partial Dynamic Stochastic General Equilibrium Model of the Czech Economy. *Mathematical Methods in Economics 2007*, pages 197 – 204, Ostrava, VŠT – Technical University of Ostrava, ISBN 978-80-248-1457-5, 2006.

[8] Santacreu, A. M.: *Reaction Functions in a Small Open Economy: What Role for Non-traded inflation?*. Reserve Bank of New Zealand, Discuss Paper 04/2005, 2005, access from internet <www.rbnz.govt.nz/research/discusspapers/dp05 04.pdf> (cited to date 31. 11. 2007).

APPLICATION OF MULTICRITERIA DECISION MAKING FOR EVALUATION OF REGIONAL COMPETITIVENESS[1]

Jan Nevima, Jaroslav Ramík

VŠB – TU Ostrava, Ekonomická fakulta
nevima.jan@tiscali.cz

Extended abstract
Regional competitiveness is the source of national competitiveness. This paper deals with multi-criteria decision making methods for evaluation of the regional competitiveness. Specific coefficients reflect economic productivity of the region in form of factors of production inside of the region (effect of one-regional unit) and are revitalized by the capacity of real employment in the region. Particularly, we deal with the coefficient of effective disposability, effectiveness of economical development, effectiveness of investments, effectiveness of revenues and effectiveness of construction works. These coefficients are evaluated in the years 2000 and 2006. The technology for evaluation of regional competitiveness is based on application of two methods of multi-criteria decision making. The first one is the classical Analytic Hierarchy Process (AHP) where relevance of criteria's significance is determined by the method of Ivanovic deviation. The second method - FVK is a multiplicative version of AHP. The results of both methods are compared with the simple averaging. On the basis of the multi-criteria techniques, we obtain a more detailed perspective of the competitiveness of the NUTS 3 regions in the Czech Republic.
Keyword. Competitiveness, region, multi-criteria decision making, Ivanovic deviation, specific coefficients, Analytic hierarchy process

1. The competitiveness of regions

The increasing significance of regions in concept of EU deserves more attention especially because of the economic efficiency of regions representing the basis of competitiveness of the country. Regional competitiveness becomes a subject of evaluation on both levels: qualitative and quantitative. The initial problem is to choose such criteria for evaluation of regional competitiveness which would become widely acceptable. Approaches to evaluation of regional competitiveness must solve a problem of non existence of any mainstream, respectively of unique methodology for monitoring and evaluating competitiveness. For this reason the evaluation of competitiveness is determined by the selection of the regional unit. This practice is especially typical for the EU where the NUTS nomenclature of territorial units for statistics is being used. Equally important significance has the length of the period, duration of cycle and especially the selection applicable factors which describe the position of the regions.
There also exist different methodologies which have higher aspiration level. One of these evaluation methods of regional competitiveness is the method according to Viturka [9]. However, this method is orientated to a long term horizon. The problem arises from the fact that we are very limited by the small number of the regional indicators in formulation of the model of regional competitiveness. Nevertheless, these indicators are continuously monitored for each unit. European system of accounts (ESA 1995) is seen as the best tool for monitoring and evaluation of competitiveness at the European level. This system is also the basic starting point in the national accounts and it includes regional accounts. Unfortunately, these accounts do not allow us for assessing the inter-regional transfer of factors of production. If we are looking for initial data, the most useful procedure is to find the smallest territorial unit for which it is possible to obtain comparable data. The smallest territorial unit which provides comparable regional information in the Czech Republic is the level of NUTS 3. At NUTS 3 level it is possible to use the concept of evaluation of the competitiveness of specific indicators as an alternative factor. This is an individual approach which allows us to use a limited number of regional indicators followed by evaluation and comparison of the competitive position of regions.

[1] This research is supported by the Czech Science Foundation GACR, No. 402/08/1015 - Macroeconomic Models of the Czech Economy and Economies of the other EU Countries.

2. Specific indicators as alternative method for evaluation of regional competitiveness

The non existence of a mainstream in evaluation of the competitiveness causes that it is not possible to use "universal" methodology for assessment of already reached degree of regional (non)competitiveness. An "alternative approach" for evaluation of regional competitiveness is to define the group of *specific economic indicators of efficiency*. The basic idea is to assess the internal sources of regional competitiveness in detail. The reason is that only by detailed analysis we can identify risk factors which prevent increasing of competitiveness of selected regions (Martin, [3]). The proposed concept of evaluation of the competitiveness through specific indicators shall be understood individually, see also [2]. The first specific *coefficient of employment* is constructed as

$$K_z = \frac{\frac{Z_0}{Z}}{\frac{S_0}{S}}, \tag{2.1}$$

Z_0 - number of employment in the region (NUTS 3);
Z - number of employment in the state (ČR);
S_0 - number of inhabitants in the region (NUTS 3);
S - number of inhabitants in the state (ČR).

Each specific coefficient compares a concrete level of the value in the region with respect to its total level in the Czech Republic (CR). Particularly, it takes into account the number of inhabitants in the region compared to the total number of inhabitants in the Czech Republic. Here, the coefficients could be used also for regions NUTS 2, however, we focus only on the level of NUTS 3.
The quarterly data are not available for the territorial units NUTS 3. Instead, we use the valuation of the macro-regional aggregates at current prices. Generally we can write the coefficient (K_x) as follows:

$$K_x = \frac{\frac{P_x}{P}}{\frac{S_x}{S}}, \tag{2.2}$$

P_x - macro-regional indicator for regions NUTS 3;
P - value of macro-regional indicator for all the CR;
S_x - number of inhabitants in the region NUTS 3;
S - number of inhabitants in the country (CR).

Coefficient of efficiency of disposability results from the comparison of the coefficient of disposability (KD) and the coefficient of employment (K_z):

$$KED = \frac{KD}{K_z}. \tag{2.3}$$

From the methodological point of view, the coefficient of efficient disposability is based on the net disposable income. The net disposable income represents the result of current revenues and expenditures (current transactions) by the primary and secondary distribution of incomes.

Coefficient of efficiency in development is compatible with the level of gross domestic product in the region and it also compares the coefficient of productivity (KP) with the coefficient of employment:

$$KER = \frac{KP}{K_z}. \tag{2.4}$$

From the methodological point of view, it is the gross domestic product which is reached in regions in current prices and its value derives from the degree of facility of production factors in the region.

Coefficient of efficiency of investment construction is derived from the gross fixed capital formation which represents the investment activity of enterprises. We compare this coefficient of efficiency of investment construction (KV) with the coefficient of employment:

$$KIV = \frac{KV}{K_z}. \tag{2.5}$$

Coefficient of efficiency of revenues considers also sales of own products and services of industrial nature. It compares the coefficient of sales (*KT*) with the coefficient of employment:

$$KET = \frac{KT}{K_z} \tag{2.6}$$

Coefficient of efficiency of building works includes works according to supply contracts. It compares the coefficient of works (*KS*) with the coefficient of employment:

$$KSP = \frac{KS}{K_z} \tag{2.7}$$

Finally, the 5 specific coefficients of efficiency described above were calculated for the year 2000 and 2006 (see Table 1).

region / spec. coef.	*KED 2000*	*KER 2000*	*KSP 2000*	*KIV 2000*	*KET 2000*	*KED 2006*	*KER 2006*	*KSP 2006*	*KIV 2006*	*KET 2006*
Praha (PHA)	1.184	1.774	2.560	1.853	1.006	1.209	1.869	2.768	2.156	0.888
Středočeský (STČ)	1.030	0.927	0.494	0.983	1.693	1.028	0.877	0.537	0.884	1.588
Ústecký (ULK)	0.937	0.885	1.114	0.987	0.857	0.948	0.880	0.711	0.753	0.620
Jihočeský (JHČ)	0.954	0.895	0.639	1.162	0.980	0.971	0.905	0.549	1.126	1.163
Plzeňský (PLK)	0.908	0.779	0.450	0.567	0.617	0.889	0.764	0.348	0.768	0.513
Moravskoslezský (MSK)	1.000	0.891	0.755	0.791	1.579	0.947	0.860	0.761	0.769	1.244
Jihomoravský (JHM)	0.939	0.876	0.637	0.870	0.982	0.941	0.861	0.630	0.687	1.111
Liberecký (LBK)	0.968	0.912	0.646	0.890	0.791	0.955	0.855	0.521	0.664	0.667
Zlínský (ZLK)	0.941	0.874	0.912	0.739	0.773	0.946	0.819	0.671	0.611	1.338
Pardubický (PAK)	0.902	0.829	0.686	0.909	0.766	0.951	0.844	0.618	0.605	0.906
Královehradecký (HKK)	0.976	0.934	1.105	0.895	0.551	0.991	0.937	1.562	0.920	0.586
Vysočina (VYS)	0.976	0.850	0.763	0.818	0.637	0.932	0.764	0.564	0.920	0.665
Olomoucký (OLK)	0.963	0.858	0.971	0.755	0.742	0.955	0.819	0.724	0.730	0.894
Karlovarský (KVK)	1.006	0.867	0.658	0.796	1.208	0.994	0.940	0.672	0.954	1.256

Table 1: Specific coefficients in 2000 and 2006

Moreover, thanks to all 5 coefficients it was possible to rank the Czech NUTS 3 regions. We started with the classical multi-criteria method by averaging the specific coefficients (i.e. criteria) for individual regions. The following Table 2 shows the result – the ranks of regions separately for the year 2000 and 2006. The same results can be found also in Table 3, column 2 and 3.

region rank/year	2000	2006	region rank/year	2000	2006
1.	PHA	PHA	8.	LBK	LBK
2.	STČ	JHM	9.	ZLK	ZLK
3.	ULK	STČ	10.	PAK	VYS
4.	JHČ	MSK	11.	HKK	JHČ
5.	PLK	PLK	12.	VYS	OLK
6.	MSK	ULK	13.	OLK	HKK
7.	JHM	PAK	14.	KVK	KVK

Table 2: Rank of regions - average of 5 specific coefficients

Here, regional competitiveness (see Nevima, [4]) means an application of the regional own production factors which are available to promote an increase of regional economic growth.

3. Ivanovic deviation and its applications for evaluation of regional competitiveness

In the previous chapter we presented specific coefficients of efficiency. Here, they become the basis for further analysis. Another technique of evaluation of regional competitiveness is Ivanovic deviation (ID), see [8]. This method belongs to the techniques of multi-criteria decision-making and its purpose is to assess the ranks of the regions, too. In comparison with simple averaging, here, we have to take into account the importance and mutual dependence of the decision-making criteria (i.e. 5 specific coefficients already mentioned). First, the criteria (i.e. specific coefficients) should be ranked according to their relative importance. This ranking is done by an expert evaluation. Here, *KER* is the most important coefficient as it reflects total economic efficiency of the region and it also includes the level of production. The second most important criterion is *KED*. The reason is that the disposable income is another source for the household consumption and transfer savings to investment. *KIV* is the third in the row as the gross fixed capital formation is an indicator of connection of expenditures for creation of the fixed assets. These assets are also included in the regional production. *KET* represents another criterion. *KET* could be interpreted as the result of realized production so it is not a source of its value but it is an indicator of its use. *KSP* is the last (but not the least) criterion in this rank.
In the following step, the weight of each criterion is based on its relative importance - ranking and it takes into account correlation coefficients with the previous (i.e. more important) criteria. Then the weighted distance of the current variant to the ideal (fictitious) one is calculated as follows, see [8]:

$$I_j = \frac{\left|x_1^f - x_{1j}\right|}{s_1} + \sum_{i=2}^{n} \frac{\left|x_i^f - x_{ij}\right|}{s_i} \prod_{k=1}^{i-1} (1 - |r_{ki}|), \tag{3.1}$$

x_i^f – value of i-th criterion of ideal (fictitious) variant (i.e. region),

x_{ij} – value of i-th criterion j-th variant,

r_{ki} – correlation coefficient i-th a k-th criterion (i.e. specific coefficient),

s_i – standard deviation i-th criterion calculates:

$$s_i = \sqrt{\frac{1}{m}\sum_{j=1}^{m} \left(x_i^j\right)^2 - \left(\bar{x}_i\right)^2} \tag{3.2}$$

m – total value of variants (m=14),
n – total number of criterions (n=5).

The rank of 14 NUTS 3 regions in the year 2000 and 2006 calculated by ID method can be seen in Table 3, columns 4 and 5 below.

The approach which is based on the application of Ivanovic deviation seems to be more relevant comparing to the results of the method of simple averaging. As we know the importance of the criteria and correlations (i.e. dependences) among the criteria we are able to determine the "distance" to the ideal region in a more realistic way. Then the final rank of regions corresponds to the different economic importance of individual criteria (i.e. specific coefficient of efficiency). Thanks to this we can consider the final rank as another contribution of this alternative approach to evaluation of regional competitiveness of the NUTS 3 regions in the Czech Republic.

4. FVK method

In this section we deal with the same problem by another alternative method. The AHP method was published already in 1980s, see [7], recently, it is considered as the "classical" methodology. On the other hand, FVK is a newly created tool extending application possibilities of the classical AHP, see [5,6]. Here, we compare and discuss the results obtained by this method with the previous Ivanovic deviation method (ID).

Comparing ID and FVK methods there are some significant differences:

- In classical AHP the weights w_k are calculated from the pair-wise comparison matrix $S = \{s_{ij}\}$ by the principal eigenvector method, see [7], whereas in FVK the weights w_k are calculated from the pair-wise comparison matrix S through the geometric mean as:

$$w_k = \frac{\left(\prod_{j=1}^{n} s_{kj}\right)^{1/n}}{\sum_{i=1}^{n}\left(\prod_{j=1}^{n} s_{ij}\right)^{1/n}} \tag{4.1}$$

where k = 1,2,…,5. The elements of the pair-wise comparison matrix S are evaluated by expert pair-wise comparisons of the relative importance of the criteria – i.e. the specific coefficients.

- The total evaluation J_i of every variant (i.e. region) i = 1,2,...,14 is calculated as the weighted average:

$$J_i = \sum_{k=1}^{n} w_k a_{ik} \tag{4.2}$$

where a_{ik} is the normalized value of the k-th specific coefficient for the i-th region described in Section 2.

- FVK method reduces some theoretical disadvantages of the method of principal eigenvector used in AHP, e.g. the rank reversal problem.

All results presented in Table 3, column 6 and 7, have been calculated by software tool named FVK. This SW has been created as an add-in of MS Excel 2003 within the GACR project No. 402060431, see [5]

Table 3, columns 6 and 7 shows the results of ranking the regions calculated by FVK method.

1. rank /method	2. average 5 coeff. 2000	3. average 5 coeff. 2006	4. Ivanovic deviation 2000	5. Ivanovic deviation 2006	6. FVK 2000	7. FVK 2006
1.	PHA	PHA	PHA	PHA	PHA	PHA
2.	STČ	JHM	STČ	STČ	STČ	JHM
3.	ULK	STČ	ULK	MSK	ULK	STČ
4.	JHČ	MSK	MSK	PAK	JHČ	MSK
5.	PLK	PLK	PLK	ULK	JHM	PLK
6.	MSK	ULK	LBK	PLK	PLK	ULK
7.	JHM	PAK	HKK	LBK	MSK	PAK
8.	LBK	LBK	JHČ	VYS	ZLK	LBK
9.	ZLK	ZLK	PAK	ZLK	PAK	ZLK
10.	PAK	VYS	ZLK	JHM	LBK	JHČ
11.	HKK	JHČ	JHM	HKK	HKK	VYS
12.	VYS	OLK	VYS	JHČ	OLK	OLK
13.	OLK	HKK	OLK	OLK	VYS	HKK
14.	KVK	KVK	KVK	KVK	KVK	KVK

Table 3: Final ranks of regions using selected methods

5. Conclusion

In this paper we have dealt with some approaches for evaluation of regional competitiveness in the Czech Republic. The final rank of regions has been presented on the basis of specific coefficients of efficiency. In Table 3 the final ranks of NUTS 3 regions in the Czech Republic applying selected methods: the simple averaging, Ivanovic deviation and FVK method is presented. It is clear that any of presented techniques does not lead to very different results. From this perspective the results are considered as robust. For example the rank of extreme cases – Praha (PHA) and Karlovarský (KVK) region remain unchanged. On the other hand each technique is unique so we can not say that some leads to more (or less) credible result than the others. Table 3 also shows that more similar rank of regions is reached in 2000/2006 when using arithmetic average and FVK method. Contrary, smaller similarity in rank is reached when applying Ivanovic deviation. Consequently, all approaches presented here could be considered as allowable alternative ways for evaluation of regional competitiveness in the Czech Republic.

The results presented here are also the part of outcomes of the project of the bilateral cooperation between the Czech Republic and the Slovak Republic under the project “Macroeconomic simulation with focus on currency transformation in the Czech Republic and Slovakia” (MEB 080853).

References

[1] BENEŠ, M. *Konkurenceschopnost a konkurenční výhoda*. Centrum výzkumu konkurenční schopnosti české ekonomiky [online]. Working paper 5/2006. [cit. 5. dubna 2008]. Dostupné na World Wide Web: http://www.econ.muni.cz

[2] FARKAŠOVÁ, E. *Analýza nerovnomerného vývoja regiónu Východné Slovensko*. Technická univerzita v Košiciach, Ekonomická fakulta, 2007. Košice, Slovensko.

[3] MARTIN, R. *A Study on the Factors of Regional Competitiveness*. A final Report for the European Commission – Directorate General Regional Policy [online]. Cambridge: University of Cambridge, 24 November 2003. [cit. 7. května 2008]. Dostupné na World Wide Web: http://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/3cr/competitiveness.pdf

[4] NEVIMA, J. *Hodnocení konkurenceschopnosti využitím vybraných kvantitativních metod na případu ČR*. Doktorská disertační práce na Ekonomické fakultě VŠB-TU Ostrava. 2008. Školitel: doc. Ing. Jana Hančlová, CSc.

[5] RAMÍK, J., PERZINA, R. *Moderní metody hodnocení a rozhodování.* Studia Oeconomica, Slezská univerzita v Opavě, OPF v Karviné, 2008. ISBN 978-80-7248-497-3.

[6] RAMÍK, J., PERZINA, R. A method for solving fuzzy MCDM problems with dependent criteria. In Proceedings of the Joint 4-th International Conference on Soft Computing and Intelligent Systems and 9-th International Symposium on advanced Intelligent Systems. Nagoya University, 2008, s. 1323-1328. ISSN 1880-3741.

[7] SAATY, T.L. Multicriteria decision making - the Analytic Hierarchy Process. RWS Publications, Pittsburgh, 1991.

[8] SCHINDLER, et al. Logistika - Soubor odborných příspěvků k metodologii a k aplikačním nástrojům. In Sborníku vědeckých prací VŠB - Technické univerzity Ostrava, řada ekonomická a hutnická. VŠB – TU Ostrava, 2001. ISBN 80-7078-915-8.

[9] VITURKA, M. *Konkurenceschopnost regionů a možnosti jejího hodnocení*. *In* Politická ekonomie, 2007, č. 5, s. 637 – 658. ISSN 0032-3233.

PRICE JUMP ANALYSIS ON PSE AND VISEGRAD INDICES*

Jan Novotny

CERGE-EI, Praha
jnovotny@cerge-ei.cz

Abstract. In this study, I employ high frequency 5 minute market data of close prices of the main indices from Prague, Warsaw, Budapest and Frankfurt and perform a detailed price jump analysis focusing on the tail behavior. The data spans June 2003 to the end of 2008. I use two definitions of price fluctuations, namely the price jump index and normalized returns. Using the first approach, I find that the tail part of the distribution of this quantity shows scale behavior. I perform a sensitivity test on the parameters used in this estimation such as the length of the moving average. Using this quantity, I test the different behavior of the high frequency returns between upward and downward movements as well as for a change in the scale behavior after the financial turmoil. To be able to compare it with the existing literature, I employ the second approach and repeat the basic analysis. To strengthen my analysis, I perform a robustness check of the same study using lower frequencies.
Keywords. Price jumps, PSE, scale behavior.

1 Introduction

The distribution of the stock market price fluctuations using high frequency data has been extensively analyzed, mainly for mature markets. Tail distributions are particularly interesting as they are directly connected with extreme events. The general understanding of the tails of price return distributions suggests that the price returns follow a distribution close to the inverse cubic one [1, 2]. The tail behavior of price jump distribution using high frequency data for a broad range of stock market indices has been studied in [3], where the authors used normalized returns. The tail distribution of the emerging markets was further studied in [4], where the authors focused on Chinese stock markets. They have shown that power-law scaling behavior is valid only for long term distributions, while for short term distributions the tail behavior is suppressed in an exponential manner. An alternative to normalized returns represents the so-called price jump index used in [5], which is the ratio of the absolute value of returns to its moving average. This quantity obeys the same distributional properties as normalized returns.

In this analysis I focus on stock market indices from the Czech Republic (PX), Hungary (BUX) and Poland (WIG20), and also on the index from Frankfurt (DAX), Germany, chosen because of its geographical closeness. In the first part, I use a price jump index to show a tail distribution for all four indices and for various frequencies as well as for the moving averages with various length. Further, I study the positive and negative jumps separately. Next, I test the change of the scaling properties during the beginning of the financial turmoil. Finally, I employ normalized returns to repeat my basic analysis of the tail behavior.

2 Data

I use data with 5 minute frequency for indices from Prague (PX50), Budapest (BUX), Warsaw (WIG20) and Frankfurt (DAX) spanning 06/03 to 12/08. Since the main purpose of this research is to study scaling of the indices with respect to the market dynamics, I have cut-off the very beginning and the very end of the trading day and thus defined the trading period as follows: i) PX: opening hours from 9:15 to 16:00 and trading period from 9:30 to 15:30; ii) DAX: opening hours from 9:00 to 17:30 and trading period from 9:30 to 17:00; iii) BUX: opening hours from 9:00 to 16:30 and trading period from 9:30 to 16:00; and, iv) WIG: opening hours from 9:00 to 16:20 and trading period from 9:30 to 16:00. I distinguish between two types of times, calendar time and trading time, where the latter one runs only during the trading periods, i.e., the last minute at the end of the trading period is followed by the first minute of the next trading period.

* Partial financial support from GAČR grant No. 402/08/1376 is gratefully acknowledged.

3 Price Jumps Analysis

In this analysis, I perform price jump analysis using two definitions of price jumps: price jump index and normalized returns.

3.1 Price Jump Index and Scale Behavior

The intuitive way to study price jumps and their distribution is to compare an absolute value of returns to the average of the same quantity [5]. I use the standard definition of returns as $r(t) = \log(R(t)/R(t-1))$, where $R(t)$ is the price level of an index. Hence, the price jump index $j_T(t)$ with the time window T at time t is defined as

$$j_T(t) = \frac{|r(t)|}{\langle |r(t)| \rangle_T} \ , \tag{1}$$

where $\langle |r(t)| \rangle_T$ is the standard moving average of the length T of time steps including a current value.

The distribution of the price jump index $j_T(t)$ should behave as $\propto s^{-\alpha_T}$, i.e.,

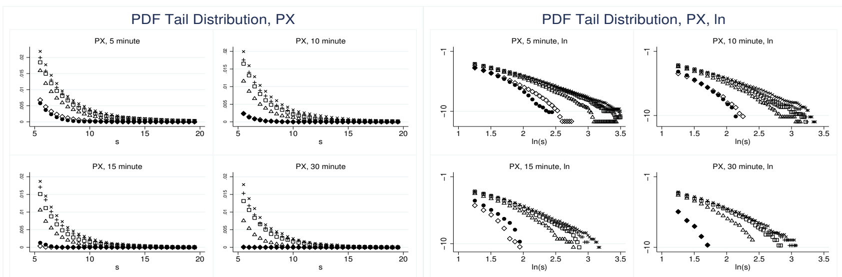
$$P(j > s; T) \sim s^{-\alpha_T} \ . \tag{2}$$

The value $\alpha_T < 3$ indicates the Levy-like behavior with an infinite volatility. As mentioned above, the high-frequency returns of the stock prices and indices should follow a process which is on average close to $\alpha \propto 4$, corresponding to an inverse cubic distribution for tails.

First, I will discuss the dependence of α on the length of time window. I take $T = 12$, 24, 100, 1000, 2000 and 5000 [1]. The time steps $T = 12$ and 24 are taken in calendar time to focus on a current day without any history. The remaining 4 time widows are taken in trading time.

The tail behavior of the probability distribution from eq. (2) for the PX index for all four frequencies and all six time windows is represented in the Figure 1a. This Figure clearly shows that the higher the time window T, the higher is the probability of the occurrence of high events. This result is in agreement with [4], where they explicitly tested for the exponential behavior of the tails.

Fig. 1. LHS: a) Tail part of the price jump index distribution for the PX index for all combinations of frequencies and time windows. RHS: b) The same distribution after performing natural logarithms of both sides of 2. Symbols used: $\bullet$ $T = 12$, $\diamondsuit$ $T = 24$, $\triangle$ $T = 100$, $\square$ $T = 1000$, $+$ $T = 2000$ and $\times$ $T = 5000$.



Further on, I focus on the long-term properties of the price distributions. To estimate a characteristic coefficient α_T of a power-law distribution, one can proceed at least in two ways. First, one can estimate it using a MLE directly from this distribution, as advocated in [1], or one can linearize eq. (2) by taking natural logarithms and use, as I have done, OLS.

The example of such distributions after linearization is depicted in the Figure 1b. The same Figures for the remaining indices can be found in [6].

Numerical results of the estimated α_T for PX can be found in Table 1. The algorithm to estimate α_T is the following: Using OLS, I have estimated α_T for each combination of frequency, T and for various tail intervals of $\ln(s)$. I have taken the result which has the highest R^2. Such an algorithm was simple and in agreement with visual observation of the linear region of tails. The results suggest that behavior significantly depends on the length of the time window. Generally, the longest time window supports the existence of extreme events (smaller α_T).

Using the longest time window, I have estimated α_T for all the indices and all frequencies. The result is summarized in Table 2. The results with the highest frequency suggest the close characteristic of PX,

[1] The length in minutes depends also on the used frequency.

Table 1. Estimated parameter α_T for index PX and for T =100, 1000, 2000 and 5000, and for 5, 10, 15 and 30 minute frequency. The value in the bracket is the standard deviation.

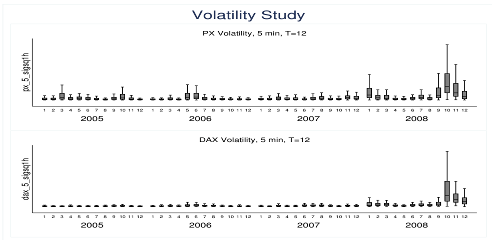
$\alpha_T(\sigma_\alpha)$	$T = 100$	$T = 1000$	$T = 2000$	$T = 5000$
PX, 5 min	4.156 (.055)	3.513 (.036)	3.943 (.038)	3.586 (.033)
PX, 10 min	4.791 (.131)	4.194 (.061)	4.330 (.072)	3.349 (.029)
PX, 15 min	5.626 (.157)	4.309 (.078)	4.674 (.164)	3.713 (.040)
PX, 30 min	5.670 (.260)	4.923 (.147)	4.668 (.108)	3.666 (.093)

Table 2. Estimated α_T for all four indices, all frequencies and $T = 5000$. The value in the bracket is the standard deviation.

$\alpha_T(\sigma_\alpha)$	5 min	10 min	15 min	30 min
PX	3.554 (.033)	3.313 (.031)	3.713 (.040)	4.167 (.085)
DAX	4.098 (.046)	3.773 (.030)	3.542 (.049)	2.973 (.032)
BUX	3.809 (.024)	3.403 (.027)	3.493 (.029)	3.465 (.059)
WIG	4.949 (.083)	4.825 (.096)	4.517 (.099)	3.927 (.063)

DAX and BUX, while WIG20 index has the most different behavior (compare standard deviations). This table also shows a change in the tail behavior on the frequency, which is most visible in the Levy-like behavior of the DAX index for 30 minute frequency.

Fig. 2. Volatility study for PX and DAX using 5 minute data and the shortest time window.



Financial Turmoil In August 2008, financial markets experienced a sudden change in market conditions. The change in the market conditions can be seen in Figure 2, which contains a volatility study for absolute returns for PX and DAX. Both indices confirm that "something" disturbing happened after August 2008, which significantly increased volatility on the markets. Hence, I estimate α month by month test for the significant change in its value after the turmoil started. Figures, including 95% CI can be found in [6]. However, results cannot statistically confirm any pattern. One of the possible explanations is insufficient amount of data, since one month represents roughly 1.5% of the data and since this study focuses on extreme events, it would need much stronger statistics.

Up/Down asymmetry I have tested up/down asymmetry by estimating α for the movements up and down separately, using all four indices, all frequencies and $T = 5000$. The results are in Table 3. At highest frequency, the hypothesis of the more extreme events in the downward direction is clearly confirmed for PX and BUX. On the other hand, WIG20 behaves quite differently, see the size of the standard deviation. The results also suggest the strong dependence on the used frequency, which can even change the manner of the asymmetry.

3.2 Normalized returns

Normalized returns represent the second possible definition of price fluctuations

$$r_T^n(t) = \frac{r(t) - \langle r(t)\rangle_T}{\sigma_T(t)},$$

where $\langle r(t)\rangle_T$ is defined as above but with true values of returns and $\sigma_T(t)$ is the standard deviation of the true values of returns.

In this case, I use $T = 5000$. The tail part of the normalized returns distribution should also follow the power-law behavior (2). The plot with the tail part of the distributions for all four indices and various frequencies can be found again in [6]. Numerical results for estimated α_T are presented in the Table 4.

Table 3. Up/down asymmetry for all four indices and all frequencies for $T = 5000$.

α (σ_α)	5 min	10 min	15 min	30 min
PX, +	3.654 (.047)	3.721 (.058)	3.475 (.040)	3.516 (.077)
PX, -	3.435 (.041)	3.065 (.031)	4.028 (.053)	4.008 (.082)
DAX, +	3.836 (.037)	3.704 (.049)	3.310 (.044)	2.720 (.044)
DAX, -	4.277 (.057)	3.792 (.043)	3.703 (.068)	3.233 (.037)
BUX, +	3.887 (.027)	3.388 (.033)	3.712 (.062)	3.712 (.078)
BUX, -	3.769 (.033)	3.365 (.046)	3.352 (.045)	3.359 (.063)
WIG, +	4.208 (.075)	4.103 (.096)	3.979 (.067)	3.277 (.081)
WIG, -	6.205 (.293)	4.811 (.157)	4.769 (.159)	4.076 (.127)

The results suggest that the scale behavior of the price jump index and normalized returns have different numerical values. However, they also show the same patterns, see, e. g., the Levy-like behavior of the DAX index for lowest frequency, which is confirmed for both definitions of price jumps, or the problems with the size of the standard deviations of the characteristic exponent estimates for the WIG index, which are much higher compared to other indices. Generally, the size of the standard deviation serves as an indicator for the accuracy of a given approach.

Table 4. Estimated coefficient α for the normalized returns (positve and negative sides) and $T = 5000$. Value in brackets means standard deviation.

$\alpha(\sigma_\alpha)$	5 minute	10 minute	15 minute	30 minute
PX +	3.659 (.066)	3.607 (.080)	3.651 (.038)	3.397 (.093)
PX -	3.277 (.070)	3.186 (.027)	3.932 (.105)	3.654 (.141)
DAX +	4.288 (.067)	4.025 (.036)	3.275 (.061)	2.699 (.115)
DAX -	4.064 (.095)	3.815 (.061)	3.830 (.088)	3.388 (.082)
BUX +	3.972 (.044)	3.636 (.060)	3.827 (.103)	4.144 (.138)
BUX -	4.060 (.058)	3.756 (.037)	3.689 (.072)	3.874 (.133)
WIG +	3.992 (.092)	4.316 (.098)	4.140 (.103)	3.472 (.079)
WIG -	5.153 (.110)	4.872 (.215)	5.636 (.234)	4.652 (.246)

4 Conclusion

I have performed both quantitative and qualitative analysis of the tail distribution of price fluctuations using high frequency data for three emerging markets (Visegrad indices: PX, BUX and WIG20) and for one mature market (DAX). The analysis confirmed a power-law behavior of tails and revealed its strong statistical dependence on the used frequency and also on the length of the reference window. I have used two different definitions of price fluctuations: the price jump index, and the normalized returns. I have also performed an analysis of the possible asymmetry between upward and downward movements. This test confirmed three of the four indices that the extreme movements down are more probable compared to movements up. I have also tested for possible different scale behavior before and after the financial turmoil beginning in August 2008. Such an analysis does not confirmed an intuitive hypothesis of the lower characteristic exponent after the financial turmoil, as was expected from the volatility behavior. Despite the existing literature and the work done here, the price jumps and its connection to economics reality still offers a field for future research.

References

1. Eryigit, M., Cukur, S., and Eryigit, R., Tail Distribution of Index Fluctuations in World Market, Physica A 388 (2009), 1879-1886
2. Stanley, H. E., and Mantegna, R. N., An Introduction to Econophysics, Cambridge University Press, 2000
3. Plerou, V. *et. al.*, Scaling of the Distribution of Price Fluctuations of Individual Companies, Phys. Rev. **E** Vol 60(6) (1999)
4. Jiang, J. *et. al.*, Empirical Study of Recent Chinese Stock Market, Physica A 388 (2009), 1893-1907
5. Joulin, A., *et al.*, Stock Price Jumps: News and Volume Play a Minor Role, arXiv: 0803.1769v1
6. http://home.cerge-ei.cz/jnovotny/PriceJumpAnalysisPSE/

AUCTIONS IN COMMUNICATION NETWORKS

Maciej Nowak[1], Petr Fiala[2], Václav Kořenář[2]

[1]University of Economics, Katowice
Faculty of Management, Department of Operations Research
[2]University of Economics, Prague
Faculty of Informatics and Statistics, Department of Econometrics
nowak@ae.katowice.pl, pfiala@vse.cz, korenar@vse.cz

Abstract. Communication networks have increased in scale and heterogeneity. There are multiple network owners and operators. It is useful to find economic mechanisms that result in efficient resource allocations among agents. Auctions for spectrum allocation are devoted to selling frequency spectrum to telecommunication companies. Several auction formats were discussed for realizing the spectrum auction objectives such as efficient outcome, preventing monopolization, and maximizing revenue. Bandwidth auction problems are characterized by sellers with slots of bandwidth and by buyers with values for bundles of slots. The allocation problem here is to assign combinations of bandwidth slots to buyers and match them with sellers so as to maximize the total surplus in the system. These problems lead to combinatorial auctions with multiple criteria. The paper presents mathematical models of auctions in communication networks.
Keywords. Communication network, spectrum allocation, bandwidth exchanges, combinatorial auctions, multi-criteria auctions

1. Introduction

Communication networks have increased in scale and heterogeneity. The demand for bandwidth in communication networks has been growing exponentially since the birth of worldwide networks and new applications, more and more bandwidth-needing, have been appearing. As a result, despite the efforts made to increase communication rates, the available capacities are often insufficient to satisfy all service requests. Situations of congestion occur frequently, meaning that some users' requests are rejected.

The pricing scheme for communications, based on a fixed charge independent of use, does not take into account the negative externalities among users (a user consuming bandwidth may prevent another request from being treated successfully), and thus constitutes an incentive to overuse the network. Designing new allocation and pricing schemes therefore appears as a solution for solving congestion problems, by inciting users to limit their consumption.

There are multiple network owners and operators. It is useful to find economic mechanisms that result in efficient resource allocations among agents. Auctions for spectrum allocation are devoted to selling frequency spectrum to telecommunication companies. Bandwidth auction problems are characterized by sellers with slots of bandwidth and by buyers with values for bundles of slots. The allocation problem here is to assign combinations of bandwidth slots to buyers and match them with sellers so as to maximize the total surplus in the system.

Selling frequency spectrum to telecommunication companies through on-line auctions was first attempted in New Zealand (1989) and in England (1990) (see [6]). Later, auctions were used for selling spectrum rights in Australia in 1993. These auctions failed to generate much revenue due to flaws in auction design. The paper [5] attributes the problems of the Dutch and Swiss 3G contests, in particular, to a lack of appreciation of the role of auction design in creating incentives for entry and discouraging collusion.

Auctions are important market mechanisms for the allocation of goods and services. An auction provides a mechanism for negotiation between buyers and sellers. In forward auctions a single seller sells resources to multiple buyers. In a reverse auctions, a single buyer attempts to source resources from multiple suppliers, as is common in procurement. Auctions with multiple buyers and sellers are called double auctions. Auctions with multiple buyers and sellers are becoming increasing popular in electronic commerce. It is well known that double auctions in which both sides submit demand or supply bids are much more efficient than several one-sided auctions combined.

The communication problems lead to combinatorial auctions. Combinatorial auctions are those auctions in which bidders can place bids on combinations of items, so called bundles. The advantage of combinatorial auctions is that the bidder can more fully express his preferences. This is particular important when items are complements. The auction designer also derives value from combinatorial auctions. Allowing bidders more fully to express preferences often leads to improved economic efficiency and greater auction revenues.

The paper presents mathematical models of auctions in communication networks. Several auction formats were discussed for realizing the spectrum auction criteria such as maximizing revenue, efficient outcome, and preventing monopolization. The criteria can be taken simultaneously. For multicriteria auctions we propose to use the Analytic Network Process (ANP) and a dynamic version the Dynamic Network Process (DNP).

2. Spectrum Allocation

Communication authority allocates spectrum licenses. The problem here is to achieve an efficient allocation of new spectrum licenses to communication companies. The mobility of clients leads to synergistic values across geographically consistent license areas. It is required that all the licenses be allocated at the same time as some companies might value certain combination of licenses more than individual licenses.

We consider a version of spectrum auctions, a forward combinatorial auction of indivisible items with one seller and multiple buyers. Let us suppose that the seller S (communication authority) offers a set R of r items (spectrum licenses), $j = 1, 2, \ldots, r$, to n potential buyers (communication companies) $B_1, B_2, \ldots, B_n$. Buyers are interested in buying the entire set or some subsets of R. Each buyer receives only one subset. The seller wants to maximize his revenue i.e. allocate the goods to the buyers with the highest bids.

Items are available in single units. A bid made by buyer B_i, $i = 1, 2, \ldots, n$, is defined as

$$b_i = \{C, p_i(C)\},$$

where

$C \subseteq R$, is a combination of items,

$p_i(C)$, is the offered price by buyer B_i for the combination of items C.

$x_i(C)$ is a bivalent variable specifying if the combination C is assigned to buyer B_i.

The forward auction can be formulated as follows

$$\sum_{i=1}^{n} \sum_{C \subseteq R} p_i(C)\, x_i(C) \rightarrow \max$$

subject to

$$\sum_{C \subseteq R} x_i(C) \le 1, \quad \forall\, i,\ i = 1, 2, \ldots, n, \qquad (1)$$

$$\sum_{i=1}^{n} \sum_{C \subseteq R} x_i(C) \le 1, \quad \forall\, j \in \square\, R,$$

$$x_i(C) \in \{0, 1\}, \ \forall\ C \subseteq R, \ \forall\, i,\ i = 1, 2, \ldots, n.$$

The objective function expresses the revenue for the communication authority. The first constraint ensures that no buyer receives more than one combination of items. The second constraint ensures that overlapping sets of items are never assigned. Approaches are proposed for solving of forward auctions (see [1]).

3. Bandwidth Exchanges

Slots of bandwidth are available of a fixed size and duration with public and private companies (sellers). Service providers or smaller companies (buyers) have values for bundles of slots. The allocation problem here is to assign combinations of bandwidth slots to buyers and match them with sellers so as to maximize the total surplus in the system. The total surplus is the total amount received from the buyers minus the total payments to be made to sellers (see [2], [4]). This problem leads to a double combinatorial auction.

For the double auction, the auctioneer is faced with the task of matching up a subset of the buyers with a subset of the sellers. The profit of the auctioneer is the difference between the prices paid by the buyers and the prices paid to the sellers. The objective is to maximize the total surplus in the system given the bids made by sellers and buyers. Constraints establish the same conditions as in spectrum allocation problem.

We present a double auction problem of indivisible items (slots of bandwidth) with multiple sellers and multiple buyers. Let us suppose that m potential sellers $S_1, S_2, ..., S_m$ offer a set R of r items, $j = 1, 2, …, r$, to n potential buyers $B_1, B_2, ..., B_n$.

A bid made by seller S_h, $h = 1, 2, …, m$, is defined as $b_h = \{C, c_h(C)\}$,

a bid made by buyer B_i, $i = 1, 2, …, n$, is defined as $b_i = \{C, v_i(C)\}$,

where
$C \subseteq R$, is a combination of items,
$c_h(C)$, is the offered price by seller S_h for the combination of items C,
$p_i(C)$, is the offered price by buyer B_i for the combination of items C.

Bivalent variables are introduced for model formulation:
$x_i(C)$ is a bivalent variable specifying if the combination C is assigned to buyer B_i,
$y_h(C)$ is a bivalent variable specifying if the combination C is bought from seller S_h.

$$\sum_{i=1}^{n} \sum_{C \subseteq R} p_i(C)\, x_i(C) - \sum_{h=1}^{m} \sum_{C \subseteq R} c_h(C)\, y_h(C) \rightarrow \max$$

subject to

$$\sum_{C \subseteq R} x_i(C) \le 1, \quad \forall i,\ i = 1, 2, …, n,$$

$$\sum_{C \subseteq R} y_h(C) \le 1, \quad \forall h,\ h = 1, 2, …, m, \qquad (2)$$

$$\sum_{i=1}^{n} \sum_{C \subseteq R} x_i(C) \le \sum_{h=1}^{m} \sum_{C \subseteq R} y_h(C), \quad \forall j \in R,$$

$$x_i(C) \in \{0, 1\},\ \forall\ C \subseteq R,\ \forall i,\ i = 1, 2, …, n,$$

$$y_h(C) \in \{0, 1\},\ \forall\ C \subseteq R,\ \forall h,\ h = 1, 2, …, m.$$

The objective function expresses the total surplus in the system. The first constraint ensures that no bidder receives more than one combination of items. The second constraint ensures that no seller sells more than one combination of items. The third constraint ensures for buyers to purchase a required item and that the item must be offered by sellers. Approaches are proposed for solving of double auctions (see [9]).

4. Multicriteria auctions

Several auction formats were discussed for realizing the spectrum auction criteria such as maximizing revenue, efficient outcome, preventing monopolization, and other criteria. The criteria can be taken simultaneously. For multicriteria auctions we propose to use the Analytic Network Process (ANP) and a dynamic version the Dynamic Network Process (DNP).

The key feature that makes combinatorial auctions most appealing is the ability for bidders to express complex preferences over bundles of items, involving complementarity and substitutability. Items are complements when a set of items has greater utility than the sum of the utilities for the individual items. Items are substitutes when a set of items has less utility than the sum of the utilities for the individual items.

Different elicitation algorithms may require different means of representing the information obtained by bidders. The paper [8] describes a general method for representing an incompletely specified valuation functions. A constraint network is a labeled directed graph consisting of one node for each bundle b representing the elicitor's knowledge of the preferences of a bidder. A directed edge (a, b) indicates that bundle a is preferred to bundle b. Figure 1 represents an example of a constraint network for bundles of three items (A,B,C).

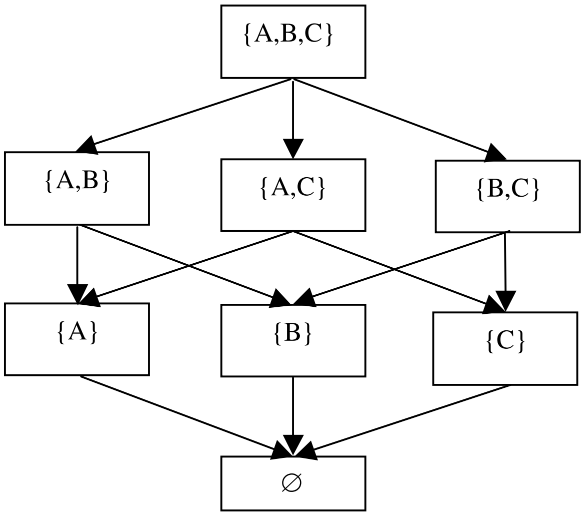


Fig. 1 Constraint network

The constraint network representation is useful conceptually, and can be represented explicitly for use in various elicitation algorithms. But its explicit representation is generally tractable only for small problems, since it contains 2^m nodes. For preference elicitation of bundles in a constraint network can be used Analytic Network Process. The ANP is the method that makes it possible to deal systematically with all kinds of dependence and feedback in the performance system (see [7]). The structure of the ANP model is described by clusters of elements connected by their dependence on one another. A cluster groups elements that share a set of attributes. At least one element in each of these clusters is connected to some element in another cluster. These connections indicate the flow of influence between the elements (see Figure 2).

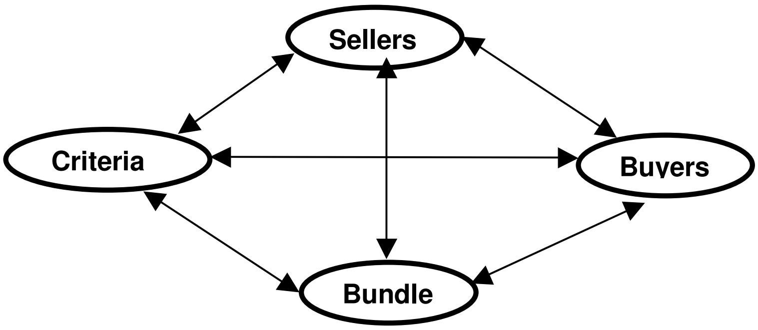


Fig. 2 Clusters and connections in multicriteria auctions

The clusters in multicriteria combinatorial auctions can be sellers, buyers, bundles of items, and evaluating criteria also. Paired comparisons are inputs for preference elicitation in combinatorial auctions. A supermatrix is a matrix of all elements by all elements. The weights from the paired comparisons are placed in the appropriate column of the supermatrix. The sum of each column corresponds to the number of comparison sets. The weights in the column corresponding to the cluster are multiplied by the weight of the cluster. Each column of the weighted supermatrix sums to one and the matrix is column stochastic. Its powers can stabilize after some iterations to limited supermatrix. The columns of each block of the matrix are identical in many cases, though not always, and we can read off the global priority of units.

We used the ANP software Super Decisions developed by Creative Decisions Foundation (CDF) for some experiments for testing the possibilities of the expression and evaluation of the multicriteria combinatorial auction models (Figure 3).

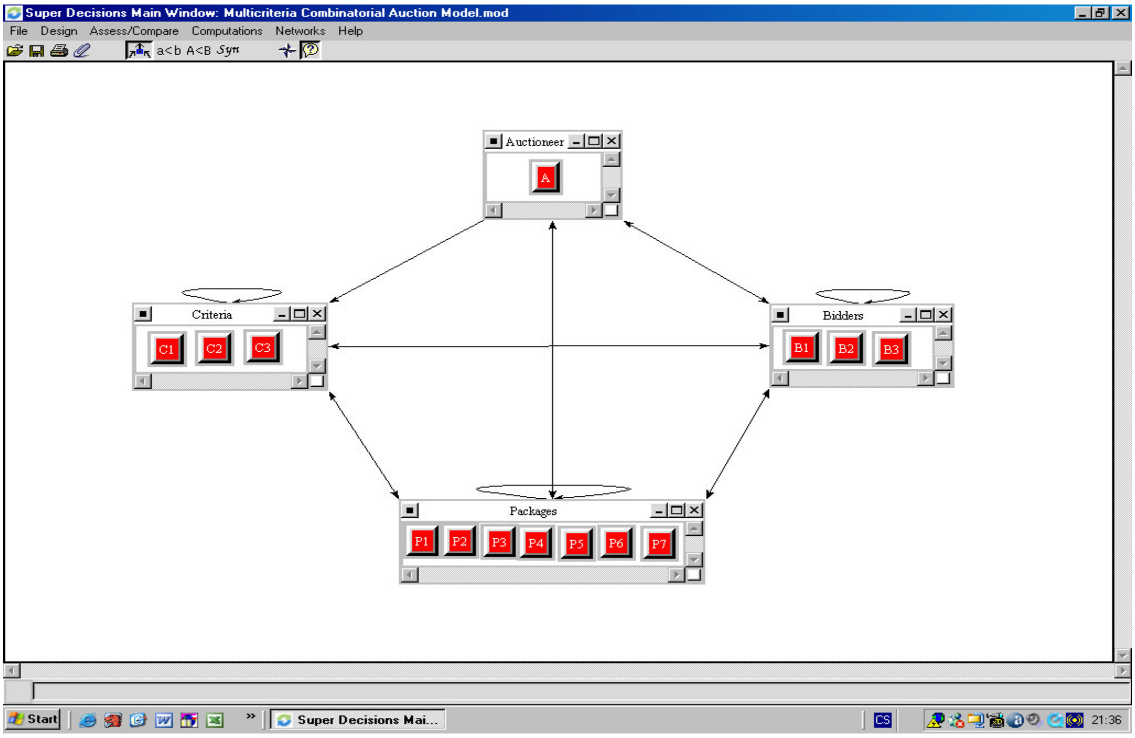


Fig. 3 Multicriteria Auction Model

The ANP have been static but for today's world analyzing is very important time dependent decision making. The DNP methods were introduced. There are two ways to study dynamic decisions: structural, by including scenarios, and functional by explicitly involving time in the judgment process. For the functional dynamics there are analytic or numerical solutions. The basic idea with the numerical approach is to obtain the time dependent principal eigenvector by simulation. The DNP seems to be the appropriate instrument for analyzing dynamic network effects (see [3]). The method is appropriate also for the specific features of multicriteria combinatorial auctions. The method computes time dependent weights for bundles of items or weights of bidders.

5. Conclusions

The paper is devoted to modeling auctions in communication networks. Modeling and solving the communication problems by combinatorial auctions is a promising for practical exploitations. Spectrum allocation and bandwidth exchange problems are formulated. The spectrum allocation problem is solved by a forward combinatorial auction. The bandwidth exchange problem is solved by a double combinatorial auction. The proposed combinatorial auction models give us an opportunity to design modifications of the auction. Several auction formats were discussed for multiple criteria. The approach is based on the ANP and a dynamic version the DNP. A flexible instrument is proposed for modeling and solving spectrum allocation and bandwidth exchange problems.

Acknowledgements

The research project was supported by Czech - Polish project MEB050834 and by Grant No. 402/07/0166 from the Grant Agency of Czech Republic „Combinatorial Auctions - Modeling and Analysis“.

References

[1] Cramton, P., Shoham, Y. and Steinberg, R. (eds.) (2006). *Combinatorial Auctions*. MIT Press, Cambridge.

[2] Eso, M., Kalagnanam, J., Ladanyi L., Li, Y. G. (2004). Winner determination in bandwidth exchanges. Technical report, IBM TJ Watson Research Center.

[3] Fiala, P. (2006). An ANP/DNP analysis of economic elements in today's world network economy. *Journal of Systems Science and Systems Engineering*, 15, 131–140.

[4] Jain, R., Varaiya, P. P. (2004). Combinatorial bandwidth exchange: mechanism design and analysis. *Communications in Information and Systems*, 3(4), 305-324.

[5] Klemperer, P. (2002). How (not) to run auctions: the European 3G Telecom auctions. *European Economic Review*, 46(4-5), 829-845.

[6] McMillan, J. (1994). Selling spectrum rights. *Journal of Economic Perspectives*, 145–162.

[7] Saaty, T. L. (2001). Decision making with Dependence and Feedback: The Analytic Network Process, Pittsburgh, RWS Publications.

[8] Sandholm, T., Boutilier, C. (2006). Preference elicitation in combinatorial auctions. In: P. Cramton, Y. Shoham and R. Steinberg (eds.). *Combinatorial Auctions*. MIT Press, Cambridge.

[9] Xia, M., Stallaert,J., Whinston, A. B. (2005). Solving the combinatorial double auction problem. European Journal of Operational Research, 164, 239–251.

FDI AND ECONOMIC GROWTH - CENTRAL EUROPE

Václava Pánková[1]

University of Economics, Prague
FIS, Department of econometrics,
pankova@vse.cz

Abstract. The relationship between foreign direct investment (FDI) inflows and economic growth is important for policy makers. A positive relationship between this two entities is theoretically expected but available empirical evidence does not always support this idea. Numerous articles bring the attempts to find the factors the presence of which boost, or reduce, the impact of FDI on GDP growth in the host country. Some authors conclude that FDI inflows exert a positive impact on economic growth only in the presence of a highly skilled labor force and trade openness and the more negative one the higher degree of corruption occurs. Others give results showing that FDI imply increase of GDP in developing countries but not in advanced economies. Bringing together, an ambiguity of relations is apparent.

In this article, a FDI and GDP relation in Central European new EU members is studied. Evidently, the difficulties arise how to evaluate relevant characteristics of Czech Rep., Hungary, Poland, Slovakia as for their degree of development. Besides, all the countries do not take part in the FDI process for a sufficiently long time to assemble representative data sets. So, all the countries are studied by the help of panel data techniques and the group is enlarged by Austria, the results of which are used as a sort of verification of credibility of movements and influences found by the models.

Keywords. foreign direct investment, economic growth, efficiency, panel data

1. Introduction

The relationship between foreign direct investment (FDI) inflows and economic growth is important for policy makers. A positive relationship between this two entities is theoretically expected but available empirical evidence does not always support this idea. Numerous articles bring the attempts to find the factors the presence of which boost, or reduce, the impact of FDI on GDP growth in the host country. In [5], the authors conclude that FDI inflows exert a positive impact on economic growth only in the presence of a highly skilled labor force and trade openness and the more negative one the higher degree of corruption occurs. Johnson [3] gives results showing that FDI imply increase of GDP in developing countries but not in advanced economies. A negative effect on growth should, according to [1], have FDI inflows into the primary sector. Bringing together, an ambiguity of relations is apparent. As we intend to study the FDI and GDP relation in central European new EU members, the difficulties arise how to evaluate relevant characteristics. Czech Republic, Hungary, Poland, Slovakia, are they developing countries or advanced enough? Besides, all the countries do not take part in the FDI process for a sufficiently long time to assemble representative data sets. So, all the countries will be studied by the help of panel data techniques and the group will be enlarged by Austria, the results of which should be used as a sort of verification of credibility of movements and influences found by the models.

Two approaches are applied. A traditional one is a formulation of a VAR model looking for eventual relations between GDP and FDI. Newly, an influence of FDI on GDP is studied by the help of an efficiency measurement, as it is e.g. in [5].

Data concerning years 1998 – 2007 (yearly, source: Eurostat), FDI, GDP and gross fixed capital formation are measured in mil. of Euro; their logarithms are fdi, gdp, cap. Productivity is measured as value added per person employed per year and is related to 100=EU27*average*; prod means relevant logarithmic values. The main two entities data are briefly presented at the Figures 1. – 2. Each of the five countries is represented by an average value of GDP and FDI respective. The relations correspond with a common expectation.

[1] Financial support of GA ČR 402/07/0049 is gratefully acknowledged by the author.

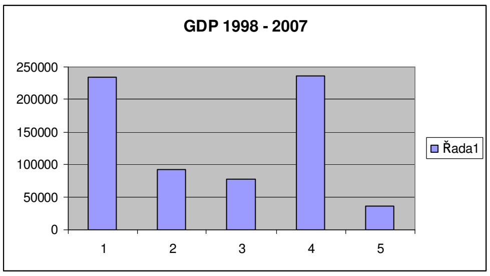


Figure 1.: Average GDP

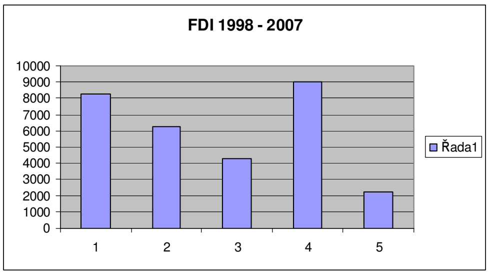


Figure 2.: Average FDI

In both Figures:
1 – Austria, 2 – Czech Republic, 3 – Hungary, 4 – Poland, 5 – Slovakia

2. Causality between FDI and GDP

As for the basic question, if FDI induces GDP growth or vice versa, we cannot take advantage of the Granger causality concept (see e.g. [2]) because of the short data series. For the same reason, the individual VAR models dealing with unique countries were not estimated and a usual inference as stationarity and cointegration could not be performed. Instead of it, an equation

$$GDP = \alpha_0 + \alpha_1 GDP_{-1} + \alpha_2 FDI_{-1} + u \tag{1}$$

was estimated for all economies contemporaneously using the SUR method. Analogically, the other "VAR" equations

$$FDI = \beta_0 + \beta_1 GDP_{-1} + \beta_2 FDI_{-1} + v \tag{2}$$

were found. The results are summarized in Table 1.

The shadowed arrays indicate unconfirmed influence. Hence, in Austria FDI do not affect the GDP but the GDP is a supporting factor of FDI inflows. A similar situations exhibit Slovakia, thought an economic reasoning will probably be different.

endogen.		Au	Cz	Hu	Pl	Sk
GDP	const	-35.3369	-125.012	89.5811	-230.182	-102.535
	GDP_{-1}	1.04481	1.19103	1.04542	0.999569	1.17068
	st.err.	0.004434	0.02184	0.02299	0.04089	0.01750
	t-prob	0.000	0.000	0.000	0.000	0.000
	FDI_{-1}	-0.0368617	-0.805675	0.817135	3.06503	0.0617358
	st.err.	0.1472	0.2950	0.3865	0.9743	0.2302
	t-prob	0.804	0.011	0.044	0.004	0.791
FDI	const	-113.783	24.8395	-13.8522	-64.8206	28.2206
	GDP_{-1}	0.0613299	0.0965960	0.0414152	0.0197367	0.119191
	st.err.	0.01395	0.01068	0.008969	0.006979	0.009115
	t-prob	0.000	0.000	0.000	0.009	0.000
	FDI_{-1}	-0.704363	-0.364205	0.325419	0.649273	-0.672293
	st.err.	0.4631	0.1443	0.1508	0.1663	0.1199
	t-prob	0.139	0.018	0.040	0.001	0.000

Table 1. Estimates of equations (1), (2)

The SUR estimates guarantee the smallest possible standard errors but the VAR models reconstructed by this way do not enable to profit from the VAR structure, at least in a case of a standard software support.

