QUANTITATIVE METHODS IN ECONOMICS

Multiple Criteria Decision Making XIX

Proceedings of the International Scientific Conference
23rd May - 25th May 2018
Trenčianske Teplice, Slovakia

Letra Edu

The Slovak Society for Operations Research
Department of Operations Research and Econometrics
Faculty of Economic Informatics, University of Economics in Bratislava

Proceedings of the International Scientific Conference

QUANTITATIVE METHODS IN ECONOMICS

Multiple Criteria Decision Making XIX

23rd May – 25th May 2018
Trenčianske Teplice, Slovakia

Programme Committee:
Brezina Ivan, University of Economics, Bratislava
Ferreira Manuel A. M., University Institute of Lisbon
Fiala Petr, University of Economics, Prague
Groshek Gerald, University of Redlands
Jablonský Josef, University of Economics, Prague
Lukáčik Martin, University of Economics, Bratislava
Mlynarovič Vladimír, Comenius University, Bratislava
Palúch Stanislav, University of Žilina, Žilina
Pekár Juraj, University of Economics, Bratislava
Plevný Miroslav, University of West Bohemia
Turnovec František, Charles University, Prague
Vidović Milorad, University of Belgrade

Organizing Committee:
Furková Andrea
Gežík Pavel
König Brian
Lukáčiková Adriana
Reiff Marian
Surmanová Kvetoslava

Referees:
Borovička Adam, University of Economics, Prague
Brezina Ivan, University of Economics, Bratislava
Brezina Ivan Jr., Pan-European University, Bratislava
Brožová Helena, Czech University of Life Sciences, Prague
Čičková Zuzana, University of Economics, Bratislava
Domonkos Tomáš, Institute of Economic Research SAS
Fábry Ján, University of Economics, Prague
Ferreira Manuel A. M., University Institute of Lisbon
Fiala Petr, University of Economics, Prague
Figurová Dana, University of Economics, Bratislava
Furková Andrea, University of Economics, Bratislava
Gežík Pavel, University of Economics, Bratislava
Chocholatá Michaela, University of Economics, Bratislava
Jánošíková Ľudmila, , University of Žilina, Žilina
Jablonský Josef, University of Economics, Prague
Janáček Jaroslav, University of Žilina, Žilina
Jánošíková Ľudmila, University of Žilina, Žilina
König Brian, University of Economics, Bratislava
Kuncová Martina, University of Economics, Prague
Kvet Marek, University of Žilina, Žilina
Lukáčik Martin, University of Economics, Bratislava
Lukáčiková Adriana, University of Economics, Bratislava
Majer Tomáš, University of Žilina, Žilina
Marek Luboš, University of Economics, Prague
Mitková Veronika, Comenius University in Bratislava
Mlynarovič Vladimír, Comenius University in Bratislava
Ostrihoň Filip, Institute of Economic Research SAS
Palúch Stanislav, University of Žilina, Žilina
Pekár Juraj, University of Economics, Bratislava
Peško Štefan, University of Žilina, Žilina
Reiff Marian, University of Economics, Bratislava
Sequeira Lopez Allan Jose, University of Economics, Bratislava
Skočdopolová Veronika, University of Economics, Prague
Surmanová Kveta, University of Economics, Bratislava
Szomolányi Karol, University of Economics, Bratislava
Zouhar Jan, University of Economics, Prague

23rd May - 25th May 2018
Trenčianske Teplice, Slovakia

Technical Editor: Marian Reiff, Pavel Gežík, University of Economics, Bratislava
Web Editor: Martin Lukáčik, University of Economics, Bratislava
Web: http://www.fhi.sk/en/katedry-fakulty/kove/ssov/papers/

Publisher: Letra Edu, s. r. o.
ISBN 978-80-89962-07-5 (book)
ISBN 978-80-89962-08-2 (pdf)
Number of copies: 90

CONTENT

CLUSTER ANALYSIS OF EUROPEAN ECONOMIES IN REGARD TO EUROPE 2020 CLIMATE CHANGE AND ENERGY OBJECTIVES

Adam P. Balcerzak, Nicolaus Copernicus University in Toruń

Abstract
The political and economic actions aiming at reducing the negative influence of economies on climate change have shaped European policies for last few years. The principals of reducing greenhouse gasses emissions have been included in long term European strategies and must be implemented at national level. In this context, the aim of the article is to conduct analysis of similarities between European member states in regard to fulfilling the climate change and energy objectives, which were declared in Europe 2020 plan. In the research five criteria provided by the European Commission were applied, which justifies application of taxonomic tools in the study. To obtain the objectives of the article the Ward's clustering method was used, which enabled to group the countries into homogeneous clusters. The research was done for the years 2005 and 2015 and was based on data form Eurostat. The study showed significant disparities between EU-15 and new member states. Additionally, the differences among EU-15 have been confirmed. The conducted research showed high stability of these disparities in time. These factors should be taken into consideration in the political decisions affecting situation of European countries in regard to energy security.

Keywords: *cluster analysis, climate change, energy objectives, Europe 2020, European Union, Ward's method*

JEL Classification: C38, F64
AMS Classification: 91C20

1 INTRODUCTION

Recent years have confirmed that building energy sustainability has become one of the most important roles of governments, which is crucial for national security (Czech, 2017). In the same time the governments responsibility for supporting green economy development has become significant criteria influencing actions in this sphere (Szyja, 2016). The European policies have been formed to high extent by political actions, which aim at reducing the negative influence of economies on climate change. The purpose of reducing greenhouse gasses emissions has been included in long term European strategies such as Europe 2020 and must be implemented at national level. This factor can become a significant economic burden for some lower developed European member states and can threaten their energy sustainability. Therefore, proposing unified European policy, which would lead to decreasing the negative consequences of energy consumption, should be based on structural characteristics of European economies. The significant differences between national energy systems should be taken into consideration here. In this contexts, the aim of the article is to conduct analysis of similarities between European member states in regard to fulfilling the climate change and energy objectives, which were declared in Europe 2020 plan.