That is why the data sample was treated as a panel set and a real VAR model comprising (1) and (2), and introducing dummy variable D for Austria, was estimated in the sense of pooled regression. The results are in Table 2.

endogen.		
	const.	44834
GDP	dummy	23831
	st.err.	29540
	t-prob	0.425
	GDP_{-1}	0.899
	st.err.	0.157
	t-prob	0.000
	FDI_{-1}	0.019
	st.err.	2.398
	t-prob	0.016
FDI	const.	2751.96
	dummy	135.502
	st.err.	2498
	t-prob	0.957
	GDP_{-1}	0.032
	st.err.	0.013
	t-prob	0.024
	FDI_{-1}	-0.138
	st.err.	0.202
	t-prob	0.498

Table 2.

The shadowed arrays again indicate statistic unsignificance. There is no confirmation of a special position of Austria in the followed context. Besides, appropriate tests show that the presence of FDI and GDP including lag 1 is relevant. Economic interpretation speaks in favor of influence in both directions ($FDI \leftrightarrow GDP$).

Besides, the two roots of companion matrix have modulus less then one, so the equation system is stable and an impulse response analyze can be performed. The results are given by the Figure 3. and show a quick return to an equilibrium position.

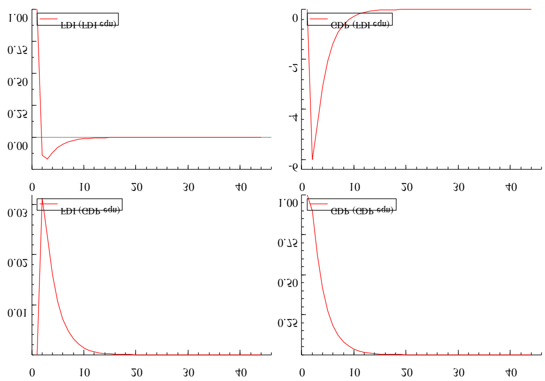
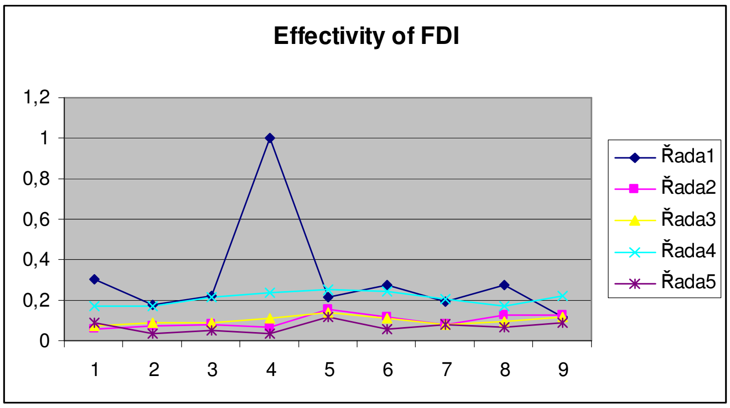


Figure 3.

3. Effectiveness of FDI influence on GDP

In latest papers e.g. [5], a stochastic frontier production function is used to study whether FDI inflows increase economic growth by the help of efficiency gains. Under an assumption of the same technology available across all countries in the sample, technical efficiency of the economies is computed as $TE = exp(\hat{u}_{it} - max\{ \hat{u}_{it} \})$ when assuming the production function to be a Cobb – Douglas one (details e.g. [4]).



Row1(Řada1) - Austria – dark blue[2]
Row 2(Řada2) - Czech Rep. – lilac
Row 3(Řada3) - Hungary - yellow
Row 4(Řada4) - Poland - azure
Row 5(Řada5) - Slovakia – bordeaux

Figure 4.

A most simple relation using logarithmic values show

$$\hat{gdp} = 7.155 + 0.526 fdi \quad . \qquad (3)$$

[2] The color – commentary relates to the fact that presentation as well as the Proceedings are electronic. The author admit that a printed version is not transparent enough.

Technical efficiency TE of FDI influence is presented by the help of Figure 4. The acute part of the Austrian line may be related to a measurement error (for a verification there was no alternative data source available), otherwise a comparability between Austria and Poland is apparent what can be a consequence of evident similarity of GDP and FDI levels (Fig. 1 and 2). The main difference between this two economies is that Poland exhibit the same macroeconomic performance with triple of inhabitants. The other three economies show similar lower results.

Having in mind an important influence of capital and labor force performance, the alternative variant of the model is

$$\hat{gdp} = 1.135 + 1.098cap + 0.00006 fdi - 0.136 prod \quad (4)$$

with strongly non-significant fdi and slightly non-significant $prod$ influence.

To be more exact when speaking about an economic growth, the increments of GDP should be used as endogeneous variable instead of level values. But, such a model fails completely.

4. Conclusions

Four central European economies were studied; their results were individualized by the choice of SUR, respective panel data, technique. A basic influence of FDI on GDP and vice versa was confirmed by relevant estimates for the group as a whole but, not always if detailed to countries. An efficiency of FDI with respect to GDP is higher and comparable for Austria and Poland. Czech Republic, Hungary and Slovakia exhibit similar but lower values. Nevertheless, the influence of FDI on GDP is marginal in comparison with an impact of gross fixed capital formation on GDP. Besides, there is no evidence about an influence of FDI on increments of GDP.

Hence, a significant contribution of FDI to an economic growth during years 1998 – 2007 cannot be validated in the followed countries. The results rather support an idea that a total economic performance is a main factor influencing the other indicators.

References

[1] ALFARO, L.: *Foreign Direct Investment and Growth: Does the Sector Matter*, Paper 14, 2003 Harvard Business School

[2] HUŠEK, R.: *Ekonometrická analýza.* , Oeconomica Praha, 2007. 368 s. ISBN 978-80-245-1300-3.

[3] JOHNSON, A.: *The Effects of FDI Inflows on Host Country Economic Growth*, CESIS Working Paper Series, Paper No. 58, 2006, Royal Institute of Technology, Sweden

[4] KUMBAKAR, S.C. And C.A.K.: LOVELL : *Stochasic Frontier Analysis*, Cambridge Univ. Press, 2000

[5] WIJEWEERA, A., VILLANO, R., DOLLERY, B.: *An Econometric Analysis of Economic Growth and FDI Inflows*, School of Economics, University of New England, Armidale, NSW, 2351 Australia (internet, 2009)

SPLIT DELIVERY VEHICLE ROUTING PROBLEM

Jan Pelikán, Jan Fábry
University of Economics Prague
Faculty of Informatics and Statistics, Department of Econometrics
pelikan@vse.cz

Abstract. Vehicle routing problem (VRP), in which splitting of the customer demand into more routes is permitted, is denoted as split delivery vehicle routing problem (SDVRP). Due to both NP hardness of the problem and many binary variables in the mathematical model of the problem, using of heuristics is necessary. Some heuristics for the problem are shown in the paper.
Keywords. vehicle routing problem, integer programming, heuristics

1. Split Delivery Routing Problem (SDVRP)

Vehicle routing problem (VRP) in the basic formulation presumes that the vehicles involved have the fixed transport capacity. It is generally supposed the number of the vehicles to be unlimited and the vehicles have the same transport capacity. VRP can be modified by adding restriction regarding to the pre-set fixed number of the vehicles, eventually withdrawing restriction relating to the same transport capacity. In VRP a communication network consisting of nodes – points and edges is given, represented e.g. by road net. Initial node (depot, production point) is denoted as node 1, remaining nodes represent customer points with prescribed volume of demand. Edges in the communication net are weighted by distances in km, eventually by transport costs connected with the edge. Demand in node i is denoted q_i>0, vehicle capacity is V>0. It is presumed that each node is served at a time, therefore $q_i \leq V$ and $\sum q_i > V$. In this case, the vehicle has to serve nodes by several routes, every route has to begin and end in node 1. The task is to minimize the sum of distances of all routes.
Mathematical model VRP supposing the homogeneous and unlimited vehicle park, non-split delivery to nodes and demand in nodes not surpassing vehicle capacity came from Miler-Tucker-Zemlin formulation of the travelling salesman problem (see [2]). Arrangement that a vehicle serves node i first and node j afterwards on a route is denoted by binary variable $x_{ij.}$

(MTZ):

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \longrightarrow \min \qquad (1)$$

$$\sum_{i=1}^{n} x_{ij} = 1, \quad j = 2,3,\ldots,n, \qquad (2)$$

$$\sum_{j=1}^{n} x_{ij} = 1, \quad i = 2,3,\ldots,n, \qquad (3)$$

$$u_i + q_j - V(1 - x_{ij}) \leq u_j, \quad i = 1,2,\ldots,n, \quad j = 2,3,\ldots,n, \quad i \neq j, \quad (4)$$

$$u_1 = 0, \quad u_i \leq V, \quad i = 1,2,\ldots,n. \qquad (5)$$

Another formulation of the VRP is based on the Dantzig-Fulkerson-Johnson formulation of the travelling salesman problem.
Mathematical model of the split delivery vehicle routing problem is based on a formulation of the VRP. As all the routes have to be identified, it is necessary to use variables indexed by three indices likewise the vehicle routing problem with various vehicle capacity. In contrast to the non-split delivery model to nodes, new variables q_i^k are introduced to represent delivery to node i in k- th route as one node i may be included in several routes (therefore the equations (2) and (3) have not valid for the solution).

$$\sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x^k{}_{ij} \longrightarrow \min \tag{6}$$

$$\sum_{i=1}^{n} x^k{}_{ij} = \sum_{i=1}^{n} x^k{}_{ji},\ \ j = 2,3,\ldots,n, k = 1,2,\ldots K, \tag{7}$$

$$\sum_{j=1}^{n} x^k{}_{1j} \le 1,\ \ i = 2,3,\ldots,n,\ k = 1,2,\ldots K, \tag{8}$$

$$u^k{}_i + q^k{}_j - V^k(1 - x^k{}_{ij}) \le u^k{}_j,\ \ i = 1,2,\ldots,n,\ \ j = 2,3,\ldots,n,\ \ i \ne j,\ k = 1,2,\ldots,K, \tag{9}$$

$$\sum_{k=1}^{K} q_i^k = q_i,\ \ i = 2,3,\ldots,n, \tag{10}$$

$$q_i^k \le q_i \sum_{j=1}^{n} x_{ij}^k,\ \ i = 2,3,\ldots,n,\ k = 1,2,\ldots K, \tag{11}$$

$$u^k{}_1 = 0,\ \ u^k{}_i \le V^k,\ \ \ i = 1,2,\ldots,n, k = 1,2,\ldots,K. \tag{12}$$

Objective function (6) and constraints (7) and (8) have the same meaning as the corresponding equations in previous VRP model (1) - (5). Vehicle balance is described, likewise in previous VRP model, by the inequalities (9) and (12) with one distinction: the delivery to the node *j* in the route *k* is q_j^k, not q_j. Equation (10) describes the fact that the node *i* is stocked up to its demand q_i by partial deliveries q_i^k supplied via various routes. Inequality (11) expresses that the node not included in appropriate route is not supplied.

The "three indices" model represents considerable increase of variables in comparison with the "two indices" model. This factor has the effect that in solving practical NP-hard problems an optimal solution can not be obtained. In this situation, one can put up with a suboptimal solution yielded by branch and bound method when the computation is aborted (if this solution is found) or various heuristics can be proposed and employed.

It holds that any feasible solution of the non-split demand problem is also an feasible solution of the split delivery demand problem. It can be proved that splitting the demand can cut the overall transport costs, especially in cases when demands in nodes are higher then half the vehicle capacity. An algorithm was proposed (see [1]) enabling the optimal solution VRP (eventually heuristic solution of the VRP) to be farther enhanced by a heuristic. This heuristic is inspired by the savings method seeking cut down transport costs by splitting the delivery to a given node and executing it by two different routes.

2. Features of SDVRP

Reducibility in SDVRP

If it holds $q_i > V$ for at least one node, we can form the reducibility problem for SDVRP and VRP. Problem is reducible if for each node *i* with $q_i > V$ $\lfloor q_i / V \rfloor$ direct routes exist in the optimal solution ([1]).

We can use the reducibily of the problem for the solution the problem, when we at first form direct routes for all nodes with $q_i > V$ and then solve the problem with reduced demands of the nodes.

Problem is reducible if the matrix C satisfies the triangular inequality.

Dror-Trudeau theorems

If the cost matrix C satisfies the triangular inequality then each two routes can have at least one split demand node in the optimal solution. So, if there are two routes which have two common nodes with split demand, if one of them is removed from one route, the costs did not increase [2].

We can define *k*-split cycle: the set of nodes $\{i_1, i_2,\ldots,i_k\}$ is *k*-split cycle if the nodes i_1, i_2 lay on the first route, the nodes i_2, i_3 lay on the second route,..., i_k, i_1 lay on the *k*-th route. It was proved that in the optimal solution does not exist *k*-split cycle.

3. Heuristics for SDVRP

Quest for starting solution SDVRP

Any solution of the corresponding VRP problem (i.e. non-split demand to node) can be used as a starting solution of the SDVRP problem. The following heuristics are the most common ones: nearest neighbour method, savings method and insert method. Besides, improving methods for this starting solution (e.g. method of node exchanges) can be used, by which this starting solution can be further enhanced. In [3] the good heuristics for TSP and VRP is proposed, named as GENIUS heuristics.

Heuristic methods proposed for vehicle routing problem (VRP) can be easily modified in the following way: The last node intended for including to the route but unable to be included for capacity reasons (incomplete vehicle capacity) will be nevertheless included with the proviso that delivery to that node (point) will be partial only, limited by remaining vehicle capacity. Further heuristics are method of node exchanges, 2-split heuristics and adding route heuristics [2].

Method of node exchanges between routes

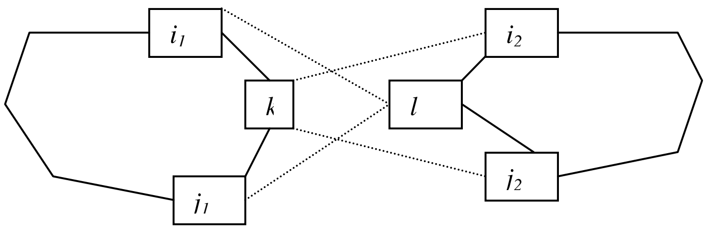
Let us have 2 routes with underused capacities s_1 and s_2. Method of node exchanges consists in substituting the node k of the first route for the node l of the second route and vice versa. In the first route, the node k is preceded by the node i_1 and succeeded by the node j_1, in the second route, the node l is preceded by the node i_2 and succeeded by the node j_2 (see Pic.1). Exchange of the nodes k and l is advantageous, if the following term is positive:

$$s_{kl} = c_{i_1,k} + c_{k,j_1} + c_{i_1,l} + c_{l,j_2} - (c_{i_1,l} + c_{l,j_1} + c_{i_1,k} + c_{k,j_2}) > 0 .$$

The exchange is possible if the vehicle capacity in both routes is sufficient, i.e.

$$q_l \le q_k + s_1 \quad \text{and} \quad q_k \le q_l + s_2.$$

This method can be used even in the case the delivery to any node has been split.

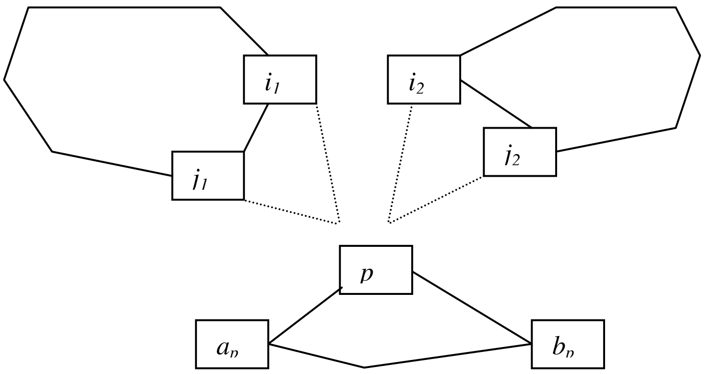


Pic.1 Node exchanges

Heuristic method of splitting delivery (2-split heuristics)

The method consists in changing nodes in three routes; the first two routes are appended by an additional node with split delivery and the very same node is removed from the third route. Let us denote this node p; in the third route, the node p is preceded by the node a_p and succeeded by the node b_p. The demanded delivery to this node is q_p.

The first and second route do not use the complete vehicle capacity, the available capacity of these routes is denoted s_1 and s_2, supposing $s_1 + s_2 \ge q_p$. On this condition the delivery q_p can be split into two parts and each of them can be included in one of these two routes if favourable. Convenience of this exchange can be calculated using distance matrix C. The node p is included into the route 1 between two given nodes in this route denoted as i_1 and j_1, accordingly it is included into the route 2 between two given nodes denoted as i_2 and j_2. The nodes (i_1, j_1) form the edge of the route 1, the nodes (i_2, j_2) form the edge of the route 2. These edges are cancelled in both routes, as well as the edges (a_p, p) and (p, b_p) in the third route. On the other hand, the inclusion of the node p into the first route implies the inclusion of the two edges (i_1, p) and (p, b_1) into this route; accordingly, the inclusion of the node p into the second route implies the inclusion of the two edges (i_2, p) and (p,b_2) into this route.



Pic. 2 : 2-split

The length difference of the added edges and the cancelled edges represents the total change of the solution value; if and only if this difference is negative, then the change is carried into effect. The length difference of the added edges and the cancelled edges is described by the term

$$c_{i_1,p} + c_{p,j_1} + c_{i_2,p} + c_{p,j_2} + c_{a_p,b_p} - (c_{i_1,j_1} + c_{i_2,j_2} + c_{a_p,p} + c_{p\ ,b_p}) .$$

Adding route heuristics

If the demanded delivery to the node p is split into two routes, then this node can be removed from both routes and a new route can be constructed with this node only. In the first route, the node i_1 precedes the node p and the node j_1 succeeds the node p, in the second route the node p is preceded by the node i_2 and succeeded by the node j_2. This change is favourable and it is carried into effect, if the following term is positive:

$$c_{i_1,p} + c_{p,j_1} + c_{i_1,p} + c_{p,j_2} - (c_{i_1,j_1} + c_{i_2,j_2} + c_{1,p} + c_{p,1}) > 0 .$$

4. Summary

The split delivery routing problem (with delivery to network nodes) is a generalization of the basic vehicle routing problem VRP and it may represent a more effective implementation of the transport to customers. As the mathematical model represents an extensive integer programming problem, the implementing of this model to a problem with a large number of nodes eventuates in the unacceptable computing time. The using of heuristic methods is therefore a proper problem-solving instrument.

The research was supported by the grant no.402/09/0041 of the Grant Agency of the Czech Republic.

Reference

[1] ARCHETTI, C., SPERANZA, M.G., HERTZ, A.: A tabu search algorithm for the split delivery vehicle routing problem. *Transportation Science*, Vol. 40, No.1, 2006

[2] DROR, M., TRUDEAU, P., *Savings by Split Delivery Routing*. Transportation Science 23 141-145

[3] GENDREAU, M., HERTZ, A., LAPORTE, G.: New insertion and postoptimization procedures for the traveling salesman problem. *Operation research*, Vol.40, No. 6, 1992

[4] JABLONSKÝ, Josef. Models for efficiency evaluation of decision making units. *Politická ekonomie*, 2004, roč. LII, č. 2, s. 206–220. ISSN 0032-3233.

[5] KUČERA, P., HOUŠKA, M., HOUŠKOVÁ BERÁNKOVÁ, M.: Selected methods for the time limited vehicle routing problem. Liberec 17.09.2008 – 19.09.2008. In: *Mathematical Methods in Economics 2008 [CD-ROM].* Liberec : TU, 2008, s. 307–312. ISBN 978-80-7372-387-3.

[6] PELIKÁN, Jan: *Diskrétní modely v operačním výzkumu.* Professional Publishing.2001. ISBN 80-86419-17-7

EXPECTATION IN E BASELINE DSGE CLOSED ECONOMY MODEL[3]

Jiří Polanský[12] and Jaromír Tonner[12] and Osvald Vašíček[1]

[1] Faculty of Economics and Administration,
Masaryk University, Brno
[2] Macroeconomic Forecasting Division,
Czech National Bank, Praha

[3] This work was supported by funding of specific research at Faculty of Economics and Administration. The views expressed here are those of the authors and do not necessarily reflect the position of the Czech National Bank.

Abstract. In our contribution, we analyze fully anticipated shocks in a medium-scale closed economy DSGE model and compare them with unanticipated shocks. Anticipated shocks might lead to a significant adjustment of agents' behavior before a period when these shocks hit the economy. From a practical policy side, modelling of anticipated shocks can be a useful tool for a policy forecasting since we admit relevance of rational expectations in reality.
Keywords. DSGE Model, monetary transmission, credit channels.

1 Introduction

The behavioral analysis of DSGE models via impulse responses to structural shocks is one of key methods for accepting a model for macroeconomic research or practical forecasting. In literature, we can often find impulse response analysis of main macroeconomic variables to various unanticipated shocks. This analysis is important, but sometimes not sufficient. For a practical use of a DSGE model, modelers need to know (and understand) the reaction of *almost all* model variables to *all* shocks. Hence in our previous work, we laid emphasis on understanding the behavior of the middle-scale closed economy DSGE model (see Tonner et al. [11]). Moreover, one might be interested in impulse responses to fully anticipated shocks. Anticipated shocks play an important role in modern economies. All agents have some personal projections about future economic development, based (at least partly) on rational expectations. When rational expectations are relevant (and people are aware of significant future shocks), agents do not wait for the period when the shocks hit the economy. Instead, they change their behavior immediately (or at least some time before the hit).

From the practical forecasting side, anticipated shocks are one of key elements of a *good* forecast. An admirable example of using anticipated shocks is the Czech National Bank case. One indisputable advantage of the new g3 model is its ability to combine anticipated and unanticipated structural shocks. Our former and current colleagues (O. Kameník, M. Andrle, J. Beneš, T. Hlédik, and J. Vlček) developed an excellent method which allow for a mix of some unanticipated and some fully anticipated shocks in the forecast.

In this paper, we investigate the behavior of a DSGE model via impulse responses to fully anticipated shocks. We compare the behavior with the unanticipated case when a shock is a surprise for agents. This contribution extends our last paper [11] where only unanticipated shocks were discussed. The rest of this paper is organized as follows. Section 2 briefly describes the model and the technique. The third section presents impulse responses and explains the differences between anticipated and unanticipated shocks. The final section concludes the contribution.

2 The Model

For the analysis, we use the closed economy middle-scale DSGE model with staggered nominal price and wage setting. This version of the model can be found directly in Fernández-Villaverde and Rubio-Ramírez [6] and [7], but the structure is very similar to other models that can be found in the literature.[4] Because of

[4] See Galí [9], Walsh [12], or Woodford [13] for more detailed description of the model and a survey of the related literature.

the limited length of this contribution, we refer to cited papers for the description of the model structure and its features and, instead, focus on the model behavior.[5]

The model economy has a standard structure with optimizing agents and rational expectations borrowed from the real business cycle theory. It is enriched with some real rigidities and the price and wage stickiness. In spite of its closed economy nature, it can be qualified as a middle-scale model. It incorporates Calvo-type sticky prices and wages, wage and price indexations, adjustment costs for investment, intensity of use of the capital stock etc. The model contains five sectors - monopolistically competitive households, monopolistically competitive intermediate goods producers, two aggregators, and a policy authority. The final equations can be found in the Appendix of [11].

The solution of linear difference models under rational expectations is outlined in famous paper of Blanchard and Kahn [3]. Because of lack of the space, and because we do not intend to rewrite equations of the solution, we refer to the paper for the technique description.

3 Behavior of the Model

The model economy is hit by five shocks - a monetary policy shock, a shock to the investment specific technology, a shock to the total factor productivity, an intertemporal preference shock, and a shock to the labor supply. In this section, we present the impulse responses to the monetary and total factor productivity shocks. For the analysis, we use a parametrization according to the estimation in [11]. The model is simulated with the IRIS Toolbox [2].

In the figures, the solid line denotes unanticipated shocks whereas the dashed line denotes fully anticipated shocks. In general, we can see that the behavior changes significantly when a shock is anticipated. In the anticipated case, agents know the timing of the shock, its nature (short- or long-lasting), and they realize (rationally) all its effects. Hence, they alter their behavior before the period of a hit.[6]

The figure 1 shows impulse responses to the monetary policy shock. With the presence of nominal rigidities in the model, the monetary policy matters. We can see from the figure that in the unanticipated case, the nominal and real interest rates rise and the inflation falls below its target. Consequently, the output, consumption, investment, hours worked, and the real wage fall below their steady-states. Also, these variables display a hump-shaped dynamics. The optimal inflation rate is significantly below the total inflation rate because of the presence of nominal rigidities in the model. Due to the presence of capital adjustment cost, the return of investment (and capital) to the steady-state is relatively long-lasting.

In the anticipated case, agents are aware of a future positive monetary policy shock. Moreover, they understand its effects (a rise of the real interest rate, a decrease of inflation, output, wages, hours worked etc.). Since they are optimizing, they lower their optimal price inflation before the shock hits the economy.[7] As in the previous case, the decrease of the overall inflation is smaller than the optimal inflation fall. The monetary authority sets interest rate consistently with new conditions in the economy.[8] This objective implies a releasing of monetary conditions via lower interest rates. Note that the interest rate increase is much smaller in the period of the hit and the return to the steady-state is faster than in the unanticipated case. Output, wages, hours worked, consumption, and investment increase initially. These variables have a hump-shaped profile and return slowly to their steady-states. Some of them jump significantly below the steady-state after some periods as a reaction to new conditions in the economy (returning of inflation, expecting of monetary policy shock, dynamics of both production inputs deviations etc.). The behavior of the capital stock is strongly affected by investment adjustment cost (see Tonner et al. [11]).

The figure 2 shows a positive total factor productivity (TFP) shock. The unanticipated shock results in a decline of investment, consumption, wages, hours worked, and output below the steady-state.[9] Simply said, an unanticipated technological shock means that agents observe (fully unexpected) change of the technology which makes them more productive. A decrease of inflation and interest rates is consistent

[5] See also Dixit and Stiglitz [4], Erceg et al. [5], Yun [14], Fuhrer [8], Abel [1], and Kim [10] for detailed explanation of model features.

[6] Note that they are maximizing their objectives (utility or profits functions). Typically when optimizing, agents try to spread the effects of a shock in a longer period.

[7] Note that the timing of deviations before the hit depends on the calibration and the overall model design. Here the effects are relatively long-lasting, but we do not perceive this feature as a model fault. One needs to realize that this model does not serve for forecasting purposes and our objective is the behavioral analysis only.

[8] The monetary authority targets inflation and output.

[9] Note that the model has the balanced growth path where all variables are either constant or grow at a constant pace in the long run. In a case of a positive technological shock, one needs to realize a permanent level shift after a positive temporary technological shock. Hence, some variables below the steady state are still growing, but with less growth rates with respect to the after-shock steady state.

with the technology improvement. On the other hand, when the shock is fully anticipated, agents know perfectly that they will be more productive in future. Moreover, they understand a permanent level shift of the shock and adapt immediately to the new situation. They raise their consumption and decrease their hours worked and investment below the steady-state. Inflation raises with respect to the demand stimulus because the economy still faces the old (before-shock) productivity. Nominal interest rate raises as a reaction to the above-target inflation and the positive output gap.[10] When the shock hits the economy, it makes inputs more productive. Almost all variables are below the steady-state. As was noted, this means (with the reference to technology shocks) that the economy is more productive and the old growth rates are now insufficient.

4 Conclusion

In our contribution, we analyze fully anticipated shocks in a medium-scale closed economy DSGE model and compare them with unanticipated shocks. When a shock is fully anticipated, agents adjust their behavior before a period when a shock hits the economy. This behavioral change might be significantly different when compared to the unanticipated case. In the paper, we present impulse responses to two model shocks (a monetary policy shock and a total factor productivity shock) with a parametrization according to [11]. This analysis extends our previous work where only unanticipated shock were assumed. From figures, one can observe a significant altering of agents' behavior. The current calibration results in a relatively long-lasting return of model variables to their steady states. These deviations depend on the parametrization of the model, but general patters might remain. From a practical policy side, modelling of anticipated shocks can be a useful tool for a policy forecasting since we admit relevance of rational expectations in reality.

[10] At first, output slightly falls as a consequence of a decreased investment, but quickly raises above its steady-state since consumption and hours worked divert substantially from the steady-state.

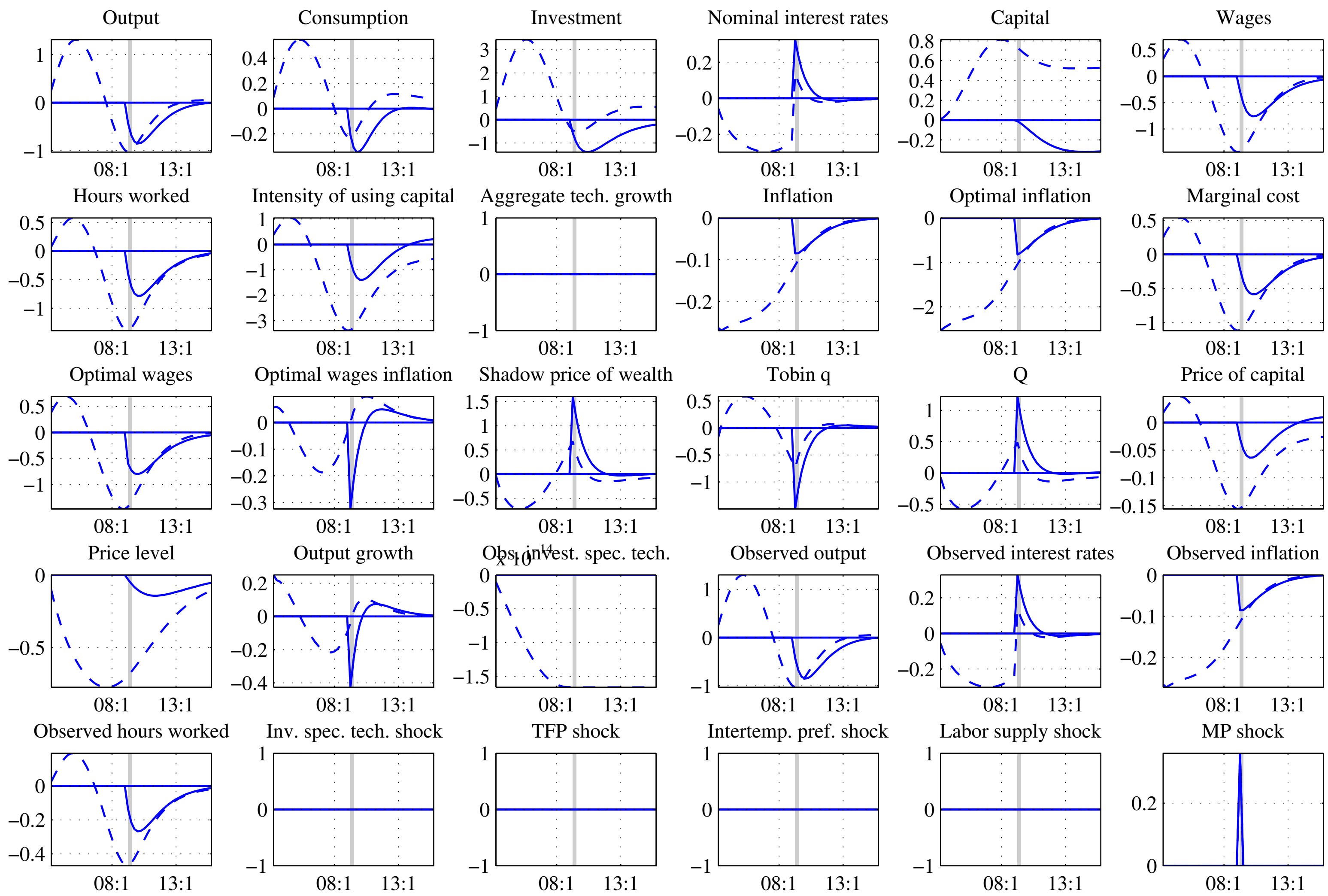


Fig. 1. One std monetary policy shock in %, YoY

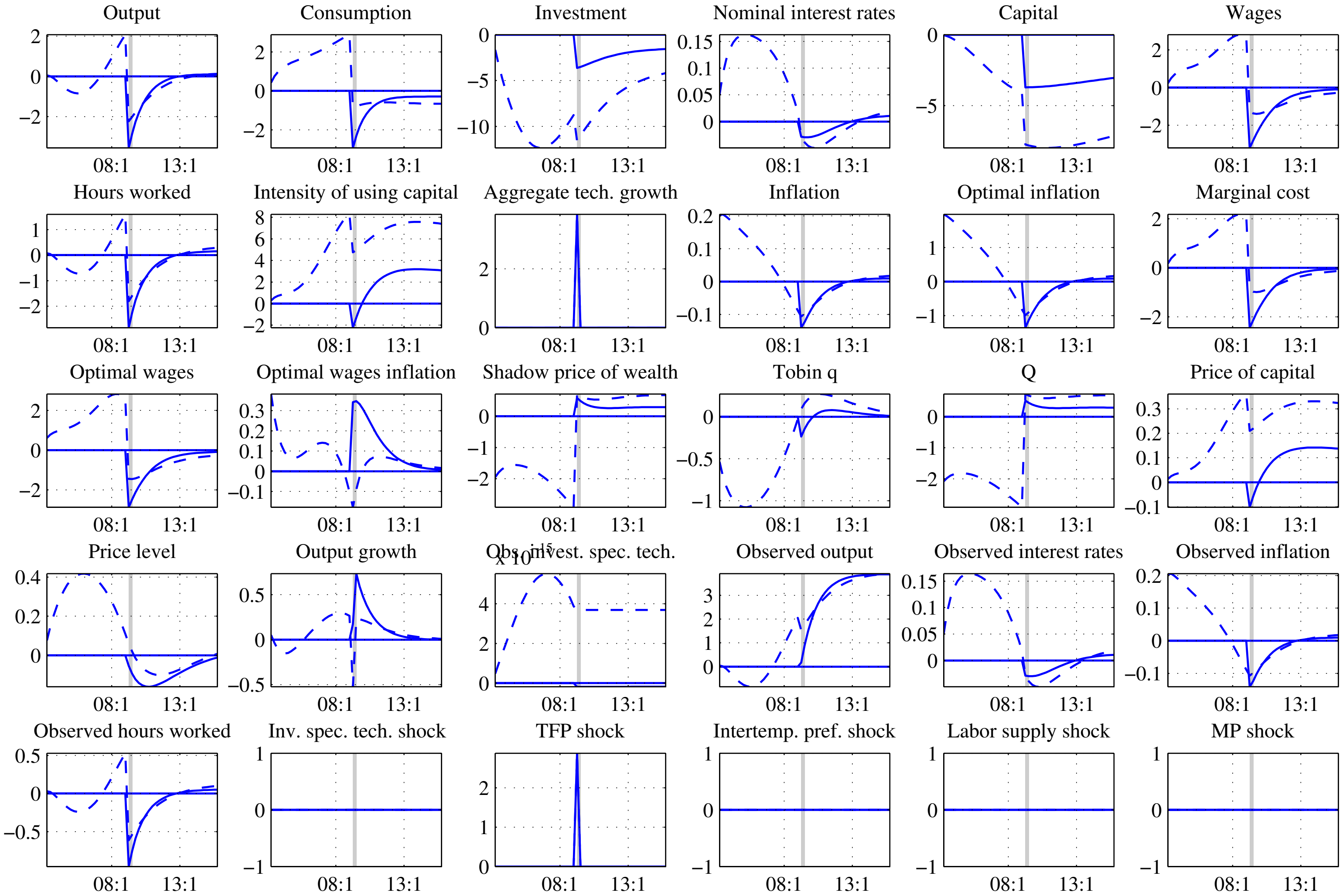


Fig. 2. One std TFP shock in %, YoY

References

1. Abel, A. B. Asset Prices under Habit Formation and Catching up with the Joneses. *The American Economic Review 80*, 2 (1990), 38–42.
2. Beneš, J. Iris Toolbox for Matlab. www.iris-toolbox.com.
3. Blanchard, O. J., and Kahn, C. M. The Solution of Linear Difference Models under Rational Expectations. *Econometrica 48* (1980), 1305–1312.
4. Dixit, A. K., and Stiglitz, J. E. Monopolistic Competition and Optimum Product Diversity. *The American Economic Review 67*, 3 (1977), 297–308.
5. Erceg, C. J., Henderson, D. W., and Levin, A. T. Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics 46* (2000), 281–313.
6. Fernández-Villaverde, J., and Rubio-Ramírez, J. F. A Baseline DSGE Model. Mimeo, 2006.
7. Fernández-Villaverde, J., and Rubio-Ramírez, J. F. How Structural Are Structural Parameters. NBER Working Paper 13166, 2007.
8. Fuhrer, J. F. Habit Formation in Consumption and Its Implications for Monetary Policy Models. *The American Economic Review 90*, 3 (2000), 367–390.
9. Galí, J. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, 1 ed. Princeton University Press, 2008.
10. Kim, J. Functional Equivalence between Intertemporal and Multisectoral Investment Adjustment Costs. *Journal of Economic Dynamics and Control 27* (2003), 533–549.
11. Tonner, J., Polanský, J., and Vašíček, O. A Baseline Model for Monetary Policy Analysis. Bulletin of Czech Econometric Society, forthcoming, 2009.
12. Walsh, C. E. *Monetary Theory and Policy*, 2 ed. MIT Press, 2003.
13. Woodford, M. *Interest and Prices: Foundations of a Theory of Monetary Policy*, 1 ed. Princeton University Press, 2003.
14. Yun, T. Nominal price rigidity, money supply endogeneity, and business cycles. *Journal of Monetary Economics 37* (1996), 345–370.

PARTICLE SWARM OPTIMIZATION FOR THE VEHICLE ROUTING PROBLEM

Daniela Ponce

University of Hradec Králové,
Faculty of Informatics and Management, Department of Information Technologies
daniela.ponce@uhk.cz

Abstract. The paper proposes a particle swarm optimization (PSO) algorithm for solving the basic vehicle routing problem (VRP). A direct route-based representation of VRP is used in the implementation of PSO, i.e. each particle represents a set of feasible routes. After initial routes have been constructed, a modified edge recombination crossover operator is used to move particles towards better solutions. The paper presents early results on application of the proposed PSO algorithm to VRP using benchmark data sets available from the literature.
Keywords. Metaheuristic algorithm, vehicle routing problem, particle swarm optimization, edge recombination crossover operator

1. Introduction

In the vehicle routing problem (VRP), the task is to specify a set of routes on which vehicles will run to serve customers, while all restrictions on customers, vehicles and routes are hold and the set is optimal. Many variants and solution methods of VRP have been formulated ([9]). In this paper, we deal with the so-called capacitated vehicle routing problem (CVRP): the customers have delivery demand only, the vehicles have equal capacities and they all start and end at a single depot. The total route cost is subject to optimization and it is only composed of distance-dependent travel costs of vehicles.

VRP is an important NP-hard optimization problem and since its formulation fifty years ago, plenty of exact, heuristic and metaheuristic solution techniques have been invented. In our contribution, we present a metaheuristic solution technique based on particle swarm optimization (PSO). PSO is a population-based algorithm that simulates spatial behavior of animals living in groups such as flock of birds or school of fish in which the individuals move toward common target without central control. Although PSO has been used for many optimization problems, its applications for VRP are rare ([6]). Recently, PSO for VRP with simultaneous pickup and delivery has been presented ([1]).

The paper is organized in the following way. First, we define the problem of capacitated vehicle routing and we mention some solution methods. Then, we introduce the particle swarm optimization and we propose its application for CVRP including the explanation of the modified edge recombination operator. Finally, we discuss the results and future work.

2. Capacitated vehicle routing problem

The capacitated vehicle routing problem (CVRP) can be formally defined in the following way. Let G = (V, A) be such a graph where $V = \{v_0, v_1, \ldots, v_n\}$ is a set of vertices representing customers, and $A = \{(v_i, v_j) | v_i, v_j \in V, i \neq j\}$ is a set of arcs. Let k_{max} denote the number of vehicles available, each with the same capacity Q. Let v_0 denote a depot where all vehicles are stationed; the $v_1, \ldots, v_n$ vertices represent customers. For each customer, delivery demand quantity q_i is defined. Finally, a matrix of travel costs associated with arcs (c_{ij}) for each $(v_i, v_j) \in A$ is also defined. A solution of CVRP is defined as a set of at most k routes respecting the following conditions:

each route starts and ends at the depot, (1)
each customer is visited exactly once by exactly one vehicle, (2)
the total demand of customers in each route does not exceed the vehicle capacity. (3)

Solution of CVRP is subject to optimization. Let the total cost of all routes be defined as the sum of travel costs spent by vehicles running on arcs that are parts of respective routes. The following condition has to be fulfilled, too:

the total cost of all routes is minimized. (4)

2.1. Solution methods for CVRP

Since the CVRP was formulated, it was solved by exacts methods. The exact methods include branch-and-bound algorithms, branch-and-cut algorithms and set-covering-based algorithms ([9]). Twenty years ago, heuristic methods were introduced which outperformed the exact methods. These methods explore rather limited regions of the solution space, find good quality solutions and require moderate computing time ([5]). The heuristic methods (now called classical) can be classified as constructive heuristics, two-phase heuristics and improvement heuristics ([5]).

Recently, metaheuristic techniques have been invented. These methods explore the most promising regions of the solution space in detail ([5]) and typically, the solutions generated during the solution search process may be of poorer quality or even infeasible ([3]). The main metaheuristics widely used for the CVRP are simulated annealing, deterministic annealing, tabu search, genetic algorithms, ant algorithms and neural networks ([3]). Very recently, particle swarm optimization and artificial immune systems have been used as yet another two promising metaheuristic algorithms for solving CVRP. So far few results can be found in the literature ([7]).

3. Particle swarm optimization

Particle swarm optimization (PSO) is a population-based algorithm ([2]). The algorithm is inspired by spatial behavior of large groups of animals (e.g. bird flocks or fish schools) which are capable to move in a coordinated manner towards common target without central control. The group is called swarm and the individuals are called particles. Each particle represents one solution of the problem. For each particle, position in the solution space and velocity driving its movement in the solution space is specified. Let $\mathbf{x}_i(t)$ denote the position of the ith particle in the solution space at time step t and let $\mathbf{v}_i(t)$ denote the velocity vector of the ith particle in the solution space at time step t. The new position of the ith particle is calculated as:

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1) \quad (5)$$

The velocity vector comprises the individual experience of the particle (so-called cognitive component) as well as the collective experience of the swarm (so-called social component). Each particle keeps information about the best position it has ever reached as personal best position. For each position, its quality is calculated using the fitness function defined on the solution space.

For each particle, its neighbors are specified according to the neighborhood structure defined. One of the basic neighborhood structures is so-called total star neighborhood where each particle is a neighbor of all other particles. Each particle also keeps information about the best position ever reached by its neighbors as global best position. Let $\mathbf{y}_i(t)$ denote the personal best position of the ith particle at time step t, and let $\hat{\mathbf{y}}_i(t)$ denote the global best position of the ith particle at time step t. The velocity of the ith particle is calculated as:

$$v_{ij}(t+1)=v_{ij}(t)+c_1r_{1j}(t)[y_{ij}(t)-x_{ij}(t)]+c_2r_{2j}(t)[\hat{y}_{ij}(t)-x_{ij}(t)] \quad (6)$$

so the new position of a particle can be formulated as:

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t)+c_1r_{1j}(t)[y_{ij}(t)-x_{ij}(t)]+c_2r_{2j}(t)[\hat{y}_{ij}(t)-x_{ij}(t)] \quad (7)$$

Cognitive and social acceleration coefficients c_1 and c_2 serve the purpose to strengthen or weaken the attraction of a particle towards its personal best and global best positions, respectively. They are usually constant during the iterations. Random coefficients $\mathbf{r}_1$ and $\mathbf{r}_2$ in the range <0;1> represent small perturbations in the attraction of a particle towards the personal best and global best positions, respectively. Typically, the c_1 and c_2 constants are set to a value in the range <0;2>, the swarm size is between 20 to 50, and the simulation is run from 500 to 5000 time steps. In the basic version of the PSO, the algorithm starts by creation and initialization of the particles of the swarm. The remaining steps are performed in the loop until the stopping criterion is met: for each particle, its personal best position is updated, then its neighbor best position is updated and finally the position of the particle is updated using the equation (7).

3.1. PSO representation of CVRP

Different coding schemas can be used for representation of CVRP solution in PSO. In their work ([1]), Ai and Kachitvichyanukul represented the solution as priorities assigned to customers. The solution is used in the constructive heuristics: the higher the priority, the earlier the customer will be placed on a route to be served.

We propose the solution instance based representation which is a sequence based representation. Let n denote the number of customers and k the number of vehicles used, $k \leq k_{max}$. The number of vehicles used is set in such a way that:

$$(k-1)*Q < \sum_{i=1}^{n} q_i \leq k*Q \tag{8}$$

Each particle represents a set of k routes. The position **x** of a particle is defined as n+k-1 dimensional vector such that:

$(\mathbf{x})_i \in \{0, 1, \ldots, n\}$ for $0 \leq i \leq n+k-1$, where (9)
$(\mathbf{x})_i = j$ for exactly one value of i, $1 \leq j \leq n$, and (10)
$(\mathbf{x})_i = 0$ for exactly (k-1) values of i. (11)

$a=(\mathbf{x})_i$ and $b=(\mathbf{x})_{i+1}$ adjacent values in the position represent the fact that the arc (v_a, v_b) is included in the route. The arc (v_0, v_b) represents the starting arc of the route of a vehicle, the arc (v_a, v_0) represents the ending arc of the route of a vehicle. For $c=(\mathbf{x})_1$ and $d=(\mathbf{x})_{n+k-1}$, the (v_0, v_c) and (v_d, v_0) arcs are also added at the beginning and the end of the corresponding routes, respectively. For example, the position **x**=(8,17,11,12,15,0,6,16,2,4,18,0,5,1,9,13,3,10,7,14) represents three routes on customers $\{v_1, \ldots, v_{18}\}$ (see Fig. 1).

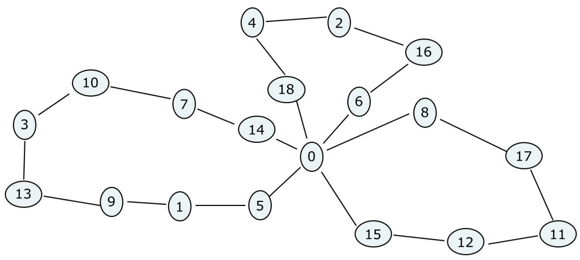


Fig. 1. Three routes on customers $\{v_1, \ldots, v_{18}\}$.

Because $(\mathbf{x})_i=0$ for exactly (k-1) values of i, the position of a particle represents routes of exactly k vehicles, all starting and terminating in the depot (the v_0 vertex), and all customers (the $v_1,\ldots,v_n$ vertices) are visited by exactly one vehicle. The representation thus ensures that the solution encoded by the position of a particle will hold the conditions (1) and (2) of the optimization problem. The condition (3) of the optimization problem is kept during the process of the particle construction. Initially, the position of each particle is constructed in such a way that the total demand of customers in a route does not exceed the capacity Q of the vehicle, if possible. During the solution search process, when the particles move in the solution space, the position is adjusted if the condition (3) does not hold. The fitness function includes a penalty function which adds a big penalty to the resulting fitness of the solution if the (3) condition is violated. The fitness function is formulated as total travel costs of the route.

As the representation is based on sequences instead of state variables, the general PSO model has to be reinterpreted. The velocity models attraction of a particle towards a position in the solution space, and in the case of state variable based representation it is computed as difference of two positions. This difference is then added to a position of the particle and thus the particle moves to a new position. In the case of sequence-based representation, subtracting and adding positions as vectors is not defined. Instead, we apply edge recombination crossover (ERX) operator on position pairs. The result of such an operation is again a new position such that it shares common parts with both operand positions. We reformulate the calculation of the new position of a particle in the following way:

$$
\mathbf{x}_i(t+1) = \begin{cases} \mathbf{y}_i(t) \otimes \mathbf{x}_i(t), \text{if } p_i(t) \in \left\langle \dfrac{u_i(t)}{u_i(t)+\hat{u}_i(t)};1 \right\rangle \\ \hat{\mathbf{y}}_i(t) \otimes \mathbf{x}_i(t), \text{if } p_i(t) \in \left\langle 0;\dfrac{u_i(t)}{u_i(t)+\hat{u}_i(t)} \right\rangle, \end{cases} \tag{12}
$$

where $\otimes$ denotes the edge recombination crossover operator, $p_i(t) \in <0;1>$ is assigned a value with uniform distribution probability, $u_i(t)$ is fitness of $\mathbf{y}_i(t)$ and $\hat{u}_i(t)$ is fitness of $\hat{\mathbf{y}}_i(t)$.

3.2. Edge recombination crossover operator

The edge recombination crossover operator was introduced as crossover operator in genetic algorithm for the travelling salesman problem ([10]). When the solution of TSP is represented as a path, the ERX operator applied to two operand (parent) solutions (e.g. see Fig. 2) yields a new offspring solution such that each segment of the offspring solution was randomly selected from one of the two parent solutions (e.g. see Fig. 3a). As the experiments have shown ([4]), the effectiveness of the ERX operator can be enhanced by giving higher priority to the segments shared by both operands (e.g. see Fig 3b).

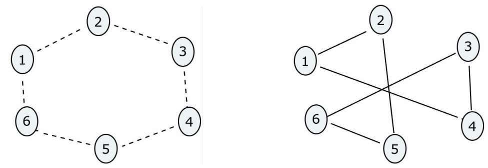


Fig. 2. Two parent paths sharing segments 1-2, 3-4 and 5-6.

Fig. 3. a) Offspring path. b) Offspring path with segments 1-2 and 5-6 included.

The application of the ERX operator is performed in the following way. Firstly, the adjacency matrix of all nodes contained in the both operand paths is constructed in such a way that for every node the nodes adjacent with the given node in either operand paths are listed (e.g. see Tab. 1.).

node	adjacent nodes
1	2,6,2,4
2	1,3,1,5
3	2,4,4,6
4	3,5,1,3
5	4,6,2,6
6	1,5,3,5

Tab. 1. Adjacency matrix for paths from Fig. 2.

Each node in the adjacency matrix has exactly four adjacent nodes (two from each operand paths). Duplicate nodes on the list of adjacent nodes mean that the segment is common for the both operand paths.

The adjacency matrix is used in the resulting path construction. Initially, a starting node n_s of the path and one node n_a from the list of its adjacent nodes are selected and they become the first segment of the resulting path. The rest of the path is constructed in a loop: the information on adjacency of n_s and n_a in the resulting path is recorded for both nodes n_s and n_a. Furthermore, the node n_s is removed from the list of adjacent nodes for all nodes in the adjacency matrix. The node n_a gets then assigned to n_s, and n_a is assigned a node selected from the list of the adjacent nodes of n_s (see Tab. 2a, 2b).

Adjacency matrix	
node	adjacent nodes
1	~~2~~,6,~~2~~,4
2	~~1~~,3,~~1~~,5
3	~~2~~,4,4,6
4	3,5,~~1~~,3
5	4,6,~~2~~,6
6	~~1~~,5,3,5

Resulting path	
node	adjacent nodes
1	2,4
2	1
4	1

Tab. 2a. Construction of the resulting path, n_s=1, n_a=4.

Adjacency matrix	
node	adjacent nodes
1	~~2~~,6,~~2~~,~~4~~
2	~~1~~,3,~~1~~,5
3	~~2~~,~~4~~,~~4~~,6
4	3,5,~~1~~,3
5	~~4~~,6,~~2~~,6
6	~~1~~,5,3,5

Resulting path	
node	adjacent nodes
1	2,4
2	1
4,	1,5
5	4

Tab. 2b. Construction of the resulting path, n_s=4,n_a=5.

The loop halts when the list of the adjacent nodes of the current n_s is empty and the adjacency matrix becomes empty. The last node of the path is then made adjacent with the node which was used to start the path.

3.3. ERX operator for vehicle routes

When we use the ERX operator for the VRP (see e.g. Fig. 4), we have generally more than one path, thus the node 0 corresponding to the depot has more than four adjacent nodes in the adjacency matrix of the operand solutions. During the construction of the resulting route set, we have to deal differently with the node 0 in the role of n_s: it cannot be removed from the list of adjacent nodes for all nodes in the adjacency matrix, as the construction would result in a single huge route passing through the depot node and all nodes. On the other hand, keeping the node 0 in the adjacency matrix, routes can be closed too early leaving some nodes not included in any route. Therefore, when selecting the next n_a node from the adjacent nodes of the current n_s node, we give less priority to the node 0 than to other nodes. We also give higher priority to duplicate adjacent nodes of the current n_s node (meaning common segments in the operand solutions). When the list of the adjacent nodes of the current n_s is empty, we close the current route, and we start another one by selecting such a node in the adjacency matrix that still has adjacent nodes listed. The application of the ERX operator halts when the adjacency matrix becomes empty.

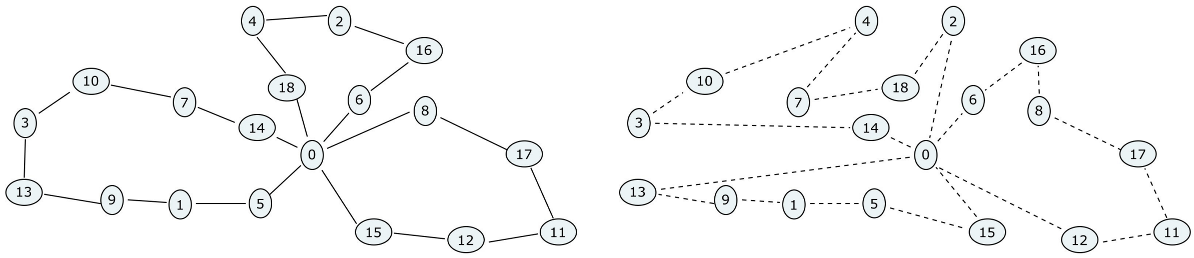


Fig. 4. Two operand route sets.

3.4. Refinement of the solution

To terminate the construction of new solution, we have to check that the solution is composed of k routes and that each route passes through the depot. If necessary (see e.g. Fig. 5), we randomly merge two routes into one or split one route into two until we get k number of routes such that they all pass through the depot.

Finally, when k number of routes is reached, we have to ensure that the condition (3) is met. We can apply some solution multiroute improvement heuristic regarding the edge exchange (see [5] for details) to balance the total demand of routes.

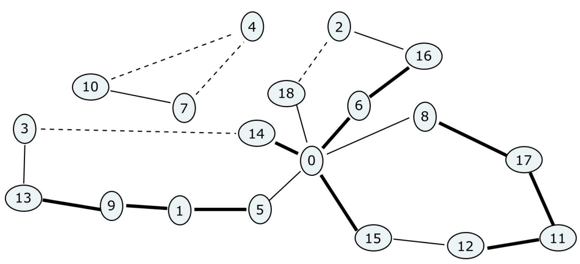


Fig. 5. The result of the ERX operator application on the operand route sets from Fig. 4.

4. Results and future work

The algorithm proposed has been run on several CVRP instances, i.e. the benchmark problems of Solomon ([8]). So far, the results are not competitive with the other results achieved. The reason is that, currently, a little problem-specific information is used to guide the movement of the particle swarm in the solution space. Instead of random selection, the problem specific features can be taken into account, e.g. during the application of the ERX operator or refinement of the solution. The PSO algorithm can be made more elaborated, too. Random replacement of a particle can be introduced as another mechanism that brings stochastic element into the search. It allows the swarm to escape from the local optimum which can attract all particles. Currently, in the calculation of the next position of a particle, fitness of the personal best and neighbor best positions is used. Instead, probabilities proportional to the social and cognitive components can be applied.

5. Acknowledgments

The research has been done on the base of the research project GAČR-402/09/0405 funded by the Grant Agency of the Czech Republic. The support of the agency is kindly acknowledged.

References

[1] T. J. AI and V. KACHITVICHYANUKUL, *A particle swarm optimization for the vehicle routing problem with simultaneous pickup and delivery*, Computers & Operations Research, Vol. 36, No. 5, 2009, pp. 1693-1702.

[2] A. P. ENGELBRECHT, *Fundamentals of Computational Swarm Intelligence*, Wiley, 2006. ISBN 978-0-470-09191-3

[3] M. GENDREAU, G. LAPORTE, J.-Y. POTVIN, *Metaheuristics for the Capacitated VRP*, In P. Toth and D. Vigo, (Eds.), The Vehicle Routing Problem (Monographs on Discrete Mathematics & Applications), Society for Industrial & Applied Mathematics, 2001, pp. 129-154. ISBN 0-89871-579-2

[4] J. HYNEK, *Genetické algoritmy a genetické programování*, GRADA, 2008, ISBN 978-80-247-2695-3. (in Czech)

[5] G. LAPORTE and F. SEMET, *Classical Heuristics for the Capacitated VRP*, In P. Toth and D. Vigo, (Eds.), The Vehicle Routing Problem (Monographs on Discrete Mathematics & Applications), Society for Industrial & Applied Mathematics, 2001, pp. 109-128. ISBN 0-89871-579-2

[6] R. POLI, *Analysis of the publications on the applications of particle swarm optimisation*, Journal of Artificial Evolution and Applications, Vol. 2008, No. 1, 2008, pp 1–10.