In the research the following hypothesis was given: The European Union economies are characterized with high and stable in time differences in regard to their possibilities to reach climate change and energy objectives taken in Europe 2020 plan. In order to measure the results of the countries in reaching the objectives five criteria have been proposed by the European Commission. Therefore, in the article the Ward's clustering method was used in order to group the countries into homogeneous clusters. The study was conducted for the years 2005 and 2015 with application of Eurostat data.

2 CLUSTER ANALYSIS AS A TOOL FOR ANALYSING CLIMATE CHANGE AND ENERGY OBJECTIVES

The phenomenon of environmental sustainability is a complex and multiple-criteria research problem. The same can be said about climate change and energy objectives, which has been also stressed by the European Commission in the Europe 2020 plan (see Balcerzak, 2015; Stanickova, 2017). As a result the Commission suggested to apply five criteria, which should be used for its assessment. Therefore, a taxonomic or multiple-criteria analysis tools provide appropriate approach for the research in this regard.

Multidimensional character is a property, which is characteristic for the majority of socio-economic phenomena (Pietrzak et al., 2014; Sekuła and Śmiechowicz, 2016; Walczak and Pietrzak, 2016; Chrupała-Pniak et al., 2017; Pietrzak, 2017; Meluzin et al., 2017; 2018). This factor is the main reason for high popularity and growing economic applications of taxonomic and multiple-criteria decision-making (MCDM) or multiple-criteria decision analysis (MCDA) tools (Keshavarz Ghorabaee et al., 2016; Mardani et al., 2016; Zavadskas et al., 2016; Pietrzak, 2016; 2018; Krylovas et al., 2016; Balcerzak, 2016; Balcerzak et al. 2017; Pietrzak and Ziemkiewicz, 2017; Stankeviciene et al., 2017; Pieloch-Babiarz, 2017; Cano et al., 2017). These methods are especially useful in the case of decision making process which is based on multiple – often conflicting – criteria, as they enable to find the best alternative, or when one needs to order or rank some object, which are characterised by many factors (Trinkūnienė et al., 2017; Mikušová 2017; Vavrek et al., 2017; Żelazny and Pietrucha, 2017; Balcerzak and Pietrzak, 2017; Zygmunt, 2017; Kruk and Waśniewska, 2017; Pietrzak et al., 2017a, 2017b; Shpolianska et al., 2017).

In order to conduct comparison of economic objects in regard to their similarity to each other one can apply cluster analysis. The use of cluster analysis allows classification of objects into relatively homogeneous groups (clusters) based on a designated similarity between objects. In that cases the metrics are used to determine the similarity, where the most commonly used distance measure is the Euclidean metric (Trąpczyński et al., 2016; Małkowska and Głuszak, 2016; Mačerinskienė and Aleknavičiūtė, 2017). On the other hand, the classification of economic objects is usually carried out using the agglomeration methods. In economics, the Ward's method (Ward, 1963) is commonly used due to the fact that a lot of economic phenomena tend to generate many clusters with a small number of objects. This method is based on the analysis of variance, where clusters are determined on the basis of the criterion of minimizing the sum of squares of distances between objects. The classification of economic objects with application of the agglomeration method enable to create a hierarchy of them, where the criterion for object selection is the similarity previously assigned. The hierarchy is started by all objects that are combined into groups according to the accepted similarity criterion. In successive levels of the hierarchy, the objects are joined in increasing groups in terms of numbers. Combining of objects into groups is performed until there is only one group to which all objects belong. At the same time, this group is the highest level in the hierarchy. The created hierarchy is presented in the form of a dendrogram, where groups can be identified at every level of the hierarchy and the economic objects can be assigned to the appropriate groups (Reiff et al. 2016).

3 EMPIRICAL RESEARCH

In accordance with the adopted objective of the article, the research on similarities between countries in regard to obtaining the climate change and energy objectives was conducted for all 28 European Union member states in 2005 and 2015. As part of the cluster analysis, the Ward's method was used, which allowed the countries to be classified into groups. The research was based on the five diagnostic variables suggested by the European Commission, where the specific data is provided by Eurostat as a subset for Euro 2020 headline indicators: http://ec.europa.eu/eurostat/data/database):

x_1 – Greenhouse gas emissions, base year 1990, Index (1990 = 100);

x_2 – Share of renewable energy in gross final energy consumption, percentage;

x_3 – Primary energy consumption per capita, million tonnes of oil equivalent (TOE) per capita;

x_4 – Final energy consumption per capita, million tonnes of oil equivalent (TOE) per capita;

x_5 – Greenhouse gas emissions in ESD sectors per capita, million tonnes CO2 equivalent per capita.

Four of the above pointed variables should be considered as de-stimulants (x_1,x_3, x_4, x_5) and one as stimulant x_2. In the preliminary stage of the research the data was standardised with application of classic standardisation formula.

The application of the Ward's method allowed to establish a hierarchical structure for the countries. Dendograms presenting this structure for the year 2005 and 2015 are given in figure 1. The analysis enabled to identify clusters of the EU member states, in which countries are similar to each other due to applied diagnostic variables. On the basis of the dendrograms received, the author has chosen in 2005 a division of countries into seven groups and in 2015 into eight groups. The large number of accepted groups indicates the existence of significant differences between countries as to the issue of reducing the negative influence of economies on climate change.