[7] J.-Y. POTVIN, *A Review of Bio-inspired Algorithms for Vehicle Routing*, In F. Baptista Pereira and J. Tavares (Eds.), Bio-Inspired Algorithms for the Vehicle Routing Problem (Studies in Computational Intelligence), Springer, 2008, pp. 1-34. ISBN 978-3-540-85151-6

[8] M. M. SOLOMON, 25, 50 and 100 customers problem instances, 1983, available at http://neo.lcc.uma.es/radi-aeb/WebVRP/Problem_Instances/CVRPTWInstances.html#Solomon, access on May 24, 2009

[9] P. TOTH and D. VIGO, An Overview of Vehicle Routing Problems, In P. Toth and D. Vigo (Eds.), *The Vehicle Routing Problem (Monographs on Discrete Mathematics & Applications),* Society for Industrial & Applied Mathematics, 2001, pp. 1-26. ISBN 0-89871-579-2

[10] D. WHITLEY, T. STARKWEATHER, D. A. FUQUAY, *Scheduling Problems and Traveling Salesman: The Genetic Edge Recombination Operator*, In J.D. Schaffer (ed.), Proc. of the 3rd Intl. Conf. on Genetic Algorithms, San Mateo, California, Morgan Kaufmann,1989, pp. 133-140.

On the Neoclassical Theory of Investment and Tobin's q

Pavel Pražák, Pavel Trojovský

Faculty of Informatics and Management, University of Hradec Králové
pavel.prazak@uhk.cz, pavel.trojovsky@uhk.cz

Abstract. The neoclassical theory of investment considers a firm's optimization behavior - the objective of the firm is to maximize the present value of the firm subject to some technological constraints. Another concept of investment was developed by Tobin who suggested that the rate of investment is a function of q, defined as the ratio of the market value of new additional investment goods to their replacement cost. It was shown by Hayashi that the two different approaches to investment can be integrated and that q can be interpreted as the value to the firm of an additional unit of capital, which is the discounted value of its future marginal revenue products. This can be demonstrated from an application of Pontryagin's maximum principle. In the present paper we consider a similar problem. We present necessary conditions for an infinite horizon optimal control problem to find an optimal investment behavior of a firm.
Keywords. Investment, Tobin-s q, optimal control problem, infinite horizon, Pontryagin-s maximum principle.

1 The Value of the Firm

The basic principle of investment comes from a common profit maximization and can be formulated as follows: the firm chooses such level of the input that the marginal productivity of this input is equal to its relative price. In the long period, all inputs of the firm can be adjusted, including such fixed inputs like capital, i.e. machines, buildings and equipment. To develop the principle of investment in more detail it is necessary to consider decisions that will influence profits in long period. The concept that allows us to do it is the value of the firm. The value of a firm can be defined as the present value of the expected future net cash flow (NCF) that the firm will generate. The general formula can be written as

$$V_t = \int_t^\infty e^{-\int_t^s r(v)dv} \Pi(s)ds, \tag{1}$$

where t, $t \geq 0$ represents time and $\Pi(s)$ is the net cash flow of the firm at time $s, s \geq t$, which represents revenues minus costs without any concern for depreciation of assets. The parameter $r(s)$ is the discount rate at time $s, s \geq t$. For purposes of this paper we may assume that this rate is a constant. Then the present value of the firm V_t at time t can be expressed in a simpler form

$$V_t = \int_t^\infty e^{-r(s-t)} \Pi(s)ds. \tag{2}$$

2 Neoclassical Investment Problem

One of the first dynamic models of the firm is the neoclassical investment model, see (Jorgenson, **?**) and (Hall, Jorgenson, **?**). Let $p(t)$, $p(t) > 0$, be the unit price of the output of the firm at time t, $w(t)$, $w(t) > 0$, the unit price of labor and finally $g(t)$, $g(t) > 0$, the unit price of investment good. Let $Q(t)$ be the quantity of output at time t, $L(t)$ the quantity of input, say labor, and $I(t)$ the rate of investment at time t. Now the net cash flow at time t can be written as

$$\Pi(t) = p(t)Q(t) - w(t)L(t) - g(t)I(t). \tag{3}$$

The firm that is owned by households wants to maximize its value V_0. If we consider the neoclassical production function F, then $Q = F(K, L)$, where K, $K = K(t)$, is the quantity of capital stock at time t and L, $L = L(t)$ is the quantity of labor at time t. It is assumed that capital depreciates at the rate δ, $\delta \in (0, 1)$. The value of the firm V_0 defined by (**??**) is then given as a function of the variables K, L and I and the problem can be formulated as follows

$$\max\{V_0(K, L, I) |\, K \in M_1, L \in M_2, I \in M_3\} \tag{4}$$

where

$$M_1 = \{K \in PC^1([0,\infty))|\ \dot{K} = I - \delta K, K(0) = K_0 > 0, K \geq 0, \delta \in (0,1)\} \tag{5}$$

$$M_2 = \{L \in PC([0,\infty))|\ L \geq 0\} \tag{6}$$

$$M_3 = PC([0,\infty)) \tag{7}$$

The current-value Hamiltonian of this problem is

$$H^c = H^c(K, L, I, q) = pF(K, L) - wL - gI + q(I - \delta K), \tag{8}$$

where we suppressed the variable t denoting time in the enrollment of the functions $K = K(t)$, $L = L(t)$, $I = I(t)$, $p = p(t)$, $w = w(t)$, $g = g(t)$ and where $q = q(t)$ is the current-value adjoint variable. Let $(\hat{K}, \hat{L}, \hat{I})$ be the optimal solution to problem (**??**). We will also use the notation $\hat{H}^c = H^c(\hat{K}, \hat{L}, \hat{I}, q)$. If Pontryagin maximum principle for infinite horizon is used, see e.g. (Seierstad, Sydsæter, **?**), then the following necessary conditions are obtained:

$$\hat{H}^c_L = pF_L(\hat{K}, \hat{L}) - w = 0, \tag{9}$$

$$\hat{H}^c_I = -g + q = 0, \tag{10}$$

and the adjoint equation is

$$\dot{q} - rq = -\hat{H}^c_K = -pF_K(\hat{K}, \hat{L}) + \delta q. \tag{11}$$

The relation (**??**) is a standard employment optimality condition - the labor is employed if the marginal labor product $p(t)F_L(\hat{K}(t), \hat{L}(t))$ is equal to unit price of labor $w(t)$, $t \in [0,\infty)$. Since (**??**) is a necessary condition for optimal solution to (**??**) we have $q(t) = g(t)$, $t \in [0,\infty)$. This condition can be substituted into (**??**) and we gain the relation

$$p(t)F_K(\hat{K}(t), \hat{L}(t)) = \delta g(t) + rg(t) - \dot{g}(t),\ t \in [0,\infty) \tag{12}$$

which means that the cost of capital has three components: the capital depreciates at rate δ; using the capital, the firm gives up the choice of selling it and invest the money at interest rate r; and the firm may get a capital gain or suffer a capital loss, i.e. the price of the capital good may change in the rate $\dot{g}(t)$. If the relative price of capital is constant, i.e. $\dot{g}(t) = 0$ and is normalized to unity, we obtain the standard rental rate expression $p(t)F_K(\hat{K}(t), \hat{L}(t)) = r + \delta$ which one gets in the static model of the firm. On the optimal path the firm should determine a desired level of the capital stock from (**??**), i.e. it should adjust the actual capital stock immediately to its desired level. The firm chooses its desired capital stock through optimal investment. What is the optimal investment rate if the initial capital stock $K(0)$ is not at its desired level? Since capital evolves continuously over time, see (**??**), and a discrete change in the capital stock is required an infinite rate of investment is necessary. It means that the optimal investment strategy given by (**??**) to obtain the desired level of capital stock is infeasible. Moreover, any discrete change of the exogenous parameter as interest rate r leads to a discrete change in the desired capital stock. Similarly as in the previous observation this will lead to an infinite investment level. The theory of investment tackles these problems by introducing capital adjustment costs.

3 Adjustment Cost and Investment

We will consider that capital is costly to install and uninstall - the more detailed discussion was given in (Pražák, **?**). The net cash flow at time t, $t \in [0,\infty)$, of the representative firm can be modified as follows

$$\Pi(t) = p(t)Q(t) - w(t)L(t) - g(t)I(t) - C(I(t), K(t)). \tag{13}$$

The key assumption in this formulation of the net cash flow of the firm is that there is a convex installation cost $C(I, K)$, particularly we will assume that the adjustment cost for installing new capital has the form

$$C(I, K) = \beta \frac{I^2}{2K}, \tag{14}$$

where β is a positive real constant. Notice that if $p(t), w(t)$ and $g(t)$ are exogenously given function and moreover if $Q(t)$ is given by a neoclassical production function then $\Pi(t) = \Pi(K(t), L(t), I(t))$. If we use (**??**) in (**??**) then the current-value Hamiltonian in the problem (**??**) is

$$H^c = H^c(K, L, I, q) = pF(K, L) - wL - gI - C(I, K) + q(I - \delta K). \tag{15}$$

Now we use Pontryagin maximum principle for infinite horizon problem. Since $H^c(K, L, I, q)$ is concave with respect to variables L and I the optimal solution $(\hat{K}(t), \hat{L}(t), \hat{I}(t))$ to (**??**) satisfies the following necessary conditions:

$$\hat{H}^c_L = pF_L(\hat{K}, \hat{L}) - w = 0, \tag{16}$$

$$\hat{H}^c_I = -g - C_I(\hat{I}, \hat{K}) + q = 0. \tag{17}$$

The relation (**??**) is again the standard employment optimality condition. If we use (**??**) for adjustment cost function we can rewrite (**??**) as

$$\frac{\hat{I}(t)}{\hat{K}(t)} = \frac{q(t) - g(t)}{\beta} \tag{18}$$

which means that optimal investment has the same sign as $q(t) - g(t)$, $t \in [0, \infty)$. This criterion can be used as a rule for regulation of optimal investment. Whenewer the control variables $\hat{I}$ and $\hat{L}$ are continuous the adjoint function q satisfies the adjoint equation

$$\dot{q} - rq = -\hat{H}^c_K = -pF_K(\hat{K}, \hat{L}) + C_K(\hat{I}, \hat{K}) + \delta q \tag{19}$$

with the transversality condition

$$\lim_{t \to \infty} e^{-rt} q(t) \hat{K}(t) = 0. \tag{20}$$

Note that $-pF_K(\hat{K}, \hat{L}) + C_K(\hat{I}, \hat{K}) = -\Pi_K(\hat{K}, \hat{L}, \hat{I})$, see (**??**), which means that adjoint equation (**??**) can be rewritten as

$$\dot{q} - (r + \delta)q = -\Pi_K(\hat{K}, \hat{L}, \hat{I}). \tag{21}$$

This is a linear differential equation and its solution $q(t)$ can be express with the following relation

$$\lim_{s \to \infty} q(s) e^{-(r+\delta)s} - q(t) e^{-(r+\delta)t} = \int_t^\infty e^{-(r+\delta)s} \left[-\Pi_K(\hat{K}(s), \hat{L}(s), \hat{I}(s)) \right] ds.$$

The economically relevant solution requires that $\hat{K}(t)$ is bounded and it means that transversality condition (**??**) gives

$$\lim_{t \to \infty} e^{-rt} q(t) = 0.$$

This observation yields that

$$q(t) = \int_t^\infty e^{-(r+\delta)(s-t)} \left[\Pi_K(\hat{K}(s), \hat{L}(s), \hat{I}(s)) \right] ds, \tag{22}$$

which denotes the present value of the marginal contribution of capital to net cash flow of the firm. A unit increase in the firm's capital stock at time t increases the present value of the firm's net cash flow by $q(t)$. It means that $q(t)$ shows how an additional unit price of capital affects the present value of net cash flow of the firm and thus it raises the value of the firm by $q(t)$. Thus $q(t)$ can be interpreted as the market value of a unit of capital for a given firm.

3.1 Marginal *q*

To check the interpretation of $q(t)$ let us find the derivative of the functional that represents the value of the firm at time t, i.e. the functional

$$V_t = V(\hat{K}(t), \hat{L}(t), \hat{I}(t)) = \int_t^\infty e^{-r(s-t)} \Pi(\hat{K}(s), \hat{L}(s), \hat{I}(s)) ds$$

with respect to function $K(t)$ that represents capital stocks. It can be proved that

$$\frac{\partial V_t}{\partial K} = \int_t^\infty e^{-(r+\delta)(s-t)} \Pi_K(\hat{K}(s), \hat{L}(s), \hat{I}(s)) ds = q(t). \tag{23}$$

Therefore the marginal $q(t)$ can be considered as the ratio of the change in the value of the firm to the added capital cost for a small increment to the capital stock.

3.2 Tobin's *q*

There is a problem with marginal q. It cannot be measured and therefore it cannot be considered for empirical work. Instead of that it is necessary to use some approximations and average values. The average q or Tobin's q is defined as

$$q^A(t) = \frac{V_t}{g(t)K(t)},$$

which represents the ratio of the total value of the firm to the replacement cost of its total capital stock, see (Tobin, **?**). The relationship between average q and marginal q was for the first time shown in (Hayashi, **?**).

3.3 Hayashi observation

Now we use the fact that F is a neoclassical production function which reflects a constant return to scale. It means that F is a homogeneous function of degree one. Let us further consider that also the adjustment cost function C is a homogeneous function of degree one such that we consider for example in (**??**). Then according to Euler's theorem for homogeneous function of degree one we have

$$F(K,L) = KF_K(K,L) + LF_L(K,L),\ C(I,K) = IC_I(I,K) + KC_K(I,K).$$

Relations (**??**), (**??**) and the previous observations allow us to express $\hat{K}\Pi_K(\hat{K},\hat{L},\hat{I})$ as follows

$$\begin{aligned}\hat{K}\Pi_K(\hat{K},\hat{L},\hat{I}) &= p\hat{K}F_K(\hat{K},\hat{L}) - \hat{K}C_K(\hat{I},\hat{K})\\ &= p[F(\hat{K},\hat{L}) - LF_L(\hat{K},\hat{L})] - [C(\hat{I},\hat{K}) - IC_I(\hat{I},\hat{K})]\\ &= pF(\hat{K},\hat{L}) - \hat{L}w - C(\hat{I},\hat{K}) + \hat{I}(q-g)\\ &= \Pi(\hat{K},\hat{L},\hat{I}) + q\hat{I}\end{aligned}$$

This relation, (**??**) and (**??**) can help us to find the following derivative

$$\begin{aligned}\dot{(q\hat{K})} &= \dot{q}\hat{K} + q\dot{\hat{K}}\\ &= [(r+\delta)q - \Pi_K(\hat{K},\hat{L},\hat{I})]\hat{K} + q[\hat{I} - \delta\hat{K}]\\ &= rq\hat{K} + \delta q\hat{K} - \Pi(\hat{K},\hat{L},\hat{I}) - q\hat{I} + q\hat{I} - \delta q\hat{K}\\ &= rq\hat{K} - \Pi(\hat{K},\hat{L},\hat{I}).\end{aligned}$$

We obtained a linear differential equation for an unknown function $q(t)\hat{K}(t)$. The solutions to this equation can be written in the following form

$$\lim_{s\to\infty} e^{-rs}q(s)\hat{K}(s) - e^{-rt}q(t)\hat{K}(t) = -\int_t^\infty e^{-rs}\Pi(s)ds,$$

where we use a truncated version for $\Pi(\hat{K}(s),\hat{L}(s),\hat{I}(s)) = \Pi(s)$. If we take into account transversality condition (**??**) and we can finally write

$$q(t)\hat{K}(t) = \int_t^\infty e^{-r(s-t)}\Pi(s)ds = V_t,$$

which means that

$$q(t) = \frac{V_t}{\hat{K}(t)} = g(t)q^A(t),\ t \in [0,\infty). \tag{24}$$

The latter expression means that if we normalize the unit price of investment good, i.e. $g(t) = 1, t \in [0,\infty)$, the marginal q and average q are equal.

4 Conclusion

The paper deals with a neoclassical investment model with adjustment cost on infinite time horizon, uses necessary conditions given from Potryagin maximum principle and offers a simple and a transparent proof of Hayashi result which states that Tobin's q and marginal q are equal providing that the firm is price taker, its production function is concave, its cost function is convex and both functions are homogeneous of degree one.

References

Hall R.E., Jorgenson D.W.: *Application of the Theory of Optimum Capital Accumulation*, in Tax Incentives and Capital Spending, ed. by G. Fromm, Washington, Brookings Institution (1973)

Hayashi F.: *Tobin's Marginal q and Average q : A Neoclassical Interpretation*, Econometrica, Vol. 50, pp. 213–224 (1982)

Jorgenson D.W.: *Capital Theory and Investment Behavior*, American Economic Review 53, pp. 247–259 (1963)

Pražák P.: *On a Continuous Time Adjustment Cost Model of the Firm*, in Mathematical Methods in Economics, pp. 426-431, Technical University of Liberec (2008)

Seierstad, A., Sydsæter K.: *Optimal Control Theory with Economic Application,* North-Holland (1987)

Tobin J.: *A General Equilibrium Approach to Monetary Theory*, Journal of Money, Credit, and Banking, Vol. 1, No. 1, pp. 15-29 (1969)

STRUCTURAL CHANGE OF THE CZECH ECONOMY*

Adam Remo and Osvald Vašíček**

Masaryk University, Faculty of Economics and Administration
Department of Economics
adamremo@seznam.cz osvald@econ.muni.cz

Abstract. The goal of this paper is to estimate timing and shape of a structural change (regime switch) of the Czech economy, which took place at the end of the nineties. The approach we have chosen is based on Dynamic Stochastic General Equilibrium (DSGE) model with Markov switching regimes. In order to keep the estimate meaningful with respect to the number of parameters (model complexity) and the length of Czech macroeconomic time series, we have chosen a small prototypical New Keynesian model. This model has been estimated by Bayesian techniques. The results do not consist of parameter estimates for different regimes only, but also of probabilities of these regimes in time. This offers very interesting view of the changes the Czech economy went through in the second half of the nineties.
Keywords. Structural change, Markov switching regimes, New Keynesian, Czech economy, DSGE.

Key words: structural change, Markov switching regimes, New Keynesian, Czech economy, DSGE

1 Introduction

Even a quick glance at main macroeconomic time series of the Czech economy raises a question:

– Did the Czech economy undergo a structural change in the late nineties?

In other words, did the dynamic behavior of the Czech economy change or did the economic development at the late nineties just reflect extreme exogenous shocks?

Prior to the estimation of the possible structural change we know that at least a switch of the monetary regime to inflation targeting took place at that time. Therefore we expect that the structural change happened and the question above have to be refined:

– Did the structural change in the late nineties consist only in a monetary regime switch or another aspects of the Czech economy changed as well?

In this paper, we try to find the answer to this question. According to this objective, we focus on estimation of a possible change in fairly general aspects of the economy. The precise assessment of a structural change would require more realistic (and complicated) model structure. At least an open economy model should be considered in this case.

2 Simple New Keynesian Model

As has been written above, we would like to find out whether the dynamics of the Czech economy changed due to other aspects then monetary policy or not. Because we do not aspire to track down concrete sources of a structural change, a simple New Keynesian model fits our requirements well enough. Moreover, estimation of more complicated model structure could be problematic with respect to the number of estimated parameters and length of the Czech time series.

The chosen model is a slightly modified version of the closed economy New Keynesian dynamic stochastic general equilibrium (DSGE) model presented in [1]. We refer the reader to the cited paper for more detailed discussion.

The economy consists of a representative household, a continuum of intermediate goods producing firms, a final goods producing firm, and a monetary authority. The representative household consumes

* This work was supported by MŠMT project Research Centers 1M0524, and funding of specific research at ESF MU.

** We appreciate the access to the METACentrum computing facilities provided under the research intent MSM6383917201.

final goods and supplies the intermediate goods producing firms with homogeneous labor. The model supposes perfect competition on labor market and that's why household and firms are wage-takers. The intermediate goods producing firms hire labor and produce goods according to constant returns on scale production function. The j-th firm faces the quadratic price adjustment costs

$$AC_t(j) = \frac{\phi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - (1+\pi^{ss}/100)\right)^2 Y_t(j)\,, \tag{1}$$

where ϕ is the measure of price stickiness, π^{ss} is steady state inflation rate of the final goods, $P_t(j)$ is price of its production in period t, and $Y_t(j)$ is its production. The monetary authority which pays attention to inflation rate and growth rate of output sets nominal interest rate according to Taylor type rule.

The linearized model in terms of stationary variables is

$$\hat{y}_t = E_t\hat{y}_{t+1} - \frac{1}{\tau}(\hat{r}_t - E_t\hat{\pi}_{t+1} - \rho_z\hat{z}_t)\,, \tag{2a}$$

$$\hat{\pi}_t = \beta E_t\hat{\pi}_{t+1} + \frac{\tau\kappa}{(1+\frac{\pi^{ss}}{100})^2}\left[\hat{y}_t + \varepsilon_{\pi,t}\right]\,, \tag{2b}$$

$$\hat{r}_t = \rho_r\hat{r}_{t-1} + (1-\rho_r)\big[\psi_1\hat{\pi}_t + \psi_2(\hat{y}_t - \hat{y}_{t-1} + \hat{z}_t)\big] + \varepsilon_{r,t}\,, \tag{2c}$$

$$\hat{z}_t = \rho_z\hat{z}_{t-1} + \varepsilon_{z,t}\,, \tag{2d}$$

where variables with hat are deviations from steady state. The first equation is IS curve, the second is Phillips curve, the third is Taylor rule, and the last one represents technology process. Parameter τ is the inverse elasticity of intertemporal substitution in consumption, β is the discount factor, and $\kappa = \frac{1-\nu}{\nu\phi}$ where ν is the inverse elasticity of demand for intermediate goods. The shocks $\varepsilon_{\pi,t}$, $\varepsilon_{r,t}$ and $\varepsilon_{z,t}$ are Gaussian with mean zero and standard deviation σ_π, σ_r and σ_z, respectively.

The model variables are connected with observed data via measurement equations

$$YGR_t = \gamma^{ss} + \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t\,, \tag{3a}$$

$$INFL_t = \pi^{ss} + \hat{\pi}_t\,, \tag{3b}$$

$$INT_t = \pi^{ss} + 100(\beta^{-1} - 1) + \gamma^{ss} + \hat{r}_t\,, \tag{3c}$$

where γ^{ss} is steady state growth rate of technology, and the observed data are

YGR – growth rate of seasonally adjusted and adjusted by working days per capita real gross domestic product. Data are obtained from Eurostat.
INFL – CPI inflation per quarter. Data are obtained from SourceOECD Statistics.
INT – 3-month Prague InterBank Offered Rate (PRIBOR, in per cent per quarter). Data are obtained from ARAD data series system of the Czech National Bank.

We used quarterly data from 2Q1996 to 3Q2007.

3 Estimation Method

This section gives a brief description of the method used to estimate model parameters and regime switches. The detailed discussion of the theory of Markov switching state space models is presented in [3] and [4].

After solving rational expectations[1], the model given by the equations (2) and (3) is put into the state space form

$$\begin{aligned} x_t &= F(\theta,S_t)x_{t-1} + G(\theta,S_t)u_t + w_t\,,\\ y_t &= H(\theta,S_t)x_t + A(\theta,S_t)u_t + \epsilon_t\,, \end{aligned} \tag{4}$$

where $w_t \sim N(0,Q(\theta,S_t))$ and $\epsilon_t \sim N(0,R(\theta,S_t))$ is a process and measurement noise, respectively, x_t is a state vector, y_t is a measurement vector and u_t is known exogenous vector. Matrices F, G, H, A, Q and R are functions of parameter vector θ and state indicator S_t. The process of state indicators $\{S_t\}_{t=0}^T$ follows a discrete-time two state Markov process with unknown transition matrix ξ. The vector θ equals $(\theta_1',\ \theta_2')'$, where θ_1 is the vector of model parameters when state 1 arises and θ_2 consists of values of model parameters when state 2 of the system arises.

[1] We used the algorithm of C. Sims described in [8] and his Matlab code *gensys* to solve rational expectations.

We carried out Bayesian estimate of the parameter vector θ, transition matrix ξ of the Markov chain, and realization $S = (S_1, S_2 \ldots, S_T)$ of the (hidden) Markov chain based on the componentwise Metropolis-Hastings sampling. This means that we simulated draws from posterior simultaneous probability

$$p(S, \xi, \theta | \mathcal{Y}), \tag{5}$$

where $\mathcal{Y}$ is set of observed data. Because the likelihood $p(\mathcal{Y}|S, \xi, \theta)$ is invariant to relabeling of states of the Markov chain, we chose priors for S_0, ξ and θ invariant to permutation of states, performed random permutation sampling variant of the sampler mentioned above, which provided us with balanced mixture of draws from all labeling-specific subspaces, and processed the posterior draws in order to attain unique labeling rather than restricting to the subspace specified by a certain formal identifiability constraint. We had decided for permutation sampling method because an a priori chosen identifiability constraint need not guarantee unique labeling.

4 Results

We simulated 100,000 draws and threw away first half of the simulated chain to get rid of effect of initial conditions. The computation was done in Matlab. We estimated all model parameters except for discount factor β. We calibrated this parameter to value 0.995, which imposed steady state real interest rate approximately equal to $0.5+\gamma^{ss}$ per cent per quarter. Prior densities for other parameters are reported in Table 1. The priors are identical among different states. Next, we chose Dirichlet distribution[2] with parameter 3 for diagonal element and 2 for off-diagonal element for each row of the transition matrix ξ of the hidden Markov chain. The probabilities of state 1 and state 2 at the initial period were set both to 0.5.

We have estimated "sharp" distribution of the states in the period under consideration. The probability of state 1 is depicted in Figure 1. The whole period is divided into two intervals corresponding to two states of the economy. Based on the estimation results, the regime switch arose during the first quarter of the year 1999. This is the quarter after the first inflation target.[3] This means that the estimated regime switch do not coincide with monetary regime switch, which took place at the beginning of 1998. This lag can be caused by subsequent changes of another aspects of the Czech economy and/or by gradual propagation of the monetary regime switch to the structure of the economy.

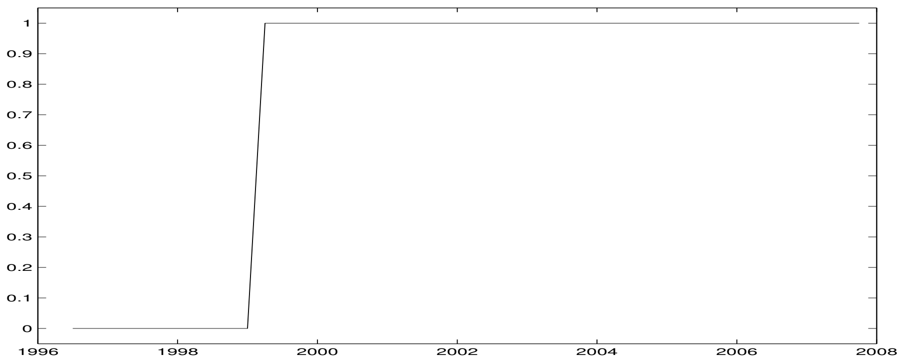


Fig. 1. Probability of state 1

The estimates of the model parameters are presented in Table 1. As can be seen, we found strong evidence of a change in steady state values. The steady state growth rate of technology decreased from 1.174 (approx. 4.7 per cent p.a.) to 0.358 per cent (approx. 1.4 per cent p.a.). The steady state inflation also decreased. It dropped from 2.025 (approx. 8 per cent p.a.) to 0.142 per cent (approx. 0.6 per cent p.a.). The steady state inflation in the second part of the period is quite low and significantly lower then inflation targets. This reflects the fact that the Czech National Bank (CNB) undershooted the inflation target systematically (see [5]). Nevertheless, the estimate is still low despite of undershooting.

[2] The Dirichlet distribution is conjugate prior in problems of this type.

[3] The first inflation target was set in December 1997 for the rate of net inflation at the end of year 1998.

Table 1. Priors and estimation results (The prior distribution functions are Beta (B), Gamma (G) and Inverse Gamma (IG).)

Param.	Prior Distribution			Posterior – State 1			Posterior – State 2		
	Dist.	Mean	Std. dev.	Mean	5%	95%	Mean	5%	95%
τ	G	2.000	0.500	2.214	1.834	2.723	3.752	3.506	4.043
κ	G	0.200	0.100	0.242	0.109	0.377	1.541	1.395	1.856
ρ_r	B	0.500	0.200	0.857	0.795	0.922	0.697	0.569	0.794
ψ_1	G	1.500	0.250	1.202	1.067	1.335	1.227	1.074	1.353
ψ_2	G	0.500	0.250	0.456	0.312	0.696	0.502	0.335	0.701
ρ_z	B	0.500	0.100	0.272	0.183	0.348	0.425	0.326	0.491
γ^{ss}	G	0.800	0.250	0.358	0.272	0.451	1.174	1.031	1.324
π^{ss}	G	1.000	0.500	0.142	0.048	0.251	2.025	1.921	2.129
σ_π	IG	1.000	0.500	1.326	1.016	1.623	2.286	1.952	2.484
σ_r	IG	0.300	0.150	0.368	0.265	0.473	0.865	0.707	0.958
σ_z	IG	1.500	0.500	1.663	1.436	1.953	2.850	2.507	3.130

The steady state values are not the only parameters which changed. The parameters influencing the shape of reaction of the Czech economy to shocks changed as well. The parameter τ decreased. This induces an increase of intertemporal elasticity of substitution in consumption and hence consumption (output in closed economy model) has become more sensitive to real interest rate. The parameter κ decreased significantly. This parameter lowers if the adjustment costs increase (parameter ϕ) or elasticity of demand for intermediate goods ($1/\tau$) decreases. Looking at estimates of τ, κ and π^{ss} and the Phillips curve (2b), the inflation seems less sensitive to output fluctuations and hence more stabilized during the second part of the period under consideration. As far as parameters of monetary policy rule are concerned, the results are ambiguous. The interest rate smoothing parameter ρ_r raised during the period but another two parameters (ψ_1 and ψ_2) remained almost unchanged.

Based on the estimates, the standard deviations of all shocks decreased in the late nineties inducing lesser volatility of the Czech economy after the structural change. The cost-push shock is the one whose influence has reduced the most if changes of parameters τ, κ and π^{ss} are taken into account (see (2b)).

5 Conclusion

This paper has aimed to detect a structural change of the Czech economy in the late nineties, its timing and shape in general aspects. It has made an inference based on Bayesian estimate of the Markov switching state space model.

The paper has found strong evidence of one structural change of the Czech economy which took place during the first quarter of the year 1999. This switch did not consist in change of steady states only but also in change in reaction of the economy to exogenous shocks. The structural change of the Czech economy did not comprise only of a change of monetary regime but also of another aspects of the economy (elasticities, degree of nominal rigidity). Besides that, volatility of the economy lessened as indicated by standard deviations of exogenous shocks included in the model.

References

1. AN, S., F. SCHORFHEIDE: Bayesian Analysis of DSGE Models. Econometric Reviews, 26, 113–172, 2007.
2. EO, Y.: Bayesian Analysis of DSGE Models with Regime Switching. ver. February 11, 2009. Available at SSRN: http://ssrn.com/abstract=1304623
3. FRÜHWIRTH-SCHNATTER, S.: Finite Mixture and Markov Switching Models. Springer, 2006, ISBN 978-0387-32909-3
4. KIM, C. J., C. R. NELSON: State-Space Models with Regime Switching. Cambridge, MIT Press, 1999, ISBN 978-0-262-11238-3.
5. ŠMÍDKOVÁ, K.: Vyhodnocení plnění inflačních cílů ČNB v letech 1998–2007. Prague, Czech National Bank, 2008, ISBN 978-80-87225-10-3
6. GEWEKE, J.: Contemporary Bayesian Econometrics and Statistics. John Wiley & Sons, Inc., 2005.
7. SAWYER, S.: The Metropolitan-Hastings algorithm and extensions. Washington University, 2006.
8. SIMS, C. A.: Solving Linear Rational Expectations Models. Computational Economics, 20, 1–20, 2001.

IS THE STABILITY OF LEVERAGE RATIOS DETERMINED BY THE STABILITY OF THE ECONOMY?[1]

Anastasiya Shamshur

Charles University, CERGE-EI
anastasiya.shamshur@cerge-ei.cz

Abstract. The choice of capital structure by firms is a fundamental issue in financial literature. According to a recent finding, the capital structure of firms remains almost unchanged during their lives meaning that leverage ratios are significantly stable over time. The stability of leverage ratios is mainly generated by an unobserved firm-specific effect that is liable for the majority of variation in capital structure [1]. However, the important fact is that the study focuses on the US economy which is relatively stable. I study how substantial changes in the economy affect the stability of firm capital structure in transition countries. Specifically, I concentrate on Central and Eastern European economies that passed through transition from central planning to market economy and privatization, Russian financial crisis, and EU membership. In addition, I investigate whether the ownership structure of the firms is responsible for the part of the unexplained variation in leverage.
Keywords. Capital structure, financing decisions, ownership.

1. Introduction.

Capital structure choice is an important decision for a firm. It is important not only from the returns maximization point of view, but also because this decision has a great impact on a firm's ability to successfully operate in a competitive environment. Current literature has suggested a number of factors that can explain about 30% of the total variation in capital structure. However, an important finding is that the capital structure of firms remains almost unchanged from their birth to death meaning that leverage ratios are significantly stable over time. The stability of leverage ratios is not affected by the process of going public, but it is mainly generated by an unobserved firm-specific effect that is responsible for the majority of variation in capital structure [1]. As Lemmon, Roberts, and Zender (2008) focus on the US economy which is relatively stable over time, their finding raises a question whether this significant stability in leverage ratios is determined by the stability of the economy the firms operate in. The impact of substantial changes in the economy on capital structure stability has not been studied yet.

In this paper I focus on countries that experienced substantial changes in all economic spheres during the transition. There are several events that could have the strong impact on firms' capital structures. These are transition from a central planning to market economy and privatization, the Russian financial crisis, and EU membership. The paper is organized as follows. In the next section I survey the literature. Section 3 describes the data sources. In section 4, I present the models and discuss the obtained results. I summarize the paper and conclude in section 5.

2. Literature.

The question about the choice of capital structure by firms is fundamental in financial literature. This literature is fairly extensive [2-4]. Scholars have identified a number of factors which are correlated with leverage. Six factors (industry median leverage, market-to-book assets ratio, tangibility, profitability, firm size and expected inflation) account for more than 27% of the variation in leverage, while another 19 factors improve the explanatory power of the model by only 2% [4]. However, traditional leverage determinants explain a minor part of the variation in leverage (at most 30%), while 60% remain unexplained [1]. This variation comes from an unobserved firm-specific time-invariant component that is responsible for persistence in leverage ratios over time. As the authors focus on the US economy, which is relatively stable, it is not clear whether leverage ratios exhibit a similar level of persistence when the economic environment rapidly changes over time. To answer this question I will refer to transition economies.

[1] I would like to thank Jan Hanousek for helpful suggestions and comments. Partial financial support from GACR grant No. 402/09/1595 is gratefully acknowledged.

There are only a few papers that attempt to study the capital structure of firms in transition economies. Some of them are concentrated on the firm-specific determinants [5, 6]. They investigate capital structure determinants and find that leverage ratios of firms in transition economies behave differently from leverage ratios of firms in Western economies. For example, asset tangibility and profitability are negatively related to leverage in transition countries, while studies on Western firms report positive relationships. Moreover, profitability and firm age are found to be the most robust determinants of target leverage ratios among transition countries [7, 8]. Some studies suggest that not only firm-specific factors matter, but also institutional and macroeconomic factors are able to explain the variation in leverage ratios [9]. However, my paper differs from existing studies on the capital structure of firms from transition economies by focusing on the question of capital structure stability and its sources.

In addition, I attempt to investigate whether the ownership structure is able to explain the part of unexplained firm-specific variation in leverage. My motivation for the inclusion of this factor into the model is based on the existing differences in ownership patterns between the US and Europe. In the US dispersed ownership prevails, while in Europe it is more concentrated. The majority ownership not only gives a right to make important strategic decisions, but also creates strong incentives to monitor managers. The controlling share owner is directly interested in firm performance and is likely take part in firm capital structure decisions. Thus, ownership structure seems to be an important determinant of firm capital structure.

3. Data.

The firm-level data is obtained from the Amadeus database constructed by Bureau Van Dijk. This database is the most comprehensive source containing financial information on public and private companies in Europe. In this study I use the Top 1 million companies and focus on ten Eastern European countries (Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia) in 1996 — 2006.

I require that all key variables have nonmissing data. I keep only firms that have leverage from zero to one interval. Firms from financial intermediation sector are excluded from the sample, since they have a specific liability structure. I exclude observations if the sum of current and non-current liabilities does not exceed the trade credit because in this case the nominator is negative[2] or if capital is negative. In addition, as I am interested in studying the firm-specific time-invariant effect, I require every firm in the sample to have a minimum 3 consecutive years of data. The construction of all of the variables used in the study is presented in the Appendix.

The resulting sample is unbalanced and the number of observations across countries varies. Estonia and the Czech Republic have the greatest coverage, while Lithuania and the Slovak Republic have the lowest. The mean leverage in all countries is in the 40% range, however, it is lower in Estonia (0.36) and higher in Hungary (0.52). The largest firms in terms of total assets are located in Poland. In terms of profitability, firms' mean return in assets is larger than their median return. This implies that firms' profitability distribution is positively skewed and most firms have low profitability, while only a few firms have very high profitability. The average age of firms in the sample is about 7 years[3].

4. Model.

4.1. The determinants of leverage in transition economies.

I study the determinants of leverage using three different models: pooled OLS, fixed effect and random effect. I start with estimating pooled OLS model [3, 4].

$$Y_{it} = \alpha + \beta X_{it-1} + \nu_t + \varepsilon_{it}, \tag{1}$$

where Y_{it} is leverage of firm i at time t; ν is a time fixed effect and ε is a random error term. The standard errors are robust to heteroskedasticity and correlation within firm [10].

As recent literature suggests that fixed effect matters and it is responsible for the majority of the variation in leverage, I also estimate the following model using fixed effect and random effect regressions.

$$\begin{aligned} Y_{it} &= \alpha + \beta X_{it-1} + \eta_i + \nu_t + u_{it}, \\ u_{it} &= \rho u_{it-1} + w_{it}, \end{aligned} \tag{2}$$

where u is stationary, w is a is a random disturbance that assumed to be possibly heteroskedastic, but serially and cross-sectionally uncorrelated, η is a firm fixed effect.

The results are reported in Table 1. The first column contains the pooled OLS regression results. The results are similar to previous works examining transition economies. The size of the firm has a positive highly

[2] According to the leverage definition; see Appendix.
[3] To save space, descriptive statistics are not reported here but available on request.

significant effect. However, it appears that tangibility, profitability and GDP growth are insignificant, while industry median leverage has a strong positive effect on leverage. Unexpectedly, the age of the firm is negatively related to the leverage ratio. On the one hand, older firms are better known on the market. They have certain reputation and lower information asymmetries, thus, it is easier for them to get debt. On the other hand, the negative sign could be due to older firms who are able to finance their operations from their internal sources and prefer to do so rather than use the external sources. Finally, the dummy for a firm's status is negatively related to leverage and highly significant, meaning that public firms tend to have lower leverage than private firms [11].

On the whole, the pooled OLS model explains only about 11% of the variation in leverage. Fixed and random effect models perform better; however, the random effect model is rejected in favor of the fixed effect model using the Hausman specification test. All the determinants are statistically significant despite tangibility, which is measured imprecisely. Larger firms tend to have higher leverage because they more diversified and face lower bankruptcy risk. Firms also prefer to borrow more when growth opportunities are high as predicted by the pecking order theory. The estimated relation between leverage level and profitability is positive. This finding is consistent with agency theory, which says that firms use more debt to avoid misuse of the free cash flow by managers [12]. Moreover, higher debt could be served as a tax shield. I also find that in the fixed effect model age is positively related to leverage. Older firms appear to be more leveraged because they have a well-established reputation which guarantees them easier access to debt.

Table 1. Determinants of Leverage in Transition Economies

Variable	Book Leverage					
	Pooled OLS		Fixed Effect		Random Effect	
Log(Total Assets)	0.010***	(0.0005)	0.002***	(0.0004)	0.004***	(0.0003)
Tangibility	0.001	(0.0008)	0.0001	(0.0002)	0.0001	(0.0002)
Profitability	-0.002	(0.0014)	0.0004**	(0.0002)	0.0003	(0.0002)
GDP growth	0.001	(0.0004)	0.002***	(0.0003)	-0.001***	(0.0002)
Industry median	0.663***	(0.0103)	0.460***	(0.0094)	0.586***	(0.0055)
Expected inflation	0.001**	(0.0004)	0.001***	(0.0003)	-0.0002	(0.0002)
Age	-0.002***	(0.0001)	0.006***	(0.0004)	-0.002***	(0.0001)
cons	0.017	(0.0112)	0.078***	(0.0025)	0.195***	(0.0063)
Hausman			389.38	(0.0000)		
Adjusted R^2	0.1072					
AR(1)			0.532		0.532	
Obs	313537		238296		313596	

Although the fixed effect model has a statistical advantage over the random effect and the pooled effect models because it has the highest adjusted R^2, there is a certain threat that fixed effect estimation kills all of the cross-sectional variation and evaluates the time-series variation in the data. However, this is acceptable when the cross-sectional and the time-series impacts are equal, but if this assumption is violated, fixed effect estimates are not able to capture the total impact of the leverage factors [13].

4.2. Ownership structure of the firm as a determinant of firm capital structure.

I started with the question of how much variation in firms' leverage is firm-specific and time-invariant in transition economies. To answer this question I run the regression of leverage on firms‘ fixed effects. The adjusted R^2 from this regression is about 65% which is larger than in the US. Then, I conduct a sensitivity analysis by considering only firms that have at least 5, 7, and 10 years of non-missing data for book assets. I find that the variation explained by firm fixed effects approaches 60% as the number of years available increases. However, it is quite surprising that even in rapidly changing transition economies, the fixed effect is responsible for the same or even a larger part of the variation in leverage. As the majority of unexplained variation comes from unobserved time-invariant firm characteristics, I suggest looking at such firm-specific leverage determinants in European countries as ownership of the firm. I distinguish between three ownership categories which are majority ownership (>50%), blocking minority ownership (>33%, but ≤50%) and legal minority ownership (>10%, but ≤33%).

The majority ownership not only gives a right to make important strategic decisions, but also creates strong incentives to monitor managers. The controlling share owner is directly interested in firm performance and is likely to choose following low-risk strategies which result in lower leverage levels. In addition, blocking minority ownership grants a right to block a number of decisions concerning major changes in the business activity of the firm. For example, blocking minority owners are able to block the decisions of major shareholder concerning changes in assets and the firm's activities. Finally, legal minority ownership gives the possibility to delay or completely block the implementation of larger shareholders decisions through lengthy court proceedings [14]. Thus, ownership structure has a potential to be an important determinant of firm capital structure.

The direct ownership data are available starting from 2004. Firms controlled by major owners are largest in terms of total assets. In fact, median total assets are significantly lower compared to their mean value. This fact suggests that total assets are positively skewed. In other words, total assets of most firms are low, while total assets of few firms are high. However, in terms of profitability, tangibility and leverage level there are no significant differences with respect to ownership concentration.

Table 2 reports the results of leverage regression that accounts for ownership structure of the firm. The first column contains estimates from the pooled OLS regression. In the second column I add ownership concentration variables to traditional leverage determinants. As can be seen from Table 2, it only slightly increases the explanatory power of the model. All the estimates stay approximately the same. At the same time, majority and legal minority ownership appear to affect the level of leverage. In the third column in addition to ownership concentration I also distinguish between foreign and domestic ownership. Only majority and legal minority domestic ownership estimates are significant and positively related to leverage. Thus, ownership structure of the firm could be considered as one of the capital structure determinants. However, it does not increase substantially the explanatory power of the model.

Table 2. Leverage and Direct Ownership

Variable	Book leverage					
	(1)		(2)		(3)	
Log(Total Assets)	-0.004***	(0.001)	-0.004***	(0.001)	-0.004**	(0.001)
Tangibility	0.0003*	(0.0001)	0.0003*	(0.0001)	0.0003*	(0.0001)
Profitability	-0.082***	(0.026)	-0.082***	(0.026)	-0.08***	(0.026)
GDP growth	0.008***	(0.002)	0.008**	(0.002)	0.007***	(0.002)
Industry median	0.562***	(0.028)	0.561***	(0.028)	0.559***	(0.028)
Expected inflation	0.009***	(0.004)	0.009***	(0.004)	0.009**	(0.004)
Age	-0.001***	(0.0002)	-0.001***	(0.0002)	-0.001***	(0.0002)
Quoted	-0.004	(0.017)	-0.003	(0.017)	-0.008	(0.019)
Majority			0.01*	(0.006)		
Monitored Majority			0.004	(0.007)		
Minority			0.01*	(0.006)		
Majority*foreign					0.009	(0.008)
Majority*domestic					0.011*	(0.006)
Monitored Majority*foreign					-0.029	(0.050)
Monitored Majority*domestic					0.004	(0.007)
Minority*foreign					0.001	(0.022)
Minority*domestic					0.011*	(0.006)
cons	0.208***	(0.029)	0.199***	(0.030)	0.205***	(0.030)
Adjusted R^2	0.1041		0.1043		0.1042	
Obs	38521		38521		37688	

5. Conclusion.

Using a comprehensive database of firms in transition countries, I examine the determinants of capital structure. First, I check how much variation in leverage is explained by traditional determinants. I took the determinants which were previously identified as relevant for both developed and transition economies. It appears that a number of core determinants are able to explain about 11% of the variation in leverage. This percent is low mostly because the majority of firms in the sample is unlisted. Listed companies are about 1% of the entire sample. The obtained coefficient estimates are in line with estimates reported in earlier studies in transition economies [9, 6]. An interesting finding is that tangibility and profitability are weak determinants of firm capital structure. At the same time, age is a highly significant determinant of capital structure in transition economies. Different models estimate different signs for this determinant, but a positive relation between leverage and age estimated by the fixed effect model is more intuitive. Older firms have a well-established reputation on the market and are able to borrow at lower cost. As a result, they have higher leverage. Moreover, the fixed effect model accounts for unobserved firm-specific time-invariant factors which are responsible for about 65% of the variation in leverage in transition economies.

Second, I investigate whether the ownership structure of the firm is able to explain at least a part of the unexplained variation in leverage. This determinant has a certain potential because major owner is directly interested in firm's performance. He has control over firms and can directly monitor and replace the management. I control for majority, monitored majority and minority ownership. Majority and minority ownership appear to be significant determinants of firm's leverage, however, they do not improve the explanatory power of the model substantially. I also distinguish between foreign and domestic owners. The estimated coefficients of ownership concentration dummies appear to be insignificant for foreign owners, while estimates of majority and minority domestic ownership are significant.

As I mentioned earlier, Eastern European firms operate in an environment that is completely different from that of US firms, but firm fixed effect is responsible for even a larger part of the variation in leverage. It seems to be not an effect of the stable economic environment firms operate in, but rather purely a firm-specific characteristic.

Appendix.

$Leverage = \frac{debt}{debt + equity}$, where $debt = total\ liabilities - trade\ credit$ [7].

GDP *growth* is a proxy for growth opportunities of the firm.

$Age = Age_t - Year\ of\ \ incorporation$

Log(total assets) is the natural log of total assets.
Majority ownership=1, firm is controlled by a majority owner, the rest are hold less than 10%.
Monitored minority ownership=1, despite of the majority owner at least one minority owner is present.
Minority ownership=1, either blocking or legal minority owner is the largest owner.
Dispersed ownership=1, all shareholders hold less than 10% of equity.

References:

[1] LEMMON, M., ROBERTS, M., ZENDER, J.: *Back to the beginning: Persistence and the cross-section of corporate capital structure*. Journal of Finance 63(4) (2008) 1575-1608

[2] TITMAN, S., WESSELS, R.: *The determinants of capital structure choice*. Journal of Finance 43 (1988) 1-21

[3] RAJAN, R., ZINGALES, L.: *What do we know about capital structure: Some evidence from international data.* Journal of Finance 50 (1995) 1421-1460

[4] FRANK, M., GOYAL, V.: *Capital structure decisions: Which factors are reliably important*? Working paper, University if Minnesota and Hong Kong University of Science and Technology (2007a)

[5] CORNELLI, F., PORTES, R., SCHAFFER, M.: *The capital structure of firms in Central and Eastern Europe*. CEPR Discussion Paper 1392 (1996).

[6] DELCOURE, N.: *The determinants of capital structure in transitional economies*. International Review of Economics and Finance 16 (2007) 400-415.

[7] HAAS, R., PEETERS, M.: *The dynamic adjustment towards target capital structure of firms in transition economies.* Economics of Transition 14(1) (2006) 133-169.

[8] NIVOROZHKIN, E.: *Financing choices of firms in EU accession countries.* Emerging Markets Review 6 (2005) 138-169

[9] JOEVEER, K.: *Sources of capital structure: Evidence from transition countries.* CERGE-EI Working Paper 306 (2006)

[10] PETERSEN, M.: *Estimating standard errors in finance panel data sets: Comparing approaches.* The Review of Financial Studies 22(1) (2009) 435-480

[11] BRAV, O.: *Access to capital, capital structure, and the funding of the firm.* The Journal of Finance 64(1) (2009) 263-308

[12] JENSEN, M.: *Agency costs of free cash flow, corporate finance, and takeovers*. American Economic Review 76 (1986) 323-329

[13] BARANCHUK, N., XU, Y.: *On the persistence of capital structure - reinterpreting what we know.* Unpublished (2009)

[14] HANOUSEK, J., KOCENDA, E., SVEJNAR, J.: *Origin and concentration: Corporate ownership, control and performance in firms after privatization*. Economics of Transition 15(1) (2007) 1-31

[15] BENA, J., HANOUSEK, J.: *Rent extraction by large shareholders evidence using dividend policy in the Czech Republic*? Finance a Uver: Czech Journal of Economics and Finance 58(3) (2008) 106-130.

OPTIMISATION OF PRODUCTION PROCESS IN FOOD-PROCESSING COMPANY

Veronika Skočdopolová, Josef Jablonský

University of Economics, Prague
Faculty of Informatics and Statistics, Department of Econometrics
veronika.skocdopolova@vse.cz, jablon@vse.cz

Abstract. The paper presents the model for a real-life problem that consists in optimisation of production process of white masses (WMs) for one of the global world companies. The problem is formulated as a standard blending problem with several special features. The optimisation criterion is minimisation of total costs of the production process. The most important constraints are requirements on composition of WMs, composition of raw materials and limits for surplus sales of materials. The results of the model offer recommendation for purchase of raw materials, own production of materials, and surplus sales of raw materials. We have used MS Excel interface cooperating with LINGO solver for its solving.
Keywords. *blending problem, food-processing industry, optimisation, LINGO*

1. Introduction

We were addressed by a food-processing company to set up a model that would optimise protein costs. The company is intent on production of diary products such as yoghurts, acidified milk drinks etc. They would like to optimise the purchase of raw milk (RM) and other dairy ingredients such as cream (CR), skimmed milk (SM), concentrated milk (SMC), powdered milk (SMP) or anhydrous milk fat (AMF). The aim of the model is to blend different white masses (WMs) according to recipes using different ingredients and to achieve the lowest costs as possible due to constraints. The problem of production WMs is a standard blending problem with several special features that can be formulated as a linear programming problem.

2. Mathematical model

The optimisation criterion is the minimization of total costs of the production process. They consist of costs of input supplies and costs of processing reduced by revenues from surplus sales. The constraints that must be fulfilled during the optimisation follows from requirements. The most important ones are as follows:

- requirements on the composition of WMs (fat and protein content, total weight of milk, limits of protein and fat contents from different sources, etc.),
- composition of ingredients (RM, CR, SM, etc.) with respect to fat, protein, available amount and price of ingredients,
- limits for surplus sales of RM and ingredients, unit prices of the sales, etc.

We have to take into consideration that raw milk cannot be used as an ingredient, but it is used for producing cream and skimmed milk. Produced and purchased skimmed milk can be used for producing concentrated skimmed milk. The company has contracts with some of the vendors to purchase certain amount of raw milk or ingredients as well as contracts with customers to resell certain amount of raw milk, cream or skimmed milk.

The results of the optimisation offer recommendation for purchase of raw milk and ingredients, own production of ingredients, surplus sales, composition of WMs, etc.

The complete model is as follows. Each part of the model is described in detail below.

Minimize

$$z = \sum_i PURC_i Rprice_i + \sum_k RM_k Mprice_k + prodRM\left(\sum_k RM_k - \sum_k spRM_k\right) + \\ + prodSMC \sum_{i \in SM} SMC1_i - \left(\sum_p priRM_p srpRM_p + \sum_q priCR_q srpCR_q + \sum_r priSM_r srpSM_r\right) \quad \textbf{(2.1)}$$

subject to

$$\begin{aligned} &Rprot_1 CR1_k + Rprot_8 SM1_k + Dprot_k^- - Dprot_k^+ = Mprot_k(RM_k - spRM_k), \quad \forall k, \\ &Rfat_1 CR1_k + Rfat_8 SM1_k = Mfat_k(RM_k - spRM_k), \quad \forall k, \\ &CR1_k + SM1_k = RM_k - spRM_k, \quad \forall k, \end{aligned} \quad \textbf{(2.2)}$$

$$\begin{aligned} &Dprot_k^+ \le deviation1 \cdot Mprot_k(RM_k - spRM_k), \quad \forall k, \\ &Dprot_k^- \le deviation1 \cdot Mprot_k(RM_k - spRM_k), \quad \forall k, \end{aligned} \quad \textbf{(2.3)}$$

$$PURC_i = \sum_j x_{ij}, \qquad i \ge 2,\ i \in CR, \quad \textbf{(2.4)}$$

$$\begin{aligned} &PURC_1 = \sum_k CR1_k, \\ &PURC_8 = \sum_k SM1_k, \end{aligned} \quad \textbf{(2.5)}$$

$$\begin{aligned} &\sum_j x_{1j} + \sum_q srpCR_q = PURC_1, \\ &\sum_j x_{8j} + \sum_r srpSM_r = PURC_8 - useSMC_8, \end{aligned} \quad \textbf{(2.6)}$$

$$\sum_k spRM_k = \sum_p srpRM_p, \quad \textbf{(2.7)}$$

$$\begin{aligned} &PURC_i = useSMC_i + useWM_i, \quad \forall i \in SM, \\ &\sum_j x_{ij} = useWM_i, \quad \forall i \in SM, \end{aligned} \quad \textbf{(2.8)}$$

$$\begin{aligned} &Rprot_i useSMC_i \ge Rprot_{17} SMC1_i, \quad \forall i \in SM, \\ &Rfat_i useSMC_i \ge Rfat_{17} SMC1_i, \quad \forall i \in SM, \end{aligned} \quad \textbf{(2.9)}$$

$$\begin{aligned} &\sum_{i \in SM} SMC1_i = \sum_j x_{17,j}, \\ &PURC_{17} = \sum_{i \in SM} SMC1_i, \end{aligned} \quad \textbf{(2.10)}$$

$$\begin{aligned} &PURC_i = \sum_j x_{ij}, \quad \forall i \in SMC, \\ &PURC_i = \sum_j x_{ij}, \quad \forall i \in AMF, \\ &PURC_i = \sum_j x_{ij}, \quad \forall i \in SMPW, \end{aligned} \quad \textbf{(2.11)}$$

$$PURC_i = \sum_j x_{ij} + \sum \frac{x_{i+6,j}}{10{,}5}, \qquad \forall i \in SMP,$$

$$\sum_j x_{ij} = useSMPW_i, \qquad \forall i \in SMPW, \tag{2.12}$$

$$\sum_j x_{ij} + \frac{useSMPW_{i+6}}{10{,}5} \le ubR_i, \qquad \forall i \in SMP,$$

$$\sum_i x_{ij} + dev1_j - dev2_j = WMweig_j WMquant_j, \qquad \forall j, \tag{2.13}$$

$$dev1_j + dev2_j \le 0{,}01(WMweig_j WMquant_j), \qquad \forall j,$$

$$\sum_i Rfat_i x_{ij} - fatDEV_j^+ = WMfat_j WMquant_j, \qquad \forall j, \tag{2.14}$$

$$\sum_i Rprot_i x_{ij} - protDEV_j^+ = WMprot_j WMquant_j, \qquad \forall j, \tag{2.15}$$

$$\sum_{i \in AMF} Rfat_i x_{ij} \le AMF\max_j WMfat_j WMquant_j, \qquad \forall j,$$

$$\sum_{i \in SMC} Rprot_i x_{ij} \le SMC\max_j WMprot_j WMquant_j, \qquad \forall j, \tag{2.16}$$

$$\sum_{i \in SMP} Rprot_i x_{ij} \le SMP\max_j WMprot_j WMquant_j, \qquad \forall j,$$

$$\sum_{i \in SM} Rprot_i x_{ij} \le SM\max_j WMprot_j WMquant_j - \sum_{i \in SMPW} Rprot_i x_{ij}, \qquad \forall j,$$

$$lbM_k \le RM_k \le ubM_k, \qquad \forall k, \tag{2.17}$$

$$lbR_i \le PURC_i \le ubR_i, \qquad \forall i, \tag{2.18}$$

$$lbRM_p \le srpRM_p \le ubRM_p, \qquad \forall p,$$

$$lbCR_q \le srpCR_q \le ubCR_q, \qquad \forall q, \tag{2.19}$$

$$lbSM_r \le srpSM_r \le ubSM_r, \qquad \forall r,$$

$$0 \le fatDEV_j^+ \le deviation \cdot WMfat_j WMquant_j, \qquad \forall j, \tag{2.20}$$

$$0 \le protDEV_j^+ \le deviation \cdot WMprot_j WMquant_j, \qquad \forall j,$$

where the index sets mean

- j – white masses that have to be produced,
- i – ingredients that can be purchased, this index set has six subsets (*CR*, *SM*, *SMC*, *SMP*, *SMPW*, *AMF*) according to different types of ingredients, each subset contents different vendor of the ingredients,
- k – raw milks that can be purchased from k different vendors,
- p – customers that buy the surplus of raw milk,
- q – customers that buy the surplus of cream,
- r – customers that buy the surplus of skimmed milk.