Hierarchical structure for the year 2015 Hierarchical structure for the year 2005

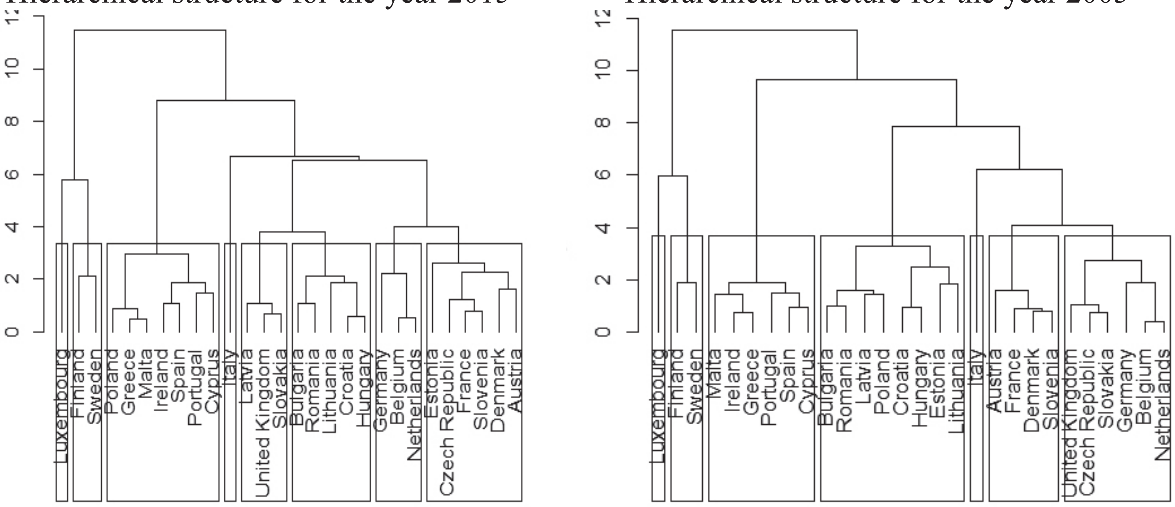


Figure 1: Dendograms for EU countries in the years 2015 and 2005

Source: own estimation based on Eurostat data with application of R-Cran package.

The determined classification of the countries was also presented in table 1. In the year 2005 single clusters were formed by two countries Luxemburg (cluster one) and Italy (cluster five).

These two countries were classified as much different from other economies. Second cluster was formed by two Scandinavian economies: Finland and Sweden that were able to obtain very good results in the sphere of reaching the climate change and energy objectives. Third cluster was made up of Southern European countries: Greece, Portugal, Spain, Malta and Cyprus. Ireland was also assigned to the third cluster. In turn, cluster fourth consisted of Central and Eastern European economies excluding Czech Republic, Slovakia and Slovenia. So in the fourth cluster one could find: Poland, Bulgaria, Romania, Hungary, Estonia, Lithuania, Latvia and Croatia. On the other hand, clusters sixth and seventh form countries with a large share of industry in the economy, which translates into much higher negative ecological impact compared with countries from cluster three and four. Cluster six make up such countries as: Austria, France, Denmark and Slovenia, and cluster seven: Great Britain, Germany, the Netherlands, Belgium, the Czech Republic and Slovakia.

During the ten years period under research some positive changes in regard to the analysed phenomenon were obtained. However, in spite of these processes in the year 2015 quite similar structure of clusters was obtained. Such countries as Luxemburg and Italy continued to form separate clusters. Therefore, they still significantly differed from other European economies. Finland and Sweden were invariably positive models of developing the country's economy with an emphasis on environmental protection. In the case of Southern Europe a similar situation to the one in 2005 occurred in 2015. The same countries created an almost identical cluster 3. In this case, in 2015 in the cluster three one could also find Poland, which left cluster four. Significant changes took place for clusters fourth, fifth, sixth and seventh. Cluster seven was separated by Germany, Belgium and the Netherlands, which were assigned to the new cluster eight. In cluster seven, as in the year 2005 one could find Great Britain and Slovakia but additionally Latvia was moved here from cluster fourth. However, Czech Republic was moved from cluster seventh to cluster sixth. Cluster sixth was again additionally formed by Austria, France, Denmark and Slovenia and Estonia (cluster four in 2005). Cluster fourth formed mostly the same countries as in 2005. However, as a result of the migration of individual countries to other clusters, cluster fourth in 2015 formed: Bulgaria, Romania, Hungary, Lithuania and Croatia.

Table 1: European Union member states and obtained clusters for the year 2005 and 2015

2005				2015			
Country	**cluster**	**Country**	**cluster**	**Country**	**cluster**	**Country**	**cluster**
Luxembourg	1	Poland	4	Luxembourg	1	Poland	3
Finland	2	Hungary	4	Finland	2	Hungary	4
Sweden	2	Croatia	4	Sweden	2	Croatia	4
Greece	3	Italy	5	Greece	3	Italy	5
Spain	3	Austria	6	Spain	3	Austria	6
Portugal	3	France	6	Portugal	3	France	6
Malta	3	Denmark	6	Malta	3	Denmark	6
Cyprus	3	Slovenia	6	Cyprus	3	Slovenia	6
Ireland	3	United Kingdom	7	Ireland	3	United Kingdom	7
Bulgaria	4	Czech Republic	7	Bulgaria	4	Czech Republic	6
Romania	4	Slovakia	7	Romania	4	Slovakia	7
Latvia	4	Germany	7	Latvia	7	Germany	8
Lithuania	4	Belgium	7	Lithuania	4	Belgium	8
Estonia	4	Netherlands	7	Estonia	6	Netherlands	8

Source: own estimation based on Eurostat data.

4 HYPOTHESIS VERIFICATION AND POLICY IMPLICATIONS

The most important result of the conducted research is the confirmation of significant differences between countries, which can be seen in the high number of obtained clusters in the first and last year of the research. The differences do not only relate to the disparities between “new” and “old” EU but also to EU-15. Additionally, the obtained clusters can be considered as quite stable in time in spite of the relatively long period under research. This result can confirm that in spite of the fact that the year 2015 can be considered as half-way point in implementation of the Europe 2020 strategy it was not possible to diminish the disparities between EU economies. This factor can be considered as confirmation of the proposed empirical hypothesis of the research.

From the policy perspective this research indicates that changing the relation between the national energetic balance and its environmental influence is a long term process, which should be taken into consideration in the case of international agreements and obligations in this regard. This factor can be especially important for Eastern and Central European economies that still face the challenge of modernizing their economies, also in regard to national energetic systems.

5 CONCLUSIONS

In the current paper the problem of reaching the Europe 2020 objectives in regard to climate change and energy objectives was taken into evaluation. The main purpose of the paper was to provide analysis of similarities between European member states in regard to fulfilling the climate change and energy objectives. The problem was analysed with application of Ward’s clustering method. The research was done for the year 2005 and 2015.

The conducted research provides important arguments in favour of the hypothesis of the research, which states that the European Union economies are characterized with high and stable in time differences in regard to their possibilities to reach climate change and energy objectives taken in Europe 2020 plan.

Form the policy perspective the conducted research confirms that in the half-way point of Europe 2020 implementation the fundamental aims of the strategies relating to the challenge of reducing the disparities between countries has not been reached in the sphere of climate change and energy objectives.

References

[1] Balcerzak, A. P. (2015). Europe 2020 Strategy and Structural Diversity Between Old and New Member States. Application of Zero-unitarization Method for Dynamic Analysis in the Years 2004-2013. *Economics & Sociology*, 8(2), pp. 190-210.

[2] Balcerzak, A. P. (2016). Multiple-criteria Evaluation of Quality of Human Capital in the European Union Countries. *Economics & Sociology*, 9(2), pp. 11-27.

[3] Balcerzak, A. P., Kliestik, T., Streimikiene, D. and Smrčka L. (2017). Non-parametric approach to measuring the efficiency of banking sectors in European Union Countries. *Acta Polytechnica Hungarica*, 14(7), pp. 51-70

[4] Balcerzak, A. P. and Pietrzak, M. B. (2017). Digital Economy in Visegrad Coutnries. Multiple-criteria decision analysis at Regional Level in the Years 2012 and 2015. *Journal of Competitiveness*, 9(2), pp. 5-18.

[5] Cano, J., Campo, E. and Gómez-Montoya, R. (2017). International Market Selection Using Fuzzy Weighing and Monte Carlo Simulation. *Polish Journal of Management Studies*, 16(2), pp. 40-50.

[6] Chrupała-Pniak, M., Grabowski, D. and Sulimowska-Formowicz, M. (2017). Trust in Effective International Business Cooperation: Mediating Effect of Work Engagement. *Entrepreneurial Business and Economics Review,* 5(2), pp. 27-50,

[7] Czech, A. (2017). Economic dimension of Polish energy security. *Oeconomia Copernicana*, *8*(3), pp. 383-399.

[8] Keshavarz Ghorabaee, M., Zavadskas, E. K., Amiri, M. and Antucheviciene, J. (2016). Evaluation by an Area-based Method of Ranking Interval Type-2 Fuzzy Sets (EAMRIT-2F) for Multi-criteria Group Decision-making. *Transformations in Business & Economics*, 15(3)(39), pp. 76-95.

[9] Krylovas, A., Zavadskas, K. E. and Kosareva, N. (2016). Multiple criteria decision-making KEMIRA-M method for solution of location alternatives. *Economic Research-Ekonomska Istraživanja*, 29(1), pp. 50-65.

[10] Kruk, H. and Waśniewska, A. (2017). Application of the Perkal method for assessing competitiveness of the countries of Central and Eastern Europe. *Oeconomia Copernicana*, 8(3), pp. 337-352.

[11] Mačerinskienė, I. and Aleknavičiūtė, R. (2017). National intellectual capital influence on economic growth in the European Union countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(4), pp. 573-592.

[12] Małkowska, A. and Głuszak, M. (2016). Pro-investment local policies in the area of real estate economics – similarities and differences in the strategies used by communes. *Oeconomia Copernicana*, *7*(2), pp. 269-283.

[13] Mardani, A., Jusoh, A., Zavadskas, E. K., Kazemilari, M., Ungku, N.U.A. and Khalifah, Z. (2016). Application of Multiple Criteria Decision Making Techniques in Tourism and Hospitality Industry: a Systematic Review. *Transformations in Business & Economics*, 15(1)(37), pp. 192-213.

[14] Meluzin, T., Pietrzak, M. B., Balcerzak, A. P., Zinecker, M., Doubravský, K. and Dohnal, M. (2017). Rumours Related to Political Instability and their Impact on IPOs: The Use of Qualitative Modeling with Incomplete Knowledge. *Polish Journal of Management Studies*, 16(2), pp. 171-187.

[15] Meluzin, T., Zinecker, M., Balcerzak, A. P., Doubravský, K., Pietrzak, M. B. and Dohnal, M. (2018), The timing of initial public offerings – non-numerical model based on qualitative trends. *Journal of Business Economics and Management,* 19(1), pp. 109-125.

[16] Mikušová, M. (2017). To be or not to be a business responsible for sustainable development? Survey from small Czech businesses. *Economic Research-Ekonomska Istraživanja*, 30(1), pp. 1318-1338.

[17] Pieloch-Babiarz, A. (2017). Determinants of payout policy and investment attractiveness of companies listed on the Warsaw Stock Exchange. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(4), pp. 675-691.

[18] Pietrzak, M. B. (2016). The Problem of the Inclusion of Spatial Dependence Within the TOPSIS Method. *Montenegrin Journal of Economics*, 12(3), pp. 69-86.

[19] Pietrzak, M. B. (2017). Structural Equation Modeling of Regional Economic Development in Polish Voivodeships in the Years 2010-2014. In: T. Loster & T. Pavelka (Eds.). *The 11th International Days of Statistics and Economics. Conference Proceedings. September 8-10, 2016*. Prague: Libuse Macakova, Melandrium, pp. 1214-1222.

[20] Pietrzak, M. B. (2018). The composition of agricultural macroregions in Eastern Poland - an empirical example of the Aggregation Problem. In: M. Papież and S. Śmiech (Eds.). *The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings*. Cracow: Foundation of the Cracow University of Economics.