The variables of the model are as follows:

- *RMk – amount of raw milk that should be purchased from vendor k,*
- *spRMk – amount of raw milk from vendor k that is not used for ingredients production and is resold as surplus,*
- *SMC1i – amount of concentrated skimmed milk produced from skimmed milk i,* $i \in SM$,
- *PURCi – amount of ingredient i that should be purchased (or produced),*

- *srpRMp – surplus of raw milk that can be sold to customer p,*
- *srpCRq – surplus of cream that can be sold to customer q,*
- *srpSMr – surplus of skimmed milk that can be sold to customer r,*
- *CR1k – amount of cream produced from raw milk,*
- *SM1k – amount of skimmed milk produced from raw milk,*
- *Dprot–k – negative deviation of protein in protein balance of raw milk k,*
- *Dprot+k – positive deviation of protein in protein balance of raw milk k,*
- *xij – amount of ingredient i in white mass j,*
- *useSMCi – amount of skimmed milk used in concentrated skimmed milk production,* $i \in SM$*,*
- *useWMi – amount of skimmed milk used for white masses production,* $i \in SM$*,*
- *dev1j – negative deviation of milk weight in white mass j,*
- *dev2j – positive deviation of milk weight in white mass j,*
- *fatDEV+j – positive deviation of fat in fat balance in white mass j,*
- *protDEV+j – positive deviation of protein in protein balance in white mass j,*
- *useSMPWi – amount of reconstituted powdered milk i.*

The input parameters of the model can be divided into two groups. The parameters given by white masses recipes are as follows:

- *AMFmaxj – maximal % of total fat content in WM i can be from anhydrous milk fat,*
- *SMCmaxj – maximal % of total protein content in WM i can be from concentrated milk,*
- *SMPmaxj – maximal % of total protein content in WM i can be from powdered milk,*
- *SMPWmaxj – maximal % of total protein content in WM i can be from reconstituted powdered milk,*
- *SMmaxj – maximal % of total protein content in WM i can be from skimmed milk,*
- *WMfatj – fat content in kg/100kg of white mass,*
- *WMprotj – protein content in kg/100kg of white mass,*
- *WMquantj – quantity of white mass i produced in chosen period,*
- *WMweigj – total weight of milk in kg/100kg of white mass.*

The other group of input parameters follows from parameters of the purchased ingredients. These are:

- *Mpricek – price in local currency per 1kg of raw milk from vendor k,*
- *Mprotk – % of protein in raw milk from vendor k,*
- *Mfatk – % of fat in raw milk from vendor k,*
- *Rpricei – price in local currency per 1kg of ingredient i,*
- *Rproti – % of protein in ingredient i,*
- *Rfati – % of fat in ingredient i,*
- *prodRM – production cost of processing of 1kg of raw milk,*
- *prodSMC – production cost of processing 1kg of concentrated milk from skimmed milk,*
- *priRMp, priCRq, priSMr – price of raw milk (cream, skimmed milk) resold to customer p (q,r),*
- *lbRMp, lbCRq, lbSMr – amount of RM (CR, SM) that has to be resold to customer p (q,r),*
- *ubRMp, ubCRq, ubSMr – maximal amount of RM (CR, SM) that can be resold to customer p (q,r),*
- *lbMk – amount of raw milk that has to be purchased from vendor k,*
- *ubMk – maximal amount of raw milk that can be purchased from vendor k,*
- *lbRi – amount of ingredient i that has to be purchased,*
- *ubRi – maximal amount of ingredient i that can be purchased,*
- *deviation – % deviation of fat or protein in each white mass,*
- *deviation1 – % deviation of protein in cream and skimmed milk production.*

The objective function (2.1) consists of four parts that are costs of ingredients purchase, costs of raw milk purchase, and costs of production of cream, skimmed milk and concentrated milk reduced by revenues from sales of raw milk, cream and skimmed milk surpluses. Relation (2.2) presents balance of protein and fat in production of cream and skimmed milk from raw milk. The balance of fat have to be fulfilled exactly, in the balance of protein there is allowed a deviation. The total amount of produced cream and skimmed milk has to be equal to the amount of purchased raw milk reduced by resold raw milk. Constraints (2.3) show the limits of positive and negative deviation of protein in the balance of protein in raw milk processing.

It follows from (2.4) that the amount of purchased cream is equal to the sum of cream used for each white mass. The "purchased" amount of produced cream and skimmed milk is equal to the sum of cream (or skimmed milk) produced from each raw milk (2.5). Constraints (2.6) presents that the "purchased" amount of produced cream is a sum of cream used for white masses production and of cream resold to customers. The "purchased" amount of produced skimmed milk reduced by amount of SM used for concentrated milk production is a sum of SM used for white masses production and of SM resold to customers. (2.7) shows that the sum of surpluses of all raw milks is equal to the sum of surpluses resold to all customers. The purchased amount of skimmed milk consists of SM used for concentrated milk (SMC) production and of SM used for WM production (2.8). Balances of protein and fat in SMC production are evident in (2.9). In (2.10) there is visible that the "purchased" amount of produced concentrated milk is equal to sum of SMC used for white mass production.

Constraints (2.11) show that the amounts of purchased SMC (AMF, reconstituted SMP) are equal to the sum of SMC (AMF, reconstituted SMP) used for each white mass. Constraints (2.12) relate to powdered (SMP) and reconstituted powdered milk (SMPW). The purchased amount of SMP consists of sum of the SMP and SMPW used for white mass production. Reconstituted powdered milk is adulterated powdered milk in the proportion 1:10,5. The sum of the SMP used for white mass production as powder and SMP in reconstituted SMP have to be lower or equal the maximal available amount of SMP. The deviation of milk in each white mass (2.13) must be lower than 1%. (2.14) and (2.15) are fat and protein balances in each white mass. Constraints (2.16) represent the maximal limits of total fat content in each white mass that can be from anhydrous milk fat or total protein content from concentrated milk, powdered milk and skimmed milk.

Last constraints are bounds. (2.17) are bounds of purchased amounts of raw milk and (2.18) are bounds of purchased amounts of each ingredient. Limits for surpluses of raw milk, cream or skimmed milk are represented by constraint (2.19). Bounds (2.20) are related to positive fat and protein deviation in fat and protein balances in each white mass.

3. Computer implementation and conclusions

For solving this model specialized optimisation solvers can be used. The company requested a system that enables repeated solving of the problem with different settings of parameters. Therefore we should create a system that could be used by people, who doesn't know anything about mathematical modelling. Due to the necessity to ensure user-friendly environment of the system we have used MS Excel interface cooperating with LINGO solver. MS Excel contains its own Solver for working with linear optimisation problems. Nevertheless this Solver can work with size-limited problems only and therefore is not applicable for this problem. That is why the system uses an external high quality professional solver LINGO (see www.lindo.com). This solver is not part of the application and has to be installed on the PC before using of the system.

Mathematical model is written in LINGO modelling language but the user doesn't get in touch with it. Users work only with MS Excel worksheets. In the MS Excel workbook there are two sheets with tables for input parameters such as fat or protein content of ingredients, prices, recipes of white masses etc. The optimisation is started from MS Excel. Results of the optimisation are entered in two other MS Excel sheets in the same workbook.

The system is now ready for operational using. The current version enables simultaneous optimisation of 30 white masses using 7 different raw milk vendors, 6 cream vendors (plus produced cream), 8 skimmed milk vendors (plus produced SM), 6 concentrated milk vendors (plus produced SMC), 6 powdered milk vendors and 3 vendors of anhydrous milk fat. There is also possibility to resell surpluses of raw milk, cream and skimmed milk to 6 different customers each. The total number of variables in the current version is 1491 and the total number of constraints is 434.

The results of this optimisation are recommendations for raw milk and ingredients purchase, for cream, skimmed milk and concentrated milk production and for raw milk, cream and skimmed milk surpluses resell. It also offers compositions of white masses. The model is quite sensitive to parameters. There is no feasible solution when any of the deviations is not allowed. The necessity of using deviations follows from deviations during measuring of fat or protein content in raw milk and ingredients.

Acknowledgements

The research is supported by the Grant Agency of the Czech Republic – grants no. 402/09/H045 and 402/09/0231.

References

[1] JABLONSKÝ, J. *Operations research: quantitative models for economic decision making (in Czech).* Professional Publishing, Praha, 2007. ISBN 978-80-86946-44-3.

[2] PELIKÁN, J. *Discrete models in operational research (in Czech).* Professional Publishing, Praha, 2001. ISBN 80-86419-17-7.

[3] SCHRAGE, L. *Optimization modeling with LINGO*. Lindo Systems,Inc., Chicago 1999. ISBN 1-893355-00-4.

[4] WALKENBACH, J.: *Microsoft Excel 2003 Programming in VBA (in Czech).* Computer Press, a.s., Brno 2006. ISBN 80-251-0911-9

RAMSEY GROWTH MODEL UNDRE UNCERTAINTY

Karel Sladký

Department of Econometrics
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
sladky@utia.cas.cz

Abstract. We consider an extended version of the Ramsey growth model under stochastic uncertainties modelled by Markov processes. In contrast to the standard model we assume that splitting of production between consumption and capital accumulation is influenced by some random factor, e.g. governed by transition probabilities depending on the current value of the accumulated capital, along with possible additional interventions of the decision maker. Basic properties of the standard formulation are summarized and compared with their counterparts in the extended version. Finding optimal policy of the extended model can be either performed by additional compensation of the (random) disturbances or can be also formulated as finding optimal control of a Markov decision process.
Keywords. Economic dynamics, the Ramsey growth model under uncertainty, Markov decision processes, optimization.
JEL Classification: C61, E21, E22

1 Classical Ramsey Growth Model

The heart of the seminal paper of F. Ramsey [6] on mathematical theory of saving is an economy producing output from labour and capital and the task is to decide how to divide production between consumption and capital accumulation to maximize the global utility of the consumption. Ramsey's original results from 1928 were revisited and significantly extended only after almost thirty years and at present the Ramsey model can be considered as one of the three most significant tools for the dynamic general equilibrium model in modern macroeconomics. In [6] the problem was considered in continuous-time setting, Ramsey suggested some variational methods for finding an optimal policy how to divide the production between consumption and capital accumulation. However, in the recent literature on economic growth models (see e.g. Le Van and Dana [2] or Majumdar, Mitra, and Nishimura [4]) the discrete-time formulation is preferred.

The Ramsey growth model in the discrete-time setting can be formulated as follows:

We consider at discrete time points $t = 0, 1, \ldots,$ an economy in which at each time t there are L_t (merely identical) consumers with consumption c_t per individual. The number of consumers grow very slowly in time, i.e. $L_t = L_0(1+n)^t$ for t with $\alpha := (1+n) \approx 1$. The economy produces at time t gross output Y_t using only two inputs: capital K_t and labour $L_t = L_0(1+n)^t$. A production function $F(K_t, L_t)$ relates input to output, i.e.

$$Y_t = F(K_t, L_t) \quad \text{with} \quad K_0 > 0,\ L_0 > 0 \quad \text{given.} \tag{1}$$

We assume that that $F(\cdot, \cdot)$ is a homogeneous function of degree one, i.e. $F(\theta K, \theta L) = \theta F(K, L)$ for any $\theta \in \mathbb{R}$.

The output must be split between consumption $C_t = c_t L_t$ and gross investment I_t, i.e.

$$C_t + I_t \le Y_t = F(K_t, L_t). \tag{2}$$

Investment I_t is used in whole (along with the depreciated capital K_t) for the capital at the next time point $t+1$. In addition, capital is assumed to depreciate at a constant rate $\delta \in (0,1)$, so capital related to gross investment at time $t+1$ is equal to

$$K_{t+1} = (1-\delta)K_t + I_t. \tag{3}$$

Preferences over consumption of a single consumer (resp. the considered L_t consumers) for the discount factor $\beta \in (0,1)$ and the considered time horizon T are expressed by means of utility function $u(\cdot)$ as

$$U^{\beta}(k_0,T) = \sum_{t=0}^{T} \beta^t u(c_t) \quad (\text{resp. } \bar{U}^{\beta}(k_0,T) = L_0 \sum_{t=0}^{T} (\alpha\beta)^t u(c_t)). \tag{4}$$

The problem is to find the rule how to split production between consumption and capital accumulation that maximizes global utility $U^{\beta}(k_0,T)$ of the consumers for a finite or infinite time horizon T.

In what follows let $k_t := K_t/L_t$ be the capital per consumer at time t, and similarly let $y_t := Y_t/L_t$ be the per capita output at time t. Recalling that the production function $F(\cdot,\cdot)$ is assumed to be homogeneous of degree one, then $f(k_t) := F(k_t,1)$ denotes the per capita production per unit time. In virtue of (2), (3) we get

$$c_t + (1+n)k_{t+1} - (1-\delta)k_t \le y_t = f(k_t), \tag{5}$$

and if we set for simplicity $\alpha \equiv (1+n) = 1$ then (5) can be written as

$$c_t + k_{t+1} + (1-\delta)k_t \le y_t = f(k_t). \tag{6}$$

In the above formulation we assume that the per capita production function $f(k)$ and the consumption function $u(c)$ fulfil some standard assumptions on production and consumption functions, in particular, that:

AS 1. The function $u(c): \mathbb{R}^+ \to \mathbb{R}^+$ is twice continuously differentiable and satisfies $u(0) = 0$. Moreover, $u(c)$ is strictly increasing and concave (i.e., its derivatives satisfy $u'(\cdot) > 0$ and $u''(\cdot) < 0$) with $u'(0) = +\infty$ (so-called Inada Condition).

AS 2. The function $f(k): \mathbb{R}^+ \to \mathbb{R}^+$ is twice continuously differentiable and satisfies $f(0) = 0$. Moreover, $f(k)$ is strictly increasing and concave (i.e., its derivatives satisfy $f'(\cdot) > 0$ and $f''(\cdot) < 0$) with $f'(0) = M < +\infty$, $\lim_{k\to\infty} f'(k) < 1$.

Since $u(\cdot)$ is increasing (cf. assumption AS 1) in order to maximize global utility of the consumers is possible to replace (6) by the (nonlinear) difference equation

$$k_{t+1} + (1-\delta)k_t - f(k_t) = -c_t \quad \text{with } k_0 > 0 \text{ given} \tag{7}$$

or equivalently for $\tilde{f}(k) := f(k) - (1-\delta)k$ by

$$k_{t+1} - \tilde{f}(k_t) = -c_t \quad \text{with } k_0 \text{ given}, \tag{8}$$

where c_t $(t = 0,1,\ldots)$ with $c_t \in [0, f(k_{t-1})]$ is selected by the decision maker. Since the considered system is purely deterministic the initial capital k_0 along with the control policy c_t fully determines development of (k_t, c_t) over time. In particular, (cf. (4), (7)) for a given initial capital k_0 policy c_t is optimal for a finite or infinite time horizon T if the global utility

$$U^{\beta}(k_0,T) = \sum_{t=0}^{T} \beta^t u(c_t) = \sum_{t=0}^{T} \beta^t u(\tilde{f}(k_t) - k_{t+1}) \tag{9}$$

attains maximum for policy $\hat{c}_t = u(\tilde{f}(\hat{k}_t) - \hat{k}_{t+1})$, i.e. *the value function*

$$\hat{U}^{\beta}(k_0,T) = \max_{(\boldsymbol{k},\boldsymbol{c})} \sum_{t=0}^{T} \beta^t u(\tilde{f}(k_t) - k_{t+1}) \quad \text{where } (\boldsymbol{k},\boldsymbol{c}) = \{k_0, c_0, k_1, c_1, \ldots, k_T, c_T\}. \tag{10}$$

Moreover, since the performance function is separable (in particular, additive) by the well-known "principle of optimality" of dynamic programming we immediately conclude that for any time point $\tau = 0,1,\ldots$ it holds

$$\hat{U}^{\beta}(k_0,T) = \sum_{t=0}^{\tau-1} \beta^t u(\hat{c}_t) + \beta^{\tau} \hat{U}^{\beta}(k_{\tau}, T-\tau) \tag{11}$$

$$\hat{U}^{\beta}(k_{\tau}, T-\tau) = \max_{k_{\tau+1}} [u(\tilde{f}(\hat{k}_{\tau}) - k_{\tau+1}) + \beta \hat{U}^{\beta}(k_{\tau+1}, T-\tau-1)]. \tag{12}$$

The following simple facts may be useful for better understanding the development of the considered economy over time.

Fact 1. i) If $f'(0) \le 1$ (and hence $f'(k) < 1$ for all $k > 0$), then by AS 2 every sequence $\{k_0, k_1, \dots, k_t, \dots\}$ must be decreasing and $\lim_{t\to\infty} k_t = 0$.

ii) If $f'(0) > 1$ (and hence, since $\lim_{k\to\infty} f'(k) < 1$, there exists some k' such that $f'(k) < 1$ for all $k > k'$), then there exists some $k^* > 0$ such that $f(k^*) = k^*$ and some $k_m \in (0, k^*)$ such that $f(k_m) - k_m = \max_k[f(k) - k]$.

iii) Supposing that $k_0 > k^*$ then elements of any sequence $\{k_0, k_1, \dots, k_t, \dots\}$ must be decreasing for all $k_t > k^*$. Furthermore, if for some $t = t_\ell$ it holds $k_{t_\ell} < k^*$ then $k_t < k^*$ for all $t \ge t_\ell$, but $\{k_t, t \ge t_\ell\}$ need not be monotonous. However, in any case $k_t \le k_{\max} = \max(k_0, k^*)$ and $f(k_t) \le f(k_{\max}) =: y_{\max}$ for all $t = 0, 1, \dots$.

In what follows we summarize some basic properties of value functions and sketch the corresponding proofs (for details see e.g. [2]).

Result 1. *Properties of the value function $\hat{U}^\beta(k_0, T)$.*

i) *(Monotonicity of value function at initial condition.)*

In case that $k_0' > k_0 > 0$ then it holds $\hat{U}^\beta(k_0', T) \ge \hat{U}^\beta(k_0, T)$.

ii) *(Continuity and differentiability of value function $\hat{U}^\beta(k_0, T)$.)*

$\hat{U}^\beta(k_0, T)$ is differentiable function at initial condition k_0.

iii) *(Concavity and continuity of value function $\hat{U}^\beta(k_0, T)$.)*

$\hat{U}^\beta(k_0, T)$ is concave and continuous with respect to initial condition k_0.

iv) *(Continuity and differentiability of value function $\hat{U}^\beta(k_0, T)$ at discount factor β.)*

$\hat{U}^\beta(k_0, T)$, as well as any feasible $U^\beta(k_0, T)$, is a continuous and differentiable function of discount factor β.

v) *(Truncation and infinite horizon.)*

There exists limits $\hat{U}^\beta(k_0) := \lim_{T\to\infty} \hat{U}^\beta(k_0, T)$, and $\hat{U}^\beta(k_0)$ converges monotonously to $\hat{U}^\beta(k_0)$ as $T \to \infty$. Note that since the discount factor $\beta < 1$ maximal global utility $\hat{U}^\beta(k_0)$ is finite also if $f'(0) < \beta^{-1}$ (cf. Fact 1i).

To verify parts i), ii), observe that if we start with initial condition $k_0' > k_0$ and follow optimal policy with respect to initial condition k_0, except of enlarging initial c_0 by $\Delta c_0 = f(k_0') - f(k_0) > 0$ (recall that by AS 1 and AS 2 $f(k_0') > f(k_0)$ and $u(\cdot)$ is increasing). For such policy we have $\tilde{U}^\beta(k_0', T) > \hat{U}^\beta(k_0, T)$, $\hat{U}^\beta(k_0', T) \ge \tilde{U}^\beta(k_0, T)$ and obviously $\hat{U}^\beta(k_0', T) > \hat{U}^\beta(k_0, T)$.
Moreover, $\hat{U}^\beta(k_0, T)$ must be continuous and differentiable function of the initial condition k_0. To this end observe that that $\hat{U}^\beta(k_0', T) - \hat{U}^\beta(k_0, T) = u(f(k_0')k_1) - u(f(k_0) - k_1)$ and $u(\cdot)$, $f(\cdot)$ are continuous and differentiable functions by assumptions AS 1 and AS 2. Using the same way of reasoning we can conclude continuity of any feasible function $U^\beta(k_0, T)$ with respect of $k_1, k_2, \dots, k_t$ and $\hat{U}^\beta(k_0, T)$ are continuous functions of k_0.

Part iii) is a consequence of concavity of instantaneous utility function $u(\cdot)$ (cf. AS 1) and per capita production function $f(\cdot)$ (cf. AS 2). Since concave function defined on an open interval must be on this interval continuous, it suffices to verify concavity. For details see [2].

Part iv) follows immediately for any feasible policy and by taking into account part i) and ii) for optimal policies.

To verify part v) observe that if $k \in [k_0, k^*]$ (cf. Fact 1ii)) and $c_t \in [0, f(k_t)]$, hence $u(c_t) > 0$ must be bounded by some $C < \infty$ implying that $\hat{U}^\beta(k_\tau, T - \tau) = \max_{c_t \in [\tau, T)} \sum_{t=\tau}^{T} \beta^t u(c_t) \le \beta^\tau C/(1 - \beta)$, and $\hat{U}^\beta(k_\tau, T - \tau)$ is decreasing in τ. Moreover, condition $\sum_{t=0}^{\infty} \beta^t u(c_t) < \infty$ is fulfilled even for unbounded $u(c_t)$ if for some $\bar{\beta} > \beta$ $u(c_t) < \bar{\beta}^t C$ (this holds if $f'(0) < \beta^{-1}$).

2 The Growth Model Under Uncertainty

To include random shocks or imprecisions into the model, we shall assume that for a given value of k_t we obtain the output y_t only with certain probability, in particular we assume that $y_t \in [f_{\min}(k_t), f_{\max}(k_t)]$ (i.e. $f_{\min}(\cdot) \le f(\cdot) \le f_{\max}(\cdot)$, AS 2 also hold for $f_\bullet(k)$). Obviously, better results can be obtained if we replace the rough estimates of y_t generated by means of $f_{\max}(k_t)$ and $f_{\min}(k_t)$ by a more detailed information on the (random) output y_t generated by the capital k_t.

To this end we shall assume that that in (6) $y_t = f(k_t)$ is replaced by $y_t = Z(k_t)$, where $Z(\cdot)$ is a Markov process with state space $\mathcal{I}_1 \subset \mathbb{R}$ and transition probabilities $p(y|k)$ from state $k \in \mathcal{I}_1$ in state $y \in \mathcal{I}_2 \subset \mathbb{R}$ such that $p(y_t|k_t) \gg p(y|k_t)$ for each $y \neq y_t = f(k_t)$ (obviously, $\sum_{y\in\mathcal{I}_2} p(y|k) = 1$ for each $k \in \mathcal{I}_1$). Moreover, we assume that the current value of the total output y_t is known to the decision maker *and then the recourse decision (intervention) may be taken to reach immediately the optimal value of* k_{t+1} for the original model (cf. (10)–(12). Such an extension well corresponds to the models introduced and studied in [8] and also in [3, 4]. Up to now we have assumed that the transition probabilities cannot be influenced by the decision maker. In what follows we extend the model in such a way that $p(y|k)$ will be replaced by $p(y|k,d)$ for $d \in \mathcal{D} = \{1, 2, \dots, D\}$ and some *cost,* denoted $c(k,d)$, *will be accrued to this decision.* Similarly, additional cost, denoted $\bar{c}(y,\bar{k})$, will be accrued to the recourse decision transferring state $y \in \mathcal{I}_2$ to the desired state $\bar{k} \in \mathcal{I}_1$ (of course, $\bar{c}(\bar{k},\bar{k}) = 0$).

So the development of the considered system over time is given by the following diagram

$$k = k_t \xrightarrow[]{\substack{c(k,d)\\ p(y|k,d)}} y = y_t \xrightarrow[]{\bar{c}(y;\bar{k})} \bar{k} = k_{t+1}$$

The above model can be also treated as a structured controlled Markov reward process X with compact state space $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ (with $\mathcal{I}_1 \cap \mathcal{I}_2 = \emptyset$), finite set $\mathcal{D} = \{0, 1, \dots, D\}$ of possible decisions (actions) in each state $k \in \mathcal{I}_1$ and the following transition and cost structure:

$p(y|k,d)$: transition probability from $k \in \mathcal{I}_1 \to y \in \mathcal{I}_2$ if decision $d \in \mathcal{D}$ is selected,

$c(k,d)$: cost of decision $d \in \mathcal{D}$ in state $k \in \mathcal{I}_1$,

$\bar{c}(y,\bar{k})$: cost for intervention, i.e. immediate transition from state $y \in \mathcal{I}_2$ in $\bar{k} \in \mathcal{I}_1$,

$r(y|k,d)$: expected value of the one-stage reward obtained in state k if decision $d \in \mathcal{D}$ is selected in state k; in particular

$$r(y|k,d) = \int_{y\in\mathcal{I}_2} p(y|k,d)[u(f(y)) - (1-\delta)k - y]\,\mathrm{d}y,$$

$\bar{r}(y|k,d)$: total expected reward earned by transition (including possible intervention) from state k to state $\bar{k}$, i.e.,

$$\bar{r}(y|k,d) = r(y|k,d) - c(k,d) - \int_{y\in\mathcal{I}_2} p(y|k,d)\bar{c}(y,\bar{k})\,\mathrm{d}y.$$

Now we have two options:

1. Follow optimal policy found for the corresponding deterministic model (cf. (12)) in such a way that we maximize by means of intervations transition rewards between consecutive states, i.e. at time points $t = 0, 1, \dots$ we maximize $\bar{r}(k_t, k_{t+1}, d)$ with respect to $d \in \mathcal{D}$ and follow the same policy as if $p(y|k_t, d) = 1$ for $y = f(k_t)$.
2. Consider the problem as finding optimal control policy of a β-discounted Markov decision chain with compact state space and finite action space.

In the latter case for the considered time horizon T let the value function $\hat{U}^\beta(k, T-\tau)$ denote expectation of the maximal (random) global utility received in the remaining τ next transitions if the considered Markov reward chain X is in state $k \in \mathcal{I}_1$ and optimal policy is followed. Then obviously

$$\hat{U}^\beta(k, T-\tau) = \max_{d\in\mathcal{D}} \int_{y\in\mathcal{I}_2} p(y|k,d)[\bar{r}(y|k,d) + \beta\hat{U}^\beta(y, T-\tau-1)]\,\mathrm{d}y, \qquad k \in \mathcal{I}_1, y \in \mathcal{I}_1 \tag{13}$$

and for τ tending to infinity, i.e. when $\lim_{T\to\infty} \hat{U}^\beta(k,T) = \hat{U}^\beta(k)$, then

$$\hat{U}^\beta(k) = \max_{d\in\mathcal{D}} \int_{y\in\mathcal{I}_2} p(y|k,d)[\bar{r}(y|k,d) + \beta\hat{U}^\beta(y)]\,\mathrm{d}y \qquad k \in \mathcal{I}_1, y \in \mathcal{I}_1. \tag{14}$$

Equation (14) demonstrates that part v) of Result 1 also holds in the considered extended version. Similarly, also part i) of Result 1 (Monotonicity of value function in initial condition), as a common property of Markov control processes holds for the considered extended version along with continuity and differentiability of value function both at initial condition k and discount factor β (cf. Result 1, parts ii) – iv)).

Unfortunately, assuming that Z is a Markov process with compact state space $\mathbb{R}$ then the model given by (13), (14) is not suitable for numerical computation. To make the model computationally tractable we it is necessary to approximate our system governed by (13), (14) by a discretized model with finite state space and estimate the resulting errors caused by such approximation.

To this end, it is necessary to assume that the values of c_t, k_t, and y_t take on only a finite number of discrete values. In particular, we assume that for sufficiently small $\Delta > 0$ there exist nonnegative integers $\bar{c}_t$, $\bar{k}_t$, and $\bar{y}_t$ such that for every $t = 0, 1, \ldots$ it holds:
$\bar{c}_t \Delta = c_t$, $\bar{k}_t \Delta = k_t$, and $\bar{y}_t \Delta = y_t$ with $\bar{k}_t \le K := k_{\max}/\Delta$ and similarly $\bar{y}_t \le Y := y_{\max}/\Delta$. Such approach was discussed in [7], [9].

Conclusion. In this article we indicated possible applications of controlled Markov processes for the analysis of extended versions of the growth models. As it was shown, many properties of the classical model can be extended to the considered more general cases.

Acknowledgement. This work was supported by the Czech Science Foundation under Grants 402/08/0107 and 402/07/1113.

References

1. O. J. Blanchard and S. Fisher: Lectures on Macroeconomics. MIT Press, Cambridge, MA 1989.
2. R.-A. Dana and C. Le Van: Dynamic Programming in Economics. Kluwer, Dordrecht 2003.
3. B. Heer and A. Maußer: Dynamic General Equilibrium Modelling. Springer, Berlin 2005.
4. M. Majamdar, T. Mitra and K. Nishinmura (eds.): Optimization and Chaos. Springer, Berlin 2000.
5. M. L. Puterman: Markov Decision Processes – Discrete Stochastic Dynamic Programming. Wiley, New York 1994.
6. F. P. Ramsey: A mathematical theory of saving. Economic Journal *38* (1928), 543–559.
7. K. Sladký: Approximations in stochastic growth models. In: Proc. 24th Internat. Conf. Mathematical Methods in Economics 2006, Univ. West Bohemia, Pilsen 2006, pp. 465–470.
8. N.-L. Stokey and R. E. Lukas, Jr.: Recursive Methods in Economic Dynamics. Harvard Univ. Press, Cambridge, Mass. 1989.
9. K. Sladký: Stochastic growth models with no discounting. Acta Oeconomica Pragensia *15* (2007), No.4, 88–98.

HABIT FORMATION AND PRICE INDEXATION IN DSGE MODELS WITH NOMINAL RIGIDITIES

Martin Slanicay and Osvald Vašíček *

Masaryk University, Faculty of Economics and Administration
Department of Economics
slanicay@mail.muni.cz osvald@econ.muni.cz

Abstract. The goal of this paper is to evaluate some characteristics commonly used in New Keynesian DSGE models, such as habit formation in consumption, full price indexation and partial price indexation. This paper estimates six DSGE models in order to determine, which model provides better data fit. One model is the most general one and others are its nested submodels. These are models of closed economy and include price and wage rigidities modelled in Calvo style. Three models include simple CRRA period utility function and other three models contain CRRA period utility function with habit formation in consumption. Two models assume traditional Calvo price setting mechanism, in which non-optimizing agents leave their price unchanged. Other two models assume partial price indexation in Calvo price setting mechanism, in which a part of non-optimizing agents adjust their prices according to past inflation. Last two models assume full price indexation in Calvo price setting mechanism, in which all non-optimizing agents adjust their prices according to past inflation. The models were estimated on data of US economy. All models were estimated using Bayesian techniques, particularly Metropolis-Hastings algorithm (using Dynare toolbox for Matlab). The data fit measure is posterior odds ratio calculated from marginal likelihood, acquired from Bayesian estimation. Results suggest that including habit formation improves significantly the data fit of the models, whereas including price indexation does not.
Keywords. New Keynesian, DSGE, habit formation, price indexation, posterior odds, Bayesian estimation.

1 Introduction

There are a lot of papers and studies concerning different DSGE models and different assumptions about them, but there has been little work on evaluating ability of different model specifications to fit macroeconomic data, in order to justify which specifications and assumptions fit the data better.
This paper estimates six DSGE models in order to determine, whether including habit formation in consumption and price indexation (in Calvo price setting mechanism) in the model provides better fit of the data. Models are following model of Jordi Galí in *Monetary Policy, Inflation, and the Business Cycle*, chapter 6 and this concept was modified in some model variants by allowing for habit formation and price indexation. Section 2 sketches the derivation of these models and their log linear equilibrium. Section 3 discusses the Bayesian estimation and results. Section 4 concludes.

2 Models

2.1 Households

We assume a continuum of households indexed by $j \in [0,1]$. Each household is specialized in different kind of work and seeks to maximize its utility function

$$E_0 \sum_{t=0}^{\infty} \beta^t U(C_t(j), N_t(j)) \tag{1}$$

where E_0 denotes expectations in period 0, β is a discount factor, $N_t(j)$ is a quantity of labor supplied and $C_t(j)$ is consumption index of household j defined as

$$C_t(j) = \left(\int_0^1 C_t(i,j)^{1-\frac{1}{\epsilon_p}} di \right)^{\frac{\epsilon_p}{\epsilon_p - 1}} \tag{2}$$

*We thank Miroslav Hloušek, Jan Čapek and Adam Remo for their helpful advices and comments.

where ϵ_p is elasticity of substitution among different types of goods, and $C_t(i,j)$ denotes a consumption of good i by household j in period t. Maximization of (1) is subject to a sequence of budget constraints, sequence of labor demand constraints and Calvo constraint on the frequency of wage adjustment. Models 1, 3 a 5 contain simple CRRA period utility function [1]

$$U_t = \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\phi}}{1+\phi} \tag{3}$$

where $\sigma > 0$ is inverse elasticity of intertemporal substitution, $\phi > 0$ is inverse elasticity of labor supply. Models 2, 4 a 6 contain CRRA period utility function with habit formation in consumption

$$U_t = \frac{(C_t - hC_{t-1})^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\phi}}{1+\phi} \tag{4}$$

where h is parameter of habit formation in consumption.

2.2 Firms

We assume a continuum of firms indexed by $i \in [0,1]$ and each firm produces its own differentiated product. Production function of the firm $i \in [0,1]$ is in the form

$$Y_t(i) = A_t N_t(i)^{1-\alpha}, \ 0 < \alpha < 1 \tag{5}$$

where A_t is exogenous technology process common to all firms, α is a paremeter of decreasing returns to scale and $N_t(i)$ is index of labor input used by firm i defined as

$$N_t(i) = \left(\int_0^1 N_t(i,j)^{1-\frac{1}{\epsilon_w}} dj \right)^{\frac{\epsilon_w}{\epsilon_w - 1}} \tag{6}$$

where $N_t(i,j)$ denotes amount of labor type j employed by firm i in period t and ϵ_w is elasticity of substitution among different labor varieties. We also define agregate wage index in the form

$$W_t = \left(\int_0^1 W_t(j)^{1-\epsilon_w} dj \right)^{\frac{1}{1-\epsilon_w}}. \tag{7}$$

Firms seek to maximize discounted sum of their current and expected future profits, taking into account their demand constraints and Calvo constraint on frequency of price adjustment.

2.3 Price Phillips Curve

Models 1 and 2 assume that the fraction θ_p of firms that are not able to reoptimize their prices simply leave their prices unchanged. Solution of firm's maximalization problem then leads (after some mathematical manipulation) to price Phillips curve in the form

$$\widetilde{\pi}_t^p = \beta E_t \widetilde{\pi}_{t+1}^p + \kappa_p \widetilde{y}_t + \lambda_p \widetilde{\omega}_t + d_t \tag{8}$$

where $\widetilde{\pi}_t^p$, $\widetilde{y}_t$ and $\widetilde{\omega}_t$ are price inflation, output and real wage in the form of log-deviation from their steady state[2], $\kappa_p = \frac{\alpha\lambda_p}{1-\alpha}, \lambda_p = \frac{(1-\theta_p)(1-\beta\theta_p)}{\theta_p} \frac{1-\alpha}{1-\alpha+\alpha\epsilon_p}$ and $d_t = \rho_d d_{t-1} + \epsilon_{d,t}$ is AR1 process for price inflation shock, which was added to this equation.

Models 3 and 4 assume that a part δ_p of those firms that are not able to reoptimize their prices adjust their prices according inflation in the previous period. This leads (after some manipulation) to price Phillips curve in form

$$\widetilde{\pi}_t^p = \beta E_t(\widetilde{\pi}_{t+1}^p - \delta_p \widetilde{\pi}_t^p) + \delta_p \widetilde{\pi}_{t-1}^p + \kappa_p \widetilde{y}_t + \lambda_p \widetilde{\omega}_t + d_t. \tag{9}$$

Models 5 and 6 assume that all of those firms, that are not able to reoptimize their prices, adjust their prices according inflation in the previous period. It can be shown, that price Phillips curve of model 5 and 6 is the specific case of those in model 3 and 4 for $\delta_p = 1$.

[1] We assume identical preferences and because all households face the same problem, we can omit index j in specification of utility function.

[2] Formally $\widetilde{x}_t = \log x_t - \log \overline{x}$, where $\overline{x}$ is a steady state value.

2.4 Wage Phillips Curve

Models 1 and 2 assume that in each period only $1-\theta_w$ fraction of households can reoptimize their wages, while θ_w fraction of households, which can't reoptimize, leave their wages unchanged.
Wage Phillips curve of model 1 is in the form

$$\widetilde{\pi}_t^w = \beta E_t \widetilde{\pi}_{t+1}^w + \kappa_{w1}\widetilde{y}_t - \lambda_w \widetilde{\omega}_t + f_t \tag{10}$$

where $\widetilde{\pi}_t^w$ is log-deviation of wage inflation from its steady state,
$\kappa_{w1} = \lambda_w \left(\sigma + \frac{\phi}{1-\alpha}\right)$, $\lambda_w = \frac{(1-\theta_w)(1-\beta\theta_w)}{\theta_w(1+\epsilon_w\phi)}$ and $f_t = \rho_f f_{t-1} + \epsilon_{f,t}$ is AR1 process for wage inflation shock, which was added to this equation.
Wage Phillips curve of model 2 is in the form

$$\widetilde{\pi}_t^w = \beta E_t \widetilde{\pi}_{t+1}^w + \kappa_{w2}\widetilde{y}_t - \nu_w \widetilde{y}_{t-1} - \lambda_w \widetilde{\omega}_t + f_t \tag{11}$$

where $\kappa_{w2} = \lambda_w \left(\frac{\sigma}{1-h} + \frac{\phi}{1-\alpha}\right)$, $\nu_w = \lambda_w \frac{\sigma h}{1-h}$.
Models 3 and 4 assume that a part δ_w of those households, which are not able to optimize their wages, adjust their wages according past inflation.
Wage Phillips curve of model 3 is in the form

$$\widetilde{\pi}_t^w = \beta E_t(\widetilde{\pi}_{t+1}^w - \delta_w \widetilde{\pi}_t^w) + \delta_w \widetilde{\pi}_{t-1}^w + \kappa_{w1}\widetilde{y}_t - \lambda_w \widetilde{\omega}_t + f_t. \tag{12}$$

Wage Phillips curve of model 4 is in the form

$$\widetilde{\pi}_t^w = \beta E_t(\widetilde{\pi}_{t+1}^w - \delta_w \widetilde{\pi}_t^w) + \delta_w \widetilde{\pi}_{t-1}^w + \kappa_{w2}\widetilde{y}_t - \nu_w \widetilde{y}_{t-1} - \lambda_w \widetilde{\omega}_t + f_t. \tag{13}$$

Models 5 and 6 assume that all households, which are not able to reoptimize their wages, adjust their wages according to inflation in the previous period. It can be shown, that wage Phillips curve of model 5 is the specific case of those in model 3 for $\delta_w = 1$ and wage Phillips curve of model 6 is the specific case of those in model 4 for $\delta_w = 1$.

2.5 Euler Equation

Euler equation of **models 1, 3** and **5** is in the form

$$\widetilde{y}_t = E_t\{\widetilde{y}_{t+1}\} - \frac{1}{\sigma}(\widetilde{i}_t - E_t\{\widetilde{\pi}_{t+1}^p\}) + u_t \tag{14}$$

where $\widetilde{i}_t$ is is log-deviation of nominal interest rate from its steady state, $u_t = \rho_u u_{t-1} + \epsilon_{u,t}$ is AR1 process for Euler equation shock, which was added to this equation.
Euler equation of **models 2, 4** and **6** is in the form

$$\widetilde{y}_t = E_t\{\widetilde{y}_{t+1}\} - \frac{1-h}{\sigma}(\widetilde{i}_t - E_t\{\widetilde{\pi}_{t+1}^p\}) + u_t. \tag{15}$$

2.6 Inflation Identity and Monetary Policy Rule

Wage inflation, price inflation and real wage are connected in an identity

$$\widetilde{\omega}_t = \widetilde{\omega}_{t-1} + \widetilde{\pi}_t^w - \widetilde{\pi}_t^p. \tag{16}$$

All models are completed with monetary policy rule in the form of modified Taylor rule

$$\widetilde{i}_t = \rho_i \widetilde{i}_{t-1} + (1-\rho_i)[\psi_p \widetilde{\pi}_t^p + \psi_y \widetilde{y}_t] + v_t \tag{17}$$

where ρ_i is a backward looking parameter for interest rate, ψ_p is an elasticity of interest rate to inflation, ψ_y is an elasticity of interest rate to output and $v_t = \rho_v v_{t-1} + \epsilon_{v,t}$ is AR1 process for nominal interest rate shock.

2.7 Relations between Alternative Models and Parameters Interpretation

The most general is the model 4, other models are his nested submodels for specific values of selected variables. Model 1 is a submodel of model 2 for $h = 0$, a submodel of model 3 for $\delta_p = 0$, $\delta_w = 0$ and a submodel of model 4 for $h = 0$, $\delta_p = 0$, $\delta_w = 0$. Model 2 is a submodel of model 4 for $\delta_p = 0$, $\delta_w = 0$. Model 3 is a submodel of model 4 for $h = 0$. Model 5 is a submodel of model 4 for $h = 0$, $\delta_p = 1$, $\delta_w = 1$ and submodel of model 6 for $h = 0$. Model 6 is a submodel of model 4 for $\delta_p = 1$, $\delta_w = 1$.

Table 1. Definition of Parameters

Parameter	Interpretation	Restriction
β	discount factor	$\langle 0, 1 \rangle$
σ	inverse elasticity of intertemporal substitution	$\langle 0, \infty)$
ϕ	inverse elasticity of labor supply	$\langle 0, \infty)$
h	degree of habit formation in consumption	$\langle 0, 1 \rangle$
α	parameter of decreasing returns to scale	$\langle 0, 1 \rangle$
θ_p	fraction of non-optimizing firms	$\langle 0, 1 \rangle$
θ_w	fraction of non-optimizing households	$\langle 0, 1 \rangle$
δ_p	degree of price indexation	$\langle 0, 1 \rangle$
δ_w	degree of wage indexation	$\langle 0, 1 \rangle$
ϵ_p	elasticity of substitution among different goods	$\langle 0, \infty)$
ϵ_w	elasticity of substitution among labor varieties	$\langle 0, \infty)$
ψ_p	elasticity of interest rate to inflation	$\langle 0, \infty)$
ψ_y	elasticity of interest rate to output	$\langle 0, \infty)$
ρ_i	backward looking parameter for interest rate	$\langle 0, 1 \rangle$
ρ_u	AR1 parameter for Euler equation shock	$\langle 0, 1 \rangle$
ρ_v	AR1 parameter for interest rate shock	$\langle 0, 1 \rangle$
ρ_d	AR1 parameter for price inflation shock	$\langle 0, 1 \rangle$
ρ_f	AR1 parameter for wage inflation shock	$\langle 0, 1 \rangle$

3 Estimation

3.1 Data Set Used for Estimation

Quarterly data of USA from 1.4.1985 to 1.10.2008 downloaded from database of FRED (Federal Reserve Economic Data), http://research.stlouisfed.org/fred2

- $\widetilde{y}_t$: detrended data (linear trend used) of log real GDP per capita (per cent). A measure of GDP is GDPC1: Real Gross Domestic Product, in Billions of Chained 2000 Dollars, Seasonally Adjusted Annual Rate. A measure of population is POP: Total Population: All Ages including Armed Forces Overseas
- $\widetilde{\pi}_t^p$: detrended data (constant trend used) of quarterly price inflation rate (in percentage points). An underlying index for price inflation is CPIAUCSL: Consumer Price Index For All Urban Consumers: All Items, Index 1982-84=100, Seasonally Adjusted.
- $\widetilde{\pi}_t^w$: detrended data (constant trend used) of quarterly wage inflation rate (in percentage points). An underlying index for wage inflation is COMPNFB: Nonfarm Business Sector: Compensation Per Hour, Index 1992=100, Seasonally Adjusted
- $\widetilde{i}_t$: detrended data (linear trend used) of quartelized interest rate. A measure of interest rate is TB3MS: 3-Month Treasury Bill: Secondary Market Rate (in percentage points), Averages of Business Days, Discount Basis.

3.2 Estimated Parametres

Except parameter β, which was calibrated at 0.99, were all parametres estimated using Metropolis-Hastings algorithm (using Dynare toolbox for Matlab).

3.3 Results

Comparison of models with habit formation in consumption and models with no habit formation in corresponding variants of Calvo constaint, based on their posterior odds (PO)

$$PO_{M2/M1} = 67.7495, \quad PO_{M4/M3} = 8.2741, \quad PO_{M6/M5} = 97.9295 \tag{18}$$

Comparison of models with no price indexation and models with partial price indexation in corresponding variants of utility function, based on their posterior odds (PO)

$$PO_{M1/M3} = 1137900.2255, \qquad PO_{M2/M4} = 9317307.5663 \tag{19}$$

Comparison of models with partial price indexation and models with full price indexation in corresponding variants of utility function, based on their posterior odds (PO)

$$PO_{M3/M5} = 15804548797308032, \qquad PO_{M4/M6} = 1335328494303031 \tag{20}$$

Table 2. Estimated Parameters

parameter	prior	M1	M2	M3	M4	M5	M6
β		0.99	0.99	0.99	0.99	0.99	0.99
σ	4.0	5.7203	4.0354	5.6958	3.8690	6.0764	3.6552
ϕ	1.5	1.6945	1.8667	1.5702	1.7437	1.3221	1.5037
h	0.6	-	0.7350	-	0.7428	-	0.7373
α	0.2	0.2733	0.2638	0.2690	0.2650	0.2505	0.2197
θ_p	0.7	0.8154	0.8283	0.8136	0.8270	0.7535	0.7571
θ_w	0.7	0.7303	0.7734	0.7168	0.7676	0.6764	0.7340
δ_p	0.5	-	-	0.1149	0.1170	-	-
δ_w	0.75	-	-	0.4633	0.3974	-	-
ϵ_p	6.0	10.4302	10.6273	10.3735	10.7219	8.1624	8.1021
ϵ_w	6.0	7.7860	9.5885	6.7885	9.4969	4.4854	7.8360
ψ_p	1.5	1.2536	1.3274	1.2622	1.3228	1.3799	1.4174
ψ_y	0.5	0.4719	0.4434	0.4824	0.4420	0.4260	0.4159
ρ_i	0.8	0.9007	0.9159	0.9029	0.9159	0.9043	0.9177
ρ_u	0.7	0.9378	0.9024	0.9366	0.8944	0.9267	0.8963
ρ_v	0.3	0.3815	0.3903	0.3813	0.3951	0.3608	0.3864
ρ_d	0.2	0.1357	0.1422	0.1019	0.1093	0.0526	0.0583
ρ_f	0.3	0.2376	0.1787	0.0877	0.0751	0.0494	0.0512

4 Conclusion

Posterior odds of model i and j ($PO_{Mi/Mj}$) is a bayesian statistic based on comparison of maximal likelihood functions. It shows us how many times more probable is model i then model j. DeJong and Dave in *Structural Macroeconometrics,* p. 242 show an interpretation of PO values

– $1:1-3:1$ - "very slight evidence"
– $3:1-10:1$ - "slight evidence"
– $10:1-100:1$ - "strong to very strong evidence"

Results suggest, that including habit formation in consumption improve significantly fit of the US economy data, see posterior odds in (18), whereas including partial price indexation does not, see posterior odds in (19). The worst model variants, regarding fit of the US data, where those with full price indexation, see posterior odds in (20).

References

1. DeJong, D. N., Dave, C.: Structural Macroeconometrics. Princeton University Press, Princeton and Oxford, 2007.
2. Galí, J.: Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework. Princeton University Press, Princeton, New Jersey, 2008.

ARE THE PRICES OF THE DRIVING FUEL IN THE SLOVAK ECONOMY MAKING ASYMMETRIC?[1]

Karol Szomolányi, Martin Lukáčik and Adriana Lukáčiková

University of Economics in Bratislava
Faculty of Economic Informatics, Department of Operation Research and Econometrics,
szomolan@euba.sk, lukacik@euba.sk, istvanik@euba.sk

Abstract. Are the prices of the driving fuel in the Slovak economy making asymmetrically? Will in the Slovak economy the same negative and positive changes of the oil price cause the different negative and positive changes of the driving fuel prices? Yes it will. The price maker reacts more strictly, when the oil price increases and, on the other hand, he is more benevolent when the oil price falls. His preferences are asymmetric. The source of the asymmetric price making is a conflict between the society that requires respectable fuel prices and the price-maker that requires making use of his dominant position by setting optimal prices. As we show, the theory explains the structure of the driving fuel price-making. To verify our statements we form an asymmetric cost function of the driving fuel price maker and derive his reaction function. The coefficients of the function are estimated by the generalised method of moments and by the two stage least square method.
Keywords. Price making policy, driving fuel, asymmetric preferences, oil, price, time inconsistency

1. Introduction

Institutions of the Slovak Republic have decided that the dominant seller of the fuel (price-maker) was counting some items towards his costs without authorisation in years 2002-2004 and he was selling fuel for such prices that he made inappropriate profit. According to our theory this conflict between the government and the seller leads to asymmetric price making of the driving fuel in Slovakia.

This statement follows from two assumptions. Firstly, optimal price of the dominant driving fuel seller is in average higher as a price respectable by society. If price-maker deviates from the respectable price, he risks the possibility of the punishment from the society. Dominant price-maker therefore has additional “threat costs”. Secondly, the fuel price making is discretional – a price maker behaves sequentially rational, as he has a possibility to react to the changes of the expected oil prices every time. His reaction is rather elastic, if he expects the oil prices oil increase and, on the contrary, his reaction is rather rigid, if he expects the oil price will decrease. This asymmetry we form by a “threat” cost function in linex specification.

In the paper we will form a theory of the asymmetric fuel price-making. We will derive and estimate a price-making reaction function of the oil price. We choose correctly the specification and verify the estimations statistically as well as economically by the confrontation with our theory.

2. Brief Literature Overview

Pfann and Palm [11] provided a nice example of an asymmetry approach in a labour market. According to the approach the firing costs of manufacturing exceed the hiring costs. As Adda and Cooper [1] state, Pfann and Palm estimated parameters of the of the cost function in linex specification using generalised method of moments (GMM) approach on data for manufacturing in the Netherlands (quarterly, seasonally unadjusted data, 1971(I)-1984(IV)) and annual data for U.K. manufacturing. They used data on both production and non-production workers and the employment choices are interdependent from the production function. For both countries they find evidence of asymmetry. They report that the costs of firing production workers are lower than the hiring costs. But, the opposite is true for the non-production workers.

Surico [15], [16] and [17] used similar approach for the analysis of U.S. and EMU monetary policies. He responded on questions, whether central bankers weight differently positive and negative deviations of inflation, output and the interest rate from their reference values. As showed Surico and before him Cukierman [5] or Ruge-Murcia [12], an asymmetry of the monetary policy for the output is “a novel source of monetary policy time inconsistency”. Time inconsistency of the economic policy problem is defined by the theories of Kydland and Prescott [10] or Barro and Gordon [2] according to which central banks systematically produce an inflation bias.

[1] This paper is supported by the Grant Agency of Slovak Republic – VEGA, grant no. 1/4652/07 “The Analysis of Actual Problems of Slovak Economy Progress before the Entrance to European Monetary Union – Econometrical Approach”.

Chari, Kehoe and Prescott [3] showed that the source of the time consistency problem is a conflict between economic agents. Actually in Slovak driving fuel price-making process there is a conflict between the society that requires respectable fuel prices and the price-maker that requires making use of his dominant position by setting optimal prices. Asymmetry in price-making process therefore means that the process is time inconsistent. The price-maker systematically produces a bias from the driving-fuel price respectable by the society.

3. The Model

Let us form threat cost function in linex specification:

$$C\left[p_t, E_{t-1}\left(c_t\right)\right] = \frac{-\gamma\left[p_t - kE_{t-1}\left(c_t\right)\right] + e^{\gamma\left[p_t - kE_{t-1}\left(c_t\right)\right]} - 1}{\gamma^2}, \tag{1}$$

where we suppose that the price-maker derive a fuel price, p_t, from the expected value of the oil price using an information set available at t–1, $E_{t-1}(c_t)$. Coefficient k contains information about an oil consumption needed to produce a litre of the fuel and γ is an asymmetry coefficient. A negative value of the coefficient γ implies that a negative value of the difference $p_t - E_{t-1}(c_t)$ causes for the price-maker higher costs than the same positive one. The linex specification nests the quadratic form as a special case, so that applying twice l'Hôpital's rule when γ tends to zero, the cost function (1) reduces to the symmetric parameterization:

$$C\left[p_t, E_{t-1}\left(c_t\right)\right] = \frac{1}{2}\left[p_t - kE_{t-1}\left(c_t\right)\right]^2$$

The fuel price-maker minimizes the cost function (1). The first-order condition is in the form:

$$E_{t-1}\left[\frac{-1 + e^{\gamma\left(p_t - kc_t\right)}}{\gamma}\right] = 0 \tag{2}$$

The condition (2) describes general a reaction function of the fuel price-maker. Whenever coefficient γ tends to zero, the reaction function (2) collapses to the linear form:

$$p_t = kE_{t-1}\left(c_t\right)$$

As the reaction function (2) is nonlinear, we perform a second-order Taylor-series expansion of the exponential terms in (2):

$$p_t - kE_{t-1}\left(c_t\right) + \frac{\gamma}{2}\left[p_t - kE_{t-1}\left(c_t\right)\right]^2 + v_t = 0 \tag{3}$$

The reminder of the approximation is v_t and contains terms of third order or higher. A litre fuel price is expressed by reaction function of the expected oil price using information set available at time t – 1. In contrast to the linear reaction function the fuel price depends on a quadratic term product half asymmetry coefficient γ.

We solve equation (3) for p_t and, prior to GMM estimation, we replace expected values with actual values. In practise we estimate the following price-making rule with coefficient restrictions:

$$p_t = kc_t - \frac{\gamma}{2}\left(p_t - kc_t\right)^2 + u_t \tag{4}$$

and without coefficient restriction:

$$p_t = \beta_1 c_t + \beta_2 p_t^2 + \beta_3 p_t c_t + \beta_4 c_t^2 + u_t \tag{5}$$

where the coefficients are given by the expressions:

$$\beta_1 = k; \quad \beta_2 = -\frac{\gamma}{2}; \quad \beta_3 = k\gamma; \quad \beta_4 = -k^2\frac{\gamma}{2}$$

and the error term:

$$u_t = \left(\beta_1 + \beta_3 p_t\right)\left[c_t - E_{t-1}\left(c_t\right)\right] + \beta_4\left[c_t^2 - E_{t-1}\left(c_t^2\right)\right] - v_t$$

depends on oil prices forecast errors and therefore is orthogonal to any variable in the information set available at time t – 1.

The coefficient restrictions are:

$$k = \beta_1 \quad \text{a} \quad \gamma = -2\beta_2 \tag{6}$$

$$k = -\frac{\beta_3}{2\beta_2} \quad \text{a} \quad \gamma = \frac{\beta_3}{\beta_1} \tag{7}$$

$$k = \sqrt{\frac{\beta_4}{2\beta_2}} \quad \text{a} \quad \gamma = -2\frac{\beta_4}{\beta_1^2} \tag{8}$$

4. Data[2] and Methodology

We dispose with weekly data about average 95-octane gasoline prices (01.02.2006 – 01.12.2009) and diesel prices (01.02.2006 – 01.19.2009) both in Slovak crowns per litre[3] from the portal natankuj.sk. The portal www.tonto.eia.doe.gov provides daily data about oil BRENT prices in American dollars per barrel. We multiplied the data by official course of the National Bank of Slovakia from the portal www.nbs.sk and from them made average weekly data of the BRENT prices in Slovak crowns per barrel.

Using (4) and (5) specification we try to explain gasoline and diesel prices by the oil prices. We modify the specifications by taking first differences of (4) and (5) due to an integration of order one of all original variables. This decision was perform on the basis of Augmented Dickey Fuller unit root test. We make use of the estimations realized by the two-stage least squares method – TSLS and the generalized method of moments – GMM (the results are displayed in the Tables 3 to 7).

Instrument set consists of a constant, first lags of first differences of gasoline, diesel and oil prices, their squares and their products. Moreover for the first differences of the gasoline (diesel) price estimation is the instrumental set richer by first differences of the square of the diesel (gasoline) prices and first differences of the product of diesel (gasoline) and oil prices. The choice of the instrumental set is verified by the corresponding tests.

The first applied test is exploiting Staiger's and Stock's [14] declaration, that in applications of TSLS, it is common for the F-statistic, testing the hypothesis that the instruments do not enter the first-stage regression, to take on a value less than 10 alternatively the instruments are weak. The results of this procedure confirm the entire instruments as relevant in both equations.

The results of the second test – Durbin-Wu-Hausman [9] test of the endogeneity of regressors are displayed in the Table 1. We perform one version of this test in two stages; the first we regress every potentially endogenous variable on all exogenous variables and instruments to obtain residual; than we include all these residuals in our original model and test for their significance. We can reject the endogeneity of regressors if we can reject the significance of these residuals.

The last test used only for TSLS is Sargan [13] test of overidentifying restrictions. This test requires the coefficient of determination R^2 of the regression TSLS residuals on all exogenous variables (instruments and controls). Under the null hypothesis that all instruments are exogenous, the test statistics $S = nR^2$ (where n is a number of observations) is distributed as χ^2 with degree of freedom equal to the degree of overidentification, which is the number of instruments minus the number of endogenous variables. The null hypothesis isn't rejected in any equation.

The parallel test is employed for GMM estimations. In the context of GMM, the overidentifying restrictions are tested via the commonly employed J statistic of Hansen [8]. The J statistic is distributed as $\chi 2$ with degrees of freedom equal to the number of overidentifying restrictions. A rejection of the null hypothesis implies that the instruments are not satisfying the orthogonality conditions required for their employment. This may be either because they are not truly exogenous, or because they are being incorrectly excluded from the regression. The results of both overidentifying tests are displayed in the Tables 2, 3 (last two rows) and 7 (the fifth and the last columns).

For testing hypotheses (6) to (8) in both models, we use the Wald test (Greene [6]) comparing specifications of nested models by assessing the significance of restrictions to an extended model with unrestricted parameters. When the test statistic exceeds a critical value in its asymptotic distribution, the Wald test rejects the null, restricted model in favour of the alternative, unrestricted model. The asymptotic distribution is χ^2 with degree of freedom equal to the number of restrictions. The results are displayed in the Tables 4 and 5. The same test we apply to verify whether specification (4) and (5) equally explain first differences of gasoline and diesel prices. Wald test does not reject the null hypothesis of (6) = (7) = (8) for all cases.