[21] Pietrzak, M. B., Balcerzak, A. P., Gajdos, A. and Arendt, Ł (2017a). Entrepreneurial environment at regional level: the case of Polish path towards sustainable socio-economic development. *Entrepreneurship and Sustainability Issues, 5*(2)*,* pp. 190-203.

[22] Pietrzak, M. B, Gajdos, A. and Arendt Ł. (2017b). Assessment of Socioeconomic Development in Poland in the years 2010-2015. In: T. Kliestik (Ed.). *17th International Scientific Conference Globalization and Its Socio-Economic Consequences. University of Zilina, The Faculty of Operation and Economics of Transport and Communication, Department of Economics, 4th – 5th October 2017. (Part IV.).* Zilina: University of Zilina, pp. 1951-1956.

[23] Pietrzak M. B., Wilk J., Kossowski T. and Bivand R. (2014). The Identification of Spatial Dependence in the Analysis of Regional Economic Development - Join-count Test Application. In: M. Papież & S. Śmiech (Eds.). *Proceedings of the 8th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena.* Cracow: Foundation of the Cracow University of Economics, pp. 135-144.

[24] Pietrzak, M. B. and Ziemkiewicz, B. (2017). The use of random fields in the Modifiable Areal Unit Problem. In: M. Papież and S. Śmiech (Eds.). *The 11th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings*. Cracow: Foundation of the Cracow University of Economics, pp. 319-328.

[25] Reiff, M., Surmanová, K., Balcerzak, A.P. and Pietrzak, M.B. (2016). Multiple Criteria Analysis of European Union Agriculture. *Journal of International Studies*, 9(3), pp. 62-74.

[26] Sekuła, A. and Śmiechowicz, J. (2016). Systems of general grants for local governments in selected EU countries against the background of the general theory of fiscal policy. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 11(4), pp. 711-734.

[27] Shpolianska, I., Dolzhenko, A. and Glushenko, S. (2017). Fuzzy models of web services selection in the development of service-oriented information systems for small and medium enterprises. *Ekonomicko-manazerske spectrum*, 11(2), pp. 64-75.

[28] Stanickova, M. (2017). Can the implementation of the Europe 2020 Strategy goals be efficient? The challenge for achieving social equality in the European Union. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(3), pp. 383-398.

[29] Stankeviciene, J., Kraujaliene, L. and Vaiciukeviciute, A. (2017). Assessment of technology transfer office performance for value creation in higher education institutions. *Journal of Business Economics and Management*, 18(6), pp. 1063-1081.

[30] Szyja, P. (2016). The role of the state in creating green economy. *Oeconomia Copernicana*, 7(2), pp. 207-222.

[31] Trąpczyński, P., Jankowska, B., Dzikowska, M. and Gorynia, M. (2016). Identification of Linkages between the Competitive Potential and Competitive Position of SMEs Related to their Internationalization Patterns Shortly after the Economic Crisis. *Entrepreneurial Business and Economics Review*, 4(4), pp. 29-50.

[32] Trinkūnienė, E., Podvezko, V., Zavadskas, K. E., Jokšienė, I., Vinogradova, I. and Trinkūnas, V. (2017). Evaluation of quality assurance in contractor contracts by multi-attribute decision-making methods. *Economic Research-Ekonomska Istraživanja*, 30(1), pp. 1152-1180.

[33] Vavrek R., Adamisin P. and Kotulic R. (2017). Multi-criteria evaluation of municipalities in Slovakia – case study in selected districts. *Polish Journal of Management Studies*, 16(2), 290-301.

[34] Walczak D. and Pietrzak, M. B. (2016). Analysis of Agrarian Structure in Poland in 1921 and 2002 based on the Example of Selected Districts. In: M. H. Bilgin, H. Danis, E. Demir & U. Can (Eds.). *Business Challenges in the Changing Economic Landscape - Vol. 1. Proceedings of the 14th Eurasia Business and Economics Society*. Springer International Publishing, pp. 461-472.

[35] Ward, J. H., Jr. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58, pp. 236-244.

[36] Zavadskas, K. E., Govindan, K., Antucheviciene, J. and Turskis, Z. (2016). Hybrid multiple criteria decision-making methods: a review of applications for sustainability issues. *Economic Research-Ekonomska Istraživanja*, 29(1), pp. 857-887.

[37] Zygmunt, A. (2017). Innovation activities of Polish firms. Multivariate analysis of the Moderate Innovator countries. *Oeconomia Copernicana*, 8(4), pp. 505–522.

[38] Żelazny, R., and Pietrucha, J. (2017). Measuring innovation and institution: the creative economy index. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(1), pp. 43-62.

Author's address

Adam P. Balcerzak, PhD.
Nicolaus Copernicus University in Toruń, Department of Economics, ul. Gagarina 13a, 87-100 Toruń, Poland
Email: adam.balcerzak@umk.pl

ANALYSIS OF AN INFLUENCE OF THE RISK MEASUREMENT APPROACH ON THE COMPOSITION OF AN INVESTMENT PORTFOLIO MADE BY REDUCED MARKOWITZ MODEL

Adam Borovička, University of Economics, Prague

Abstract
One of the most important criteria in a portfolio making is the risk associated with the investment. Risk can be measured from historical returns of the investment instruments in several ways (variance, semivariance, standard deviation, average absolute negative deviation, etc.). Each measure looks at the risk from the perspective of another characteristic; it has its informative value. Then it seems that a choice of the risk measurement approach should influence the composition of a portfolio. But is it generally really so? Are not there cases where this does not apply? Answers these questions are notable for real investment decision making because a choice of the "right" risk measure can be very difficult task. In case of no influence, the portfolio selection process becomes easier. To analyze this effect, a methodical approach based on the Markowitz model is proposed. The effect is studied from a few perspectives, namely investment horizon and kind of specified investment instruments (unit trusts). This complex analysis of influence of a risk measure on the portfolio composition is being performed on the case of open unit trusts offered by Česká spořitelna.

Keywords: *influence (effect), Markowitz model, open unit trust, portfolio making, risk measure*

JEL Classification: C44, G11
AMS Classification: 90B50

1 INTRODUCTION

Risk of the investment can be defined as the probability of occurrence of losses relative to the expected return of the investment. Then it is natural that risk (with an expected return) is usually perceived as the most important criterion in a portfolio making. We know several concepts for measuring the investment risk. A completely intuitive question is which risk measure to choose. This choice should potentially influence the composition of an investment portfolio. Therefore, it is very important to properly deal with this task.

The analysis of influence of the risk measure to the composition of a portfolio is studied on the capital market with open unit trusts because this segment of financial market becomes more and more popular in the Czech Republic during the last two decades. The main aim of this article is to show (through this case), whether a portfolio composition is generally dependent on the type of a risk measure. If the answer is yes, what risk measure should be used. Does any measure bring a better solution (portfolio with lower risk) then the others? These are other important questions that will be answered in this article because these facts can significantly influence the investment decision making process. To meet the aims, the methodological approach for "risk effect" analysis is proposed. This approach is based on the Markowitz model (for making a portfolio) which is adequately modified to incorporate the prospective specified needs of a particular investment situation. This mathematical programming approach is user-friendly and is able to take into account various demands on the portfolio. The analysis includes most used risk measurement approaches based on the historical returns of the open unit trusts −variance, semivariance, standard deviation, average absolute negative deviation and SRRI. To make a complete and representative analysis, a few investment horizons are specified, and various kinds of selected open unit trusts are used. Comparison of these empirically determined portfolios originated from such diverse situations can help properly studied an influence of the investment risk.

The structure of the article is organized as follows. After the introduction, in Section 2, the entire proposed methodological concept for "influence" analysis is described, including selected risk measures. In Section 3, the analysis is performed for the capital market with open unit trusts. Section 4 summarizes the paper and provides some ideas for the future research.

2 PROPOSED METHODOLOGICAL APPROACH FOR THE RISK INFLUENCE ANALYSIS

In this section, the methodological concept of the analysis of an influence of the risk measure on the composition of an investment portfolio is described. This approach is based on the reduced Markowitz model extended by some necessary conditions. In the second part, selected risk measures are briefly described, and their pros and cons are specified.

2.1 Methodological background for a portfolio making

To make an investment portfolio, the principles of modern portfolio theory of Harry Markowitz (Markowitz 1952 and 1959) can be applied. It has a few benefits. At first, the Markowitz model is easily comprehensible for a wider range of users. Secondly, this model takes into account two most important characteristics of the investment – risk and return. The model can be easily extended by additional conditions (according to the needs of a real-life portfolio making problem). Lastly, the model is simple to solve using the well-known methods.

The investment in open unit trusts is rather longer-term because of their features and specifics with their trading. The potential investor usually tries to properly balance a level of expected return and risk. Most investors are rather more conservative. Thus, they try to minimize the risk connected with the investment under a chosen lower limit of the expected return. Various applied risk measures do not work with interdependencies among assets returns. Then the Markowitz model can be denoted as reduced. Such a mathematical model for making a portfolio may generally look as follows

$$\begin{aligned} & z_r(\mathbf{x}) = \sum_{i=1}^{n} r_i x_i \to \min \\ & \sum_{i=1}^{n} v_i x_i \geq V \\ & \mathbf{x} \in X \end{aligned} \qquad , \qquad (1)$$

where $\mathbf{x} = (x_1, x_2, ..., x_n)^T$ is a vector of variables $x_i\,(i = 1, 2, ..., n)$ representing a share of the i-th open unit trust in the portfolio, $v_i\,(i = 1, 2, ..., n)$ denotes a return of the i-th fund, $r_i\,(i = 1, 2, ..., n)$ marks a risk of the i-th unit trust, V is a lower limit of the portfolio return. The set X includes a standard "portfolio" constraint $\sum_{i=1}^{n} x_i = 1$ and other conditions representing additional demands on the portfolio (e.g. lower/upper limit for the share of fund). The value V

is usually determined subjectively. This decision can be affected by the highest possible return attainable under all specified investment conditions. The extreme return is determined by a standard one-objective optimization model formulated for maximizing portfolio return $z_v(\mathbf{x}) = \sum_{i=1}^{n} v_i x_i$ on the set *X*. Then let denote the highest return $z_v^* = z_v(\mathbf{x}_\mathbf{v}^*)$, when it holds the following

$$\mathbf{x}_\mathbf{v}^* = \underset{\mathbf{x} \in X}{\arg\max}\, z_v(\mathbf{x}) . \tag{2}$$

All mathematical models are linear in a portfolio selection process. The set of feasible solutions is limited (among others thanks by a portfolio condition or non-negativity conditions). Then the optimal solution can be easily found by simplex method. Through these models, a portfolio can be made for various types of risk measures, time periods (investment horizons), or specified investment instruments.

2.2 Risk measures

We can meet several types of risk measures. Very popular techniques are based on the historical returns. They actually represent a specific characteristic of the investment instrument, or the portfolio.