---

[2] We are grateful to the employees of the portal openiazoch.zoznam.sk for their help in the data processing.

[3] Data in euros per litre after the entrance of the Slovak economy to the European Monetary Union at 01.01.2009 are multiplied by the official converse course 30.126 crowns per euro

5. Results

Table 1: Durbin-Wu-Hausman test of the endogeneity – original models with residuals

Parameter	Gasoline equation		Diesel equation	
β_1	0.0075	**	0,005100	**
	(0.000488)		(0.000414)	
β_2	0.021	**	0.015912	**
	(0.000736)		(0.000560)	
β_3	-0.000325	**	-0,000147	**
	$(3.04\cdot10^{-5})$		$(2.28\cdot10^{-5})$	
β_4	$1.29\cdot10^{-6}$	**	$2.04\cdot10^{-7}$	
	$(1.94\cdot10^{-7})$		$1.52\cdot10^{-7}$	
δ_c	-0.00387	**	-0.003439	**
	(0.000665)		(0.000549)	
δ_{p2}	-0.00521	**	-0.001167	*
	(0.000812)		(0.000602)	
δ_{pc}	0.000217	**	$6.38\cdot10^{-5}$	**
	$(3.34\cdot10^{-5})$		$(2.45\cdot10^{-5})$	
δ_{c2}	$-8.47\cdot10^{-7}$	**	$7.80\cdot10^{-8}$	
	$(2.02\cdot10^{-7})$		$(1.58\cdot10^{-7})$	
R^2	0.998		0.998	
D-W	1.98		1.81	
endogeneity	16.5		21.85	

Note. Parameters δ are parameters of residuals from regression of subscript endogenous variable on all exogenous variables and instruments. The standard errors of parameters are in the parentheses. The superscript ** and * denote the rejection of the null hypothesis of the *t*-test at the 1% and 5%. R^2 denotes the coefficient of determination and *D-W* the value of Durbin-Watson statistics. The values of F-statistics in line endogeneity are the significance test of residuals.

The estimates of the different forward-looking models and results of the test of overidentifying restrictions of the gasoline price equation and diesel price equation are in the Table 2 and 3.

Table 2: Comparison of estimates of different forward-looking models – gasoline price equation

$l =$	0		0		+1		+2		+4		+8	
	GMM		TSLS		GMM		GMM		GMM		GMM	
β_1	0.009492	**	0.007503	**	0.008349	**	0.005895	**	0.003439	**	0.004512	**
	(0.001040)		(0.000888)		(0.000714)		(0.000575)		(0.001082)		(0.001291)	
β_2	0.023796	**	0.021082	**	0.021346	**	0.018289	**	0.015388	**	0.016352	**
	(0.002035)		(0.001340)		(0.001108)		(0.000830)		(-0.000110)		(0.000617)	
β_3	-0.000446	**	-0.000327	**	-0.000353	**	-0.000215	**	-0.000110	**	-0.000145	**
	$(8.27\cdot10^{-5})$		$(5.54\cdot10^{-5})$		$(5.02\cdot10^{-5})$		$(3.42\cdot10^{-5})$		$(4.04\cdot10^{-5})$		$(2.96\cdot10^{-5})$	
β_4	$2.03\cdot10^{-6}$	**	$1.31\cdot10^{-6}$	**	$1.42\cdot10^{-6}$	**	$5.98\cdot10^{-7}$	*	$2.46\cdot10^{-7}$		$2.94\cdot10^{-7}$	*
	$(5.59\cdot10^{-7})$		$(3.53\cdot10^{-7})$		$(3.63\cdot10^{-7})$		$(2.40\cdot10^{-7})$		$1.77\cdot10^{-7}$		$(1.34\cdot10^{-7})$	
$J(7)$/S	4.790		13.114		6.275		5.589		3.615		4.335	
p	0.686		0.069		0.508		0.588		0.823		0.714	

Table 3: Comparison of estimates of different forward-looking models – diesel price equation

$l =$	0		0		+1		+2		+4		+8	
	GMM		TSLS		GMM		GMM		GMM		GMM	
β_1	0.006784	**	0.005100	**	0.008134	**	0.004886	**	0.003410	**	0.003413	*
	(0.000541)		(0.000780)		(0.002150)		(0.000945)		(0.000830)		(0.001567)	
β_2	0.018737	**	0.015912	**	0.020816	**	0.017080	**	0.015809	**	0.014749	**
	(0.000932)		(0.001056)		(0.002148)		(0.000948)		(0.000507)		(0.000716)	
β_3	-0.000244	**	-0.000147	**	-0.000349	**	-0.000178	**	-0.000131	**	-9.71·10-5	**
	(3.57·10-5)		(4.30·10-5)		(0.000104)		(4.50·10-5)		(2.27·10-5)		(3.36·10-5)	
β_4	7.33·10-7	**	2.04·10-7		1.60·10-6	*	4.68·10-7	*	4.06·10-7	**	3.76·10-8	
	(2.70·10-7)		(2.86·10-7)		(6.31·10-7)		(3.08·10-7)		(1.54·10-7)		(2.28·10-7)	
$J(7)$/S	7.290		13.235		2.913		5.093		6.043		3.802	
p	0.399		0.067		0.893		0.649		0.535		0.802	

Note. The standard errors of parameters are in the parentheses. The superscript ** and * denote the rejection of the null hypothesis of the *t*-test at the 1% and 5%. $J(7)$/S denotes the Hansen [8] or Sargan [13] test of overidentifying restrictions and p is p-value of these tests. It is compared with χ^2 with degree of freedom equal to 7. In the first line there is the lag of weeks used for computing of regressors in forward looking equations.

The estimates of the coefficients k and γ from the restrictions given by (6)-(8) in the gasoline and diesel price equation are in the Table 4 and 5.

Table 4: Estimates of parameters k and γ in gasoline price equation – Wald test of restrictions

	l =	0	0	+1	+2	+4	+8
		GMM	TSLS	GMM	GMM	GMM	GMM
(6)	k	0.009492 **	0.007503 **	0.008349 **	0.005895 **	0.003439 **	0.004512 **
		(0.001040)	(0.000888)	(0.000910)	(0.000714)	(0.001082)	(0.001291)
	γ	-0.047591 **	-0.042163 **	-0.042692 **	-0.036577 **	-0.030776 **	-0.032705 **
		(0.004070)	(0.001340)	(0.002216)	(0.001661)	(0.001757)	(0.001235)
(7)	k	0.009374 **	0.007750 **	0.008265 **	0.005870 **	0.003582 **	0.004443 **
		(0.000938)	(0.000825)	(0.000751)	(0.000673)	(0.001111)	(0.000747)
	γ	-0.046997 **	-0.043549 **	-0.042260 **	-0.036424 **	-0.032055 **	-0.032208 **
		(0.003769)	(0.003293)	(0.002548)	(0.002942)	(0.002959)	(0.004187)
(8)	k	0.009238 **	0.007882 **	0.008145 **	0.005718 **	0.004000 **	0.004237 **
		(0.000880)	(0.000822)	(0.000835)	(0.001025)	(0.001331)	(0.000960)
	γ	-0.045080 **	-0.046527 **	-0.040635 **	-0.034409 **	-0.041639 **	-0.028838
		(0.003985)	(0.007650)	(0.004193)	0.009649	(0.017740)	(0.023826)

Table 5: Estimates of parameters k and γ in diesel price equation – Wald test of restrictions

	l =	0	0	+1	+2	+4	+8
		GMM	TSLS	GMM	GMM	GMM	GMM
(5)	k	0.006784 **	0.005100 **	0.008134 **	0.004886 **	0.003410 **	0.003413 *
		(0.000541)	(0.000780)	(0.002150)	(0.000945)	(0.000830)	(0.001567)
	γ	-0.037474 **	-0.031824 **	-0.041632 **	-0.034159 **	-0.031618 **	-0.029499 **
		(0.001863)	(0.002111)	(0.004296)	(0.001896)	(0.001013)	(0.001432)
(6)	k	0.006517 **	0.004626 **	0.008390 **	0.005200 **	0.004129 **	0.003291 **
		(0.000634)	(0.001048)	(0.001641)	(0.001033)	(0.000593)	(0.000991)
	γ	-0.036002 **	-0.028870 **	-0.042946 **	-0.036359 **	-0.038285 **	-0.028445 **
		(0.003532)	(0.004979)	(0.002362)	(0.003302)	(0.005721)	(0.006511)
(7)	k	0.006253 **	0.003580	0.008772 **	0.005236 **	0.005070 **	0.001596
		(0.001003)	(0.002400)	(0.001286)	(0.001596)	(0.000917)	(0.004857)
	γ	-0.031843 **	-0.015681	-0.048421 **	-0.039236 *	-0.069893	-0.006446
		(0.009896)	(0.019075)	(0.009429)	(0.017280)	(0.047010)	(0.043115)

Note. The standard errors of parameters are in the parentheses. The superscript ** and * denote the rejection of the null hypothesis of the Wald test of restrictions at the 1% and 5%. The results are obtained for the different specifications (6), (7) and (8).

The estimates of the coefficients k and γ using specification (4) are in the Table 6.

Table 6: Estimates of parameters k and γ from restricted model

l	*method*	*first differences of the gasoline prices*			*first differences of the oil prices*		
		k	γ	$J(9)/S$	k	γ	$J(9)/S$
				p			p
0	GMM	0.007069 **	-0.039760 **	9.219	0.006504 **	-0.037383 **	8.159
		(0.000519)	(0.001523)	0.417	(0.000452)	(0.001183)	0.519
0	TSLS	0.007172 **	-0.040644 **	14.651	0.005196 **	-0.033282 **	16.665
		(0.000653)	(0.002040)	0.101	(0.000604)	(0.001252)	0.054
+1	GMM	0.008444 **	-0.043171 **	6.755	0.005944 **	-0.036245 **	9.634
		(0.000750)	(0.002219)	0.664	(0.000430)	(0.000828)	0.381
+2	GMM	0.005972 **	-0.036825 **	5.875	0.004603 **	-0.032317 **	7.784
		(0.000544)	(0.001263)	0.754	(0.000373)	(0.000611)	0.556
+4	GMM	0.002268 **	-0.028537 **	5.021	0.004832 **	-0.033456 **	7.636
		(0.000870)	(0.001311)	0.832	(0.000724)	(0.001372)	0.572
+8	GMM	0.004382 **	-0.032622 **	4.638	0.002798 **	-0.028429 **	8.362
		(0.000553)	(0.000937)	0.865	(0.000385)	(0.000577)	0.498

Note. The standard errors of parameters are in the parentheses. The superscript ** and * denote the rejection of the null hypothesis of the t-test at the 1% and 5%. $J(9)/S$ denotes the Hansen [8] or Sargan [13] test of overidentifying restrictions and p is p-value of these tests. It is compared with χ^2 with degree of freedom equal to 9. In the first column there is the lag of weeks used for computing of regressors in forward looking equations.

6. Conclusions

The most of the estimates of the coefficients are significant at 1% level and have signs that correspond to our theory. Using (6)-(8) specification we estimated coefficients of the threat cost function k and γ. The negative value of the γ coefficient implies that the driving fuel price making is asymmetric with respect to expected oil prices. Different restrictions (6), (7) and (8) generate the same value of the coefficients. It follows that the theory explains the structure of the price-making process.

A conflict between Slovak Government and the dominant driving fuel price-maker is a source of the asymmetric price-making of the fuel with respect to expected oil price. In order to make use of the dominant position of the price-maker, he tends to overvalue expected increase and undervalue expected decrease of the oil price.

The asymmetric behaviour of the price-maker is in line with the time-inconsistency problem theory. Dominant price-maker probably systematically produces a bias from the driving-fuel price respectable by the society. For its quantification we need information from one side of the conflict. Our conclusions are only from estimated coefficients with use of the data. We lack information about average optimal driving fuel price and oil price ratio for the price-maker or about driving fuel production technology. Or we lack information about respectable driving fuel price and oil price ratio for the society.

References

[1] ADDA, J. and COOPER, R. *Dynamic Economics: Quantitative Methods and Applications.* MIT Press, 2003, ISBN 0-262-01201-4

[2] BARRO, R. J. and GORDON, D. B. A Positive Theory of Monetary Policy in a Natural Rate Model. *Journal of Political Economics*, vol. 91(4), pages 589-610, University of Chicago Press, 1983

[3] CHARI, V. V., KEHOE, P. J. and PRESCOTT, E. C. Time Consistency and Policy. In Barro, R. J. *Modern Business Cycle Theory.* Harvard University Press, 1989, ISBN 0-674-57860-0

[4] CHOCHOLATÁ, M. The Flexible-Price Monetary Models of the SKK/EUR and the CZK/EUR Exchange Rates. *APLIMAT 2007: 6th international conference,* February 2007, pages 171-178

[5] CUKIERMAN, A. Are Contemporary Central Banks Transparent about Economic Models and Objectives and What Difference Does it Make? *Federal Reserve Bank of St. Louis Review* 84, pages 15-45, 2002

[6] GREENE, W. H. *Econometric Analysis*. 5th ed. Upper Saddle River, NJ: Prentice Hall, 2002, pages 486 – 488. ISBN 0-13-066189-9

[7] HANČLOVÁ, J The Relationship Between Pressure of Demand in the Labour Market and the Wage Rate in the Czech Republic. *Econometrics of Wages*, Applied Econometrics Association, Brussels, May 27-30, 2002, pages 1-10.

[8] HANSEN, L. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* No. 50, Issue 3, journal of Econometric Society, 1982, pages 1029–1054.

[9] HAUSMAN, J.: Specification Tests in Econometrics. *Econometrica* No. 46, Issue 3, journal of Econometric Society, 1978, pages 1251 – 1271.

[10] KYDLAND, F. and PRESCOTT, E. Rules Rather than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy*, Vol. 85, 1977, pages 473-491

[11] PFANN, G. and PALM, F. Asymmetric Adjustment Costs in Non-linear Labour Models for the Netherlands and U.K. Manufacturing Sectors. *Review of Economic Studies*, vol. 60, 1993, pages 397–412

[12] RUGE-MURCIA, F. J. Does the Barro-Gordon Model Explain the Behavior of US Inflation? A Reexamination of the Empirical Evidence. *Journal of Monetary Economics*, vol. 50, 2003, pages 1375-1390

[13] SARGAN, J. The Estimation of Economic Relationships using Instrumental Variables. *Econometrica* No. 26, Issue 3, journal of Econometric Society, 1958, pages 393 – 415.

[14] STEIGER, D. and STOCK, J. H. Instrumental Variables Regression with Weak Instruments. *Econometrica* No. 65, Issue 3, journal of Econometric Society, 1997, pages 557 – 586.

[15] SURICO, P. The Fed's Monetary Policy Rule and US inflation: the Case of Asymmetric Preferences. *Journal of Economics Dynamics and Control*, vol. 31, 2007, pages 305-324

[16] SURICO, P. The Monetary Policy of the European Central Bank. *Scandinavian Journal of Economics* vol. 109(1), 2007, pages 115-35

[17] SURICO, P. Measuring the Time Inconsistency of US Monetary Policy. *Economica*, vol. 75(297), 2008, pages 22-38

RANDOM GRAININESS IN STOCHASTIC DIFFERENCE EQUATIONS

Blanka Šedivá *

University of West Bohemia, Faculty of Applied Sciences,
Department of Mathematics,
sediva@kma.zcu.cz

Abstract. This paper studies a simple extension of the linear first-order difference equations. Our generalization is based on variable, and random lengths of time steps. Our interest is devoted to magnitudes of solutions of these equations. We illustrate our approach and results one simple model.
Keywords. Stochastic difference equation, random grain.

1 Introduction

The twentieth century witnessed the rebirth of discrete-time modeling and difference equations. While the continuous calculus had been widely developed, its discrete counterpart was full of open problems. It was thus natural to study discrete phenomena with respect to the massive continuous patterns. This established numerical analysis as an important and integral part of modern mathematics and created many corresponding discrete structures. However, many blank pages remain intact and they mostly correspond to the phenomena without appropriate correspondence in the continuous settings. Our paper tries to fill one of these gaps. Namely, we study discrete models with variable steps and concentrate on those in which the variability of the step is governed by a certain probability distribution. Note that such a probabilistic model has no counterpart neither in the continuous world (no time steps) nor in the computer science or numerical analysis (the phenomena are naturally deterministic).

On the other hand, most of the economic (demographic, psychologic, political...) decisions are not regular and thus less predictable. However, this lack of regularity is very scarcely studied. Moreover, when the regularity question is addressed, it is approximated by aggregate variables and constant time steps. Since the study of rigidities (degrees of commitment, lengths of institutional tenures...) has become a central question in certain areas of macroeconomics, institutional economics and others (see Taylor [9]), it would be desirable to study also models where the lengths of intervals behave randomly.

Applicable theory in this direction would allow to form rules more effectively since it would answer the following two questions. Firstly, how long should the policymaker (central bank, CEO, etc.) govern. Secondly, how much freedom should he receive in determination of his own tenure.

Some attempts have been already made to answer this question in the different (mainly game theoretical) settings. See e.g. Lagunoff and Matsui [6], Cho and Matsui [2], Rogoff [4] or Libich and Stehlik [7] or Pospíšil, J.,Stehlík, P.,Šedivá, B. [8].

2 Economics Models

In this section we present four basic micro- and macroeconomic models which can be viewed as the motivation for our theory. More details and different problems of similar form can be found in Gandolfo [3].

Formal economic modeling began in the late 19th century with the use of differential calculus to describe and predict economic behavior. Economics became more mathematical as a discipline throughout the first half of the 20th century, but it was not until the Second World War that new techniques would allow the use of mathematical formulations in almost all of economics. This rapid systematizing of economics alarmed critics of the discipline as well as some esteemed economists.

Firstly we are interested in non-stochastic mathematical models. This models may be purely qualitative (for example, models involved in some aspect of social choice theory) or quantitative (involving rationalization of financial variables, for example with hyperbolic coordinates, and/or specific forms of functional relationships between variables). In some cases economic predictions of a model merely assert the direction of movement of economic variables, and so the functional relationships are used only in a qualitative sense: for example, if the price of an item increases, then the demand for that item will decrease.

* This work was supported by the Research Plan LC06075

The cobweb model. The cobweb model or cobweb theory explains why prices (p_t is price in period t) could be subject to periodic fluctuations in certain types of markets. It is an economic model of cyclical supply (S_t) and demand (D_t) in which there is a lag in producers' response to a change of price. Assume that both functions are linear $D_t = a_D + b_D p_t$ and $S_t = a_S + b_S p_{t-1}$. A last assumption is the market-clearing assumption, this means $D_t = S_t$, whence the finally equation is

$$b_D\, p_t - b_S\, p_{t-1} = a_S - a_D, \tag{1}$$

where p_t is price in period t and b_D, b_S, a_D, a_S are parameters from demand and supply reaction linear functions of model.

Keynesian macroeconomics models. Keynesian model of income determination theory stated that people will put aside the same level of income or save the same amount of money at all possible rates of interest. As an example of multiplier analysis, we shall describe a foreign trade multiplier. Imports M are function of income Y and exports X are assumed to be exogenous. In open economy, aggregate supply is the sum of national product Y and imports M , aggregate demand is national consumption C plus investment I plus export X . In the formal model there are four linear functions and the equilibrium condition $Y + M = C + I + X$. After substitution and elimination of depending functions we obtain

$$Y_t - (b + h - m)Y_{t-1} = a + I_0 + X_0 - M_0 + \triangle I + \triangle X, \tag{2}$$

where Y_t is the national product in period t and M_t is import, I_t national investment and X_t export; X_0, I_0, M_0 are initial condition and b, h, a, m are parameters of model.

3 Models with nonconstant steps

Models from previous section (e.g. (1) or (2)) shows that various phenomena in different fields of economics lead to simple discrete models of the form

$$\begin{aligned} \Delta x_j &= p_j x_j + f_j, \qquad j = 0, 1, 2, \ldots, N, \\ x_0 &= A \end{aligned} \tag{3}$$

where x_j is a function of real variable j, the difference operator Δ is defined by $\Delta x_j = x_{j+1} - x_j$ and p_j, f_j are real constants. The condition $x_0 = A$ is initial condition for discrete model. Using the equality $\Delta x_j = x_{j+1} - x_j$, one can rewrite (4) into

$$\begin{aligned} x_{j+1} &= (1 + p_j)x_j + f_j, \qquad j = 0, 1, 2, \ldots, N, \\ x_0 &= A \end{aligned} \tag{4}$$

This form is widely used and accepted. However, it implicitly suggests that all considered phenomena change in constant time intervals. However, this assumption goes against the intuition and economic theory. Indeed, our experience and standard circumstances of economic life imply that real-life rigidities and commitment periods vary in time. Moreover, in many fields (central banking, price setting, ...), the policymakers seek the algorithm to set the optimal length of tenures or policies. Standard discrete models with constant steps are unable to provide answers to such questions.

First of all, we have to adjust the notation to the new settings. We take into account an arbitrary finite ordered set $\mathbf{T} := \{t_0, t_1, t_2, \ldots, t_N\}$ and denote $x_j = x(t_j)$. In order to describe the length of the time gaps, we define the *graininess function* $\mu : \mathbf{T} \setminus \{t_N\} \to (0, +\infty)$:

$$\mu_j = t_{j+1} - t_j. \tag{5}$$

Finally, the relative difference will be denoted by delta in the upper index, i.e.:

$$x_j^{\Delta} := \frac{x_{j+1} - x_j}{\mu_j}. \tag{6}$$

Thus, one can replace the constant version (4) by

$$\begin{aligned} x_j^{\Delta} &= p_j x_j + f_j, \qquad j = 0, 1, 2, \ldots, N, \\ x_0 &= A \end{aligned} \tag{7}$$

or equivalently

$$x_{j+1} = (1 + \mu_j p_j)\, x_j + \mu_j f_j, \qquad j = 0, 1, 2, \ldots, N, \tag{8}$$
$$x_0 = A$$

Obviously, the latter form straightforwardly implies the existence of time transformation $\tau : \mathbf{T} \to \{0, 1, \ldots, N\}$,where $\tau(t_j) := j$, which enables to transform the non-constant model to the one with constant steps, i.e.

$$\hat{x}_{j+1} = (1 + \hat{p}_j)\, \hat{x}_j +, \hat{f}_j \qquad j = 0, 1, 2, \ldots, N, \tag{9}$$
$$x_0 = A$$

where $\hat{x}_j = x_j$, $\hat{p}_j = \mu_j p_j$, $\hat{f}_j = \mu_j f_j$.

This transformation allows us to exploit the wide range of results for (4) (see Agarwal [1] or Kelley, Peterson [5] or Pospíšil, J.,Stehlík, P.,Šedivá, B. [8]). We can write explicit form of the solution of (8) and to obtain solution in the following form

$$x_j = x_0 \prod_{i=0}^{j-1} (1 + \mu_j p_j) + \sum_{i=0}^{j-1} \mu_j f_j \left(\prod_{k=i+1}^{j-1} (1 + \mu_k p_k) \right). \tag{10}$$

One can also study the behavior of the solution of (7) as N tends to infinity,

$$x_{j+1} = (1 + \mu_j p_j)\, x_j + \mu_j f_j, \qquad j\ in \mathbf{N}_0, \tag{11}$$
$$x_0 = A$$

and the stability theory for the infinite counterpart of (4) provides basic results for the stability and boundedness of (11). The complete analysis of problem with infinite horizont can be found e.g. in Agarwal [1].

In order to motivate our choice of random steps in the following section, we restrict our attention only to the simple case when p_j, f_j and $\mu_j > 0$ are constant. We can show that for the restrict model hold (for proof see e.g. Pospíšil, J.,Stehlík, P., Šedivá, B. [8]) following lemma

LEMMA 1: Let us suppose that $p_j = p \in \mathbf{R}$, $\mu_j = \mu > 0$ and $f_j = f$ for all $j \in \mathbf{N}_0$.

(a) *If $p = 0$ then (11) converges if and only if $f = 0$.*
(b) *If $p = -\frac{2}{\mu}$ and $f_j = f$ is constant then the solution oscillates between two values x_0 and $\mu f - x_0$.*
(c) *If $p \in (-\frac{2}{\mu}, 0)$ then the solution converge to $-\frac{f}{p}$.*
(d) *If $p \notin \left\langle -\frac{2}{\mu}, 0 \right\rangle$ and $x_j \neq \frac{-f}{1+\mu p}$ then the solution diverge.*

4 Analysis of the model

We will study one of the simplest examples of equation(10). Let $p_j = -2$, $f_j = 0$ and $\mu_j = 1 + \alpha u_j$ for $j = 0, 1 \ldots$, where $\alpha \in (0, 1)$ is a *graininess parameter* (or *graininess diameter*) and $u_j \in \{-1, 1\}$ is a random variable equal to 1 or -1, each with probability 1/2. This model has all the properties described in the Introduction, it is a discrete time model with random time steps. Though being simple, it encounters an interesting behavior, as we will show below.

We have

$$x_{j+1} = -\,(1 + 2\alpha u_j)\, x_j, \quad j = 0, 1, \ldots. \tag{12}$$
$$x_0 = A$$

If u_j was zero, we would get the oscillation between x_0 and $-x_0$. In this case we have a random behavior such that in each time step we have either $p \in (-\frac{2}{\mu}, 0)$ (convergent behavior) with probability 1/2 or $p \notin [-\frac{2}{\mu}, 0]$ (divergent behavior) also with probability 1/2. We oscillate with time step μ between values $1 - \alpha$ and $1 + \alpha$. We study the qualitative behavior of the solution depending on the value of the graininess parameter α.

Using (10) we can write

$$x_j = (-1)^j x_0 \prod_{i=0}^{j-1} (1 + 2\alpha u_i) \, . \tag{13}$$

In what follows we consider only the absolute value of the solution keeping in mind that the solution oscillates due to the term $(-1)^j$.
LEMMA 2: Let x_j, $j = 0, 1, \ldots$ be the solution of (12) and let X_j^p be the p-percentile of x_j. Then

1. *function $X^p(\alpha) := \max_j X_j^p$ is an increasing function of α and*
2. *if we denote by $\iota(\alpha)$ the index where X_j^P attains its maximum, then $\iota(\alpha)$ is a decreasing function of α.*

The proof of this see (Pospíšil, J.,Stehlík, P., Šedivá, B. [8]). In Figure (1) we can see the simulated and exact 99%-percentiles for different values of graininess parameter α. The behavior of the solution could be explained the following way. The bigger the value of parameter α, the larger "peak" of the solution for lower times, but also the faster convergence towards zero.

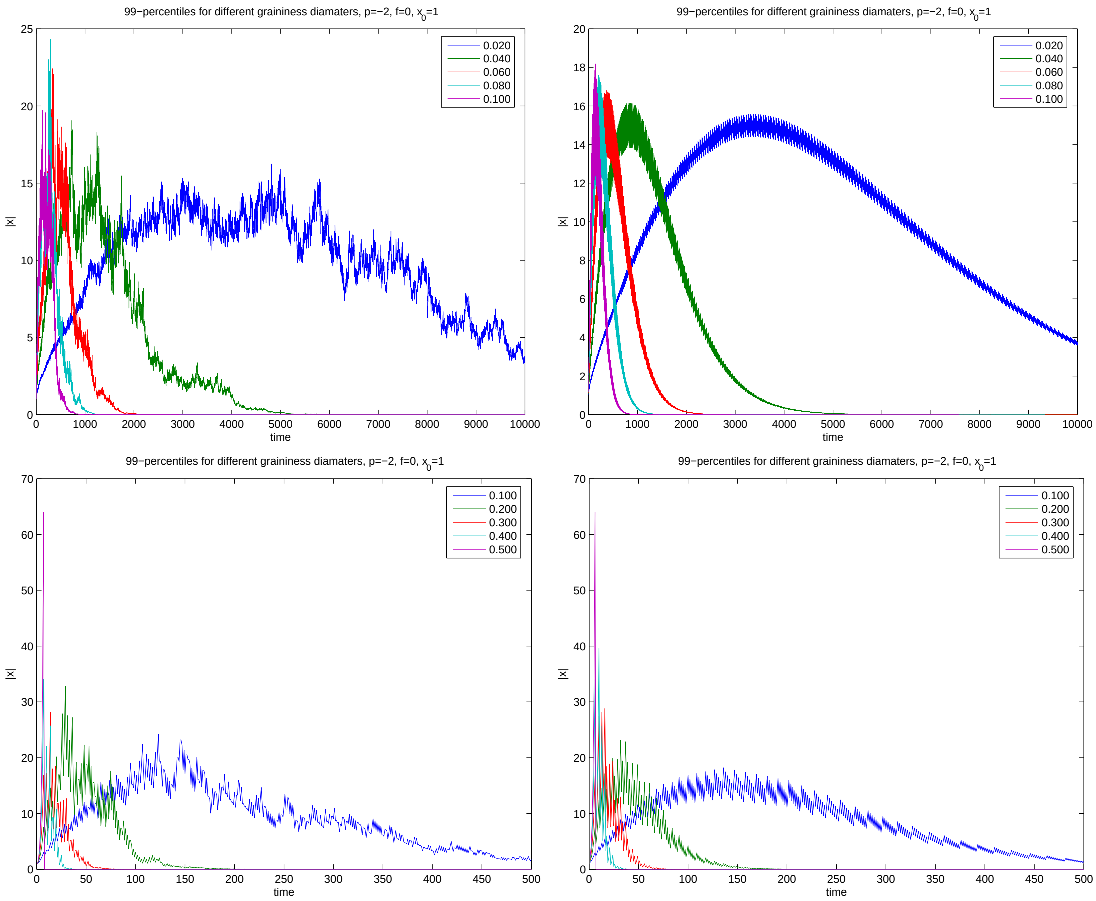


Fig. 1. Simulated and exact 99%-percentiles for different values of graininess parameter α.

5 Conclusion

This paper studies a simple extension of the linear first-order difference equations. Our generalization is based on variable, and random lengths of time steps. We will study one of the simplest examples of equation $x_{j+1} = (1 + \mu_j p_j) x_j + \mu_j f_j$, with $p_j = -2$, $f_j = 0$ and $\mu_j = 1 + \alpha u_j$ for $j = 0, 1 \ldots$. We suppose $\alpha \in (0, 1)$ and $u_j \in \{-1, 1\}$ is a random variable equal to 1 or -1, each with probability $1/2$. We study the qualitative behavior of the solution depending on the value of the graininess parameter α.

Acknowledgements

The author has been supported by the Research Plan LC06075.

References

1. R.P. Agarwal, *Difference equations and inequalities*, second ed., Monographs and Textbooks in Pure and Applied Mathematics, vol. 228, Marcel Dekker Inc., New York, 2000.
2. I. Cho and A. Matsui, *Time consistency in alternating move policy games*, Japanese Economic Review (2005), no. 56(3), 273–294.
3. G. Gandolfo, *Mathematical methods and models in economic dynamics*, North-Holland Publishing Company, 1971.
4. Rogoff K., *The optimal degree of commitment to an intermediate monetary target*, Quarterly Journal of Economics (1985), no. 100, 1169–1189.
5. W.G. Kelley and A.C. Peterson, *Difference equations*, second ed., Harcourt/Academic Press, San Diego, CA, 2001, An introduction with applications.
6. R. Lagunoff and A. Matsui, *Asynchronous choice in repeated coordination games*, Econometrica (1997), no. 65, 1467–77.
7. J. Libich and P. Stehlik, *Macroeconomic games on time scales*, Proceedings of Dynamic Systems and Applications 11, To Appear.
8. J. Pospíšil, P. Stehlík, and B. Šedivá, *Stochastic difference equations: Random graininess in stochastic difference equations and its applications in macroeconomics-in proceeding*, University of West Bohemia, Faculty of Applied Sciences (2009).
9. J.B. Taylor, *Staggered wage setting in a macro model*, American Economic Review (1979), no. 69, 108–113.

PROBABILISTIC PROPERTIES OF THE CONTINUOUS DOUBLE AUCTION - UNIFORM CASE

Martin Šmíd

Institute of Information Theory and Automation,
Academy of Sciences of the Czech Republic *†
mailto:smid@utia.cas.cz

August 5, 2009

Abstract

We study probabilistic properties of a zero intelligence model of a limit order market, very similar to those of [4] and [8]. We (recursively) describe the distributions of the order books and the best quotes. Based on these theoretical results, a procedure for statistical inference of the model may be designed and the evolution of the process may be simulated more efficiently then by the crude simulation of all the events.
Keywords. Continuous double auction, limit order markets, distribution, simulation, statistical inference, price increment tails.

1 Introduction

Recently, several zero-intelligence models of limit order markets[1] have been introduced: [4] shows that even a simple model assuming a Poisson orders' arrival and the uniform distribution of the limit prices generates fat-tailed price increments. [6] computes the rate of the tail by means of a mean-field approximation. [8] introduce a model, similar to that of [4], including, in addition, order cancelations.[2] A generalization of this model incorporating the statistical properties of real-life order books is made by [5]. Even though many stylized facts cannot be explained by the zero-intelligence models (see [1]) they may be regarded as a good first approach.

Despite the great effort of the authors, no exact probabilistic description of any of the zero-intelligence has been published yet, which, among others, disallows statistical inference of the models. In our recent work [7] we formulated a general model, covering fully the models of [4] and [8] (the latter after a discretization) and partially the one by [5] and we recursively described its distribution. The purpose of the present paper is to apply these results to a simpler (uniform) model of the continuous double auction.

In particular, we consider a model with Poisson order flows, constant cancelation rate and continuous uniform distribution of limit prices (Section 2) which may be easily transformed both to the Maslov's model [4] (by sending the cancelation rate to infinity) and to the model of [8] (by a slight redefinition and the rounding of the prices, see [7], Section 3). The distribution of the present model is described in Section 3). Further (Section 4), we discuss applications of our results. Finally, we conclude the paper (Section 5).

*This work is supported by the grants No. 402/09/0965 and No. 402/07/1113 of the Czech Science Foundation and by project No. LC06075 of the Ministry of Education, Youth and Sports.

†This work is supported by the grants no. 402/06/1417 and 402/07/1113 of the Czech Science Foundation.

[1] For a detailed description of a limit order market, see e.g. [8].

[2] In a subsequent work [2] it is argued that, despite the radical assumption of economic agents acting as gas particles, zero intelligence models replicate several stylized facts found in real-life limit markets.

2 Definitions

2.1 Inputs

In the model we study, the orders of all the four types (i.e. buy/sell market/limit orders) arrive with (possibly different) constant intensities, the intensity of the limit orders' cancelations is constant and the limit prices are uniformly distributed. In particular,

- the arrivals of sell limit orders form a marked Poisson process x with an intensity ι and with marks $\pi_i \sim \mathrm{U}(-h,h)$, $u_i \sim \mathrm{Exp}(\upsilon)$, standing for the absolute limit price, lifetime of the order respectively, where h is a finite positive constant[3] and where all the marks are mutually independent and independent of the arrival times,
- the arrivals of sell market orders form a Poisson process $\bar{x}$ with an intensity $\bar{\iota}$,
- the process y of arrivals of buy limit orders is Poisson with an intensity ω and with independent, mutually independent marks $(r_i, \zeta_i)_{i=1}^{\infty}$ (absolute limit prices, cancelation times respectively) such that $r_i \sim \mathrm{U}(h,h)$ and $v_i \sim \mathrm{Exp}(z)$ for some constant z,
- the process $\bar{y}$ of arrivals of buy market orders is Poisson with an intensity $\bar{\omega}$,
- the random elements x, y ,$\bar{y}$, $\bar{x}$ are mutually independent.

2.2 Dynamics of the System

We describe the state of the market at a time τ by a tuple

$$\Xi_\tau = (A_\tau, B_\tau) \tag{1}$$

where A_τ and B_τ are simple atomic measures (collections of points on the real line) describing the sell order book, buy order book respectively, each atom (point) standing for a waiting limit order with the (absolute) limit price equal to its location. We denote

$$a_\tau \stackrel{\triangle}{=} \min\{\pi : \pi \text{ is an atom of } A_\tau\} \wedge h \tag{2}$$

the value of the (best) ask (we put $\min \emptyset = \infty$) and

$$b_\tau \stackrel{\triangle}{=} \max\{\pi : \pi \text{ is an atom of } B_\tau\} \vee -h \tag{3}$$

the value of the (best) bid (we put $\max \emptyset = -\infty$).

We assume our process to start by a single limit order on each side, i.e. both A_0 and B_0 contain a single (deterministic) point π_0, ρ_0 respectively, such that $\rho_0 < \pi_0$. Further, we assume the lifetime u_0 (v_0) of the starting sell (buy) order to be exponentially distributed with parameter υ, (z) such that u_0, v_0 and $(x, \bar{x}, y, \bar{y})$ are mutually independent.

We let the process Ξ (possibly) jump only at the times $(\tau_i)_{i\in\mathbb{N}}$, $(\bar{\tau}_i)_{i\in\mathbb{N}}$, $(\sigma_i)_{i\in\mathbb{N}}$, $(\bar{\sigma}_i)_{i\in\mathbb{N}}$, denoting the jump times of x, $\bar{x}$, y, and $\bar{y}$ respectively, or at $(\eta_i)_{i\in\mathbb{N}}$, $(\zeta_i)_{i\in\mathbb{N}}$ where $\eta_i \stackrel{\triangle}{=} \tau_i + u_i$ and $\zeta_i \stackrel{\triangle}{=} \sigma_i + v_i$, $i \geq 0$ (i.e. the cancelation times of sell limit orders, buy limit orders respectively, we put $\tau_0 = \sigma_0 = 0$).

If τ is one of the possible jump times of Ξ then the change of Ξ at τ is defined as follows:

- if $\tau = \tau_i$ for some $i > 0$ and $\pi_i > b_{\tau-}$ then π_i is added into A
- if $\tau = \eta_i$ for some $i \geq 0$ and π_i is present in $A_{\tau-}$ then π_i is removed from A
- if $\tau = \bar{\tau}_i$ for some $i > 0$ and $B_{\tau-} \neq 0$ then $b_{\tau-}$ is removed from B
- if $\tau = \sigma_i$ for some $i > 0$ and $\rho_i < a_{\tau-}$ then ρ_i is added into B
- if $\tau = \zeta_i$ for some $i \geq 0$ and ρ_i is present in $\mathbb{N}_{\tau-}$ then ρ_i is removed from B
- if $\tau = \bar{\sigma}_i$ for some $i > 0$ and $A_{\tau-} \neq 0$ then $a_{\tau-}$ is removed from A
- if τ does not fit any of the conditions above then both A and B are left unchanged at τ.

[3] Later we define our model also for infinite h

It follows from the definition of the input processes that

$$\mathbb{P}[\text{any pair of } \tau_i, \bar{\sigma}_i, \sigma_i, \bar{\tau}_i, \eta_i, \zeta_i, i \in \mathbb{N}, \text{ coincide}] = 0 \tag{4}$$

hence our definition is correct with probability one. Further, since

$$\mathbb{P}[\text{any pair of } \pi_i, \rho_i, i \in \mathbb{N}, \text{ coincide}] = 0 \tag{5}$$

it is guaranteed that A_τ and B_τ are simple (with no overlapping points) at each $\tau \geq 0$.

For a "formula based" definition of Ξ, see [7].

3 Distributions

In the present Section, we describe probabilistic properties of Ξ and of the process of the best quotes

$$\xi_\tau \triangleq (a_\tau, b_\tau). \tag{6}$$

An uninterested reader may skip this section; however, (s)he will need to return here for a notation sometimes.

We start with a nearly obvious statement:

Proposition 1 *(Markov properties of Ξ)*

(i) For any deterministic $0 \leq s_1 < \cdots < s_n < s$,

$$\mathcal{L}\left(\Xi_{s+\bullet} \,|\Xi_s, \Xi_{s_1}, \dots, \Xi_{s_n}\right) = \mathcal{L}\left(\Xi_{s+\bullet} \,|\Xi_s\right) \tag{7}$$

(here, for any random elements X, Y, symbol $\mathcal{L}(X\,|Y)$ denotes the conditional distribution of X given Y).

(ii) Relation (7) keeps holding even if $s_1 < \cdots < s_n < s$ are optional times with respect to the filtration generated by Ξ.

Proof. Statement (i) follows from the fact that all the inter-jump times and the orders' lifetimes are exponential and that the value of Ξ at a jump is fully determined by the value at the last jump and the type of jump time. Part (ii) stems from the fact that Ξ is pure jump type process (see [3], chp. 12). □

To go on, we need to introduce some notation: For any interval $\beth$ such that $\beth = [s,t)$ or $\beth = [s,t]$, $0 \leq s \leq t$, denote

$$\tilde{a}^{\beth} : \beth \to \mathbb{R}, \qquad \tilde{a}^{\beth}_\tau \triangleq \max_{\theta \in \beth \cap [\tau,\infty)} a_\theta, \quad \tau \in \beth, \tag{8}$$

and introduce a function

$$\kappa^{\beth}(\mathfrak{p}) \triangleq \sum_{j=0}^{J^{\beth}} \frac{\iota^\star}{\upsilon}[\mathfrak{p} \wedge \tilde{a}^{\beth}_{[j-1]} - \tilde{a}^{\beth}_{[j]}]^+[1 - \exp\{-\upsilon(t - \varsigma^{\beth}_j)\}], \qquad \mathfrak{p} \in \mathbb{R}, \tag{9}$$

where $\iota^\star \triangleq (2h)^{-1}\iota$, $J^{\beth}$ is the number of jumps of $\tilde{a}^{\beth}$ on $\beth$, $\varsigma^{\beth}_1 < \cdots < \varsigma^{\beth}_{J^{\beth}}$ denote the jump times themselves (we put $\varsigma^{\beth}_0 \triangleq s$) and where

$$\tilde{a}^{\beth}_{[j]} \triangleq \tilde{a}^{\beth}_{\varsigma^{\beth}_j}, \qquad 0 \leq j \leq J^{\beth}, \tag{10}$$

(we put $\tilde{a}^{\beth}_{[-1]} \triangleq h$). Further, denote $0 = \vartheta_0 < \vartheta_1 < \dots$ the sequence of jump times of ξ.

3.1 Order Books

Until the end of subsection 3.1, fix $s \leq t$ fulfilling one of the conditions

- both s and t are deterministic,
- $s = \vartheta_k$ and $t = \vartheta_i$ for some (deterministic) $k, i \in \mathbb{N}$, $k \leq i$,

and agree to write $\tilde{a}$ instead of $\tilde{a}^{[s,t]}$, J instead of $J^{[s,t]}$, etc.

The following two Propositions describe the distribution of the order books given Ξ_s, $\xi_{[s,t]}$ where the latter symbol denotes the trajectory of ξ restricted to interval $[s,t]$.

Proposition 2 *(conditional distribution of A_t)*

$$A_t = \delta_{a_t} + \sum_{j=1}^{J} \delta_{\tilde{a}_{[j]}} e_j + \sum_{j=1}^{K} \delta_{\alpha_j^s} d_j + L. \tag{11}$$

where δ_q denotes the Dirac measure concentrated in q (i.e. a single point $\{q\}$) and

- $e_1, \dots, e_J$ *are binary variables such that*

$$\mathcal{L}\left(e_j \left| \Xi_s, \xi_{[s,t]} \right.\right) = \text{Alternative}\left(\exp\left\{-\upsilon\left(t - \varsigma_j\right)\right\}\right), \qquad 1 \le j \le J, \tag{12}$$

- $\alpha_1 < \dots < \alpha_K$ *are all the atoms of A_s whose location is greater then $\tilde{a}_s$,*

- $d_1, \dots, d_K$ *are binary variables with*

$$\mathcal{L}\left(d_j \left| \Xi_s, \xi_{[s,t]} \right.\right) = \text{Alternative}\left(\exp\left\{-\upsilon\left(t - s\right)\right\}\right), \qquad 1 \le j \le K, \tag{13}$$

- *L is (conditionally on $\Xi_s, \xi_{[s,t]}$) a non-homogenous Poisson process on (a_t, h) with the intensity given by distribution function κ,*

- $e_1, \dots, e_J, d_1, \dots, d_K, L$ *are conditionally independent given $(\Xi_s, \xi_{[s,t]})$.*

Proof. See [7]. □

Remark 1 *Using the symmetry of A and B, an analogous formula for the distribution of B may be obtained as a corollary of Proposition 2.*

3.2 Best Quotes

Until the end of subsection 3.2, fix deterministic integers $0 \le k < i$ and agree to abbreviate $\tilde{a}^{[\vartheta_k, \vartheta_i)}$ as $\tilde{a}$, $J^{[\vartheta_k, \vartheta_i)}$ as J etc.

Our present goal is to specify the conditional distribution of ϑ_{i+1} and $\xi_{\vartheta_{i+1}}$ given Ξ_{ϑ_k} and $\xi_{[\vartheta_k, \vartheta_i]}$. For a better intuitive understanding, we add an additional step to the definition: we consider an supplementary variable χ_{i+1}, coding the type of the event happening at the time ϑ_{i+1}. Denoting the possible values of χ_{i+1} symbolically by $\mathfrak{a}^+$, $\mathfrak{a}^-$, $\mathfrak{b}^+$ and $\mathfrak{b}^-$, we define it as

$$\chi_i \triangleq \begin{cases} \mathfrak{a}^+ & \text{if } \Delta a_{\vartheta_i} > 0, \Delta b_{\vartheta_i} = 0, \\ \mathfrak{a}^- & \text{if } \Delta a_{\vartheta_i} < 0, \Delta b_{\vartheta_i} = 0, \\ \mathfrak{b}^+ & \text{if } \Delta a_{\vartheta_i} = 0, \Delta b_{\vartheta_i} > 0, \\ \mathfrak{b}^- & \text{if } \Delta a_{\vartheta_i} = 0, \Delta b_{\vartheta_i} < 0. \end{cases} \tag{14}$$

Denote

$$\varsigma_i^{\mathfrak{a}^-} = \iota^\star(a_{\vartheta_i} - b_{\vartheta_i}), \qquad \varsigma_i^{\mathfrak{a}^+} = \mathbf{1}_{[a_{\vartheta_i} < h]}(\bar{\omega} + \upsilon), \tag{15}$$

$$\varsigma_i^{\mathfrak{b}^-} = \mathbf{1}_{[b_{\vartheta_i} > -h]}(\bar{\iota} + z), \qquad \varsigma_i^{\mathfrak{b}^+} = \omega^\star(a_{\vartheta_i} - b_{\vartheta_i}), \tag{16}$$

$$\varsigma_i = \varsigma_i^{a+} + \varsigma_i^{a-} + \varsigma_i^{b+} + \varsigma_i^{b-}, \qquad \omega^\star \triangleq (2h)^{-1}\omega, \tag{17}$$

($\iota_x^\star$ was introduced at the definition of $\kappa^\daleth$).

Proposition 3 *(distribution of ϑ_{i+1})*

$$\mathcal{L}\left(\Delta\vartheta_{i+1} \left| \xi_{[\vartheta_k, \vartheta_i]}, \Xi_{\vartheta_k} \right.\right) = \text{Exp}\left(\varsigma_i\right). \tag{18}$$

Proposition 4 *(distribution of χ_{i+1})*

$$\mathcal{L}\left(\chi_{i+1}\left|\Delta\vartheta_{i+1},\xi_{[\vartheta_k,\vartheta_i]},\Xi_{\vartheta_k}\right.\right)=\left(\frac{\varsigma_i^{\mathfrak{a}^+}}{\varsigma_i},\frac{\varsigma_i^{\mathfrak{a}^-}}{\varsigma_i},\frac{\varsigma_i^{\mathfrak{b}^+}}{\varsigma_i},\frac{\varsigma_i^{\mathfrak{b}^-}}{\varsigma_i}\right). \tag{19}$$

Proposition 5 *(distribution of $a_{\vartheta_{i+1}}$)*

(i) If $\chi_{i+1}=\mathfrak{a}^+$ then

$$\begin{aligned}&\mathbb{P}[a_{\vartheta_{i+1}}>\mathfrak{p}|\Delta\vartheta_{i+1},\chi_{i+1},\xi_{[\vartheta_k,\vartheta_i]},\Xi_{\vartheta_k}]\\&\qquad=\mathbf{1}_{[\mathfrak{p}<h]}\exp\left\{-\kappa(\mathfrak{p})\right\}\left(1-\exp\{-\upsilon(\vartheta_{i+1}-\vartheta_k\}\right)^{K(\mathfrak{p})}\\&\qquad\qquad\times\prod_{1\le j<J,\tilde{a}_{[j]}\le\mathfrak{p}}\left(1-\exp\{-\upsilon(\vartheta_{i+1}-\varsigma_{j+1})\}\right)\end{aligned}$$

where $K(\mathfrak{p})$ is the number of points of A_{ϑ_k} belonging to interval $(\tilde{a}_{\vartheta_k},\mathfrak{p}]$,

(ii) If $\chi_{i+1}=\mathfrak{a}^-$ then

$$\mathcal{L}\left(a_{\vartheta_{i+1}}\left|\Delta\vartheta_{i+1},\chi_{i+1},\xi_{[\vartheta_k,\vartheta_i]},\Xi_{\vartheta_k}\right.\right)=\mathrm{U}\left(b_{\vartheta_i},a_{\vartheta_i}\right) \tag{20}$$

(iii) If $\chi_{i+1}\in\{\mathfrak{b}^-,\mathfrak{b}^+\}$ then

$$\mathcal{L}\left(a_{\vartheta_{i+1}}\left|\Delta\vartheta_{i+1},\chi_{i+1},\xi_{[\vartheta_k,\vartheta_i]},\Xi_{\vartheta_k}\right.\right)=\delta_{a_{\vartheta_i}}. \tag{21}$$

Proof of Propositions 3-5. See [7]. □

Remark 2 *The formula for the distribution of $b_{\vartheta_{i+1}}$ is symmetric.*

4 Applications

Even if the formulas describing the distribution are quite complicated, several useful and practical applications may be constructed based on them.

4.1 Statistical inference of the model

The problem of the inference of our model (i.e. testing its validity and estimating its parameters) may be broken into five parts: an inference of all the four input processes. Since the flows of market orders may be studied by standard techniques,[4] we do not discuss them in the present work. To test and infer the distributions of $(\tau_\nu)_{\nu\in\mathbb{N}}$ and $(p_\nu)_{\nu\in\mathbb{N}}$, may also apply standard statistical tools. On the other hand, we are getting into difficulties with cancelations rates υ and z because the sample of observed cancelations of the orders is censored (some orders are executed before they could be canceled) - in fact, this is the point where our theoretical results may help: based on our formulas, the cancelation rates (together with the market limit orders' arrival rates, if we are lacking detailed enough data to infer them directly) may be estimated. For details, see Section 4.2. of [7].

4.2 Efficient Simulation

Since we do not know analytic formulae for the (unconditional) distributions of Ξ and ξ, a Monte Carlo simulation is needed when working with the processes. However, our knowledge of the (conditional) distributions may help us to speed up the simulations significantly.

Contrary to the "crude" simulation, when each variable defining the system is generated, our procedure allows us to omit a vast majority of the steps of such a simulation: Thanks to our results, it suffices to generate the process ξ only and to draw the values of Ξ only at the times when they are actually needed.[5] Our procedure is as follows:

[4] Of course, this is true only if we are able to distinguish the moves of bid, ask respectively, caused by cancelations from those due to market orders.

[5] In fact, our procedure "is able" to generate the order books only at the times of jumps of ξ; however, if we wanted to draw all the values of the order books, we may use the crude simulation since the last jump of ξ.

For each $i \in \mathbb{N}_0$:

1. Generate $\Delta\vartheta_{i+1}$ and χ_{i+1} from their conditional distribution given the $(\xi_{[s_i,\vartheta_i]}, \Xi_{s_i})$ where s_i is the time of the last generation of a value of Ξ (by Proposition 3 and 4)
2. Generate $a_{\vartheta_{i+1}}$ and $b_{\vartheta_{i+1}}$ from their conditional distributions given $(\chi_{i+1}, \Delta\vartheta_{i+1}, \xi_{[s_i,\vartheta_i]}, \Xi_{s_i})$ (by Proposition 5 and its symmetric counterpart)
3. If needed, generate $B_{\vartheta_{i+1}}$ and/or $A_{\vartheta_{i+1}}$ from their conditional distributions given $(\xi_{[s_1,\vartheta_{i+1}]}, \Xi_{s_i})$ (by Proposition 2 and its symmetric counterpart).

5 Conclusion

We examined theoretically some zero-intelligence models of limit order markets and showed some applications of our results. In particular, we outlined a way of a rigorous statistical inference of such models and we proposed a way of an efficient Monte Carlo simulation of the models. Further generalizations of the model are subjects of our future research.

References

[1] J. P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart. Fluctuations and response in financial markets: the subtle nature of 'random' price changes. *Quantitative Finance*, 4(2):176–190, 2004. Preprint at `http://www.science-finance.fr/papers/0307332.pdf`.

[2] J. D. Farmer, P. Patelli, and I .I. Zovko. The predictive power of zero intelligence in financial markets,. *Proc. Nat. Acad. Sci.*, 102:2254–2259, 2005. Preprint at `http://www.pnas.org/cgi/content/full/102/6/2254`.

[3] O. Kallenberg. *Foundations of Modern Probability*. Springer, New York, second edition, 2002.

[4] S. Maslov. Simple model of a limit order driven market. *Physica A*, 278:571–578, 2000.

[5] S. Mike and J. D. Farmer. An empirical behavioral model of liquidity and volatility. *Journal of Economic Dynamics & Control*, 32(1):200–234, 2008.

[6] F. Slanina. Mean-field approximation for a limit order driven market. *Phys. Rev.*, 64:056136, 2001. Preprint at `http://xxx.sissa.it/abs/cond-mat/0104547`.

[7] M. Šmíd. Probabilistic properties of the continuous double auction, 2008. Submitted to Kybernetika, Available at SSRN: `http://ssrn.com/abstract=1107114`.

[8] E. Smith, J. D. Farmer, L. Gillemot, and S. Krishnamurthy. Statistical theory of the continuous double auction. *Quantitative finance*, 3(6):481–514, 2003. `http://www.santafe.edu/~baes/jdf/papers/doubleauctionfinal.pdf`.

ANALYSIS OF DISCRETE-EVENT SIMULATION RESULTS USING TRADE-OFF INFORMATION

Tadeusz Trzaskalik[1], , Maciej Nowak[2], Petr Fiala[3]

[1,2]Karol Adamiecki University of Economics in Katowice, Department of Operations Research
[3]University of Economics in Prague, Department of Econometrics
ttrzaska@ae.katowice.pl, nomac@ae.katowice.pl, pfiala@vse.cz

Abstract. In the paper a technique for analyzing the results of discrete-event simulation is proposed. A finite number of scenarios (decision alternatives) is analyzed. The simulation is used for generating distributional evaluations of alternatives with respect to attributes. It is supposed that the decision maker (DM) uses multiple attributes to evaluate alternatives under consideration and tries to find a solution preferred to all other solutions (most preferred solution). To solve a multiattribute problem, one have to assess information about DM's preferences. Three general approaches to solving multicriteria problems can be considered: prior articulation of preferences, interactive articulation of preferences and posterior articulation of preferences. In the paper we use the second approach and assume that the preference information is obtained in a stepwise manner. The technique proposed in the paper uses the information about trade-offs among attributes. In each iteration a candidate alternative is proposed to the DM. If he/she is satisfied with the proposal, the procedure ends. Otherwise, the DM is asked to specify a criterion to be improved and a lexicographic order of criteria that can be decreased starting from the one, that can be decreased first. Relations between distributions of trade-offs is exploited to generate a new proposal. The procedure continues until the decision maker is satisfied with the proposition. An application in the production process control problem is presented to illustrate the technique.
Keywords. Discrete-event simulation, Multiattribute analysis, Decision-making under risk, Trade-off information, Stochastic Dominance

1. Introduction

Computer simulation is an important tool in a wide variety of disciplines: economics, engineering, biology, chemistry, the social science and others. Through computer simulation one can observe the behavior of real-life systems that are too difficult to examine analytically. One of the main advantages of simulation consists in the possibility to analyze problems that include uncertainties. A simulation model transforms distributions of input variables into distributions of output variables, which are used to measure the efficiency of the system under consideration. In many complex real-life decision problems, simulation modeling is the only way to evaluate decision alternatives with respect to criteria.

This paper considers the problem of analyzing the results of simulation experiments. We focus on discrete-event systems, i.e systems in which the state variables change instantaneously through jumps at discrete points in time (Rubinstein, Kroese [1]). Sophisticated statistical methods are usually proposed for analyzing the data obtained from the simulation model (Robinson [2], Nowak [3]). Such techniques, however, are usually useless when a multiple criteria are used for evaluating decision alternatives or scenarios. In such case additional information about the preferences of the decision-maker (DM) is required to solve the problem. Three main approaches are usually used for acquiring such information: prior articulation of preferences, posterior articulation of preferences, and interactive articulation of preferences. Prior articulation of preferences means that the decision problem is solved in two steps. First, preference information is collected. Next, this information is used for constructing complete or partial order of alternatives. Multiattribute utility theory (Keeney, Raiffa [4]) is a classical example of this approach. The posterior articulation of preferences is based on opposite assumptions. This approach tries to find a good approximation of nondominated frontier in order to provide information about the possible solutions to the DM. After that the DM is asked to evaluate solutions presented to him/her and this information is used for generating the final solution. ADBASE technique proposed by Steuer [5] is an example of such approach.

Interactive approach is probably the technique, that is used most often for solving multiattribute problems. It assumes that the DM is able to provide preference information with respect to a given solution or a given set of solutions (local preference information). The interaction takes place between the DM, an analyst and a computer model of the decision problem. The DM has to express his/her local preferences with respect to a series of solutions, which are presented to him/her in a stepwise manner and are partly the result of his/her previous answers. Various advantages of the interactive approach are mentioned. First, limited amount of preference information is required from the DM as compared to other techniques. Second, The DM has to express his/her preferences evaluating well-defined solutions which are known to exist and be feasible. It is also pointed out that the DM is more closely involved in the solution process, and as a result the final solution has a better chance to be practically implemented.