Variance, semivariance, standard deviation

Very often used measure of the risk is variance (or standard deviation) of the investment instrument's return. According to Rényi (1972), the variance, or standard deviation, of return of the *i-th* open unit trust is computed as follows

$$\sigma_i^2 = \frac{\sum_{j=1}^{m} (x_{ij} - \overline{x}_i)^2}{m}, \quad \text{or} \quad \sigma_i = \sqrt{\frac{\sum_{j=1}^{m} (x_{ij} - \overline{x}_i)^2}{m}}, \tag{3}$$

where $x_{ij}\ (i = 1, 2, ..., n; j = 1, 2, ..., m)$ is the *j*-th return, $\overline{x}_i\ (i = 1, 2, ..., n)$ is an average return and *m* is the number of observations of historical returns of the *i*-th open unit trust. The risk of the investment portfolio z can be characterized by variance, or standard deviation, formulated as follows (Alexander and Francis 1986)

$$r_p = \sigma_p^2 = \sum_{k=1}^{n} \sum_{l=1}^{n} x_k x_l \sigma_{kl}, \quad \text{or} \quad r_p = \sigma_p = \sqrt{\sum_{k=1}^{n} \sum_{l=1}^{n} x_k x_l \sigma_{kl}} \tag{4}$$

where $\sigma_{kl}\ (k, l = 1, 2, ..., n)$ is a covariance (variance) of return of the *k*-th and *l*-th open unit trust, $x_k\ (k = 1, 2, ..., n)$, or $x_l\ (l = 1, 2, ..., n)$ denotes a share of the *k*-th, or *l*-th fund in the portfolio. Finding a global optimum can be complicated because of nonlinearity of this risk function. To eliminate this drawback, the formula can be simplified by an elimination of the covariances by the following weighted sum approach

$$z_r(\mathbf{x}) = \sum_{i=1}^{n} \sigma_i^2 x_i, \quad \text{or} \quad z_r(\mathbf{x}) = \sum_{i=1}^{n} \sigma_i x_i . \tag{5}$$

Elimination of the covariances is not a drawback for a portfolio making on the capital market with open unit trusts because the character of a development of prices is often the same. From my point of view, the main drawback of the variance, or standard deviation concept, is an inclusion of both deviations from the mean (positive and negative). However, the positive deviations are desirable. Analogous situation occurs with covariances. This disadvantage can be partly weakened by the concept of semivariance (Markowitz 1959). This risk measure of the *i*-th open unit trust can be formulated as follows

$$semo\sigma_i^2 = \frac{\sum_{x_{ij} < \bar{x}_i} (x_{ij} - \bar{x}_i)^2}{m} , \tag{6}$$

where *m* is the number of historical returns less than the average return. The risk of the whole investment portfolio can be expressed in the linear shape $z_r(\mathbf{x}) = \sum_{i=1}^{n} semo\sigma_i^2 x_i$. In my opinion, the square roots are unnecessary. Then a better way can be average absolute negative deviation.

Average absolute negative deviation
This measure inspired by Konno and Yamazaki (1991) is calculated for the *i-th* fund as follows

$$AAND_i = \frac{\sum_{x_{ij} < \bar{x}_i} (\bar{x}_i - x_{ij})}{m} , \tag{7}$$

where *m* is the number of historical returns less than the average return. This measure indicates the average negative deviation from the mean. Against the previous concepts, its calculation is easy, an interpretation of its value is comprehensible. A portfolio risk is also specified as a weighted sum $z_r(\mathbf{x}) = \sum_{i=1}^{n} AAND_i x_i$.

Synthetic Risk and Reward Indicator (SRRI)
This indicator was defined in 2009 by the Committee of European Securities Regulators (CESR) to provide an assessing fund's risk for the investors. Česká spořitelna does not provide many information about a methodology for the calculation of this indicator. However, more information can be found in the document from CESR (Committee of European Securities Regulators 2018). This indicator gets integer values on a scale from 1 to 7 (1-the lowest risk, 7-the highest risk). Each grade corresponds with particular scale of volatility which is measured by a standard deviation of the historical returns covering the last 5 years of the life of the fund. Then a portfolio risk is standardly measured by a weighted sum $z_r(\mathbf{x}) = \sum_{i=1}^{n} SRRI_i x_i$, where $SRRI_i \ (i = 1, 2, \ldots, n)$ is the value of SRRI indicator of the *i*-th open unit trust.

2.3 Risk influence analysis

Due to the aforementioned mathematical models (1) and (2), a portfolio can be made for various types of risk measures in a different time period, and for different specified investment instruments. Then we can make a complex analysis of an influence of risk measure on the investment portfolio composition in several situations through a comparison of solutions of these models. Moreover, to compare a portfolio from the perspective of a risk level, the relative relations between a risk level and its minimum and maximum value are computed. The lower level of these indicators the better. Then the portfolios order to a non-declining sequence according to the values of both indicators. Both orders are averaged to make a ranking of the portfolios from the lowest to the highest (relative) risk. A simply modified concept can consider the absolute differences between risk level of a portfolio and its extremal values. In general, it is possible to say that a risk influence is analysed by such an empirical analysis.

3 ANALYSIS OF INFLUENCE OF THE RISK MEASURE ON THE CZECH CAPITAL MARKET WITH OPEN UNIT TRUSTS

For the analysis of an influence of the risk measurement approach, the open unit trusts are selected. This type of investment is very popular in the Czech Republic. Česká spořitelna is one of the biggest asset managers in the Czech Republic. It namely offers bond, mixed and stock open unit trusts that have been the most popular in recent years (see AKAT ČR 2018). Moreover, I have a personal investment experiences with these funds.

3.1 Introduction to investment situation

Imagine a typical real-life situation when a client of Česká spořitelna has some free financial resources. One way how to expand these funds is to invest in open unit trusts. The portfolio of Česká spořitelna is namely created by bond, mixed and stock open unit trusts (Investiční centrum 2018a). We select 5 bond funds (*Trendbond*, *Sporoinvest*, *Sporobond*, *High Yield dluhopisový*, *Korporátní dluhopisový*), 7 mixed funds (*Vyvážený Mix*, *Konzervativní Mix*, *Fond životního cyklu 2020*, *Fond životního cyklu 2030*, *Fond řízených výnosů*, *Dynamický Mix*, *Akciový Mix*) and 3 stock funds (*Top Stocks*, *Sporotrend*, *Global Stocks*). It is possible to say that the open unit trusts are suitable for a longer time investment. These open unit trusts have a sufficiently long history for all performed analysis. The client can make a portfolio only from one group of funds, as well as a portfolio may be more diverse. The main characteristics of the investment are (expected) return and risk. Return is usually estimated on the basis of the historical development. As mentioned above, risk is usually measured from the historical returns. Then it is represented by standard deviation, variance, semivariance, average absolute negative deviation or SRRI. For the analysis of an influence of the risk measure, a few groups of open unit trusts and several investment horizons (time periods) are considered.