The main goal of this paper is to propose a new technique for analyzing results of discrete-event simulations. We assume that a finite number of possible scenarios (decision alternatives) are taken into account. The alternatives are evaluated with respect to a set of attributes. The simulation model is used for generating distributional evaluations of alternatives with respect to attributes. We use interactive approach for aiding the DM to choose the final solution of the problem. The paper is structured as follows. The problem is defined in section 2. Section 3 provides basic information about trade-offs and the way in which they can be used in multiattribute analysis. The procedure is presented in section 4. Next section provides a numerical example. The last section gives conclusions.

2. Formulation of the problem

The problem considered in this paper can be defined as follows:

- A finite set of decision alternatives $A = \{a_1, ..., a_m\}$ is considered.
- Alternatives are evaluated with respect to n attributes.
- The attributes are defined in such a way that a larger value is preferred to a smaller one. If the DM wants to minimize values of attributes, the data have to be transformed to negative values.
- The model describing the behavior of the system has at least one random input variable.
- Stochastic simulation is used for generating distributional evaluations of alternatives with respect to attributes.
- The DM tries to find a solution preferred to all other solutions (the most preferred solution).

Simulation runs produce sequences of observations for each alternative and each attribute. This data can be used for generating distributions describing the efficiency of each alternative. Let X_{ip} be the evaluation of alternative a_i with respect to attribute p, and $F_{ip}(x)$ a cumulative probability distribution function defined as follows:

$$F_{ip}(x) = \Pr(X_{ip} \leq x)$$

The attributes are supposed to be probabilistically independent, and are also supposed to satisfy the preference independence condition. Thus, the overall comparison of two alternatives can be decomposed into one-attribute comparisons of probability distributions. In this paper we will use stochastic dominance rules for modeling preferences of the DM in relation to each attribute. First stochastic dominance and second stochastic dominance are defined as follows:

Definition 1. (FSD – First Degree Stochastic Dominance)

X_{ip} dominates X_{jp} by FSD rule ($X_{ip} \succ_{\text{FSD}} X_{jp}$) if and only if
$F_{ip}(x) \neq F_{jp}(x)$ and $F_{ip}(x) - F_{jp}(x) \leq 0$ for all $x \in$ R.

Definition 2. (SSD – Second Degree Stochastic Dominance)

X_{ip} dominates X_{jp} by SSD rule ($X_{ip} \succ_{\text{SSD}} X_{jp}$) if and only if
$F_{ip}(x) \neq F_{jp}(x)$ and $\int_{-\infty}^{x} \left(F_{ip}(y) - F_{jp}(y)\right)dy \leq 0$ for all $x \in$ R.

Hadar and Russel [6] show that the FSD rule is equivalent to the expected utility maximization rule for all decision makers preferring larger outcomes, while the SSD rule is equivalent to the expected utility maximization rule for risk-averse decision makers preferring larger outcomes. In our procedure we will use FSD and SSD rules for comparing evaluations of alternatives with respect to attributes, as well as for comparing distributions of point-to-point trade-offs among attributes.

3. Trade-offs among attributes in decision-making problems under risk

A trade-off is usually defined for a particular solution and for a selected pair of the attributes. It specifies the amount by which the value of one attribute increases while that of the other one decreases when a particular solution is replaced by another given solution.

Let us start with a decision making problem under certainty. For a pair of alternatives a_i and a_j and a pair of attributes p and q, a point-to-point trade-off is the ratio of a relative value increase in one attribute (p) per unit of value decrease in the reference attribute (q) when the alternative a_i is replaced by the alternative a_j.

$$T_{ji}^{pq} = \frac{X_{jp} - X_{ip}}{X_{iq} - X_{jq}}$$

Let us assume that the DM analyzed alternative a_i and decided that the evaluation with respect to attribute p should be improved, while the evaluation with respect to attribute q may be decreased. In this case we will look for alternatives a_j such that $X_{jp} > X_{ip}$ and $X_{jq} \le X_{iq}$, and choose the one for which the increase of attribute p gets the maximal value. If such alternatives do not exist, then an alternative maximizing point-to-point trade-off can be proposed.

The situation complicates when a decision problem under risk is considered. In such case a point-to-point trade-off is a random variable. It's distribution is a mixture of distributions of four random variables: X_{ip}, X_{iq}, X_{jp}, and X_{jq}. Let N be the number of simulation runs performed for each alternative, while x_{ipl} – value of attribute p obtained in l-th simulation run performed for alternative a_i. The procedure, that can be used for generating distribution of T_{ji}^{pq} consists of following steps:

1. Perform simulation runs for alternative a_i.
2. Perform simulation runs for alternative a_j.
3. Assume $k = 1$, $l = 1$.
4. Let $r = (k - 1) \times l$
5. Calculate t_{jir}^{pq} according to the following formula:

$$t_{jir}^{pq} = \begin{cases} \dfrac{x_{jpl} - x_{ipk}}{x_{iqk} - x_{jql}} & \text{if } x_{jpl} > x_{ipk} \text{ and } x_{iqk} > x_{jql} \\ M & \text{if } x_{jpl} > x_{ipk} \text{ and } x_{iqk} \le x_{jql} \\ 0 & \text{if } x_{jpl} = x_{ipk} \\ -M & \text{if } x_{jpl} < x_{ipk} \end{cases}$$

   where M is a "big number".
6. Assume $l = l + 1$. If $l \le N$, go to 4, otherwise go to 7.
7. Assume $k = k + 1$, $l = 1$. If $k \le N$, go to 4, otherwise go to 8.
8. Generate the distribution of T_{ji}^{pq} assuming equal probabilities for t_{jir}^{pq}, $r = 1, \dots, N \times N$.

The procedure takes into account various situations that can occur, when values of attributes are compared. Formula for calculating t_{jir}^{pq} is valid both for classical situation ($x_{jpl} > x_{ipl}$ and $x_{iql} > x_{jql}$) and for other cases. The procedure compares results obtained in l-th simulation run performed for alternative a_i with the results obtained in all simulation runs performed for alternative a_j. Thus, finally we obtain $N \times N$ observations for variable T_{ji}^{pq}. This observations are used for generating probability distribution of point-to-point trade off among attributes p and q.

4. The procedure

The main ideas of the procedure are as follows:

- a candidate for most preferred solution is presented to the DM at each iteration,
- if the DM is satisfied with the proposal – the procedure ends,
- otherwise – the DM is asked to select the attribute to be improved and the attributes that can be decreased, ordered lexicographically starting with the one to be decreased first,
- information about relations between distributions of trade-offs is used to generate a new candidate.

The procedure starts with the identification of the first proposal. We propose to use SD rules and min-max criterion. The first proposal is identified in the following steps:

1. Identify SD relations between distributional evaluations for each pair of alternatives and for each attribute.
2. For each alternative compute:
$$\bar{d}_i = \max_{p \in \{1,\ldots,n\}} \{d_{ip}\}, \quad \text{where} : d_{ip} = \mathrm{card}(D_{ip}),\ D_{ip} = \{a_j : X_{jp} \succ_{\mathrm{FSD}} X_{ip} \vee X_{jp} \succ_{\mathrm{SSD}} X_{ip}\}$$
3. Choose the alternative for which $\bar{d}_i$ is minimal

The main idea of the procedure is to use distributions of point-to-point trade-offs in order to identify a new proposal. In the initial phase of the procedure the DM is asked to specify the data to be presented. For each attribute he/she may choose one or more scalar measures. Both expected outcome measures (mean, median, mode) and variability measures (standard deviation, semideviation, probability of getting outcomes not greater or not less than a target value) can be chosen. Let a_s be the current proposal, while $\mathrm{A}^{(l)}$ – the set of alternatives considered in iteration l ($\mathrm{A}^{(1)} = \mathrm{A}$). Following steps are performed in iteration l:

1. Present the proposal a_s to the DM and ask whether he/she is satisfied with the proposal. If the answer is "yes" – end the procedure, otherwise go to step 2.
2. Ask the DM to specify the attribute that should be improved first and to set the order of the remaining attributes, starting from the one that can be decreased first. Let p be the number of the attribute that the DM would like to improve, while $\{q_1, q_2, \ldots, q_{n-1}\}$ is the order of the attributes that can be decreased.
3. Identify the set of alternatives satisfying the requirements formulated by the DM:
$\mathrm{A}^{(l+1)} = \{a_i : a_i \in \mathrm{A}^{(l)}, a_i \neq a_s, \neg X_{sp} \succ_{\mathrm{FSD}} X_{ip}, \neg X_{sp} \succ_{\mathrm{SSD}} X_{ip}\}$
4. Assume: $\mathrm{B} = \mathrm{A}^{(l+1)}$, $k = 1$.
5. Generate probability distributions of trade-offs $T_{is}^{pq_k}$ for each i such that $a_i \in \mathrm{B}$.
6. Compare distributions of trade-offs with respect to FSD/SSD rules and identify the set of non-dominated distributions. If the number of non-dominated distributions is equal to 1, assume the corresponding alternative to be the new proposal and go to the next iteration.
7. Identify the alternatives with dominated trade-offs and exclude them from the set B.
8. If $k < n - 1$, assume $k := k + 1$ and go to 5, otherwise go to 9.
9. The trade-offs for each pair of attributes have been compared, and the set of potential new proposals B still consists of more than one alternative. As the analysis of trade-offs hasn't provided a clear recommendation for the new proposal, analyze the relations between alternatives with respect to attributes. Start from attribute X_p and identify the set of alternatives with non-dominated evaluations according to SD rules. If the number of such alternatives is equal to 1, assume the corresponding alternative to be a new proposal, otherwise exclude dominated alternatives from B and analyze relations with respect to other attributes using a reversed lexicographic order of attributes: $q_{n-1}, q_{n-2}, \ldots, q_1$. Continue until B consists of one alternative. If all attributes have been considered and B still consists of more than one alternative, assume any of them to be a new proposal.

5. Application

To illustrate the procedure let us consider a production process control problem assuming that Just-in-Time (JIT) approach is used for scheduling production system. The work flow is controlled by Kanban cards. Production may proceed differently according to a lot size, number of Kanban cards used, and the decision rule for choosing the waiting job to process. The problem that arises consists in deciding which rule should be used, how many Kanbans should be allocated for each operation, and what lot size should be applied. In general, smaller lot-sizes reduce work-in-progress, but also increase the number of machine set-ups. Increasing the number of allocated Kanbans improves machine utilisation, but may also increase average work-in-progress level. Finally, the performance of a scheduling rule depends on the performance measure that is used. Thus, the choice of the best triplet involving the Kanban lot size, the decision rule, and the number of Kanbans constitutes a multicriteria problem. Nowak et al. [7] proposed a procedure for this problem that uses stochastic dominance rules and multiattribute aggregation based on the methodology proposed in ELECTRE-I (Roy, Bouyssou [8]). In this paper we will show, how trade-offs can be used for solving this problem.

Four scheduling rules are considered: the first come – first served (FCFS) rule, the shortest processing time (SPT) rule, the same job as previously (SJP) rule, the shortest next queue (SNQ) rule. The FCFS rules gives priority to a job that has been queuing at the station for the longest time. The SPT rule chooses the job with shortest planned processing time. The SJP rule assumes that the job which is the same as previously processed on the station should be chosen, and finally SNQ rule gives priority to the job for which the queue at the next station is the shortest. Four values of lot size are considered: 5, 10, 15, and 20, while the number of kanbans is assumed to be between 2 and 5. Thus, 64 triplets of parameters are considered. For example alternative a_1 is defined by triple: lot-size = 5, decision rule = FCFS, number of kanbans = 2, while alternative a_{64} is defined by triple: lot-size = 20, decision rule = SNQ, number of kanbans = 5. Three criteria are used for evaluating alternatives' performances:

1. Makespan – measured in seconds – to be minimized;
2. Average work-in-progress level – measured by the average number of jobs queuing at stations – to be minimized;
3. Number of set-ups – the whole number of set-up done on all stations – to be minimized.

First, an exemplary production plan is analyzed. Series of simulation experiments are done for each alternate parameter triplets. As smaller values are preferred to larger ones, the results of simulation runs are transformed to negative values. Next, distributional evaluations with respect to three criteria are constructed and FSD and SSD relations between distributional evaluations are identified. The first proposal is alternative a_{53} (lot-size = 20, decision rule = SPT, number of kanbans = 2). The DM decides that means should be presented for each attribute in the dialog phase of the procedure.

Iteration 1.

1. The data are presented to the decision maker (table 1).

Table 1
Data presented to the DM at iteration 1

	Mean		
	Makespan	Work-in-progress	Number of set-ups
a_{53}	278813	1075	3397
min	261996	470	1513

2. The DM is not satisfied with the proposal.
3. According to the DM the evaluation of a_{53} with respect to the third attribute (number of set-ups) should be improved.
4. The DM specifies also that the value of work-in-progress may be decreased first, while the value of makespan may be decreased in the second order.
5. The alternatives with evaluations that are not dominated by the evaluation of alternative a_{53} with respect to attribute "number of set-ups" are identified:
   $A^{(2)} = \{ a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, a_{11}, a_{12}, a_{13}, a_{14}, a_{15}, a_{16}, a_{17}, a_{21}, a_{22}, a_{23}, a_{25}, a_{26}, a_{29}, a_{37}, a_{38}, a_{41} \}$
6. The set of potential new proposals is:
   $B = A^{(2)}$
7. Point-to-point trade-offs for the pair of attributes "number of setups" and "work-in progress" are analyzed. As distributions of trade-offs for the pairs (a_1, a_{53}), (a_2, a_{53}), (a_4, a_{53}), (a_{13}, a_{53}), (a_{14}, a_{53}), (a_{29}, a_{53}) , (a_{37}, a_{53}) are dominated, alternatives a_1, a_2, a_4, a_{13}, a_{14}, a_{29}, and a_{37} are excluded form the set of potential new proposals:
   $B = B \setminus \{ a_1, a_2, a_4, a_{13}, a_{14}, a_{29}, a_{37} \} = \{ a_3, a_5, a_6, a_7, a_8, a_9, a_{10}, a_{11}, a_{12}, a_{15}, a_{16}, a_{17}, a_{21}, a_{22}, a_{23}, a_{25}, a_{26}, a_{38}, a_{41} \}$
8. As the set B consists of more than one alternative, relations between trade-offs distributions for the next pair of attributes: "number of setups" and "makespan" are analyzed. As distributions of trade-offs for the pairs (a_3, a_{53}), (a_7, a_{53}), (a_8, a_{53}), (a_{11}, a_{53}), (a_{12}, a_{53}), (a_{15}, a_{53}), (a_{16}, a_{53}), (a_{17}, a_{53}), (a_{21}, a_{53}), (a_{23}, a_{53}), (a_{25}, a_{53}), (a_{26}, a_{53}), (a_{38}, a_{53}), (a_{41}, a_{53}) are dominated, so alternatives a_3, a_7, a_8, a_{11}, a_{12}, a_{15}, a_{16}, a_{17}, a_{21}, a_{23}, a_{25}, a_{26}, a_{38}, and a_{41} are excluded form the set of potential new proposals:
   $B = B \setminus \{ a_3, a_7, a_8, a_{11}, a_{12}, a_{15}, a_{16}, a_{17}, a_{21}, a_{23}, a_{25}, a_{26}, a_{38}, a_{41} \} = \{ a_5, a_6, a_9, a_{10}, a_{22} \}$
9. As relations between trade-offs for each pair of attributes have been analyzed and the set of potential new proposals still consists of more than one alternative, relations between the potential proposals with respect to the attribute "number of setups" are analyzed. The evaluation of a_5 dominated according to FSD rule evaluations of all other alternatives. Thus, a_5 is chosen to be the new proposal.

The procedure is continued in the same way, until the DM is satisfied with the proposal.

6. Conclusions

In many cases, the DM faced with a candidate solution is able to answer the simplest questions only: which attribute should be improved and which attributes can be decreased. In such a situation trade-offs can be used for generation of a new proposal. When the evaluations of alternatives with respect to attributes are characterized by random variables, a point-to-point trade-off is characterized by a random variable as well.

In this paper a new interactive procedure based on the treatment of trade-offs has been proposed. The procedure requires a limited amount of preference information from the DM. The procedure presented in this work can also be applied for mixed problems, i.e. problems in which evaluations with respect to some attributes take the form of probability distributions, while the remaining ones are deterministic. The proposed technique may be useful for various types of problems in which simulation modeling is used for analyzing the efficiency of the system under different circumstances. It has been designed for problems with up to moderate number of alternatives (not more than hundreds) and can be applied in such areas as, for example, inventory models, evaluation of investment projects, labor planning, and many others.

References:

[1] RUBINSTEIN, R. Y. and KROESE D.P. *Simulation and the Monte Carlo Method.* John Wiley & Sons, Hoboken, 2008. ISBN 978-0-470-17794-5.

[2] ROBINSON, S. *Simulation. The Practice of Model Development and Use*. John Wiley & Sons, Hoboken, 2008. ISBN 0-470-84772-7.

[3] NOWAK, M. *Computer Simulation in Decision Problems.* University of Economics, Katowice, 2007, ISBN 978-83-7246-977-7 (*in polish*).

[4] KEENEY, R. L. and RAIFFA H. *Decisions with Multiple Objectives*. Cambridge Univrsity Press, Cambridge, 1993. ISBN 0-521-44185-4.

[5] STEUER, R. *Multiple Criteria Optimization: Theory, Computation and Application.* Wiley, New York, 1986.

[6] HADAR, J. and RUSSEL, W.R. *Rules for ordering uncertain prospects.* The American Economic Review No. 59, 1969, pp. 25-34.

[7] NOWAK, M., TRZASKALIK, T., TRZPIOT, G., and ZARAS, K. *Inverse Stochastic Dominance and its Application in Production Process Control*. In: Multiple Objective and Goal Programming. Recent Developments. Physica-Verlag, Heidelberg, New York, 2002, pp. 362-376, ISBN 3-7908-1409-1.

[8] ROY, B., BOUYSSOU, D. *Aide Multicritere a la Decision: Methodes et Cas.* Economica, Paris, 1993.

DUALITY IN LINEAR PROGRAMMING WITH FUZZY PARAMETERS AND MATRIX GAMES WITH FUZZY PAY-OFFS

Josef Vícha[+], Jaroslav Ramík[*]
Silesian University in Opava
[+]Mathematical Institute, [*]School of Business Administration
Josef.Vicha@math.slu.cz, ramik@opf.slu.cz

Abstract. Generalized model of a fuzzy matrix game is introduced. A dual of linear programming problems with fuzzy parameters is presented and it is shown that a two person zero sum matrix game with fuzzy pay-offs is equivalent to a primal–dual pair of such fuzzy linear programming problems. Further, certain difficulties with similar approaches reported in the literature are discussed. Illustrative example is presented and discussed.

Keywords. Fuzzy numbers, Fuzzy matrix games, Fuzzy linear programming, duality

1. Introduction

In matrix game theory the well-known result asserts that every two person zero sum matrix game is equivalent to a pair of linear programming problems that are dual to each other. Thus, solving such a game amounts to solving any one of these two mutually dual linear programming problems and obtaining the solution – the optimal strategy of the other by using linear programming duality theory.

In recent years some attempts have been made to extend the results of crisp game theory to the fuzzy games, see [5] and the references therein. The motivation for this extension is the fact that in some real situation the pay-offs of the players in a two-person game are not exact values (e.g. amount of money to be won), however, the pay-offs could be uncertain, more or less possible (e.g. a possible amount of viewers of a TV-station, see the example in Section 3). Also, the motivating force behind these extensions is the advancement in the duality theory for fuzzy linear programming problems. Recently, new approaches have emerged for studying duality in fuzzy linear programming, see [3], [4]. The new generalized model for a matrix game with fuzzy goals and fuzzy payoffs by using fuzzy relation approach based on [3], proposed by Vijay et al. [5] is very general. The aim of this paper is to present a simplified and more practical case of this general approach and demonstrate the main results on a simple numerical example.

2. Generalized model of a fuzzy matrix game

In this section we present a generalized model of game theory. We begin by describing two person zero sum matrix game, see [2].

Let m and n be two positive integers, greater than one. Let $A \in \mathbf{R}^{m\times n}$ be an mxn, real matrix, $e^T = (1,\dots,1)$ be a vector of ones whose dimension is m or n which depends on the context. By a crisp *two person zero sum matrix game G* we mean the triplet $G = \{S^m, S^n, A\}$ where $S^m = \{x \in \mathbf{R}^m_+, e^T x = 1\}$ and $S^n = \{y \in \mathbf{R}^n_+, e^T y = 1\}$. In terminology of the matrix game theory, S^m (respectively S^n) is called the (*mixed*) *strategy space for Player 1* (respectively *Player 2*) and A is called the *pay-off matrix*. It's a convention to assume that Player 1 is a maximizing player and Player 2 is a minimizing player. Further, for $x \in S^m$, $y \in S^n$, the scalar $x^T Ay$ is the *pay-off to Player 1*, and as the game G is zero sum, the *payoff to Player 2* is $-x^T Ay$.

Definition 1: A triplet $(\bar{x}, \bar{y}, \bar{v}) \in S^m \times S^n \times \mathbf{R}$ is called a *solution of the game G* if

$$\bar{x}^T A y \geq \bar{v}, \quad \forall y \in S^n,$$
$$x^T A \bar{y} \geq \bar{v}, \quad \forall x \in S^m.$$

Where $\bar{x}$ is called the *optimal strategy for Player 1*, $\bar{y}$ is called the *optimal strategy for Player 2* and $\bar{v}$ is called *the value of the game G.*

Given a game G by Definition 1, it is customary to associate the following pair of *primal-dual linear programming problems* (P) and (D) with Player 1 and Player 2.

(P) max v
subject to

$$\sum_{i=1}^{m} a_{ij} x_i \geq v \text{ , } j = 1,2,...,n,$$

$$\sum_{i=1}^{m} x_i = 1,$$

$x_i \geq 0$, $i = 1,2,...,m.$

(D) min w
subject to

$$\sum_{j=1}^{n} a_{ij} y_j \leq w \text{ , } i = 1,2,...,m,$$

$$\sum_{j=1}^{n} y_j = 1,$$

$y_j \geq 0$, $j = 1,2,...,n.$

In this context the following theorems are well-known, see [2].

Theorem 1: Every two person zero-sum matrix game G has a solution.

Theorem 2: $(\bar{x},\bar{y},\bar{v}) \in S^m \times S^n \times \mathbf{R}$ is a solution of the game G if and only if $(\bar{x},\bar{v})$ is an optimal solution to (P), $(\bar{y},\bar{v})$ is an optimal solution to (D) and $\bar{v}$ is the common optimal value of (P) and (D).

A model with deterministic values of payoff matrix A may be too crude, since in real situations these values are often chosen in an arbitrary way, e.g. estimation of numbers of TV viewers, see the illustrative example in Section 3. An alternative approach is based on introducing into the model a more adequate representation of expert understanding of the nature of the parameters in an adequate form of fuzzy numbers. Then we obtain a model of a fuzzy matrix game with fuzzy pay-offs and relate it to two fuzzy linear programming problems.

A *fuzzy number* is a fuzzy set of the real numbers $\mathbf{R}$ given by the membership function $\mu : \mathbf{R} \to [0,1]$ satisfying the following conditions:

- there exists $x_0 \in \mathbf{R}$ such that $\mu(x_0) = 1$,
- the membership function μ is quasi-concave, i.e. for every α, $0 < \alpha \leq 1$: $\{x \in \mathbf{R} | \mu(x) \geq \alpha\}$ is a convex set,
- $\{x \in \mathbf{R} | \mu(x) > 0\}$ is a bounded set.

The set of all fuzzy numbers is denoted as $F_N(\mathbf{R})$. Notice that usual real numbers, here we call them *crisp numbers*, can be understood as particular cases of fuzzy numbers, i.e. $\mathbf{R} \subset F_N(\mathbf{R})$.

Let $\tilde{A} = \{\tilde{a}_{ij}\}$ be an mxn payoff matrix with entries as fuzzy numbers. Then a generalized model for a *zero sum matrix game with fuzzy payoffs*, denoted by FG, is defined as:

$$FG = \{S^m, S^n, \tilde{A}, \tilde{P}, \tilde{Q}, v, w\}.$$

Here $\tilde{P}$ and $\tilde{Q}$ are fuzzy relations defined by their membership functions as follows:

$$\mu_{\tilde{P}}(\tilde{r},\tilde{s}) = \sup_{x,y\in X}\{\min(\mu_{\tilde{r}}(x),\ \mu_{\tilde{s}}(y)) \mid x \le y\}, \tag{1}$$

$$\mu_{\tilde{Q}}(\tilde{r},\tilde{s}) = \inf_{x,y\in X}\{\max(1-\mu_{\tilde{r}}(x), 1-\mu_{\tilde{s}}(y)) \mid x \le y\}, \tag{2}$$

where $\tilde{r}$ and $\tilde{s}$ are fuzzy numbers. Moreover, v is called *the value of the game G for Player 1,* and w is called *the value of the game G for Player 2.* Fuzzy relation (1) is called the "*optimistic*" fuzzy relation, fuzzy relation (2) is called the "*pessimistic*" fuzzy relation, see [4].

Suppose $\tilde{r}$, $\tilde{s} \in F_N(\boldsymbol{R})$ and $\alpha \in (0,1]$. Then the ordering relation between $\tilde{r}$ and $\tilde{s}$ on the level α is given as

$$\tilde{r}(\tilde{P})_\alpha\tilde{s} \quad \text{if} \quad \mu_{\tilde{P}}(\tilde{r},\tilde{s}) \ge \alpha\,, \tag{3}$$

$$\tilde{r}(\tilde{Q})_\alpha\tilde{s} \quad \text{if} \quad \mu_{\tilde{Q}}(\tilde{r},\tilde{s}) \ge \alpha\,. \tag{4}$$

Now, using these fuzzy relations we can define a solution of *FG*.

Definition 2: Let $\alpha \in (0,1]$. The 4-tuple (x^*,y^*,v^*,w^*) is called an *α–solution of the game FG*, if

$$v^*(\tilde{P})_\alpha(x^{*T}\tilde{A}y),\ \forall y\in S^n, \tag{5}$$

$$(x^T\tilde{A}y^*)(\tilde{Q})_{1-\alpha}w^*,\ \forall x\in S^m. \tag{6}$$

Here, v* (respectively w*) is termed as the *α-value of the game FG for Player 1* (respectively Player 2) and x^* (respectively y^*) is called an *α-optimal (mixed) strategy for Player 1* (respectively Player 2).

From (5) and (6) it is clear that the problems of Player 1 and Player 2 can be formulated as the following pair of optimization problems:

(FSP) max v
subject to

$$v(\tilde{P})_\alpha(x^T\tilde{A}y),\ \forall y\in S^n\,,$$
$$e^Tx = 1\,,$$
$$x \ge 0.$$

(FSD) min w
subject to

$$(x^T\tilde{A}y)(\tilde{Q})_{1-\alpha}w,\ \forall x\in S^m\,,$$
$$e^Ty = 1\,,$$
$$y \ge 0.$$

Since S^m and S^n are convex polytopes, it is sufficient to consider only the extreme points of S^m and S^n. Therefore, solving the above two optimization problems is equivalent to solving the following two linear programs, where by $\tilde{A}_{.j}$ we denote the j-th column of $\tilde{A}$, and by $\tilde{A}_{i.}$ the i-th row of $\tilde{A}$.

(FP) max v
subject to

$$v(\tilde{P})_\alpha(x^T\tilde{A}_{.j})\,,\ j = 1,2,...,n,$$
$$e^Tx = 1,$$
$$x \ge 0.$$

(FD) min w
subject to

$$(\tilde{A}_{i.}y)(\tilde{Q})_{1-\alpha}w\,,\ i = 1,2,...,m,$$
$$e^Ty = 1\,,$$
$$y \ge 0.$$

It was shown, see [3], that (FP) and (FD) can be transformed to two linear programming problems, as follows:

(FLP) max v
subject to

$$\sum_{i=1}^{m} a_{ij}^{R}(\alpha)x_i \geq v,\ j = 1,\dots,n,$$
$$e^T x = 1,$$
$$x \geq 0.$$

(FLD) min w
subject to

$$\sum_{j=1}^{n} a_{ij}^{R}(\alpha)y_j \leq w,\ i = 1,\dots,m,$$
$$e^T y = 1,$$
$$y \geq 0.$$

Here, we define: $a_{ij}^{R}(\alpha) = \sup\{t \in \mathbf{R} \mid \mu_{\tilde{a}_{ij}}(t) \geq \alpha\}$ and $a_{ij}^{L}(\alpha) = \inf\{t \in \mathbf{R} \mid \mu_{\tilde{a}_{ij}}(t) \geq \alpha\}$.

Consequently, for solving fuzzy matrix game *FG*, we have to solve two LP problems (FLP) and (FLD). Moreover, if, for given aspiration level α, ($x^*(\alpha)$, $v^*(\alpha)$) is an optimal solution of (FLP), then $x^*(\alpha)$ is an α-optimal strategy of Player 1 with $v^*(\alpha)$ as an α-value of Player 1. A similar interpretation can also be given to the optimal solution ($y^*(\alpha)$, $w^*(\alpha)$) of the problem (FLD, i.e, $y^*(\alpha)$ is an α-optimal strategy of Player 2 with $w^*(\alpha)$ as an α-value of Player 2. It is clear that that (FLP) and (FLD) are dual to each other in the usual sense of LP duality. Consequently, we obtain $v^*(\alpha) = w^*(\alpha)$.

In Definition 2, the fuzzy relations $\tilde{P}$ and $\tilde{Q}$ are dual to each other, according to the definition of duality introduced in [3]. Hence, we can exchange their positions in (5), (6), then we can investigate the following couple of dual LP problems:

(FLP*) max v
subject to

$$\sum_{i=1}^{m} a_{ij}^{L}(\alpha)x_i \geq v,\ j = 1,\dots,n,$$
$$e^T x = 1,$$
$$x \geq 0.$$

(FLD*) min w
subject to

$$\sum_{j=1}^{n} a_{ij}^{L}(\alpha)y_j \leq w,\ i = 1,\dots,m,$$
$$e^T y = 1,$$
$$y \geq 0.$$

In (FLP) problem Player 1 is optimistic whereas Player 2 in (FLD) is pessimistic. Here, in (FLP*) and (FLD*), it is vice-versa.

3. Illustrating example

Consider two television stations A and B. TV-station A (i.e. Player 1) is new on the market, whereas TV-station B is old. The managers of TV-station A make decision which kind of movie to broadcast in given day of week at main watching time. They choose from *thriller* (Strategy 1) *crime* (Strategy 2) and *comedy* (Strategy 3). From the market research department the matrix $\tilde{A}$ of ratings with fuzzy entries is obtained. Entries of this matrix are fuzzy numbers of viewers (in millions), who would watch TV-station A at the specified time. This matrix is given as follows

$$\tilde{A} = \begin{bmatrix} \tilde{13} & \tilde{8} & \tilde{30} \\ \tilde{22} & \tilde{28} & \tilde{20} \\ \tilde{19} & \tilde{7} & \tilde{35} \end{bmatrix},$$

The entries of this payoff matrix are considered as the (linear) triangular fuzzy numbers given by

$$\tilde{13} = (10, 13, 15), \tilde{8} = (5, 8, 12), \tilde{30} = (20, 30, 35),$$
$$\tilde{22} = (20, 22, 25), \tilde{28} = (24, 28, 31), \tilde{20} = (18, 20, 25),$$
$$\tilde{19} = (14, 19, 21), \tilde{7} = (2, 7, 10), \tilde{35} = (30, 35, 40).$$

For example, suppose that both TV stations choose strategy 3 characterized by fuzzy number $\tilde{35}$ and by the aspiration level $\alpha = 0.6$. It means that station A will be watched possibly between 33 and 37 millions of viewers and the traditional TV station B will loose the same amount of viewers, i.e. between 33 and 37 millions, with the possibility measure at least 0.6, see also Figure 1.

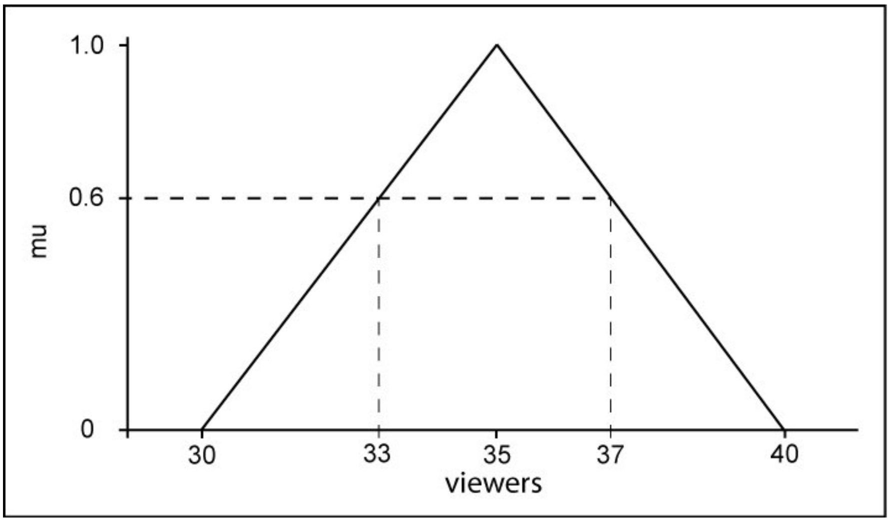


Figure 1: Pay-off $\tilde{a}_{33}$ with α = *0,6.*

Let $\alpha \in (0, 1]$. In order to calculate the solution of the game (for a given α) we have to solve following linear programming problem *(FLP)* for Player 1 and *(FLD)* for Player 2.

(FLP) max v
subject to

$$\begin{aligned} (15-2\alpha)x_1 + (25-3\alpha)x_2 + (21-2\alpha)x_3 &\geq v, \\ (12-4\alpha)x_1 + (31-3\alpha)x_2 + (10-3\alpha)x_3 &\geq v, \\ (35-5\alpha)x_1 + (25-5\alpha)x_2 + (40-5\alpha)x_3 &\geq v, \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

(FLD) min w
subject to

$$\begin{aligned} (15-2\alpha)y_1 + (12-4\alpha)y_2 + (35-5\alpha)y_3 &\leq w, \\ (25-3\alpha)y_1 + (31-3\alpha)y_2 + (25-5\alpha)y_3 &\leq w, \\ (21-3\alpha)y_1 + (10-3\alpha)y_2 + (40-5\alpha)y_3 &\leq w, \\ y_1, y_2, y_3 &\geq 0. \end{aligned}$$

In the following table we provide the optimal solution for the two above problems *(FLP)* and *(FLD)* for certain sample values of α.

In particular, for $\alpha = 0.7$, the 0.7-optimal mixed strategy for TV-station A is $x_1*(0.7) = 0$, $x_2*(0.7) = 0.9235$, $x_3*(0.7) = 0.0765$ and $v^*(0.7) = 22.65$ is a 0.7-value of Player 1. TV-station A should broadcast the *crime* with the probability of 92% and the *comedy* with the probability of 8%. A similar interpretation is given for TV-station B. The optimal strategy for TV-station B is $y_1*(0.7) = 0.8197$, $y_2*(0.7) = 0$, $y_3*(0.7) = 0.1803$ and 0.7-value of Player 2, $w^*(0.7) = 22.65$, being equal to the 0.7-value of Player 1. Station B should broadcast the *thriller* with the probability of 82% and the *comedy* with the probability of 18%.

From the following table it follows that when increasing the aspiration level α, TV-station A should decrease the probability of the *crime* (i.e. Strategy 2) and at the same time increase the probability of the *comedy* (i.e. Strategy 3). On the other hand, TV-station B should decrease the probability of the *comedy* (Strategy 3) and at the same time the probability of the *thriller* (Strategy 1) should be increased.

α	$x_1*(\alpha)$	$x_2*(\alpha)$	$x_3*(\alpha)$	$y_1*(\alpha)$	$y_2*(\alpha)$	$y_3*(\alpha)$	$v*(\alpha) = w*(\alpha)$
1.0	0.0000	0.8889	0.1111	0.8333	0.0000	0.1667	21.67
0.9	0.0000	0.9006	0.0994	0.8287	0.0000	0.1713	21.99
0.8	0.0000	0.9121	0.0879	0.8242	0.0000	0.1758	22.32
0.7	**0.0000**	**0.9235**	**0.0765**	**0.8197**	**0.0000**	**0.1803**	**22.65**
0.6	0.0000	0.9348	0.0652	0.8152	0.0000	0.1848	22.98
0.5	0.0000	0.9459	0.0541	0.8108	0.0000	0.1892	23.31
0.4	0.0000	0.9570	0.0430	0.8065	0.0000	0.1935	23.65
0.3	0.0000	0.9679	0.0321	0.8021	0.0000	0.1979	23.98
0.2	0.0000	0.9787	0.0213	0.7979	0.0000	0.2021	24.32
0.1	0.0000	0.9894	0.0106	0.7937	0.0000	0.2063	24.66
0.0	0.0000	1.0000	0.0000	1.0000	0.0000	0.0000	25.00

4. Conclusion

In this paper, zero sum matrix game with fuzzy pay-offs (*FG*) problem has been investigated, general solution of this problem has been defined and discussed. The introduced concepts are demonstrated on a simple example solved by Excel-Solver.

References

[1] BELMAN, R. E., ZADEH, L. A. Decision making in a fuzzy environment. *Management Science*, 17, 1970, 141–164. ISSN 0025-1909.

[2] OWEN, G. *Game theory*. Academic Press, San Diego, 1995, ISBN 01-2531-151-6.

[3] RAMÍK, J. Duality in fuzzy linear programming with possibility and necessity relations. *Fuzzy sets and systems*, 157, 2006, 1283–1302, ISSN 0165-0114.

[4] RAMÍK, J., Interval and Fuzzy Linear Programming (Demonstrated by Example). In: *Proceedings of the XXII. Conference MME 07*, Ostrava 4.–6.9.2007, 135-144. ISBN 83-7246-844-3.

[5] VIJAY, V., MEHRA, A., CHANDRA, S. BECTOR, C. R. Fuzzy matrix game via a fuzzy relation approach. *Fuzzy Optimization and Decision Making*, 6, 2007, 299–314, ISSN 1568-4539.

This research was supported by the grant project of GACR No. 402090405

MATHEMATICAL STRUCTURES OF ROUGHT SETS

Milan Vlach

Charles University, School of Mathematics and Physics
E-mail: milan.vlach@mff.cuni.cz

Abstract. The mathematical formalism of the original rough set theory deals with situations in which some subsets of a given fixed set can be identified only within the limits given by an equivalence relation. In disregard of the fact that this formalism uses notions of ordinary set theory to define its objects, one can surprisingly often read in the literature that the theory of rough sets is an extension of classical set theory. The main purpose of this talk is to show that rough sets are just ordinary subsets of an ordinary set, and to demonstrate that some widely known and well established algebraic, logical, and topological structures are identical or closely related to those introduced, discovered and rediscovered in the rough set literature.
Keywords. Approximation operators, Definable sets, Rough sets, Partition topology, Uniform topology, Alexandrov topology, Stone lattices, Modal logics.

1 Introduction

Throughout the paper we assume a modest familiarity with the standard concepts and basic facts of the ordinary set theory. Mostly we follow the terminology and notation of the book by P. R. Halmos [19].

If $\mathcal{B}$ is a binary relation on a set X, then the *converse* $\mathcal{B}^{-1}$ of $\mathcal{B}$ is defined by $\mathcal{B}^{-1} = \{(x, y) : (y, x) \in \mathcal{B}\}$. If $\mathcal{B}_1$ and $\mathcal{B}_2$ are binary relations on X, then the *composite* of $\mathcal{B}_1$ and $\mathcal{B}_2$ is denoted by $\mathcal{B}_1 \circ \mathcal{B}_2$; it is defined to be the set of all pairs (x, z) such that for some y it is true that $(x, y) \in \mathcal{B}_2$ and $(y, z) \in \mathcal{B}_1$. If $\mathcal{B}$ is a binary relation on a set X and if x is an element of X, then $\mathcal{B}(x)$ will denote the set $\{y \in X : (x, y) \in \mathcal{B})\}$. A reflexive binary relation on X is called an *equivalence* if it is transitive and symmetric. A *partition* of a nonempty set is a collection of its nonempty subsets that are disjoint from each other and whose union is the whole set.

Let E be an equivalence relation on a nonempty set $\mathcal{U}$. Using this simple structure $\langle \mathcal{U}, E\rangle$ some authors define rough sets as certain subsets of $\mathcal{U}$, some define them as certain pairs of subsets of $\mathcal{U}$, and some as certain collections of subsets of $\mathcal{U}$. Here we follow the view that rough sets given by E should be defined not as subsets of $\mathcal{U}$ but as subsets of the power set of $\mathcal{U}$.

Equivalence relations appear naturally and have an important role in almost every field of mathematics. Therefore, one can expect that various notions and results of rough set theory have their counterparts in other well established and more developed areas of mathematics. The main purpose of this paper is to show how some widely known and well established algebraic, logical, and topological notions and results are closely related to those introduced and discovered in the rough set literature.

2 Pawlak's rough sets

2.1 Definable sets

Intuition leads us to considering as exactly describable or definable subsets with respect to a given equivalence relation only those that are equivalence classes or unions of some equivalence classes. Formally, we obtain the family of definable subsets considered by Pawlak [2], [3] as follows. Let E be an equivalence relation in $\mathcal{U}$ and let $\mathcal{D}(E)$ be the partition of $\mathcal{U}$ induced by E.

Definition 1. *A subset X of $\mathcal{U}$ is* E-definable *if it is either empty or a member of $\mathcal{D}(E)$ or the union of two or more members of $\mathcal{D}(E)$.*

It turns out that the collection of E-*definable* subsets of $\mathcal{U}$ is identical with the collection of subsets $\mathcal{U}$ for which $f_E(X) = g_E(X)$ where f_E and g_E are the set-to-set functions defined by

$$f_E(X) = \bigcup\{A \in \mathcal{D}(E) : A \subset X\},$$

$$g_E(X) = \bigcup\{A \in \mathcal{D}(E) : A \cap X \neq \emptyset\}.$$

It is now common to say that the ordered pair $(\mathcal{U}, E)$ is Pawlak's *approximation space*, and that the values $f_E(X)$ and $g_E(X)$ are the E-*lower approximations* and E-*upper approximations* of X.

2.2 Rough sets

There often exist different subsets of $\mathcal{U}$ with the property that their lower approximations are the same and at the same time their upper approximations are also the same. Such subsets are mutually indistinguishable in terms of their approximations. It is therefore natural to use the lower and upper approximations defined by equivalence relation E for introducing a binary relation (let us denote it by $\sim_E$) on the power set of $\mathcal{U}$ by requiring that $X \sim_E Y$ holds for subsets X and Y of $\mathcal{U}$ if and only if both $f_E(X) = f_E(Y)$ and $g_E(X) = g_E(Y)$. It can easily be verified that this relation is an equivalence relation on the power set of $\mathcal{U}$. The equivalence classes of $\sim_E$ are called the *rough sets* in $\langle \mathcal{U}, E \rangle$. To avoid possible misunderstanding it is useful to keep in mind that definable sets are subsets of $\mathcal{U}$ whereas rough sets are sets of subsets of $\mathcal{U}$.

As an illustration, consider the following collection of rough sets from a recent paper by Järvinen [16]. Let $\mathcal{U}$ be the set $\{a, b, c\}$ and let E be the equivalence on $\mathcal{U}$ whose only equivalence classes are $\{a, c\}$ and $\{b\}$. The corresponding lower and upper approximations are given in Table 1.

X	$\emptyset$	$\{a\}$	$\{b\}$	$\{c\}$	$\{a,b\}$	$\{a,c\}$	$\{b,c\}$	$\mathcal{U}$
$f_E(X)$	$\emptyset$	$\emptyset$	$\{b\}$	$\emptyset$	$\{b\}$	$\{a,c\}$	$\{b\}$	$\mathcal{U}$
$g_E(X)$	$\emptyset$	$\{a,c\}$	$\{b\}$	$\{a,c\}$	$\mathcal{U}$	$\{a,c\}$	$\mathcal{U}$	$\mathcal{U}$

Table 1.

It is clear from the table that the collection of rough sets consists from the following six sets:

$$\{\emptyset\}; \{\{a\}, \{c\}\}; \{\{b\}\}; \{\{a, b\}, \{b, c\}\}; \{\{a, c\}\}; \{\mathcal{U}\}.$$

Because the rough sets are equivalence classes, we can represent every particular rough set by one of its members. Rough sets can also be represented by the ordered pairs $(f_E(X), g_E(X))$ where X is an arbitrary member of the rough set in question. This is possible because every equivalence class of $\sim_E$ is uniquely determined by the pair $(f_E(X), g_E(X))$ where X can be any member of the class.

3 Links to other fields

Various equivalence relations appear in almost every area of mathematics. It is therefore no surprise that some notions and results appearing in the literature on rough set theory have their counterparts in other fields.

3.1 Links to topology

The relationship between the theory of rough sets and theory of topological spaces was recognized by many authors already in the early days of rough set theory. Here we would like to attract attention also to the seldom noticed fact that topologies induced by approximation operators of rough set theory are the uniform topologies.

Uniformities A *quasiuniformity* for $\mathcal{U}$ is a nonempty collection ϱ of subsets of $\mathcal{U} \times \mathcal{U}$ satisfying the following conditions:

- Each member of ϱ contains the diagonal.
- The intersection of any two members of ϱ also belongs to ϱ.
- If R is a member of ϱ and S is a subset of $\mathcal{U} \times \mathcal{U}$ such that $R \subset S$, then S also belongs to ϱ.
- For each $R \in \varrho$ there is an $S \in \varrho$ such that $S \circ S \subset R$.

The quasiuniformity ϱ for $\mathcal{U}$ is called a *uniformity* for $\mathcal{U}$ if the following additional condition is satisfied: If R is in ϱ, then the inverse R^{-1} of R is also in ϱ. If ϱ is a quasiuniformity or uniformity for $\mathcal{U}$, then the pair $(\mathcal{U}, \varrho)$ is said to be a *quasiuniform* or *uniform* space, respectively. Every quasiuniformity ϱ on $\mathcal{U}$ yields a topology for $\mathcal{U}$ by taking as open sets the sets A with the property: if $x \in A$ then there is R in ϱ such that $\{y : (x, y) \in R\} \subset A$.

Topologies from equivalences Let D be a partition of $\mathcal{U}$. It can easily be seen that the collection of all sets that can be written as unions of some members of D together with the empty set is a topology for $\mathcal{U}$. This topology is called the *partition topology* generated by D. The partition topologies are very special. They are characterized by the fact that every open set is also closed, and vice versa. Except for the partition topologies generated by trivial partitions topologies, the partition topologies are not T_0 because every nontrivial partition has some member which contains two or more elements neither of which can be separated from the other. Moreover, the partition topologies are Alexandrov topologies, which means that the intersection of the members of every (not only finite) collection of open sets is also open.

Because every equivalence E in $\mathcal{U}$ defines a partition of $\mathcal{U}$, it also generates a topology for $\mathcal{U}$; namely, the partition topology generated by the partition $\mathcal{D}(E)$. We denote it by τ_E and, if there is no danger of misunderstanding, we omit references to E. In order to see clearly how Pawlak's approximation spaces $\langle \mathcal{U}, E\rangle$ are intimately related with the topological spaces $\langle \mathcal{U}, \tau_E\rangle$, we observe that:

- A subset X of $\mathcal{U}$ is E-definable if and only if it is either empty or it can be written as the union of some members of the partition induced by E.
- A subset X of $\mathcal{U}$ is τ_E-open if and only if it is either empty or it can be written as the union of some members of the partition induced by E.

Moreover, if E is an equivalence on $\mathcal{U}$, then the collection of subsets of $\mathcal{U}\times\mathcal{U}$ that include E is a uniformity for $\mathcal{U}$ and the topology for $\mathcal{U}$ induced by this uniformity coincides with topology τ_E. Hence the difference between Pawlak's approximation space $\langle \mathcal{U}, E\rangle$ and the topological space (U, τ_E) is only terminological. In particular,

- X is definable if and only if it is open.
- The lower approximation of X is the interior of X.
- The upper approximation of X is the closure of X.
- X is definable if and only if its interior is equal to its closure.

For translation of some other terms of Pawlak's terminology into the standard language of general topology, see [7].

3.2 Links to algebra

The algebraic aspects of rough set theory have been studied by many authors in several past decades and there exist excellent papers dealing in great detail and deepness with the relationship between the theory of rough sets and some parts of algebra. The interested reader finds many of these result, for example, in [10], [15], [16], [17], [18], [22], [23] and the literature therein. Therefore, here we only mention some reasons why this relationship is so intimate.

First we notice that the power set $\mathcal{P}(\mathcal{U})$ of $\mathcal{U}$ is partially ordered by the relation of set inclusion, and that this partially ordered set is a complete Boolean lattice. Second, the set of all functions mapping the power set $\mathcal{P}(\mathcal{U})$ of a nonempty set $\mathcal{U}$ into itself inherits from $\mathcal{P}(\mathcal{U})$ a set algebra and inclusion. Namely, $f_1 \cup f_2, f_1 \cap f_2$ and $f_1 \subset f_2$ are defined by the requirement that, for all $X \in \mathcal{P}(\mathcal{U})$, $(f_1 \cup f_2)(X) = f_1(X) \cup f_2(X), (f_1 \cap f_2)(X) = f_1(X) \cap f_2(X), f_1(X) \subset f_2(X)$. Moreover, this set is also a semigroup with respect to the composition of functions and the identity function. Third, the set of ordered pairs of functions from $\mathcal{P}(\mathcal{U})$ into itself can be partially ordered by the relation $\leq$ where $(f_1, f_2) \leq (g_1, g_2)$ means that, for every subset X of $\mathcal{U}$, $f_1(X) \subset g_1(X)$ and $f_2(X) \subset g_2(X)$.

Given that rough sets can be represented by ordered pairs $(f_E(X), g_E(X))$ of lower and upper approximations, the relation $\leq$ gives a natural partial order $\preceq$ on the collection of rough sets in $\langle \mathcal{U}, E\rangle$ by

$$(f_E(X), g_E(X)) \leq (f_E(Y), g_E(Y)).$$

For example, in the previous example from Järvinen [16], we have

$$\{\{a\}, \{c\}\} \preceq \{\{a, c\}\} \text{ and } \{\{a\}, \{c\}\} \preceq \{\{a, b\}, \{b, c\}\}$$

because, for the corresponding pairs of lower and upper approximations, we have $(\emptyset, \{a, c\}) \leq (\{a, c\}, \{a, c\})$ and $(\emptyset, \{a, c\})$ $(\{b\}, \mathcal{U})$. It is worth mentioning that this partially ordered set of rough sets is a complete Stone lattice.

3.3 Links to logic

The definition of lower approximation $f_E(X)$ guarantees that if $x \in f_E(X)$, then $x \in X$. In other words, the lower approximation of X consists from the elements that surely or necessarily belong to X. However X may contain elements that do not belong to X. On the other hand, the definition of upper approximation $g_E(X)$ only guarantees that some members of $g_E(X)$ are in X. That is, if $x \in g_E(X)$ then it is only possible that x belongs to X. These interpretation suggests that there is a close connection between the lower and upper approximation operators f_E and g_E of rough set theory and some operators of modal logics.

It turns out (see, for example, [24], [25] and [26]) that the operators $\Box$ and $\Diamond$ used in the modal logic $S5$ for a modality of universal character (necessarily) and its existential dual (possibly) are exact counterparts of the lower approximation operator f_E and upper approximation operator g_E. Consequently, we can translate the notions and results from the rough set theory to those of the modal logic $S5$ and vice versa in a similar way as we can do so for the rough set theory and theory of topological spaces.

References

1. I. Düntsch, G. Gediga. Approximation operators in Qualitative Data Analysis, *TARSKI, LNCS* Vol. 2929, pp. 214-230, Springer-Verlag, Berlin, Heidelberg, 2003.
2. Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data.* Dordrecht: Kluwer Academic Publishers, 1991.
3. Z. Pawlak, Some Issues on Rough sets, *Transactions on Rough Sets I, LNCS*, Vol. 3100, pp.1-58, Springer-Verlag, Berlin, Heidelberg, 2004.
4. P. C. Hammer. Extended topology: Structure of isotonic functions, *J. reine angew. Math.*, Vol.213, pp.174-186,1964.
5. S. Gniłka. On extended topologies. I: Closure operators, *Ann. Soc. Math. Pol., Ser. I, Commentat. Math*, Vol.34, pp.81-94,1994.
6. P. C. Hammer. Semispaces and the topology of convexity. *Proc. Symposium in Pure Mathematics*, Vol.VII, A. M. S. pp.305-=316,1963.
7. M. Vlach. Rough sets, neighborhoods, and set approximations. *Proceedings of the 7th Czech-Japan Seminar on Data Analysis and Decision Making under Uncertainty*, pp.155-157, Eds.: H. Noguchi, H. Ishii, M. Inuiguchi, Hyogo, Japan, 2004, pp. 155–157, 2004.
8. B. M. R. Stadler, P. F. Stadler. Basic properties of closure spaces, *J. Chem. Inf. Comput. Sci.*, Vol.42, pp.577-585,2002.
9. U. Wybraniec-Skardovska. On a generalization of approximation space, *Bulletin of The Polish Academy of Sciences: Mathematics* Vol.37, pp.51-62,1989).
10. J. Pomykała, J. A. Pomykała. The Stone Algebra of Rough Sets, *Bulletin of the Polish Academy of Sciences, Mathematics*, Vol.36, pp.495512,1988.
11. M. Gehrke, E. Walker. On the Structure of Rough Sets, *Bulletin of the Polish Academy of Sciences, Mathematics*, Vol.40, pp.235245,1992.
12. J. Järvinen. Approximations and rough sets based on tolerances, *Lecture Notes in Computer Science*, Vol.2005, pp.182189, Springer-Verlag, Heidelberg, 2001.
13. P. C. Hammer. Extended Topology: Structure of Isotonic Functions, *J. reine angew. Math.*, Vol.213, pp.174-186,1984.
14. J. A. Pomykała. About tolerance and similarity relations in information systems, *Lecture Notes in Computer Science*, Vol.2475, pp.175182, Springer-Verlag, Berlin, Heidelberg, 2002.
15. Y. Y. Yao. On generalizing rough set theory, *LNAI*, Vol.2639, pp.44-51, 2003.
16. J. Jarvinen. The ordered set of rough sets, *Lecture Notes in Computer Science*, Vol. 3066, pp.49-58, Springer-Verlag, Berlin, Heidelberg, 2004.
17. Y. Y. Yao. Constructive and algebraic methods of the theory of rough sets Information Sciences , Vol. 109, pp. 21-47, 1998.
18. J. Jarvinen, On the structure of rough approximations, *Fundamenta Informaticae*, Vol. 50, pp.135-153, 2002.
19. P. R. Halmos. *Naive set theory*, Springer-Verlag, New York, 1974.
20. M. Vlach. Approximation Operators in Optimization Theory, *Zeitschrift für Operations Research*, Vol.25, pp. 15-23, 1981.
21. M. Vlach. Closures and neighbourhoods induced by tangential approximations, *Selected Topics in Operations Research and Mathematical Economics*, pp. 119-127, Springer Verlag, Berlin, 1984.
22. J. Jarvinen, Lattice Theory for Rough Sets, *Transactions on Rough Sets VI, LNCS 4374*, pp. 400498, Springer Verlag, Berlin Heidelberg, 2007.
23. M. Banerjee, M. K. Chakraborty. Algebras from rough sets, *Rough-Neural Computing: Techniques for Computing with Words*, pp. 157184, Springer-Verlag, Berlin, 2004.
24. Y. Y. Yao, T. Y. Lin. Generalization of rough sets using modal logic, *Intelligent Automation and Soft Computing* Vol. 2, pp. 103-120, 1996.

25. E. Orlowska. Logical aspects of learning concepts, *International Journal of Approximate Reasoning*, Vol. 2, pp. 349-364, 1988.
26. D. Vakarelov. A modal logic for similarity relations in Pawlak knowledge representation systems, *Fundamenta Informaticae*, Vol. XV, pp. 61-79, 1991.

CAPITAL MARKET EFFICIENCY AND TSALLIS ENTROPY

Miloslav Vošvrda

Institute of Information Theory and Automation of the AS CR
Econometrics
vosvrda@utia.cas.cz

Abstract. The concept of the capital market efficiency is a central notion in the financial markets theory. This notion is generally useful to describe a capital market in which is relevant information. If relevant information is completely processed by the capital market price mechanism then such capital market is called to be efficient. Thus the capital market efficiency accentuates the informational efficiency of capital markets. It means that in the efficient capital market investors cannot expect achieving of enormous returns by long time. In other words, the capital market is efficient if the fluctuations of returns in time are unpredictable. Thus a time series of such fluctuations can be generated by some derivations of Brownian motion. Considering long range macroeconomic forces which are non-extensive, we use Tsallis's approach for maximum entropy formulating. By this way it is possible to obtain fat tailed distributions by power laws optimizing Tsallis's entropy and to give some theoretical base to some stylized facts on capital markets. For the estimating of the capital market efficiency is used the notion Tsallis entropy (information gain).

Keywords. Capital market efficiency, expectation, Famma's approach, Tsallis entropy, macroeconomic equilibrium

JEL C01, C13, C46

1. Introduction

The capital market efficiency (CME) has roots dating back to the turn of the last century. The term CME is used to introduce a capital market in which relevant information is absorbed into its price generating system to obtain the price of capital assets. In this definition is formerly emphasized the informational CME way. If capital markets are sofisticated and competitive then economics indicators indicate that investors cannot expect achieving superior profits. If it is assumed that a capital market is in equilibrium state then it is expected that for this capital market the equilibrium price of security will be fair price. This construction is a base for trust establishing in fair functioning of the capital market mechanism. The CME is mainly constructed on the probability calculus. The price analyzing of different securities on capital markets are executed by very sofisticated methods. The base for analyzing a time series of prices is the random walk model. Thus short-term returns have fluctuations with a random nature. It means that the security prices have incorporated the market relevant information at any moment of time immediately. After information processing and risk assessing, the market mechanism generates equilibrium prices. Any deviations of these equilibrium prices should be lock-up and unpredictable. After many theoretical and empirical investigations it is sured to admit the random walk model for the time series of security prices ([3],[4],[6],[9]).

Efficient capital market is closely related to the concept of the market information. The market information is clearly essential in financial activities or trading. This one is arguably the most important determinant of success in the financial life. We shall suppose that strategic investor decisions are formulated and executed on the basis of price information in the public domain, and available to all. Let us assume that investors are considered rational. We shall further assume that information once known remains known - no forgetting - and can be accessed in real time. The ability to retain information, organize it, and access it quickly, is one of the main factors, which will discriminate between the abilities of different economic agents to react on changing market conditions. We restrict ourselves to very simple situation not differentiating between agents on the basis of their information processing abilities. Thus as time passes, new information becomes available to all agents, who continually update their information. We shall use triples $(\Omega, \Im, P)$ for expressing of a probability space and the following expression $E\left[X \middle| \Im\right]$ for the conditional expectations. The $\{\omega \in \Omega\}$ is a set of the elementary market situations. The $\Im$ is some σ-algebra of the subsets of Ω and P is a probability measure on the $\Im$. This structure gives us all the machinery for static situations involving randomness. For dynamic situations, involving randomness over time, we need a sequence of σ- algebras $\{\Im_t, t \geq 0\}$, which are increasing $\Im_t \subset \Im_{t+1}$ for all t with $\Im_t$ representing the information available to us at time t. We always suppose that all σ-algebras to be complete. Thus $\Im_0$ represents an initial information (if there is none, $\Im_0 = \{\emptyset, \Omega\}$). On the other hand, a situation that all we ever shall know is represented by the following expression

$\Im_\infty = \lim_{t\to\infty} \Im_t$. Such a family { $\Im_{t\geq 0}$ } is called a filtration; a probability space endowed with such a filtration, $(\Omega, \Im, \{\Im_t\}, P)$ is called a filtered probability space which is also called a stochastic basis. Suppose that $(\mathrm{S}, \mathcal{A})$ is a measurable space, $\{X_t(\omega)\}_{t\geq 0}$ is a sequence of independent random elements defined on a probability space $(\Omega, \Im, \{\Im_t\}, P)$ taking values in $(\mathrm{S}, \mathcal{A})$ and having the distribution $P = \boldsymbol{P}^X$. Let $\mathcal{M}$ be the set of all measurable functions from $(\mathrm{S}, \mathcal{A})$ to $(□, \Im(□))$, where $\Im(□)$ stands for the Borel σ-algebra of the subsets of □ .

2. Capital Market Efficiency and Expectations

Let $C = (\Omega, \Im, \{\Im_{t\geq 0}\}, P)$ be a capital market with distinguished flows $\{\Im_{t\geq 0}\}$ of σ-algebras filtered probability space. We also call $\{\Im_{t\geq 0}\}$ an information flow, and an expression $\{S_t\}_{t\geq 0} \in \mathcal{M}$ is a security price process.

Definition 1

A capital market $(\Omega, \Im, \{\Im_{t\geq 0}\}, P)$ is called an efficient if there exists *P* such that each security price sequence $S = \{S_t\}_{t\geq 0}$ satisfies the following condition: the sequence

$$\{S_t\}_{t\geq 0} \tag{3.1}$$

is a *P*-martingale, i.e., the variables S_t are $\Im_t$ - measurable and

$$E_P\left[|S_t|\right] < \infty,\ E_P\left[S_{t+1} \middle| \Im_t\right] = S_t,\ t \geq 0. \tag{3.2}$$

If a sequence $\{\xi_t\}_{t\geq 1}$ is the sequence of independent random variables such that $E_P\left[|\xi_t|\right] < \infty$, $E_P[\xi_t] = 0,\ t \geq 1,\ \Im_t^\xi = \sigma(\xi_1, ..., \xi_t), \Im_0^\xi = \{\varnothing, \mathrm{S}\}$, and $\Im_t^\xi \subseteq \Im_t$ then, evidently, the security price sequence $S = \{S_t\}_{t\geq 0}$ where

$$S_t = \xi_1 + ... + \xi_t \quad \text{for} \quad t \geq 1, \text{ and } S_0 = 0,$$

is a martingale with respect to $\Im^\xi = \{\Im_t^\xi\}_{t\geq 0}$, and

$$E_P\left[S_{t+1} \middle| \Im_t\right] = S_t + E_P\left[\xi_{t+1} \middle| \Im_t\right]. \tag{3.3}$$

If a sequence $\{S_t\}_{t\geq 1}$ is a martingale with respect to the filtration $\{\Im_{t\geq 0}\}$ and $S_t = \xi_1 + ... + \xi_t$ with $\xi_0 = 0$ then $\{\xi_t\}_{t\geq 1}$ is a martingale difference, i.e.,

$$\xi_t \text{ is } \Im_t\text{-measurable, } E_P\left[|\xi_t|\right] < \infty,\ E_P\left[\xi_t \middle| \Im_{t-1}\right] = 0. \tag{3.4}$$

We have $E[\xi_t \xi_{t+k}] = 0$ for each $t \geq 0$ and $k \geq 1$, i.e., the variables $\{\xi_t\}$ are uncorrelated, provided that $E\left[|\xi_t|^2\right] < \infty$ for $t \geq 1$. In other words, square-integrable martingale belongs to the class of random sequences with orthogonal increments:

$$E\left[\Delta S_t \cdot \Delta S_{t+k}\right] = 0, \tag{3.5}$$

where $\Delta S_t \equiv S_t - S_{t-1} = \xi_t$ and $\Delta S_{t+k} = \xi_{t+k}$. Thus, the capital market efficiency is nothing else than the martingal property of security price processes in it.

3. Capital Market Efficiency and Maximum Entropy

Consider now S_t as a security price process that is represented by the following form

$$S_t = S_0 \cdot X_t, \tag{4.1}$$

where X_t is a solution of the nonlinear Fokker-Plank equation [12]

$$\frac{\partial}{\partial t} g_t(x) = -\frac{\partial}{\partial x}\left[a(x,t)\cdot g_t(x)\right] + \frac{\partial^2}{\partial x^2}\left[b(x,t)\cdot g_t^{2-q}(x)\right] \tag{4.2}$$

with both $a(\cdot)$ and $b(\cdot)$ as a drift force and diffusion coefficient respectively and a probability density function $g_t(x)$ of X_t at time $t \geq 0$. Let us consider by Borland's modification [14] a stochastic diffusion coefficient in equation (4.2) by the following form

$$c(x,t) = \sqrt{2\cdot b(x,t)\cdot g_t(x)}\,. \tag{4.3}$$

This modification yields the modification of (4.2) as follows

$$\frac{\partial}{\partial t} g_t(x) = -\frac{\partial}{\partial x}\left[a(x,t)\cdot g_t(x)\right] + \frac{\partial^2}{\partial x^2}\left[\frac{c^2(x,t)}{2}\cdot g_t^{1-q}(x)\right]. \tag{4.4}$$

A stochastic differential equation for X_t has the following form

$$dX_t = a(X_t,t)\,dt + \frac{c^2(X_t,t)}{2}\,dW_t(\omega) \tag{4.5}$$

where $W_t(\omega)$ is a standard Wiener process.

Definition 2

A generalized entropy ${}_qH_t$ (Tsallis entropy) of the capital market $\left(\Omega, \Im, \{\Im_{t\geq 0}\}, \mathrm{P}\right)$ at time *t* is introduced by the following indicator

$${}_qH_t = \frac{1 - \int\limits_{x\in X_t} g_t^q(x)\,dx}{q-1}\,. \tag{4.6}$$

A parameter *q* is called an entropic index. A maximization of (4.6) with the following constraints

$$\int g_t(x)\,dx = 1, \quad \frac{3-q}{5-3q}\int x_t^2 \frac{g_t^q(x)}{\int g_t^q(y)\,dy}\,dx = \sigma^2 \tag{4.7}$$

with $\sigma^2 = \int x_t^2 g_t(x)\,dx$ being the second-order moment leads to the following form of the probability density function $g_t(x)$ [6]:

$$g_t(x) = \left[1 - \frac{(1-q)\cdot x^2}{\sigma^2(5-3q)}\cdot\left\{\left[\frac{q-1}{\pi}\cdot\frac{1}{\sigma^2(5-3q)}\right]^{1/2}\cdot\frac{\Gamma\left[\frac{1}{q-1}\right]}{\Gamma\left[\frac{3-q}{2q-2}\right]}\right\}^{1-q}\right]^{1/(1-q)}. \tag{4.8}$$

The expression (4.8) is denoted as a *q*-Gaussian, because for $q \to 1$ the Gaussian distribution is recovered. The parameter *q* can be used as a measure of non-Gaussianity. The probability density function $g_t(.)$ is, for $q > 1$, leptokurtic, i.e. a distribution with long tails relative to Gaussian. The 3-D versions of the *q*-Gaussian are presented in Fig.1 and Fig.2 for σ=1 and *q*=1.2 respectively.

Next, let F_t and G_t be two probability distributions on $\mathbb{R}$ with f_t and g_t as their probability density functions respectively. An information gain is a non-commutative measure of the difference between two probability distributions F_t and G_t. Generalized information divergence associated with two probability distributions F_t and G_t is defined as follows

$$I_q(F_t,G_t) = -\int\limits_{\mathbb{R}} f_t^q(x)\frac{\left[g_t(x)\right]^{1-q} - \left[f_t(x)\right]^{1-q}}{1-q}\,dx\,. \tag{4.11}$$

After some operations on (4.11) we get the following expression

$$I_q\left(F_t,G_t\right)=\frac{1}{1-q}\left[1-\int_{\mathbb{R}}\left(\frac{f_t(x)}{g_t(x)}\right)^q g_t(x)\,dx\right]. \tag{4.12}$$

Note: For q=1, the important quantity (4.12) is related to the Shannon entropy which is called Kullback-Leibler divergence. The Tsallis entropy is appropriate for non-equilibrium systems replacing exponential Boltzmann factors by power-law distribution[13].

Because $I_q\left(F_t,G_t\right)\neq I_q\left(G_t,F_t\right)$ unless $F_t=G_t$, the generalized information divergence $I_q\left(F_t,G_t\right)$ is thus adjusted on the following expression

$$I_q^S\left(F_t,G_t\right)=\left(I_q\left(F_t,G_t\right)+I_q\left(G_t,F_t\right)\right)/2. \tag{4.13}$$

Values of the criterion $I_q^S\left(F,G\right)$ for any F, G belong to $\langle 0,1\rangle$. If $I_q^S\left(F,G\right)$ is equal 0 then

$$F=G, \tag{4.14}$$

i.e., the stochastic processes have the same information gain.

Let F_t be an uniform probability distribution on $\mathbb{R}$. If analyzed stochastic proces $\{X_t,t\in T\}$ in period *T* has a probability distribution G_t on $\mathbb{R}$ and $I_q^S\left(F_t,G_t\right)=0$ then G_t is an information-gain-free about an interior dependence inside $\{X_t,t\in T\}$. As an illustration of the behavior of this criterion $I_q^S\left(F_t,G_t\right)$, let us put both the uniform probability distribution on $\langle -5,5\rangle$ for F_t and the G_t to be *q*-Gaussian for each *t* on $\langle -5,5\rangle$. This one is presented in the following Fig. 3 for different levels of *q*= (1.1, … ,1.6) and σ = (1, ..., 5).

Definiton 3

Let G_{dS_t} be a probability distribution of the proces $\{dS_t,t\in T\}$. Let F_t be an uniform distribution on $\mathbb{R}$. The capital market is called an efficient if for any $dS_t, t\in T$ and for *q* > 1 the following expression $I_q^S\left(F_t,G_{dS_t}\right)=0$ holds.

Theorem 1

A capital market *C* is the efficient if and only if any $dS_t\in C$ for all $t\in T$ and $G_{dS_t}=F_t$.

Proof

$(\Rightarrow)$ Let C be an efficient. Any return $dS_t\in C$, $t\in T$ is by (4.5) a standard Wiener process. Hence $I_q^S\left(G_W,G_{dS_t}\right)=0$ for $t\in T$, and therefore G_{dS_t} has not any information gain about the dependence in $\{X_t,t\in T\}$.

$(\Leftarrow)$ Draw any $dS_t\in C$, $t\in T$ and $G_{dS_t}=F_t$. Then $I_q^S\left(F_t,G_{dS_t}\right)=0$ for any $dS_t\in C$, $t\in T$. Hence C is the efficient.

4. Simulation and Empirical Analysis

Now we generate realizations of the size 500 from the following theoretical models on the one hand with and on the other without a correlation dependence inside. Considered theoretical models are

Normal Noise with μ=0, and σ = 1	ε_t
Random Walk	$S_t(\omega)=S_{t-1}(\omega)+\varepsilon_t$
Random 2-Walks	$S_t(\omega)=S_{t-1}(\omega)-S_{t-2}(\omega)+\varepsilon_t$
Moving Average (2)	$S_t(\omega)=\varepsilon_t+0.1\cdot\varepsilon_{t-1}-0.4\cdot\varepsilon_{t-2}$
Autoregressive Model (2)	$S_t(\omega)=0.7\cdot S_{t-1}(\omega)-0.3\cdot S_{t-2}(\omega)+\varepsilon_t$

Normal Noise with μ=0, and $\sigma^2 = 3^2$	$3 \cdot \varepsilon_t$
AutoRegressive Moving Average (2,2)	$S_t(\omega) - 0.2 \cdot S_{t-1}(\omega) + 0.5 \cdot S_{t-2}(\omega) = \varepsilon_t + 0.7 \cdot \varepsilon_{t-1} - 0.3 \cdot \varepsilon_{t-2}$
Deterministic Process	$S_t(\omega) = 0.9 \cdot S_{t-1}(\omega)$

Empirical data were obtained from the following daily observations of market indices PX50, DAX, BUX, SAX, WIG, S&P 500, SGX, Nikkei 225, NASDAQ, and Dow Jones Industriel Average for the period 2006.11.14 till 2008.10.23. The security price returns model $\ln\left[S_t(\omega)/S_{t-1}(\omega)\right]$ is used for a transformation of the observations. Let F_t be a uniform distribution on $\langle 0,1 \rangle$. A smooth estimation of the empirical probability distributions by Bernstein polynomials [15] is used. We estimate criterions $I_q^S(F_t, G_t)$ and $I_1^S(F_t, G_t)$ for particular models with q=1.2, and σ=1. Results are introduced in Table 1 and in Fig. 4 and Fig. 5.

Tab. 1 Results of the criterions $I_q^S(F_t, G_t)$ and $I_1^S(F_t, G_t)$ for different models.

The Criterions $I_q^S(F_t, G_t)$ and $I_1^S(F_t, G_t)$				
model	F_t	G_t	$I_q^S(F_t, G_t)$	$I_1^S(F_t, G_t)$
1	Uniform	Uniform	0.000000	0.000000
2	Uniform	Random Walk	0.260039	0.229439
3	Uniform	Random 2-Walks	0.098187	0.002911
4	Uniform	MA(2)	0.047324	0.073004
5	Uniform	AR(2)	1.000000	0.049511
6	Uniform	Normal Noise with μ=0, and σ = 3	0.061074	0.056901
7	Uniform	ARMA(2,2)	0.973694	0.011284
8	Uniform	Deterministic Process	1.000000	1.000000
9	Uniform	PX50, Prague, Czech Republic	0.541033	0.616403
10	Uniform	DAX, Frankfurt, Germany	0.542492	0.63618
11	Uniform	BUX, Budapest, Hungary	0.549074	0.577513
12	Uniform	SAX, Bratislava, Slovak	1.000000	0.596843
13	Uniform	WIG, Warsaw, Poland	0.517336	0.229463
14	Uniform	S&P 500, USA	0.575915	0.63618
15	Uniform	SGX,Singapore, Singapore	0.546653	0.520874
16	Uniform	Nikkei 225, Tokyo, Japan	0.559004	0.577505
17	Uniform	NASDAQ, USA	0.560131	0.63618
18	Uniform	Dow Jones Industriel Average, USA	0.605994	0.616401

5. Concluding Remarks

Results show that the current non-extensive approach could be more useful tool in the analysis of CME. This approach is more sensitive on detecting of any dependencies in the capital market processes. It is demonstrated that $I_q^S(F_t, G_t)$ as a measure of the divergence on capital markets pursues this fact better than $I_1^S(F_t, G_t)$. By comparison results of $I_q^S(F_t, G_t)$ from empirical and simulation data it is shown that capital markets have had a remarkable fail in an efficiency of the incorporating of the market relevant information. It is possible to obtain any farther more accurate results after estimating of the entropy index q. Consequently, the next important step is to construct an estimator $\hat{q}$ for the entropy index q. The estimator $\hat{q}$ will give a possibility to involve more pregnantly the leptokurticity of the probability distributions into the measure $I_{\hat{q}}^S(\cdot,\cdot)$.

Acknowledgments

Financial support from Czech Science Foundation 402/09/0965 and MSMT LC06075 is acknowledged.

References

[1] BORLAND, L. *Microscopic dynamics of the nonlinear Fokker-Plank equation: Phenomenological model.* Phys. Rev. E. 1998, 57,6634

[2] KRAFT, J. *Changes of macroeconomic environment at the turn of century as determinant factor for operation of state administration companies.* Acta economia No. 7, Ekonomická teória a praxa - dnes a zajtra, Bánská Bystrica, EF UMB, 2002, str. 39 – 45. ISBN 80-8055-592-3
[3] COOTNER, P.H. *The Random Character of Stock Market Price.*, MIT Press, Cambridge, MA, 1964
[4] FAMA ,E. *The Behavior of Stock Market Prices.* Journal of Business, 1965, 38, pp. 34-105.
[5] KENDALL,l M. *The Analysis of Economic Time Serie.* Journal of the Royal Statistical Society, Series A, 96, 1953, pp. 11-25.
[6] QUEIROS DUARTE, S.M. *On non-Gaussianity and dependence in financial time series: a nonextensive approach.* Quantitative Finance, Vol. 5, No. 5, October 2005, pp.475-487
[7] OSBORNE, M.F.M. *Brownian Motion in the Stock Market.* Operations Research 7, 1959, pp. 145-173
[8] PLATEN,E. *Arbitrage in continuous complete markets.* Adv. Appl. Probab.,34, 2002, pp.540-558
[9] RISKEN,H.*The Fokker-Planck Equation: Methods of Solution and Applications.* 2nd sd. 1989, Springer-Verlag, Berlin,
[10] ROBERTS, H. *Stock Market `Patterns' and Financial Analysis: Methodological Suggestion.* Journal of Finance, 1959, 44, pp. 1-10.
[11] SHIRYAEV, A.N. *Probability*. 2nd ed., Springer-Verlag, New York, 1995
[12] TSALLIS, C., Bukmann, D.J. *Anomalous diffusion in the presence of external forces* . Phys. Rev. E. 54, 1996, R2197,
[13] TSALLIS, C. *Possible Generalization of Boltzmann-Gibbs Statistics*, Journal of Statistical Physics, Vol. 52, Nos. ½, 1988,
[14] TSALLIS, C., ANTENEODO, C., BORLAND, L., OSORIO, R. *Nonextensive Statistical Mechanics and Economics,* Physica A 324, 89 -100, 2003
[15] BABU G.J.,BOYARKY A., CHAUBEY Y.P., GORA P. *A New Statistical Methods for Filtering and Entropy Estimation of a Chaotic Map from Noisy Data,* International Journal of Bifurcation and Chaos, Vol.14, No.11,3989-3994,2004

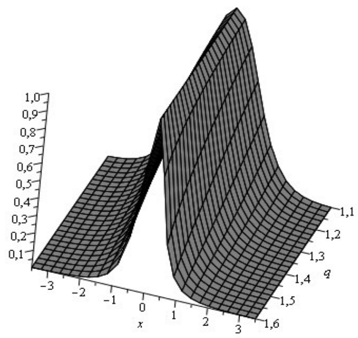


Fig. 1 q-Gaussian, σ=1

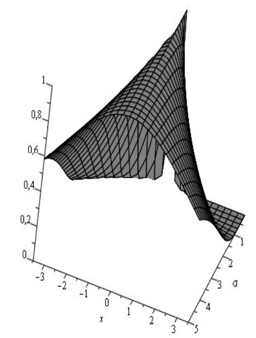


Fig.2 q-Gaussian, q=1.2

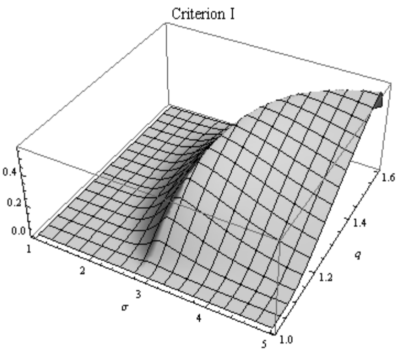


Fig.3 3D-figure of the criterion I_q

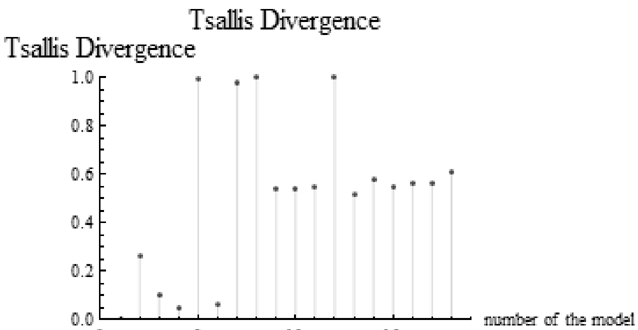


Fig. 4 Tsallis divergence among different models and uniform model

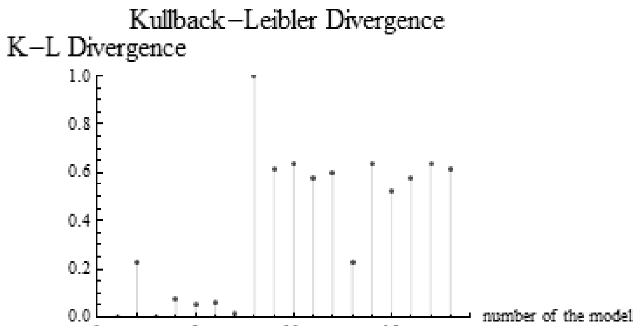


Fig.5 Kullback-Leibler divergence among different models and uniform model

THE MULTI-LEVEL FACTOR IN INSURANCE RATING TECHNIQUE

Alicja Wolny-Dominiak,

Katowice University of Economics, Poland
Faculty of Economics, Department of Statistical and Mathematical Methods in Economics
woali@ae.katowice.pl

Abstract:. Insurance companies specialising in non-life insurance create their own rating systems for setting fair premiums for every risk for different kinds of insurance portfolios. The most popular rating technique is to estimate the relativities of a number of rating factors in a multiplicative or an additive model. Generally, these relativities are discreet (e.g. sex) or continuous (e.g. age, engine capacity). In recent years the standard practice in insurance companies is to use two types of the estimation methods: minimum bias methods or generalized linear models (GLM). One of the practical problem in using this methods is when you take into consideration discreet rating factors with many levels but without an ordering (e.g. car model, geographic zone). Ohlsson and Johansson call this kind of factors as multi-level factors. Nelder and Verrall show that the multi-level factors can be interpreted as random effects in GLM. In this paper first we analyse the estimation process of the credibility estimator in comparing with classical credibility theory where the multi-level factors are estimated by means of minimum square error (MSE). Then we present the iterative algorithm which is use to predict the random risk parameter associated with the multi-level factor where the GLM technique is applied. Finally the numerical example is shown with different assumption of the loss distributions, e.g. Tweedie, Gamma, Inverse Gaussian.
Keywords: Rating technique, multi-level factor, credibility stimator, risk parameter

1. Introduction

Insurance companies specialising in non-life insurance create their own rating systems for setting fair premiums for every risk for different kinds of insurance portfolios. The most popular rating technique is to estimate the relativities of a number of rating factors in a multiplicative or an additive model. Generally, these relativities are discreet (e.g. sex) or continuous (e.g. age, engine capacity). In recent years the standard practice in insurance companies is to use two types of the estimation methods: minimum bias methods or generalized linear models (GLM). One of the practical problems in using these methods is when you take into consideration discreet rating factors with many non ordered levels (e.g. car model, geographic zone). Ohlsson and Johansson [4] call this kind of factors *multi-level factors* (MLF). Nelder and Verrall [3] showed that the multi-level factor can be interpreted as random effects in GLM. In the first part of this paper the overview of the iterative algorithm for the random risk parameter estimation is presented with GLM technique applied in the ratemaking process. In the second part this algorithm is implemented in R software and the numerical example is shown for illustration. In the last part a few simulation procedures are proposed which could be useful in analyzing different scenarios in rate making process. The code of all procedures implemented in R software is included in the appendix.

2. Discreet rating factors with many non ordered levels

In the ratemaking process in casualty insurance very often there are discrete rating variables measured on the nominal scale (variables with many different realization – categories, which cannot be ordered) e.g. car type or region in car insurance. This kind of rating variable is called *multi-level factor* [4]. The most common approach in ratemaking to multi–level factors is to perform the independent ratemaking for every single category (level) of the factor. The disadvantages of this approach are the high computation complexity and much time needed for building the rating. This is because we need to perform the whole ratemaking procedure many times, for every level of the factor independently (e.g. a few thousands times). It is possible to speed up the process when we use the GLM model. For this purpose we introduce the multi–level factor as a random variable obtaining the so called HGLM – hierarchical generalized linear model.

Let Y_i be the response variable (e.g. claim sizes, risk premiums) for i-th ordinary factor, $i = 1,...,I$, $X_1,...,X_I$ - ordinary factors, $\beta_1,...,\beta_I$ - the parameters in GLM, w_{ik} - the exposure weights (i.e. number of claims) for the i*th* ordinary factor and k*th* level of the multi-level factor, $k = 1,...,K$. Moreover we assume that the error for the k th level of the multi-level factor is the outcome of a some random variable U_k (the risk parameter). The GLM model can be presented in the following (extended) form [3]:

$$\eta' = \eta + v,$$

Where $\eta = \sum_{i=1}^{I} \beta_i X_i$ denotes the systematic part in the model and $v = v(u)$ is the error term (a strictly monotonic function of u). Therefore the MLF is modelled by a random risk parameter.

In HGLM models the variable Y_i is assumed to have a distribution of the form like in a Tweedie model with $1 \le p \le 2$ which is typical in non-life insurance. Then [4]:

$$E(Y_{ik} \mid U_k = u_k) = \mu_i u_k,\ k = 1,\dots,K,$$

where μ_i is the mean given by the ordinary rating factors and capture the fixed effects. The U_k 's are supposed to be independent and identically distributed with common variance $Var(U_k) = \sigma_U^2$ and the u_k 's are tariffs for all levels of the multi-level factor. The fixed effects and random effect should be multiplicative. According to theorem 2.1 and the other results presented in [4], the estimation of parameters μ_i of ordinary factors and the estimation of parameters u_k of the multi-level factor, are realized simultaneously. The procedure can be presented as follows (PROC1):

STAGE1 Assume $\hat{u}_k = 1$ for all $k = 1,\dots,K$

STAGE2 Estimate μ_i for all $i = 1,\dots,I$ using the standard GLM procedure with log-link function and $\log(\hat{u}_k)$ as the offset-variable

STAGE3 Estimate $\phi\alpha$ using the estimator

$$\phi\alpha = \frac{\sum_{k=1}^{K}\sum_{i=1}^{I} w_{ik}\mu_i^{2-p}\left(\frac{y_{ik}}{\mu_i} - \bar{u}_k\right)^2}{\sum_{k=1}^{K}(I_k - 1)} \frac{\sum_{k=1}^{K}\sum_{i=1}^{I} w_{ik}\mu_i^{2-p}(\bar{u}_k - 1)^2}{\sum_{k=1}^{K}\sum_{i=1}^{I} w_{ik}\mu_i^{2-p}}$$

STAGE4 Estimate $\bar{u}_k$: $\bar{u}_k = \dfrac{\sum_{i=1}^{I} \dfrac{w_{ik}\mu_i^{2-p} y_{ik}}{\mu_i}}{\sum_{i=1}^{I} w_{ik}\mu_i^{2-p}}$, $k = 1,\dots,K$

STAGE5 Estimate $\hat{u}_k$: $\hat{u}_k = \dfrac{\sum_{i=1}^{I} \dfrac{w_{ik} y_{ik}}{\mu_i^{p-1}} + \phi\alpha}{\sum_{i=1}^{I} w_{ik}\mu_i^{2-p} + \phi\alpha}$, $k = 1,\dots,K$

STAGE6 Return to STAGE2 with the new offset-variable $\log(\hat{u}_k)$

We took the data set from [1] to illustrate how the procedure PROC1 works. There are 8 942 claims in this data set, aggregated for two rating factors: age (variable X_1) and the character of the use of the vehicle (variable X_2). As the result of aggregation we obtained the following table representing the input data (with mean value of claims and the number of claims, respectively).

Table 2.1 Input data

X1	X2	Y(i)	w(i)	X1	X2	Y(i)	w(i)
17-20	A	250,48	4	35-39	A	153,62	50
17-20	B	274,78	6	35-39	B	201,67	198
17-20	C	244,52	8	35-39	C	238,21	87
17-20	D	797,8	2	35-39	D	256,21	65
21-24	A	213,71	30	40-49	A	208,59	76
21-24	B	298,6	90	40-49	B	202,8	465
21-24	C	298,13	21	40-49	C	236,06	278
21-24	D	362,23	5	40-49	D	352,49	106
25-29	A	250,57	76	50-59	A	207,57	35
25-29	B	248,56	75	50-59	B	202,67	284
25-29	C	297,9	201	50-59	C	253,63	243
25-29	D	342,31	97	50-59	D	340,56	43
30-34	A	229,09	67	60+	A	192	93
30-34	B	228,48	265	60+	B	196,33	218
30-34	C	293,87	121	60+	C	259,79	98
30-34	D	367,46	34	60+	D	342,58	33

Assuming that the distribution of the mean value of claims is identical for every category of the multi–level factor, we chose randomly the values of variable $Y_i \sim Gamma$ and $w_i \sim Poisson$ repeating sampling 10 times. As a result of the procedure PROC1 (the code for the procedure presented in Appendix A) on the generated data set (presented in Table 2.1) we obtained the following estimated values for $\hat{u}_k$, $k = 1,\ldots,10$ for different values of the parametr p in the Tweedie distribution (with the stop-criterium $\varepsilon < 0{,}0000001$):

Table 2.2 Estimates for the values of $\hat{u}_k$

p=1	p=1.2	p=1.5	p=2
1,00031	1,00224	0,99356	1,04326
0,99956	1,00076	0,99744	1,00509
1,00029	0,99575	1,00510	0,99425
0,99969	0,99938	0,99947	1,00474
0,99962	1,00256	1,00426	0,98715
1,00039	0,99976	0,99617	1,01639
0,99961	0,99760	1,00210	0,97485
1,00002	1,00102	0,99884	0,97986
1,00015	0,99929	0,99866	1,00684
1,00010	1,00089	1,00403	0,98690

These results can be used to adjust the predictions for the value of claims, based on the values of the rating factors for every category of the multi–level factor. For example, the prediction of variable Y_i, $k = 1$, $p = 1$ for the given data set is presented in Table 2.3.

Table 2.3 Mean value of claims predictions

X1	X2	p=1	p=1.2	p=1.5	p=2	X1	X2	p=1	p=1.2	p=1.5	p=2
17-20	A	244,85	256,46	255,11	274,85	35-39	A	180,64	181,01	179,76	188,87
17-20	B	255,63	268,19	266,66	283,70	35-39	B	188,60	189,30	187,90	194,95
17-20	C	309,98	326,24	323,18	348,67	35-39	C	228,69	230,27	227,72	239,60
17-20	D	398,92	418,26	416,26	440,80	35-39	D	294,31	295,22	293,31	302,91
21-24	A	261,74	260,70	256,84	270,10	40-49	A	195,75	195,53	193,90	202,94
21-24	B	273,27	272,63	268,46	278,79	40-49	B	204,38	204,47	202,68	209,48
21-24	C	331,37	331,64	325,37	342,64	40-49	C	247,83	248,73	245,64	257,45
21-24	D	426,45	425,18	419,08	433,18	40-49	D	318,93	318,89	316,39	325,48
25-29	A	232,56	232,61	232,57	241,30	50-59	A	198,38	198,63	197,02	207,89
25-29	B	242,80	243,25	243,10	249,06	50-59	B	207,12	207,72	205,94	214,59
25-29	C	294,43	295,91	294,63	306,10	50-59	C	251,16	252,68	249,60	263,73
25-29	D	378,90	379,37	379,48	386,99	50-59	D	323,22	323,94	321,48	333,42
30-34	A	224,52	223,50	221,33	231,61	60+	A	194,35	194,77	192,24	203,93
30-34	B	234,40	233,73	231,35	239,06	60+	B	202,91	203,68	200,95	210,50
30-34	C	284,24	284,32	280,39	293,81	60+	C	246,05	247,77	243,54	258,70
30-34	D	365,79	364,51	361,15	371,45	60+	D	316,65	317,65	313,68	327,07

3. Simulation procedures for ratemaking with multi–level factor

The important advantage of the algorithm presented in the previous part of the paper is that the estimation of the parameters for the multi–level factor is performed simultaneously with the estimation for the parameters for all other predictors. It significantly speeds up the procedure, especially when the number of categories of the multi–level factor is large e.g. the vehicle type (a few thousands categories). Thus using simulations is a very effective approach e.g. to evaluate the mean squared error of the model or to analyze how the choice of the parameter p in the Tweedie distribution affects the rating or how it depends on the choice of the variable Y_i distribution in GLM.

Now we present two Monte Carlo simulation procedures, where the input data sets are generated from different theoretical distributions [2].

We propose the following procedure for estimating the mean squared error of $\hat{u}_k$ (PROC2 in Appendix B) [s denotes number of iterations in the simulation]:

STAGE1 Randomize $w_{ik}^s \sim Poisson(1, nCl)$ -> **STAGE2** $w_{ik} := w_{ik}^s$ -> **STAGE3** Randomize $Y_{ik}^s \sim Tweedie(w_{ik}, cv, Y_{ik})$ **->** **STAGE4** $Y_{ik} := \sum_{l=1}^{w_{ik}} Y_{ik}^s$, where w_{ik} is the observed number of claims for (i,k) -> **STAGE5** Estimate $\hat{u}_k^s$ according to PROC1 -> **STAGE6** Calculate MSE_sim_k

We compute the mean squared error:

$$MSE\_sim_k = \frac{1}{S}\sum_{s=1}^{S}(\hat{u}_k - \hat{u}_k^s)^2$$

As the result of the procedure PROC2 for the data presented In Table 2.1 and $s = 1,\ldots,1000$ we obtain:

Table 3.1 MSE resulting from the procedure PROC2

k	MSE_sim_k
1	0,09409
2	0,10005
3	0,10144
4	0,09739
5	0,10141
6	0,09689
7	0,10540
8	0,10034
9	0,10249
10	0,10088

We propose the following procedure PROC3 – Appendix C) to analyze the influence of choice of parameter p value on the $\hat{u}_k$:

STAGE1 Randomize $p^s \sim Uniform((1,2)$ -> **STAGE2** Estimate $\hat{u}_k^s$ according to PROC1 -> **STAGE3** Calculate MSE_sim_k

Table 3.2 Deciles of the estimated values of $\hat{u}_k$ resulting from procedure PROC3

Decile	p	u(1)	u(2)	u(3)	u(4)	u(5)	u(6)	u(7)	u(8)	u(9)	u(10)
0%	1,00251	0,97373	0,99935	1,00865	0,98615	1,01049	0,98618	1,02240	0,99869	1,01247	1,00173
10%	1,11170	0,98425	1,00387	1,01218	0,99063	1,01449	0,98961	1,02850	1,00403	1,01660	1,00702
20%	1,19981	0,98913	1,00633	1,01458	0,99307	1,01662	0,99198	1,03116	1,00660	1,01912	1,00957
30%	1,29423	0,99269	1,00874	1,01691	0,99538	1,01929	0,99432	1,03381	1,00915	1,02145	1,01197
40%	1,39991	0,99646	1,01131	1,01962	0,99802	1,02199	0,99702	1,03589	1,01141	1,02418	1,01431
50%	1,50799	0,99874	1,01379	1,02221	1,00050	1,02489	0,99971	1,03830	1,01396	1,02654	1,01688
60%	1,60342	1,00143	1,01638	1,02458	1,00303	1,02722	1,00215	1,04112	1,01684	1,02908	1,01971
70%	1,70092	1,00479	1,01873	1,02699	1,00539	1,02925	1,00461	1,04320	1,01881	1,03150	1,02181
80%	1,79993	1,00819	1,02117	1,02930	1,00783	1,03174	1,00677	1,04579	1,02159	1,03383	1,02454
90%	1,89558	1,01237	1,02359	1,03157	1,01024	1,03375	1,00919	1,04848	1,02403	1,03603	1,02696
100%	1,99945	1,02399	1,02752	1,03453	1,01413	1,03827	1,01205	1,05361	1,02895	1,03965	1,03160

We observe that for the given data set the differences for different values of p are not very significant, which clearly shows Table 3.3.

Tabela 3.3 Relative differences between the max and min value of $\hat{u}_k$

u(1)	u(2)	u(3)	u(4)	u(5)	u(6)	u(7)	u(8)	u(9)	u(10)
4,5374%	2,6528%	2,4581%	2,6773%	2,4846%	2,5138%	2,7816%	2,7866%	2,5435%	2,7482%

4. Conclusions

In the paper we presented the procedure for ratemaking with multi–level predictor variable, which introduces error term to GLM. The advantage of this approach is that it enables the simultaneous estimation of the $\hat{u}_k$ parameters for every category of the multi–level factor. The disadvantage is that we have to make some additional assumptions (e.g. p values taken from the interval [1,2]). The iterative algorithm for estimating $\hat{u}_k$, implemented in R software, is a powerful tool for performing simulations and assessing the influence of other parameters used in the model on the resulting rating system. It is also possible to analyze how the results depend on the preassumed distribution of claims, which will be the object of our further research in the future.

Bibliography

[1] MILDENHALL S., *A Systematic Relationship Between Minimum Bias and Generalized Linear Models*, PCAS, LXXXVI, str. 394-487, 1999.
[2] MIELCZAREK B. *Metody próbkowania w symulacji Monte Carlo*, Prace Naukowe Instytutu Organizacji i Zarządzania Politechniki Wrocławskiej, 187-199, 2007
[3] NELDER, J. A. VERRAL, R. J. *Credibility Theory and Generalized Linear Models.* ASTIN Bulletin, Vol. 27:1, 71-82, 1997
[4] OHLSSON, E. JOHANSSON, B. *Credibility theory and GLM revised.* Research Report 2003: 15, Mathematical Statistics, Stockholm University, 2003

Appendix A – R implementation of procedure PROC1

```
model.glm=glm(formula = avCl ~ X1+X2+offset(log(of)), family=Gamma(link="log"), weights=nCl)
mi.glm=fitted.values(model.glm)
p=1
K=10
u_k=c()
u_k_1=c()
predict=c()
sigma2_k=c()
I_M=c(32,32,32,32,32,32,32,32,32,32)
for (k in 0:(K-1)){
sum=sum1=0
j=1
for (i in (1+32*k):(32+32*k)){
sum=sum+((data$nCl[i]*(mi.glm[j]^(2-p)))*data$avCl[i])/mi.glm[j]
sum1=sum1+data$nCl[i]*(mi.glm[j]^(2-p))
j+1}
u_k_1=c(u_k_1,(sum/sum1))}
for (k in 0:(K-1)){
sum=0
j=1
for (i in (1+32*k):(32+32*k)){
sum=sum+(data$nCl[i]*(mi.glm[j]^(2-p)))*((data$avCl[i]/mi.glm[j])-u_k_1[k+1])^2
j=j+1
}
sigma2_k=c(sigma2_k,(1/(I_M[k+1]-1))*suma)}
sum=sum1=0
for (k in 1:K){
    sum=sum+((I_M[k]-1)*sigma2_k[k])
    sum1=sum1+(I_M[k]-1)}
sigma2=sum/sum1
```

```
sum=sum1=0
for (k in 0:(K-1)){
j=1
for (i in (1+32*k):(32+32*k)){
     sum=sum+data$nCl[i]*(mi.glm[j]^(2-p))*((u_k_1[k+1]-1)^2)
     sum1=sum1+data$nCl[i]*(mi.glm[j]^(2-p))
j=j+1}}
sigma2_u=(sum-(K*sigma2))/sum1
psi_alfa=sigma2/sigma2_u
for (k in 0:(K-1)){
sum=sum1=0
j=1
for (i in (1+32*k):(32+32*k)){
     sum=sum+(data$nCl[i]*data$avCl[i])/(mi.glm[j]^(p-1))
     sum1=sum1+data$nCl[i]*(mi.glm[j]^(2-p))
j=j+1}
u_k=c(u_k,((sum+psi_alfa)/(sum1+psi_alfa)))}
predict=c()
for (k in 0:(K-1)){
for (i in 1:32){
     predict=c(predict,mi.glm[i]*u_k[k+1])
}}}
```

Appendix B – R implementation of procedure PROC2

Generating the number of claims using Poisson distribution – function rpospois() from the vGAM package
Generating the mean value of claims using Tweedie distribution – function rtweedie() from the TWEEDIE package

```
S=1000              ## the number of simulation
for (s in 1:S){
cv=2                ## the coefficient of variation
K=10                ## the number of levels for multi-level factor
avCl=c()            ## the average value of claims
nCl=c()             ## the number of claims
mse_k_sim=c()
u_k_sim=c()
predict_sim=c()
p=1.2
for (k in 0:(K-1)){
for (i in (1+32*k):(32+32*k)){
gener.nCl=rpospois(1, data$w_i[i])
nCl=c(n, gener.nCl)
skala=data$Y_i[i]*((cv)^2)
gener.avCl=rtweedie(gener.nCl, 2, skala, 10)
avCl=c(avCl, mean(gener.avCl)) }}
PROC1
u_k_sim=rbind(u_k_sim,u_k)
predict_sim=rbind(predict_sim,predict)
mse_sim_k=c()
for (k in 1:10){
mse_sim_k=c(mse_sim_k,(sum((u_k[,k]-u_k_mc[,k])^2))/S)}
```

Appendix C – R implementation of procedure PROC3

```
u_k_matrix=NULL
for ( sim in 1:1000){
p=runif(1,1,2)
PROC1
u_k_matrix=rbind(u_k_macierz, u_k)}
```

COMPANY FINANCIAL PERFORMANCE PREDICTION ON ECONOMIC VALUE ADDED MEASURE BY SIMULATION METHODOLOGY

Zdeněk Zmeškal, Dana Dluhošová*

VSB - Technical University of Ostrava, Faculty of Economics, Department of Finance
Zdenek.Zmeskal@vsb.cz

Abstract. Measurement, analysis, planning of company and sector financial performance is undoubtedly one of crucial instrument and problem of financial managing. There are methodologies of financial performance described and verified in the paper. Calculation and decomposition of EVA measure is described according to relative value spread approach. Typology of stochastic processes and methods of statistical estimation of parameters are explained. Since EVA measure is decomposed by pyramidal method as non-linear function of financial ratios, forecasting is carry out by Monte-Carlo simulation procedure due to Choleski algorithm. Verification of EVA on monthly data of the machinery company was investigated and particular case study of EVA measure prediction is explained. Stochastic processes of Arithmetic or Geometric mean reversion processes are estimated by Mean Square Estimation method. Simulation Monte Carlo method is applied and presented including graphical representation. Parameters (mean, median, quintiles) of forecasted distribution function are calculated and results discussed. Results can serve for decision-making and give reliability intervals of future EVA value. Possibility of EVA forecasting by simulation Monte Carlo method was verified.
Keywords. economic value added, prediction, mean-reversion process, Monte-Carlo simulation

1. Introduction

Measurement, analysis, planning of company and sector financial performance is undoubtedly one of crucial instrument of financial managing. Since the planning and prediction is one of the financial decision-making problems, prediction of distribution function of financial performance measures is very important. Managing and forecasting of non-financial institution financial performance can be characterised by long-term period (months, years). Furthermore it belongs to capital budgeting and financial risk hedging. Several methodologies should be applied. One of them is well-known CorporateMetrics methodology, see Lee (1999). Measuring a firm's performance is one of the key problems not only on the macroeconomics level management, but also on the management level of individual firms. Approaches for measuring a firm's performance have passed through evolution and reflect both technical-economic type of economy, information possibilities and also the level of knowledge of economical systems management. Traditional measures based on accounting profitability are ROE, ROA, ROC, ROI, RONA and measures based on financial cash flow include CFROI, NPV, CROGA. Relatively new group of measures is based on economic profit category include Economic Value Added (EVA) and Market Value Added (MVA).

Intention of the paper is explanation and verification of EVA measure prediction based on estimated random processes and Monte Carlo simulation methodology. The paper is structured as follows: (i) description of EVA measure calculation approaches, (ii) description of chosen random processes and Monte-Carlo simulation; (iii) verification of methodology described.

2. Characteristics and calculation of economic value added

EVA is derived from the basic rule, that a firm has to create at least such value as has been invested. This cost of capital or internal rate of return concerns both equity and debt. So the debt holders have the right to receive their interest and shareholders' dividends, i.e. to cover adequate rate of return, which could compensate their risk. EVA is actually the way, how shareholders measure their profit after covering alternative cost of capital. There have been written a lot of papers regarding this topic, e.g. see Rappaport (1986), Stewart (1991), Ehrbar (1998), Grant (1997), Young et al. (2000), Damodaran (2002), Dluhošová (2003).

General concept of EVA, as a measure of financial performance, expresses the difference between profit and cost of capital, which reflects a minimal rate of return of capital invested (equity and debt). Data set availability

---

The paper is supported by the research project GACR 402/08/1234. This support is gratefully acknowledged.*

and the way of the cost of capital calculation determine EVA calculation method. Moreover, it is also important, if we want to calculate an absolute or a relative value. Broadly speaking, there are two basic concepts of calculation: operating profit concept and value spread concept.

EVA calculation based on operating profit is defined as follows: $EVA = NOPAT - WACC \cdot C$. It implies from the definition, that basic elements for EVA calculation are net operating profit after taxation *NOPAT*, value of total capital invested *C* and weighted cost of capital *WACC*.

Other version of EVA calculation can be expressed using so-called value spread. Value spread reflects economical profit, which can be calculated as a difference between achieved profitability and cost of capital. EVA calculation based on value spread can be formulated as follows: $EVA = (ROC - WACC) \cdot C$, where *ROC* is return on capital invested. In narrow sense EVA can be calculated by formula, $EVA = (ROE - r_E) \cdot E$, where *ROE* is return on equity, r_e is market cost of equity. In this case, only *ROE* is input parameter. It is important for the owners the $ROE - r_E$ spread to be as large as possible or at least positive. Only in this case investment to the firm brings more than an alternative investment. In this situation, value of EVA is independent on equity level and relative firm's performance should be measured.

Pyramidal (Du Pont) decomposition of financial performance measures is useful. There are many variants of decomposition. Assuming *ROE* decomposition in following way $ROE = \frac{EAT}{S} \cdot \frac{S}{A} \cdot \frac{A}{E}$, where *EAT* is net profit, *S* is sales, , *A* is assets and *E* is equity, then narrow sense EVA calculation is calculated as follows, $EVA = \left(\frac{EAT}{S} \cdot \frac{S}{A} \cdot \frac{A}{E} - r_e \right) \cdot E$.

Calculation of cost of capital influence significantly results. There are several approaches of calculation. If in economy exist liquid capital market, and then CAPM model and arbitrage model should be applied. Under opposite conditions, construction model is to be used.

Construction model is written in such way,

$$r_e = \frac{WACC \cdot \frac{ERP}{A} - (1-t) \cdot \frac{I}{BC + B} \cdot \left(\frac{ERP}{A} - \frac{E}{A} \right)}{\frac{E}{A}}$$

, where *ERP* are resources to be paid (interested debt), *A* are total assets, *I* are interests, *BC* are bank credits, *B* are bonds, *E* is equity. Simultaneously, *WACC* are calculated as if we financed company only with equity $WACC = r_f + r_{size} + r_{entrepreneurial} + r_{finstab}$, where r_f is risk free rate, r_{size} is risk premium for share liquidity, $r_{entrepreneurial}$ is risk premium for trade risk, $r_{finstab}$ is risk premium resulting from financial stability.

3. Description of estimation and simulation EVA measure prediction methodology

Main goal of prediction is to give future estimation of EVA probability distribution function, which should be derived and composed from particular financial ratios stochastic processes. The problem is to be solved in simplified (linear) examples explicitly by closed form solution. Mostly however, due to non-liner computation of EVA and distribution types, methodology of random simulation has to be applied.

Procedure of EVA prediction is following.

- Determination of corporation financial outputs on EVA basis for given period.
- Determination of EVA calculation, including particular financial ratios as random factors.
- Estimation and prediction of stochastic models of financial ratios.
- Firstly, we can use models for particular ratios based on Ito process, e.g. Brown models, mean reverting models (Vasicek, CIR, HW, Schwartz).
- Secondly, simultaneous co-integration equation models can be applied, e.g. vector autoregressive models, error correction models, vector error correction models, see Kim et al (1999).
- Determination of EVA probability distribution as function of particular financial ratios by simulation methodology, e.g. Choleski decomposition, eigenvalue decomposition.
- Computation of distribution function parameters, e.g. expected value, median, percentiles, value at risk, shortfall value.

Ito process represents the generalised stochastic process, which includes Wiener, Brown and mean-reversion processes. It is defined as follows, $dx = a(x;t) \cdot dt + b(x;t) \cdot dz$, where *a* is trend, *b* is standard deviation, *dt* is interval, dz is specific Wiener process defined $dz \equiv z_T - z_0 = z \cdot \sqrt{dt}$, z is random variable of standardised normal distribution.

Furthermore, for selected models stochastic differential equation will be introduced and equation for simulation of variable derived in closed form equation will be presented.

Arithmetic Brown model, $dx = a \cdot dt + \sigma \cdot dz$, $x_t = x_{t-1} + a \cdot \Delta t + \sigma \cdot \sqrt{\Delta t} \cdot z$.

Geometric Brown model, $dx = a \cdot x \cdot dt + \sigma \cdot x \cdot dz$, $x_t = x_{t-1} \cdot EXP\left[\left(a - 0{,}5 \cdot \sigma^2\right)\Delta t + \sigma \cdot \sqrt{\Delta t} \cdot z\right]$.

Arithmetic mean-reversion (Ornstein-Uhlenbeck, Vasicek) model,

$$dx = a \cdot (b - x)dt + \sigma \cdot dz \text{ , } x_t = x_{t-1} \cdot e^{-a \cdot \Delta t} + b \cdot \left(1 - e^{-a \cdot \Delta t}\right) + \sigma \cdot \sqrt{\frac{\left(1 - e^{-2a \cdot \Delta t}\right)}{(2a)}} \cdot z \text{ .}$$

Geometric mean-reversion (Schwartz) model,

$$dx = a \cdot (b - \ln x) \cdot x \cdot dt + \sigma \cdot x \cdot dz \text{ , } x_t = \exp\left\{\left[\ln(x_{t-1}) \cdot e^{-a \cdot \Delta t}\right] + \left\{\left[b - \left(\frac{\sigma^2}{2 \cdot a}\right)\right] \cdot \left(1 - e^{-a \cdot \Delta t}\right)\right\} + \sigma \cdot \sqrt{\frac{\left(1 - e^{-2a \cdot \Delta t}\right)}{(2a)}} \cdot z\right\} \text{ .}$$

For computation of probability distribution of function of random variables several methods should be used. Assuming the normal distribution of random variables, Choleski decomposition and Eigen-value decomposition can be applied. For estimation mean square error method (MSE), maximal likelihood estimation (MLE) and method of moments can be used.

4. Verification of EVA prediction by simulation methodology

Prediction of EVA measure will be verified on machinery company monthly data. Prediction is performed for future twelve months and simulation procedure on Choleski decomposition is applied. Estimation of random financial ratios models (Arithmetic and Geometric mean reversion) is done from historical monthly time-series of 60 months by MSE method. EVA measure is expressed by following non-linear function, $EVA = \left(\frac{EAT}{S} \cdot \frac{S}{A} \cdot \frac{A}{E} - r_e\right) \cdot E$.

4. 1 Input data and financial ratios models statistical estimation

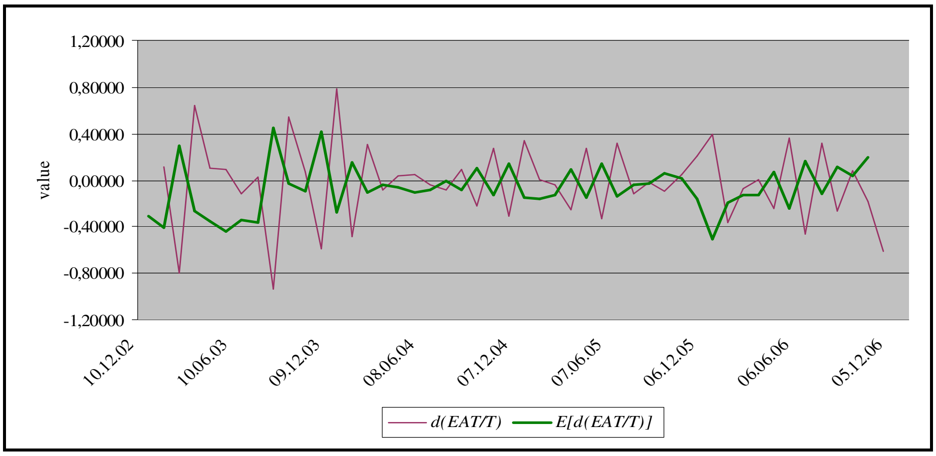
From EVA equation is apparent, that following financial ratios are variables: $\frac{EAT}{S}, \frac{S}{A}, \frac{A}{E}, r_E, V_E$. Because, equity is non-stationary, this variable is transformed on return $V_E = \frac{\Delta E}{E}$ and subsequently estimated. There are in Appendix 1 input data time-series. Since, ratios $\frac{S}{A}, \frac{A}{E}, r_E$ must be positive, geometric mean reversion process (GMRP) is applied. And rest of ratios $\frac{EAT}{S}, V_E$ are modelled by arithmetic mean reversion process (AMRP). Tab. 2 shows model types and estimated parameters of models are significant on 95% probability level. Tab. 3 presents estimated covariance matrices of residuals. On Graphs comparison of estimated and real variables of financial ratios are presented.

	Parameters		σ	Process
	a	b		
EAT/S	0,86829	0	0,39287	AMRP
S/A	0,87785	-2,02913	0,61155	GMRP
A/E	0,17999	1,02386	0,17485	GMRP
dV_e	0,63427	0	0,29376	AMRP
r_e	0,97463	-3,97267	0,47334	GMRP

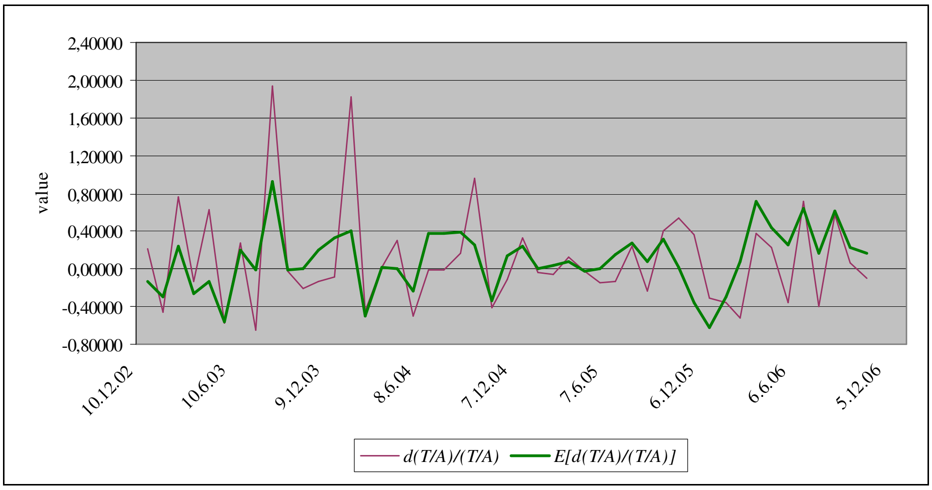
Tab. 2 Estimated parameters of financial ratios models

	EAT/S	S/A	A/E	V_e	r_e
EAT/S	0,07124				
S/A	0,05482	0,17621			
A/E	-0,02661	-0,01330	0,02567		
dV_e	0,04506	0,04430	-0,02358	0,04728	
r_e	-0,05899	-0,04224	0,02364	-0,02958	0,09858

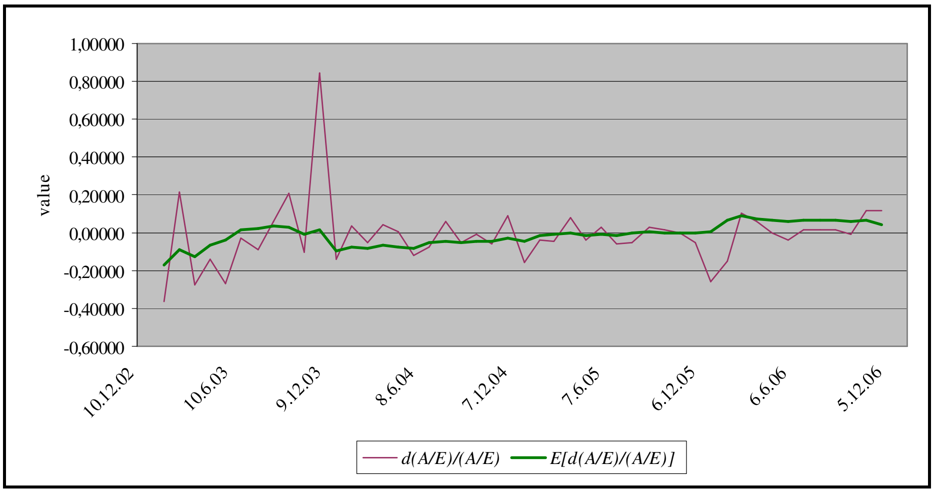
Tab. 3 Covariance matrices of residuals



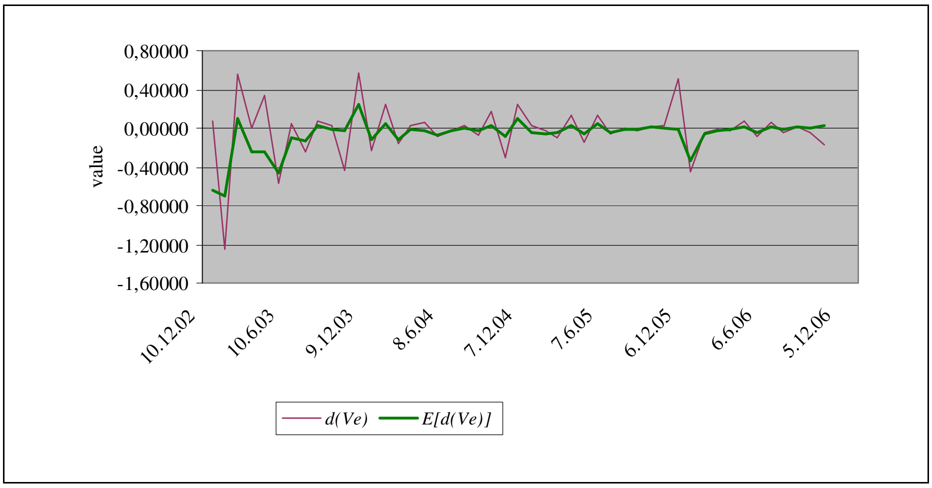
Graph 1 Estimation of EAT/S ratio



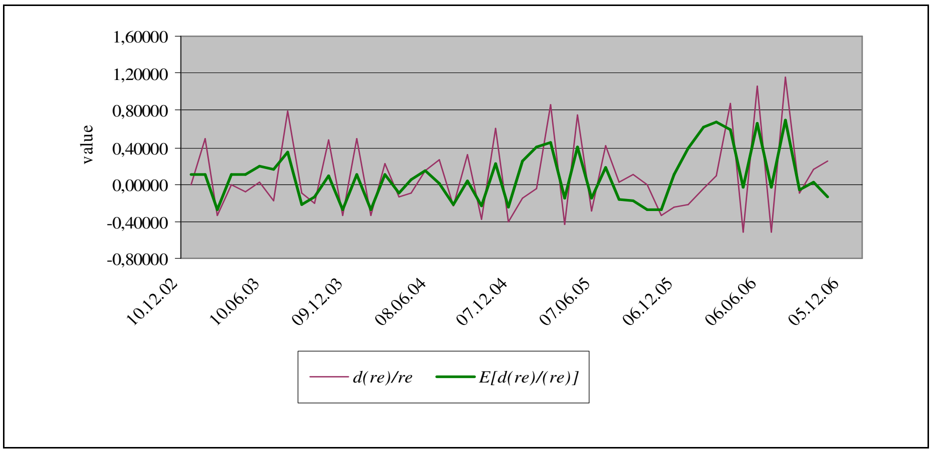
Graph 2 Estimation of T/A ratio



Graph 3 Estimation of A/E ratio



Graph 4 Estimation of V_E ratio



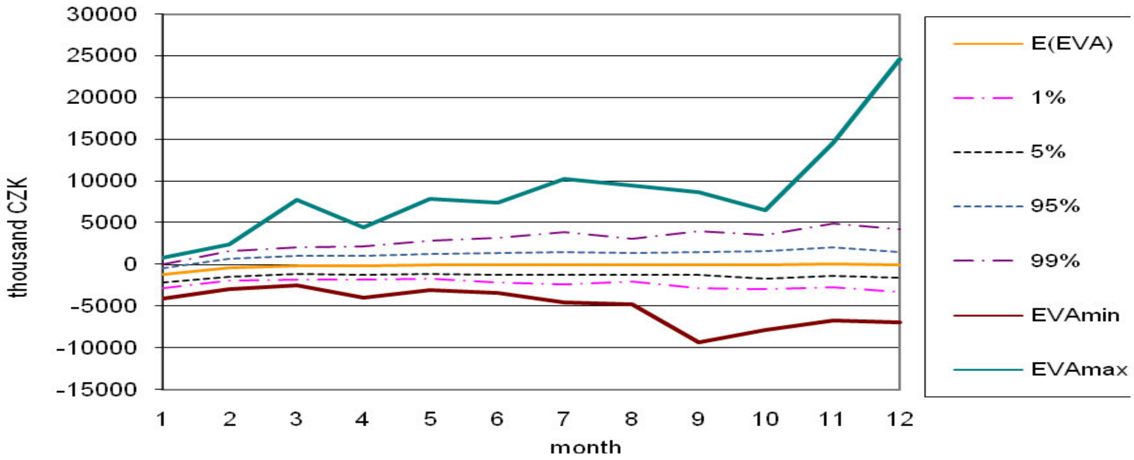
Graph 5 Estimation of r_E ratio

4. 2 Simulation and results of EVA measure prediction

Results of EVA prediction in a form of probability distribution are on Tab. 4 and Graph 6. .

th. CZK	1. month	2. month	3. month	4. month	5. month	6. month	7. month	8. month	9. month	10. month	11. month	12. month
E(EVA)	0	-469	-228	-151	-45	-38	-47	-60	-36	-64	30	-62
1%	-2761	-1868	-1755	-1744	-1647	-2063	-2296	-2016	-2815	-2922	-2676	-3262
5%	-2101	-1418	-1119	-1166	-1099	-1159	-1235	-1163	-1260	-1601	-1360	-1547
95%	-400	691	1000	1024	1288	1402	1516	1313	1465	1626	2020	1492
99%	48	1621	2089	2210	2891	3214	3923	3074	4001	3544	4930	4297
EVAmin	-4068	-3009	-2510	-3964	-3107	-3429	-4521	-4797	-9311	-7888	-6697	-6905
EVAmax	811	2403	7710	4410	7879	7383	10230	9484	8632	6462	14610	24626
median	-1213	-509	-305	-199	-139	-130	-135	-112	-101	-87	-82	-79

Tab. 4 Eva results for future twelve months



Graph 6 Graphical representation of EVA probability distribution function

From results implies that even if expected EVA value is negative, the trend is positive. Similar trend and values are in a form of median. Low percentiles (1%, 5%) are negative and high percentiles (95%, 99%) are positive.

5. Conclusion

In the paper methodology of EVA measure prediction was described and applied. Firstly, pyramidal decomposition of EVA was defined, including particular financial ratios as random variables. Secondly, particular financial ratios random models were modelled by arithmetic and geometric mean reversion processes. Parameters of models were estimated by MSE method. Finally, future EVA probability distribution functions for 12 months was determinate by simulation Monte-Carlo methodology by virtue of Choleski decomposition. Methodology was verified on real data of the machinery company.

Approach described and verified allows comprehensively model random development of EVA measure. Even if historical time-series was relatively short, it was possible carry out statistically significant estimate of financial ratios. It was verified, that the methodology can be applied on condition of companies in transitive economies as well.

References

[1] COPELAND, T.E, WESTON, J.F. *Financial theory and corporate finance,* Addison-Wesley, Reading, 1988.

[2] DLUHOŠOVÁ, D. Performance analysis of the transitive economy companies and industries. In *Business economics, management and marketing.* Athens: Athens institute for education and research, 2003

[3] DLUHOSOVA, D. An analysis of financial performance using the EVA method. FINANCE A UVER- CZECH JOURNAL OF ECONOMICS AND FINANCE, Vol. 54, pp : 541, 2004

[4] EHRBAR, A. *EVA*: *The real key to creating wealth.* New York: John W, Wiley & Sons, Inc., 1998.

[5] GRANT, J.L. *Foundations of Economic Value Added.* Pennsylvania: New Hope, 1997.

[6] KIM, J., NEALE, A. M., MINA, J. *Long-term Technical Document.* RiskMetrics Group. J.P. Morgan.

[7] LEE, A.Y.: *CorporateMetrics Technical Document*. Riskmetrics Group, J.P. Morgan, 1996.

[8] LONGERSTAEY, J., SPENCER, M. *RiskMetrics - Technical Document.* Riskmetrics Group. J.P. Morgan. 1996.

[9] RAPPAPORT, A. *Creating Shareholder Value.* New York: Free Press, 1986.

[10] STERN, J. M., SHILEY, S, ROSS, I. *The EVA chalenge: implementing value added change in an organization*, Wiley Finance, 2001.

[11] STEWART, G.B. *The Quest for Value.* New York: HarperCollins, 1991.

[12] YOUNG, S.D., O´BYRNE, S.F. *EVA and Value-Based Management. Practical guide implementation*. Mc Graw Hill, 2000.

[13] ZMESKAL, Z. Application of the fuzzy - stochastic Methodology to Appraising the Firm Value as a European Call Option, European Journal of Operational Research, Vol. 135/2, pp 303-310

[14] ZMESKAL, Z. *Financial models*. VSB-TU Ostrava, 2004.

[15] ZMESKAL, Z Hedging strategies and financial risks. FINANCE A UVER Vol.: 54, Issue: 1-2, pp 50-63,: 2004

[16] ZMESKAL, Z**.** Application of the American real flexible switch options methodology a generalized approach. FINANCE A UVER-CZECH JOURNAL OF ECONOMICS AND FINANCE Vol. 58 Issue**:** 5-6 pp 261-275, 2008

[17] ZMESKAL, Z**.** Soft Approach to Company Financial Level Multiple Attribute Evaluation. Conference Information: 21st International Conference on Mathematical Methods in Economics, SEP 10-12, 2003 Prague, CZECH REPUBLIC. PROCEEDINGS OF THE 21ST INTERNATIONAL CONFERENCE MATHEMATICAL METHODS IN ECONOMICS 2003 pp 273-280 Published: 2003

EAT/S	*S/A*	*A/E*	V_e	r_e	*EAT/S*	*S/A*	*A/E*	V_e	r_e
0,420485	0,340437	-0,19003	0,716998	-0,11373	0,1594	0,08309	-0,02378	0,07056	-0,39732
-0,39766	-0,15761	0,304237	-0,55775	0,383784	0,117671	-0,03971	-0,03569	0,031494	-0,45117
0,344399	0,527316	-0,15118	0,452002	-0,05239	-0,12807	-0,08648	0,07981	-0,06217	0,40455
0,367864	0,12537	-0,07638	0,252921	-0,11373	0,18355	0,046312	-0,02059	0,104894	-0,28091
0,448457	0,768621	-0,22479	0,581927	-0,18623	-0,18109	-0,00714	0,039521	-0,09177	0,352658
0,323775	-0,01638	-0,04449	-0,1128	-0,16964	0,183657	-0,15376	-0,04357	0,09028	-0,13832
0,368225	0,070621	-0,10873	0,142868	-0,33857	0,02825	-0,27359	-0,05238	-0,00869	0,235694
-0,56984	-0,64732	0,022938	-0,12589	0,437794	0,024612	-0,02997	0,023814	0,007221	0,182951
0,098577	1,004204	0,18064	0,035342	0,128984	-0,07266	-0,3182	0,0099	-0,02199	0,287583
0,104399	-0,01393	-0,09314	0,044817	-0,08069	-0,0048	0,093203	0,00181	0,001316	0,279832
-0,49216	-0,20876	0,835571	-0,41093	0,376542	0,1912	0,518488	-0,05255	0,027325	-0,05239
0,374004	-0,3385	-0,04061	0,317077	-0,05239	0,55877	0,71537	-0,26703	0,526874	-0,36106
-0,219	-0,4116	0,111742	-0,12343	0,383784	0,146444	0,314022	-0,21059	-0,11132	-0,60523
0,151679	1,411102	0,025355	0,205925	-0,05239	0,115953	-0,05058	0,014063	0,007161	-0,67354
0,022395	0,054068	0,111715	-0,03922	0,111126	0,128163	-0,59591	-0,01224	0,005388	-0,57922
0,069354	0,002119	0,085889	0,046257	-0,04758	-0,10678	-0,34026	-0,06166	-0,02055	0,291155
0,117862	0,303595	-0,04352	0,09685	-0,14853	0,289458	-0,20783	-0,10076	0,067127	-0,4833
0,074479	-0,26658	-0,01734	-0,00681	0,005128	-0,22031	-0,61105	-0,05443	-0,04458	0,390864
-0,00653	-0,38402	0,09989	-0,01115	0,261547	0,157636	0,080373	-0,05436	0,041219	-0,48804
0,097279	-0,39602	0,000642	0,037448	-0,00873	-0,15042	-0,56427	-0,05026	-0,02899	0,473062
-0,13569	-0,22225	0,037247	-0,05452	0,286558	-0,03032	-0,0499	-0,06956	-0,00331	-0,01955
0,162319	0,700805	-0,02034	0,147971	-0,13715	-0,22369	-0,15848	0,057542	-0,04847	0,144636
-0,1854	-0,08477	0,115212	-0,21875	0,385787	-0,80964	-0,26812	0,074955	-0,20234	0,386266
0,197688	-0,25613	-0,11248	0,13298	-0,15851					

Appendix 1 Input data of financial ratios

MULTIPLE MARGINALIZATION IN SERIAL AND PARALLEL SUPPLY CHAINS OPERATING IN A NON-LINEAR ENVIRONMENT

Jan Zouhar

University of Economics, Prague
Faculty of Informatics and Statistics, Dept. of Econometrics
zouharj@vse.cz

Abstract. Double (or multiple) marginalization is often identified as the main source of a decentralized supply chain's (SC's) inefficiency. In its core lies the fact that if the agents constituting the SC choose their output prices according to the golden rule of profit maximization (that normally applies to a single firm that produces independently and sells directly to the end consumer), the prices in the SC tend to spiral up to an inefficient (equilibrium) level where both the consumer surplus and the SC's total profit are diminished. The aim of our research is to quantify this effect. In our previous papers, we analyzed the properties of SC's operating with linear cost and demand functions; the simple linear form enabled us to derive some explicit formulas that yield the discussed results. In a non-linear setting, meaningful general formulas can no longer be obtained. Therefore, we developed numeric algorithms to find equilibrium prices in SC's with non-linear demand and cost functions, and the analysis is carried out through numeric experiments with these algorithms. The aim is to establish to what extent the results from the linear model hold in a non-linear environment, and what the effect of the cost and demand functions' convexity/concavity is.
Keywords. Supply chains, multiple marginalization, serial architecture, parallel architecture, equilibrium prices.

1. Introduction

Double (or *multiple*) *marginalization* is often mentioned as one of the weaknesses of the supply chain as a form of a production system. The effect of multiple marginalization occurs when there's a lack of coordination between in the agents constituting the supply chain in setting their output prices. More precisely, it can be shown (see [1], [2]) that if all the agents behave as if they were a single firm that produces independently and sells directly to the end consumer, and choose their output prices according to the golden rule of profit maximization, the prices tend to spiral up to an equilibrium point. The equilibrium combination of prices yields a much lower profit than what could be achieved if the supply chain was coordinated (or *centralized*). Moreover, the final product of the supply chain is sold at high prices and in low quantities, thus reducing the consumer surplus. Therefore, we can observe that the equilibrium point of the non-coordinated (*decentralized*) supply chain is inefficient.

In our previous paper [8], we studied the properties of the equilibrium point of a decentralized supply chain operating in a linear environment; i.e., the function of demand for the chain's final product and the cost functions of individual agents were all linear. In this simplified setting, we managed to derive some explicit formulae that describe the equilibrium prices, profits and quantity demanded of the final product. Using these formulae, we can quantify the impact of multiple marginalization (in terms of the loss of profit-generating efficiency) and analyze some other phenomena, such as the distribution of the total profit between the individual agents or positive externalities of cost savings among the agents.

In this paper, we carry out a similar analysis for supply chains operating in a non-linear environment; i.e. the demand and/or cost functions are non-linear (typically smooth) real functions. With such a vague specification of cost and demand functions, meaningful general formulas can no longer be obtained. Therefore, in order to study the patterns in supply chains' behaviour in various settings, we developed generic numeric algorithms that find equilibrium prices in supply chains with non-linear demand and cost functions, and the analysis is carried out through numeric experiments with fictitious supply chains using these algorithms. The main aim of these experiments is to establish to what extent the results from the linear model hold in a non-linear environment, and what the effect of the cost and demand functions' convexity/concavity is.

In section 2, we outline the main model framework we use for studying the effect of multiple marginalization (the same framework is used regardless of linearity/non-linearity of cost and demand functions). In section 3, we summarize the main findings we drew earlier from the linear model. Section 4 discusses the numeric algorithms that find the equilibrium behaviour of a supply chain with non-linear functions. Section 5 presents an example of one of the numeric experiments, which demonstrates the way the results were obtained; section 6 brings an overview of the results themselves.

2. Model Framework

First of all, we'll outline the main model framework we use for studying the effect of multiple marginalization. A supply chain in this framework is characterized by its following properties:

- *number of agents*,
- *inverse function of demand* for the final product,
- *cost functions* of individual agents,
- *architecture* of the chain.

Agents are numbered from 1 to n, where n is the total number of agents. The numbering of agents reflects the order of agents in the production process; e.g. agent n is the agent who sells the final production to end consumers. The inverse demand function expresses the quantity demanded of the supply chain's final production (q) as a function of its price (i.e. the price charged by the n-th agent, p_n):

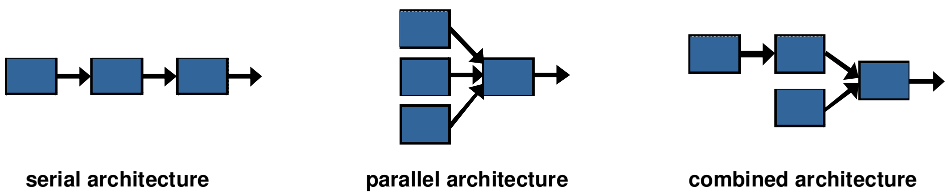
$$\delta : R_0^+ \to R_0^+ \text{ , } q = \delta(p_n) \text{ .} \tag{1}$$

The total production cost of the i-th agent (tc_i) is a function of production quantity. Without loss of generality, we can assume that each agent produces the same amount of output q, i.e. each agent in the supply chain contributes to the final product with one unit of their output. Using this convention, we can define the cost function of i-th agent (γ_i) as

$$\gamma_i : R_0^+ \to R_0^+ \text{ , } tc_i = \gamma_i(q) \text{ .} \tag{2}$$

The architecture of the supply chain explains how the agents are linked together. We distinguish three types of a supply chain's architecture (see *figure 1*): *serial architecture*, consisting only from one-to-one links, *parallel architecture*, made up by one many-to-one link, and *combined architecture*, which combines both kinds of links. A many-to-one link represents a situation where there are several agents (at the beginning side of the link) supplying *different* components of the final product that are later assembled together by the agent at the end side of the link.

Figure 1: Serial, Parallel and Combined Supply Chains



It follows from (1) and (2) that an agent's total cost can be expressed as a function of price p_n:

$$tc_i = \gamma_i(\delta(p_n)) \text{ ,} \tag{3}$$

and the total profit of the supply chain (π_{SC}) can therefore be calculated as the difference between total revenue and the sum of all costs along the chain:

$$\pi_{SC} = p_n \cdot \delta(p_n) - \sum_{i=1}^{n} \gamma_i(\delta(p_n)) \text{ .} \tag{4}$$

From (4) it can be seen that with given demand and cost functions, the total profit of the chain is determined by the price of the final production. The output prices charged by the remaining agents only influence the way the total profit is distributed between the agents (without a direct impact on). However, in a decentralized price-

In order to quantify the effect of multiple marginalization, we can use a gauge for the efficiency loss due to multiple marginalization (see [6],[8]). This gauge compares the total profit of a decentralized supply chain operating under the assumptions of multiple marginalization (π_{DSC}) and the maximum profit attainable by the supply same chain if it were completely coordinated (π_{CSC}). The resulting indicator is called the *efficiency* of a supply chain (ε) and is calculated as the ratio of the two profits ($\varepsilon = \pi_{DSC} / \pi_{CSC}$).

3. Linear Model of Multiple Marginalization: A Summary

In [8], we studied a simplified case where functions in (1) and (2) were all linear. Under these assumptions, we derived general formulae for equilibrium prices, quantity demanded and the agents' profits in all three types of supply chain architecture. Similarly, we found the profit of the centralized supply chain (π_{CSC}). From these formulae we can observe the following (for a more detailed record including the underlying formulae, see [8]):

a) The efficiency of a supply chain falls rapidly with the number of agents.

b) The efficiency of a serial chain is lower than that of a parallel chain with the same number of agents. The efficiency of a combined chain is between the two, depending on the actual structure.

c) Efficiency is related *only* to the structure of the chain (number of agents and architecture); it is not influenced by the parameters of the (linear) demand and cost functions. That means that two different supply chains with the same structure always exhibit the same efficiency, no matter what the demand and cost functions are (as long as they are linear).

d) There exist strong positive externalities of cost savings in the supply chain: if one of the agents reduces the unit cost of his production process, the profit of all agents increases. Moreover, the profit increases in the same proportion for all agents.

e) The distribution of profit between the agents depends only on the structure of the chain; i.e., the proportion of the total profit that is won by a given agent is determined by their position in the supply chain structure, regardless of the layout of costs inside the chain.

It further follows from **c** that a supply chain's efficiency is not affected by cost savings – as the profit of the decentralized chain rises due to the decrease in costs, so does the profit of its centralized counterpart, and both increase in the same proportion. Similarly, **e** implies that if one the production processes is moved from one agent to another one (resulting in a decrease in costs on one side and a corresponding increase on the other one), the individual profit of neither agent changes (nor does any other profit in the supply chain or the chain's efficiency).

4. Finding Equilibrium Prices in a Decentralized Supply Chain

As it was already mentioned, the aim of the research described in the following text is to find out whether (or to what extent) the findings of the linear model hold in case we let go the assumption of linearity. With general non-linear cost and demand functions, it's not possible to adopt the analytic approach we used for the linear model. Instead, we decided to use computer experiments and analyze numerous fictitious supply chains varying with respect to the shape of cost and demand functions. Before these experiments could have taken place, we had to develop numeric algorithms capable of finding the equilibrium combination of prices in a decentralized supply chain.

For the sake of simplicity, we studied only serial and parallel supply chains. Both the results from the linear model and common sense suggest that the properties of combined chains lie somewhere between those of the two extreme cases (i.e. serial and parallel chains).

4.1. Algorithm for a Serial Supply Chain

In a serial supply chain, an agent's equilibrium price has to maximize the agent's profit, given the input price (price of the preceding agent's output) and the firm-specific demand for the agent's output. The firm-specific demand is a function that explains the total response of the quantity of the chain's product demanded to a change in the output price by the agent in question. This response, however, includes the reaction of all the ensuing agents – a change of a price in the middle of a supply chain induces a chain of re-optimizing the output prices of those who follow in the production process. Therefore, the algorithm for finding the equilibrium prices in a serial chain goes from the last (n-th) agent, whose firm-specific demand function is directly the function of demand for the chain's final product. To find firm-specific demand curves for the preceding agents, a concept of price reaction curves is adopted: a price reaction curve of the i-th agent (denoted p_i^{opt}) is a function of the i-th agent's input price (p_{i-1}) and describes the price p_i that the i-th agent chooses in order to maximize their individual profit, given that all the ensuing agents behave in the same way (i.e., try to maximize their individual profit). By compounding price reaction curves of ensuing agents, we can derive the firm-specific demand function for an arbitrary agent in the supply chain. Going from the end of the supply chain towards the beginning, we can devise price reaction curves from agent n down to agent 1; agent 1 has no variable input price, and thus the firm-specific demand function alone determines the equilibrium price of this agent. The equilibrium prices of successive agents follow immediately from price reaction curves. The whole algorithm goes as follows:

Step 1: Formulate profit function of agent n: $\pi_n(p_n, p_{n-1}) = (p_n - p_{n-1}) \cdot \delta(p_n) - \gamma_n(\delta(p_n))$.

Step 2: Find reaction curve p_n^{opt} given by $p_n^{opt}(p_{n-1}) = \operatorname*{argmax}_{p_n \in R_0^+} \pi_n(p_n, p_{n-1})$.

Step 3: Set $i = n-1$.

Step 4: Formulate ψ by function composition: $\psi = p_{i+1}^{opt} \circ p_{i+2}^{opt} \circ \ldots \circ p_n^{opt} \circ \delta$.

Step 5: Formulate profit function of i-th agent: $\pi_i(p_i, p_{i-1}) = (p_i - p_{i-1}) \cdot \psi(p_i) - \gamma_i(\psi(p_i))$.

Step 6: If $i = 1$, go to *step 9*, else go to *step 7*.

Step 7: Find the reaction curve p_i^{opt} given by $p_i^{opt}(p_{i-1}) = \underset{p_i \in R_0^+}{\arg\max}\ \pi_i(p_i, p_{i-1})$.

Step 8: Go to *Step 4*.

Step 9: Find equilibrium price of agent 1 from $p_1^{eq} = \underset{p_1 \in R_0^+}{\arg\max}\ \pi_1(p_1, 0)$.

Step 10: Set $i = i+1$.

Step 11: Find equilibrium price of i-th agent from $p_i^{eq} = p_i^{opt}(p_{i-1}^{eq})$.

Step 12: If $i < n$, go back to *step 10*, else finish.

In a software implementation of the algorithm, the crucial steps are steps 2 and 4 – approximating the price reaction functions. Typically, one calculates values of the reaction function in sufficiently many points (this involves finding the profit-maximising output price for varying input prices) and approximates these values by a polynomial of a chosen degree.

4.2. Algorithm for a Parallel Supply Chain

In a parallel supply chain, the way the prices affect one another is a bit more complex, and so is the algorithm that finds the equilibrium. Due to limited extent of this paper, we won't discuss the algorithm in detail; its thorough description is the sole purpose of our paper [9] (in Czech only). The complication here is of course the parallel position of the agents in the first level of the supply chain. It can be shown that the selection of output prices by these agents can be modelled as a non-cooperative game, the solution to which is the Nash equilibrium combination of prices. The main task of the algorithm for parallel chains is to find a combination of prices of the first-level agents that satisfies the conditions of Nash equilibrium.

Both of the mentioned algorithms were implemented in Matlab and work with virtually any smooth function in the role of supply and demand. They were used in computer experiments, the aim of which was to analyze the impact of convexity/concavity of cost and/or demand functions. The analysis was carried out in the following way: we prepared data for several supply chains that were all identical as far as their structure was concerned, only the cost and/or demand functions varied slightly between the individual cases; for instance, the cost function stretched from those exhibiting economies of scale to those representing diseconomies of scale in production, while being quite similar in absolute measure. The equilibrium values were calculated for all of them, and the comparison provided us with some clue as to what the impact of the shape of the cost function is. Section 5 gives an example of one of such experiments.

5. Numeric Experiments: An Example

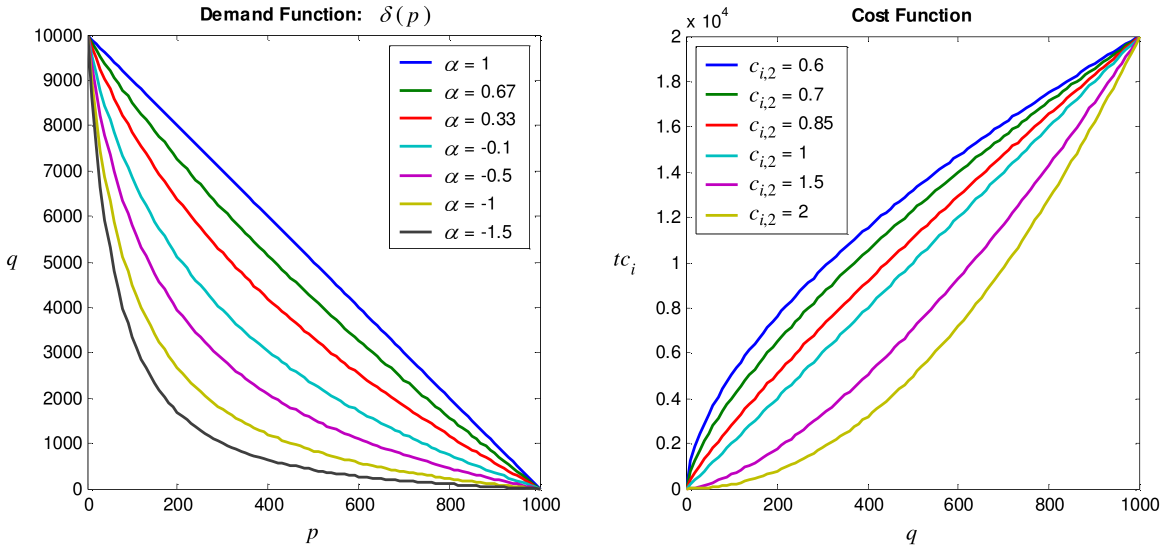
In this example, we compare the properties of 42 serial supply chains, all of them consisting of 4 successive agents. We consider 7 alternative demand functions and 6 different cost functions (in each chain, agents have identical cost functions]; the shapes of these functions are depicted in *graph 1*. The demand functions are in the form

$$\delta(p) = a \cdot (p + 100)^{\alpha} + d\,, \tag{5}$$

where the convexity of the function is controlled by the parameter α that takes on the following values: 1, 0.67, 0.33, -0.1, -0.5, -1, -1.5; parameters a and d are set so that the demand function intersect the p and q axes at values 1,000 and 10,000, respectively (see *graph 1*). The cost functions are given by

$$\gamma_i(q) = c_{i,1} \cdot q^{c_{i,2}}\,. \tag{6}$$

The convexity/concavity of the cost function results from the value of the parameter $c_{i,2}$ that ranges from 0.6 to 2 (see *graph 1*). Values lower than 1 yield a concave cost function that exhibits economies of scale; values greater than one result in convex curves, reflecting the costs of a production process with diseconomies of scale. Altogether, we have six different cost functions; three of them represent various degrees of economies of scale, two of them exhibit diseconomies of scale, and one of them is linear (with proportional scaling).

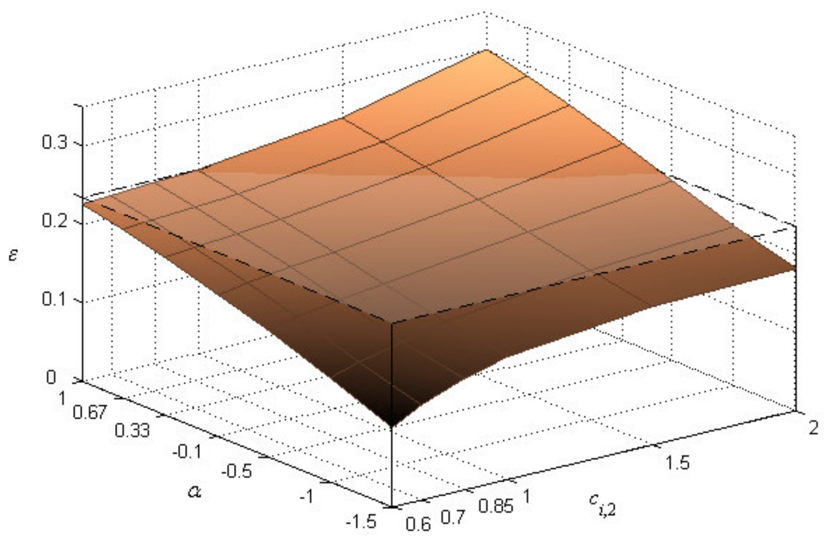


Graph 1: Example – Demand and Cost Functions

By combining the 7 demand functions and 6 cost functions, we obtain 42 different serial supply chains. To all these supply chains we applied the algorithm from section 4.1 to find the equilibrium values for the decentralized chain. The profit of a centralized chain is simply the maximum of function (4); this is a function of a single variable p_n the maximum of which can be found using standard numeric procedures.

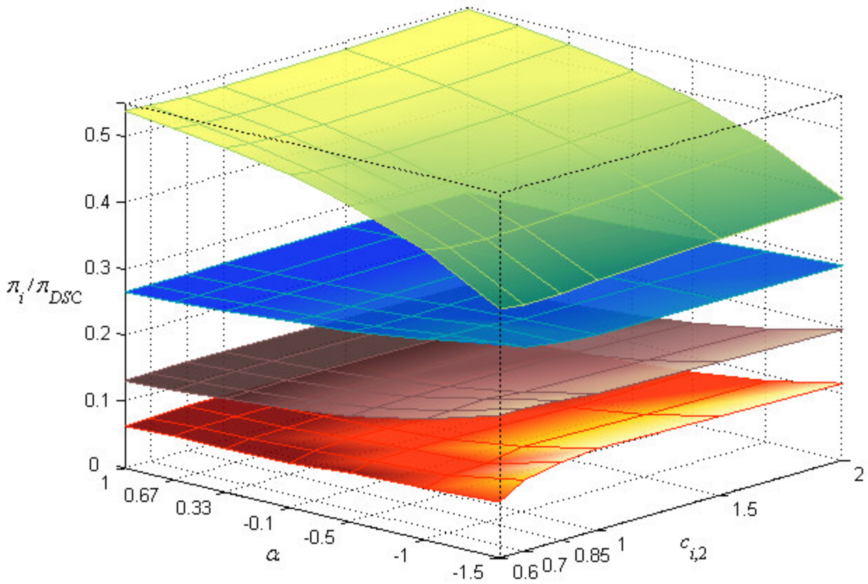
If we compare the results for different supply chains, we can observe the impact of convexity/concavity of demand and cost functions on key indicators such as the efficiency or the distribution of profit between the agents. *Graph 2* shows the efficiency plot for all the 42 supply chains; the horizontal axes show the demand and cost function parameters α and $c_{i,2}$ that control the shape of the demand and cost functions, respectively, and the vertical axis measures the efficiency (ε). The transparent surface in the chart represents the efficiency level from the linear model (note that in the linear model the efficiency is not dependent on the cost and demand functions).

Graph 2 reveals the way the degree of economies of scale affects the efficiency of a supply chain. Consider an arbitrary point on the brown efficiency surface. As we move from this point parallel with axis $c_{i,2}$ towards higher values of $c_{i,2}$, the efficiency of the chain increases. In graph 1, we can see that the lower the values of $c_{i,2}$, the higher the degree of economies of scale. This implies that supply chains with production processes exhibiting diseconomies of scale are more efficient than those with (positive) economies of scale. Analogously, we can analyze the effect of convexity of the demand function. Moving parallel with axis α towards lower values of α decreases the efficiency; thus, we can conclude that convex demand function yield lower efficiency values.



Graph 2: Example – Efficiency in a Serial Supply Chain

In a linear setting, it holds that in a serial chain with 4 agents, the ratio of the profit shares of agents 1, 2, 3 and 4 is 8 : 4 : 2 : 1. *Graph 3* shows the profit shares of the individual agents (expressed as a percentage of the total profit). The profit shares of agent 1 in various supply chains are represented by the topmost green surface; below are the profit shares of agent 2 (blue surface) and so on. The ratio from the linear model can be seen at the point where $\alpha = 1$, $c_{i,2} = 1$. We can observe that moving parallel with the axis $c_{i,2}$ has little effect on profit shares. Therefore, we can infer that there's no significant impact of economies of scale on the distribution of profit in a decentralized supply chain. However, if we move along the axis α from value 1 (linear demand) to lower values (convex demand), the profit shares get closer to one another. Even though the ratio changes, the ranking of profits remains the same in all studied supply chains (agent 1 realized the biggest profit etc.).



Graph 3: Example – Profit Shares of Individual Agents in a Serial Supply Chain

6. Numeric Experiments: A Summary

The example from section 5 is only one of the numerous experiments that have been carried out with non-linear demand and cost functions. Here is a brief summary of the most important results (all refer to *decentralized* supply chains):

- Many of the findings from the linear model hold in a non-linear environment as well, e.g. the impact of the chain's size and architecture on its efficiency or the existence of strong positive externalities of cost savings.
- Convex demand functions and economies of scale of production processes tend to slightly diminish efficiency (note that in a linear setting, the parameters of cost and demand functions do not matter as far as efficiency is concerned).
- Unlike in linear model, the shape of cost and demand functions affects the distribution of profit between the agents; on the other hand, the ranking of individuals profits from the linear model holds.
- If a production process exhibits economies of scale, it is possible to increase the total profit of a decentralized chain by moving this process towards the end of the chain.

7. Conclusions

In this paper, we briefly outlined the methods we used for studying multiple marginalization in supply chains that operate in a non-linear environment, and discussed the most important results. The results were compared to the linear model (the object of our previous research, [8]), where a general analytic description of the studied indicator can be found. Within the selected model framework (section 2), we'd like to think we have carried out a thorough and complete analysis. However, there can be many meaningful modifications to the described model framework that are worth studying. For instance, Corbett and Karmakar in [3] study a linear model of supply chains where a many-to-one link represents agents supplying *the same* component, so that the output of these agents is *substitutable* (as opposed to the *complementary* relation we used in this paper). In our current research, we study the same model as in this paper where the price selection is limited to a discrete scale.

References:

[1] BADASYAN, N., GOEREE, J., MORGAN, J., SERVÁTKA, M. et al.: *Vertical Integration of Successive Monopolists: A Classroom Experiment*. Caltech Working Papers, 2003, 18 p.

[2] BRESNAHAN, T., REISS, P.: *Dealer and Manufacturer Margins*. Rand Journal of Economics, 1985, Vol. 16, No. 2, pp. 253-268.

[3] CORBETT, C., KARMAKAR, U.: *Competition and Structure in Serial Supply Chains with Deterministic Demand*. Management Science, 2001, Vol. 47, No. 7, pp. 966–978.

[4] FIALA, P.: *Modelování dodavatelských řetězců*. Praha, Professional Publishing, 2005, 168 p.

[5] JABLONSKÝ, J.: *Models for efficiency evaluation of decision making units*. Politická ekonomie, 2004, roč. LII, č. 2, s. 206–220. ISSN 0032-3233.

[6] TAYUR, S., ERHUN, F., KESKINOCAK, P.: *Spot Markets for Capacity and Supply Chain Coordination*. 2000, s. 32. Available at WWW: http://citeseer.ist.psu.edu/erhun00spot.html.

[7] WATERSON, M.: *Vertical Integration, Variable Proportions and Oligopoly*. The Economic Journal, 1982, Vol. 92, No. 365, pp. 129–144.

[8] ZOUHAR, J: *Efficiency of Decentralized Supply Chains*. Mathematical Methods in Economics 2007 [CD-ROM]. Ostrava, VŠB-TU Ostrava, 2007. ISBN 978-80-248-1458-2

[9] ZOUHAR, J: *Numerické metody nalezení rovnovážných cen v modelu vícenásobné marginalizace pro paralelní dodavatelský řetězec*. Medzinárodný seminár mladých vedeckých pracovníkov (zborník). Praha, Oeconomica, 2008, 9 p. ISBN 978-80-245-1405-5.

STATISTICAL COMPARISON OF ECONOMIC SITUATION EVALUATION IN MUNICIPALITIES

Miroslav Žižka, Vladimíra Hovorková Valentová, Petra Rydvalová

Technical University of Liberec
Faculty of Economics, Department of Business Administration and Department of Economic Statistics
miroslav.zizka@tul.cz, vladimira.valentova@tul.cz, petra.rydvalova@tul.cz

Abstract. Problems of quantification of regional disparities are a frequently discussed topic today. The solvers from the Faculty of Economics are engaged in its solving within the Ministry for Regional Development research project No. WD-30-07-1. Based on the analysis of Strategy of Regional Development of the Czech Republic for the years 2007 to 2013 and Programme of Development of Individual Regions there were identified twenty six indicators that are used for measuring regional disparities on different levels from municipalities to regions. In the next phase of the research a municipality was determined as a basic unit for investigation. The obtained data about all 6,240 municipalities were analyzed by factor analysis. As a result eight factors which have a significant influence on disparities of the municipalities were recognized. These factors are: unemployment, domicile attractiveness, population density and allocation, age structure, civil amenities, economic structure, sustainable development and economic activity. In the second phase of the research all mayors of municipalities in the Czech Republic were contacted through a questionnaire. The aim was to evaluate the degree of relationship between the results of statistical data analysis and mayors´ opinions on economic situation in their municipalities. Therefore eight chi-square tests of independence for the above eight factors were done. The tests inform whether there is a relationship between the new methodology of research of municipalities' development dynamics and mayors´ opinions on economic strength or weakness of their municipality. The tests were supplemented by measures of a relationship between the two variables. At the end of the paper authors try to explain why there is a more striking agreement in the variables in case of some factors and why there is no relationship between the two variables for other factors.
Keywords. alternative variable, confidence interval for π, regional disparity, factor analysis, test of independence.

1 Introduction

Generally, the term disparity means an inequality, disharmony or difference. In reference to regional problems, it means differences on the socio-economic level of particular territorial units, either administrative regions (e.g. regions, municipalities) or natural micro-regions.

The quantification of regional disparities falls into important spheres of a regional policy according to the law on the support of regional development [8]. The provision of the law says that central authorities should enable balanced development of a state, and should balance the differences among the levels of particular territorial units. This is the reason why it is necessary to create a suitable methodology that will enable measurement of the regions' current socio-economic level, and will consequently change this level after the implementation of the projects that focus on the improvement of the situation in a region.

The paper proposes the methodology of an economic situation evaluation in municipalities which was created by the research team of the Faculty of Economics in Liberec under the support of the Ministry for Regional Development project No. WD-30-07-1. The data about municipalities obtained by the Czech Statistical Office (CZSO) were used. After processing these data, a questionnaire survey in municipalities was carried out. The aim of this questionnaire survey was to learn mayors´ or other representatives´ of the municipal councils opinions on the situation in the locality. The second part of the paper follows the comparison of results of both research methodologies.

2 Approaches to quantification of regional disparities

Namely the Ministry for Regional Development, regions of the Czech Republic, the CZSO and other research workplaces deal with the measurement of regional disparities. The Ministry for Regional Development elaborated The Strategy of Regional Development of the Czech Republic in 2000. This document was innovated first in 2003 and then in 2006. The regions requiring a concentrated support from the state were defined within the framework of this document. These regions were classified as structurally depressed and economically weak regions, and regions with the above-average unemployment. A number of used indicators and their weights kept changing. Nine indicators were used in 2000; just four were used in 2006 – more in [7]. A determination of other problematic spheres is in the regions' hands. The methodologies of the regions are described in the programmes of development of particular territorial units in regions. The analyses of these methodologies [7] were conducted and it is possible to assert that these methodologies are totally incomparable in reference to the administrative determination of economically weak regions, number of used indicators, and calculated procedures.

The Czech Statistical Office elaborated publications called "Regional differences in demographic, social and economical development in 2000 – 2005" for every region. The differences were evaluated in the level of administrative regions of municipalities in four spheres: D – demographic environment, S – social environment, E – economic environment, I – infrastructure.

A uniform methodology based on the calculation of a synthetic indicator for each administrative district and for each of four observed spheres was used in all the regions except three regions (Liberecký, Středočeský and Vysočina). Fifty-two indicators were used in total; the weights were set up by the method of paired comparisons. An aggregate indicator was framed as a weighted average of evaluations represented by school grades (from 1 to 5) for particular indicators within the frame of a sphere. More information is in [1].

There are other research workplaces dealing with the problems of disparity evaluations in the Czech Republic, e.g. VŠB-TU Ostrava, University of West Bohemia in Plzeň, and Technical University of Liberec (TUL). The research team from VŠB-TU suggested a two-step system of monitoring and evaluating disparities. More is in [4].

D. Martinčík from the University of West Bohemia suggested the duodevigintigon, which consists of 3 basic spheres (macroeconomic output, growth potential, quality of life), for monitoring the socio-economic level in particular regions. Each sphere is divided into 7 indicators. The first and the last sphere indicators are common to their neighbouring spheres. The simple averages for particular spheres are calculated for the final comparison [5]. By this method it is possible to compare a region with another one and with an average level of the Czech Republic.

The basic idea of the research team of the Faculty of Economics at TUL was based on the hypothesis that no completely economically weak municipality exists. This team tried to detect basic factors of municipality development dynamics with the help of a multiple data analysis. The results of the analysis were used while creating the cartograms of natural regions that have to solve problems with low development dynamics in a certain problematic sphere. Furthermore, a municipalities' database was elaborated. Thanks to the database values particular factors can be defined and compared to average values. Moreover, a number of factors in which a municipality achieves development dynamics above or under the average can be specified.

3 Research Methodology

The statistical data analysis was realized in the first phase of the research from June 2007 to March 2008. The base for the analysis was 26 indicators used in Strategies of Regional Development. A factor analysis seemed to be a suitable method for the data processing. It allows a reduction of the originally large number of indicators to a smaller one without losing too much information.

The factor analysis is a method whose main aim is to reduce a set of observable variables in terms of a small number of latent factors. The essential assumption of this method is that a number of unobserved latent variables (factors) that explain the correlations among observed variables exists. A detailed description of this method is in [2].

Steps of the factor analysis:

1. Preparation of the data set for the analysis – Data of 26 indicators for 6,240 municipalities were pasted into the MS Excel 2007 file.
2. The data were imported into the statistical programme STATGRAPHICS CENTURION XV.
3. Making decisions on the number of factors – The factors were extracted by the principal component method. The number of them is based on the eigenvalue of a matrix criterion. The factors with the eigenvalue greater than 1 are considered as statistically significant. Furthermore, the factors were required to explain at least 60% of variation because this is a limit that it is to be considered as satisfactory in social sciences.

4. Rotation of factors – The reason for realizing it is to simplify the interpretation of data. The Varimax method was chosen.
5. Evaluation of factor weights – A simple rule exists here: the greater the magnitude of the factor weights the more important the factor is in the interpretation of the factor matrix. The squared weight represents a part of the total variation of the variable explained by a factor.
6. Identification of factors – It belongs to the most difficult steps of the analysis. It results from the factor weights. Indicators with larger weights are considered to be more important with the reference to their influence on the economic weakness or strength of municipalities, and it is necessary to name them correctly.
7. Estimation of the factor score – It means values of a factor for particular municipalities. These values can be transmitted into maps and allow determination of areas with similar properties.

The technique and results of the analysis were presented in [7]. Originally, we worked with 26 indicators. This number was reduced to 8 factors which can explain nearly 64% of the total variation of the original data. The factors are: unemployment, domicile attractiveness, population density and allocation, age structure, civil amenities, economy structure, sustainable development, economic activity.

In the second phase of the research, the researchers carried out the questionnaire survey in all the municipalities of the Czech Republic. The questionnaire was tested on a sample of 100 municipalities in the Liberec region – see [3]. After finishing the factor analysis, the questionnaire was revised because its structure was required to copy 8 factors in relation to their meaning. At the same time, the knowledge from the pilot survey concerning the questions formulation and their understanding were included to the questionnaire. Most of the respondents were municipality mayors [6]. The questionnaire was sent in a form of a book in June, 2008. 1,357 questionnaires came back from the total of 6,249 sent ones. The rate of return was less than 22%. Not every questionnaire could be analysed because of formal mistakes. Just 1,078 questionnaires were analysed. This number had to be reduced to 994 because of other mistakes.

4 The comparison of results of economic level evaluation methodology and results from the questionnaire survey in municipalities

First, it must be stressed again that one of the main aims of the questionnaire survey in municipalities was to discover how mayors (or other members of the municipal councils) perceived the actual economic situation in their municipalities, and compare their opinions with the results provided by the new methodology of economic level evaluation in municipalities.

We selected the questions from the questionnaire in which the respondents evaluated each of the 8 factors named above individually. These questions were:

- Do you consider the unemployment to be a problem in your municipality?
- Do you consider your municipality to be attractive for permanent living?
- Do you consider a settlement allocation in your municipality to be a problem?
- Do you consider the age structure in your municipality to be a problem?
- Do you consider the state of civil amenities in your municipality to be a problem?
- Do you consider the economic structure in your municipality to be a problem?
- Do you consider the state of the environment in your municipality to be a problem?
- Do you consider the economic activity in your municipality to be a problem?

Possible answers were “yes” or “no”.

In order to find out whether both variables – results from the methodology and opinions of municipality mayors – are mutually dependent, we used the chi-squared test of independence for each factor. We tested the null hypothesis that there is no relationship between the two variables, i.e. the results from the methodology and mayors´ evaluation of a given factor are independent. The alternative hypothesis states that there is a relationship. The test statistic known as Pearson´s χ^2 was modified using Yates´ correction for continuity. If the null hypothesis was rejected, we computed the Pearson´s R. This statistic measures the degree of association between the two variables. All the null hypotheses were tested on the 0.05 level. We used a statistical programme STATGRAPHICS CENTURION XV for all the computations.

4.1. Factor 1 - unemployment

The data about the unemployment are displayed in the contingency table – see table 1.

Is unemployment in the municipality a problem? (Methodology)	Do you consider unemployment as a problem in your municipality? (Mayors)		Total
	NO	YES	
NO	393	121	514
YES	211	269	480
Total	604	390	994

Table 1: Frequency Table for results of the methodology and mayors´ opinions on unemployment

The table 1 shows that the agreement between the methodology and mayors occurs in 662 cases. It contains 393 municipalities where unemployment is a problem, and the methodology found them to be problematic with regard to unemployment too. The rest of the 269 municipalities do not have a problem with unemployment, and the methodology identified them to be without problems with unemployment as well.

We tested the null hypothesis that there is no relationship between the results obtained by the new methodology and mayors´ opinions on unemployment. The P-value for this test is 0, so the null hypothesis is rejected.

The Pearson´s R was computed. Its value is 0.333. The association between the two variables is quite weak and positive. It means that the agreement of the both "opinions" occurs more often than disagreement.

The results of the test were supplemented with the 95% confidence interval for π (a proportion of agreement of the two variables). The 95% confidence interval shows that the values of π fall between 63.6 and 69.5%.

4.2. Factor 2 – domicile attractiveness

757 mayors answered the question about attractiveness of their municipality positively. But when we focus on monitoring the number of agreement and disagreement, we can find out that there are 551 cases with agreement between the methodology and mayors. There are 68 mayors who consider their municipalities to be unattractive for permanent living, and the methodology gives the same result. The rest of 483 mayors think that their municipalities are attractive for permanent living, and the methodology makes the same conclusion.

The test of independence and the computed P-value 0.041 showed that there is dependence between the results of the methodology and mayors´ opinions. The value of Pearson´s R is quite surprising because it is negative. The value is – 0.067. It means that there are more cases in which the methodology says no and mayors yes, and vice versa.

We supplement the test of independence with the 95% confidence interval for π (a proportion of agreement of the two variables) again. The proportion of agreement is between 54.2 and 58.5 % at the 95% confidence level.

4.3. Factor 3 – population density and allocation

The questionnaire survey provided the information that 181 mayors consider settlement allocation to be a problem, and 813 mayors have a contrary opinion. Another information resulting from the survey shows that 403 mayors identify this factor to be unproblematic, and at the same time the methodology makes the same conclusion. Problems with the settlement allocation are in 109 municipalities according to the results of the methodology and mayors´ opinions.

The test of independence led to a rejection of the null hypothesis that supposes independence of the two variables. So, the test proved that there is dependence between the results of methodology and mayors´ opinions. The P-value is 0.021. The value of Pearson´s R (0.076) informs us that the dependence between the two variables is very weak and positive. It means that there are more cases in which both the "opinions" say yes together, and vice versa.

We can add the information about the proportion of an agreement of the two variables at the 95% confidence level again. The proportion is between 48.3 and 54.6%.

4.4. Factor 4 – age structure

The age structure is considered to be problematic in 431 municipalities and unproblematic in 563 ones according to mayors´ opinions. 594 cases were recorded in which the methodology and mayors have the same view of the factor of age structure. There are 368 municipalities without troubles with age structure. 226 municipalities identified this factor as problematic together with the methodology.

The test of independence showed that there exists the dependence between both views of this factor because P-value equals 0. The value of Pearson´s R is 0.179. So, the dependence between the two variables is weak and positive.

The proportion of the same opinions of the methodology and mayors is between 56.7 and 62.9% at the 95% confidence level.

4.5. Factor 5 – civil amenities

611 mayors answered the question about the state of civil amenities in their municipality "yes" and 383 mayors said "no". We found 239 cases in which the methodology and mayors identified the state of civil amenities in a municipality as unproblematic. 347 municipalities have problems with this factor and the methodology came to the same conclusion.

The test of independence leads to a rejection of the null hypothesis because the P-value is 0. The dependence between the two variables is weak and positive. This information is provided by Pearson´s R that equals 0.187.

The 95% confidence interval for proportion of an agreement of methodology and mayors is 55,8 to 62.0%.

4.6. Factor 6 – economic structure

The economic structure seems to be a problem in 407 municipalities. The rest of the municipalities (587) do not have any problem. There are 214 cases in which the methodology and mayors consider the economic structure to be unproblematic. 306 municipalities have a problem with this factor and the methodology confirms that.

We accept the null hypothesis in the test of independence because P-value is greater than given significance level α. P-value equals 0.163. So, the dependence between the two variables was not proved by the test.

We can notice that the proportion of the same opinions of methodology and mayors is between 49.1 and 55.4% at the 95% confidence level.

4.7. Factor 7 – sustainable development

226 mayors feel the state of environment in their municipalities to be a problem while 768 mayors are satisfied with the environment in their municipalities. The methodology and mayors identified this factor as problematic in 127 municipalities. The methodology and mayors do not find this factor problematic in 355 municipalities.

The null hypothesis in the test of independence is not rejected at 0.05 level because P-value is greater than it. P-value equals 0.572. So, the result is that there is no relationship between the two variables. It means that mayors´ opinion is not connected with results of the methodology.

We add the information about the 95% confidence interval for π (a proportion of agreement of the two variables) again. π is between 45.4 and 51.7% at the 95% confidence level.

4.8. Factor 8 – economic activity

We found out that the economic activity is a problem for 509 mayors, and 485 mayors considered this factor to be unproblematic. At the same time we learnt that this factor was identified as unproblematic by mayors and the methodology in 198 municipalities. 310 municipalities do not have a problem with the economic activity according to the results of the methodology and opinions of mayors.

The test of independence leads to an acceptance of the null hypothesis because of the large P-value (0.623). So, the test is evidence that the relationship between evaluation of this factor by the methodology and mayors do not exist. The last note is related to the 95% confidence interval for π (a proportion of agreement of the two variables). The proportion is between 47.9 and 54.3%.

5 Conclusion

When summarizing the results of our research we can say that the dependence between the results of the methodology and mayors´ evaluation was proved at factors No. 1 to 5, i.e. unemployment, domicile attractiveness, population density and allocation, age structure, civil amenities. All the dependences were positive and weak or quite weak. The exception is factor 2 – domicile attractiveness. The dependence of the two variables was negative at this factor. The mayors´ evaluation and evaluation of the methodology are independent at factors 6, 7 and 8, i.e. economic structure, sustainable development and economic activity.

It is necessary to find reasons why some “opinions” are dependent and others are independent. These causes can be divided into two spheres. The first cause can be an interpretation of particular factors. As written above (chapter 3) the interpretation and the name of the factors are the most difficult tasks of the factor analysis. The factor analysis gives the information which indicators have a dominant weight in the given factor. The evaluation of the methodology is based on statistical data collected in municipalities. On the other hand, the municipality mayors can express the same term in a different way. This idea is supported by the results of our analysis. The positive dependence was proved at the factor of unemployment, civil amenities, and age structure. These terms are generally well known. On the other hand, the independence was accepted at the factors of economic structure, sustainable development and economic activity. Knowledge of professional terms plays an important role here. It is necessary to realize that municipality mayors are mostly not specialists in economics, demography, environment etc. We have to take into account the fact that the questionnaire is quite extensive as well. A few first questions could be answered more carefully than the last ones.

The second cause of a different view of the factors can be found in the emotive sphere. We suppose this influence at the evaluation of factor 2 – domicile attractiveness. Though, the official data show low attractiveness for permanent living, municipality mayors can inhibit this fact by their relationship to the surrounding countryside etc. and thus they may feel the municipality is attractive for living.

The analysis shows that we should modify the questionnaire. We will append the interpretation of the factors and exclude some less important questions. The questionnaire will not be so long, and the answers could be more exact. We are planning to realize the new survey in 1% municipalities which will be selected randomly.

Bibliography

[1] HARTMAN, M. et al. *Regionální rozdíly v demografickém, sociálním a ekonomickém vývoji Královéhradeckého kraje v letech 2000 až 2005.* Hradec Králové: ČSÚ – krajská správa, 2006. ISBN 80-250-1371-5.

[2] HEBÁK, P., et al. *Vícerozměrné statistické metody (3).* 2nd ed. Praha: INFORMATORIUM, 2007. ISBN 978-80-7333-001-9.

[3] JÁČ, I. Vyhodnocení dotazníkového projektu v rámci inovačního řešení disparit. *E+M Ekonomie a Management,* 2008, roč. 11, č. 2, pp. 31-40. ISSN 1212-3609.

[4] KUTSCHERAUER, A. Disparity a jejich vliv na územní rozvoj země. *Working Papers REGIONÁLNÍ DISPARITY,* 2009, č. 4, pp. 5-14. ISSN 1802-9450.

[5] MARTINČÍK, D. Ekonomicko-sociální úroveň krajů – komplexní srovnávací analýza. *E+M Ekonomie a Management,* 2008, roč. 11, č. 1, pp. 14-25. ISSN 1212-3609.

[6] PRSKAVCOVÁ, M. and ŘEHOŘOVÁ, M. Metodika šetření hospodářské výkonnosti obcí České republiky. *E+M Ekonomie a Management,* 2008, roč. 11, č. 4, pp. 77-83. ISSN 1212-3609.

[7] RYDVALOVÁ, P., ŽIŽKA, M. *Konkurenceschopnost a jedinečnost obce.* 1st ed. Liberec: Technická univerzita v Liberci, 2008. ISBN 978-80-7372-423-8.

[8] Zákon č. 240/2000 Sb., o podpoře regionálního rozvoje