3.2 Typical longer-term investment

As the investment in open unit trusts is rather longer, a historical period from 2008 to 2018 is selected. This period should represent a longer-term development of returns because contains various situations on the capital market (huge falls, big rises or "calmer" time). Then the typical investment situation is analyzed when the expected (average) return (and risk) is computed from yearly returns from this time period. All data about returns and risk measured by selected techniques are shown in the following table (Tab. 1).

Table 1: Data for a typical longer-term investment

Fund	**Return**	**Variance**	**Semivar.**	**St. dev.**	**AAND**	**SRRI**
Trendbond	0.93 %	39.11 %	20.07 %	6.25 %	4.17 %	4
Sporoinvest	0.25 %	2.69 %	2.66 %	1.64 %	1.18 %	1
Sporobond	3 %	15.12 %	10.38 %	3.89 %	2.72 %	2
High Yield dluh.	6.42 %	602.22 %	365.40 %	24.54 %	13.21 %	3
ČS korporátní dluh.	3.91 %	214.56 %	140.85 %	14.65 %	8.6 %	3
Vyvážený Mix	2.29 %	93.75 %	216.33 %	9.68 %	11.45 %	3
Konzervativní Mix	1.47 %	42.99 %	91.55 %	6.56 %	7.15 %	3
Fond živ. cyklu 2020	1.56 %	163.50 %	258.94 %	12.79 %	10.27 %	3
Fond živ, cyklu 2030	1.77 %	290.23 %	690.74 %	17.04 %	19.48 %	4
Fond řízených výnosů	0.30 %	7.35 %	2.85 %	2.71 %	1.39 %	3
Dynamický Mix	2.51 %	201.86 %	351.39 %	14.21%	12.5 %	4
Akciový Mix	3.42 %	327.46 %	549.52 %	18.10 %	16.25 %	4

Top Stocks	15.30 %	1214.12 %	1228.83 %	34.84 %	24.13 %	6
Sporotrend	6.77 %	2991.65 %	1357.79 %	54.7 %	27.08 %	6
Global Stocks	6.24 %	340.28 %	585.96 %	18.45 %	15.71 %	5

Returns of the open unit trusts are calculated from daily prices of their allotment certificates which are collected from the document on the web site (Investiční centrum 2018b). The SRRI data is presented in (Investiční centrum 2018a). It is used for all analyses. It should be remembered that SRRI indicator is computed for a shorter time period. This fact should be taken into account in an interpretation (and comparison) of the results.

To diversify the portfolio from the perspective of number of open unit trusts, the minimum share is stated to the level of 30 %. According to my investment experiences, portfolios usually contain at least four open unit trusts. The investment strategy is based on the minimizing risk with some demanded level of return that is determined as 85 % of its optimal value. This bold demand is affected by the fact that the return is also important criterion in a portfolio making. Then the investment portfolio for this time period ("longer-time" investment strategy) is made by means of the model (1) in the following particular form

$$
\begin{aligned}
& z_r(\mathbf{x}) = \sum_{i=1}^{n} r_i x_i \to \min \\
& \sum_{i=1}^{n} v_i x_i \geq 0.85 * z_v^* \\
& x_i \geq 0.3 \qquad i = 1, 2, ..., 15 \\
& \sum_{i=1}^{15} x_i = 1
\end{aligned}
, \tag{8}
$$

where $x_i \, (i = 1, 2, ..., 15)$ represents a share of the i-th open unit trust in the portfolio (values of the index i are related to the funds in order from Tab. 1), $r_i \, (i = 1, 2, ..., 15)$ is a risk of the i-th open unit trust represented by a particular risk measure, $v_i \, (i = 1, 2, ..., 15)$ denotes a return of the i-th fund, z_v^* is the highest possible level of return determined by model (2) with the set X containing the following conditions from model (8)

$$
\begin{aligned}
& x_i \leq 0.3 \qquad i = 1, 2, ..., 15 \\
& \sum_{i=1}^{15} x_i = 1
\end{aligned}
. \tag{9}
$$

Models are solved for each risk measure for all open unit trusts, as well as for each group separately. In the case of all 15 funds, the portfolio is different for each risk measure (Tab. 2).

Table 2: Longer-term investment – portfolios for all open unit trusts

Fund	**Variance**	**Semivar.**	**St. dev.**	**AAND**	**SRRI**
Sporoinvest	x	x	18.49 %	19.12 %	30 %
Sporobond	30 %	30 %	30 %	30 %	10.10 %
High Yield dluh.	x	30 %	x	20.88 %	30 %
Korporátní dluh.	17.68 %	18.15 %	x	x	x
Top Stocks	22.32 %	21.85 %	30 %	30 %	29.90 %
Global Stocks	30 %	x	21.51 %	x	x

Although the risk is minimized, the portfolios are rather (but not only) created by open unit trusts with a higher risk. It is caused by a higher value of required minimum return $0.85 * z_v^*$ and the well-known fact that the funds with a higher risk usually provide a higher return. As we can see in Tab. 1, the values of risk measures are relatively different, their distance from maximum (or minimum) values are different, respectively. Even in some cases (e.g. Top Stocks and Sporotrend for variance and semivariance values), a proportion between risk levels are opposite. Then it is possible to expect that the composition of the investment portfolio will be dissimilar. Another assumption is confirmed (for all analyzed cases) that all portfolios provide the same level of return (85 % of maximum/ideal value z_v^* = 6.88 %). All portfolios contain 4 open unit trusts which is not surprise with regard to condition (9) included in the models. The scale of open unit trusts sharing in the portfolios is quite low. All funds partake in more portfolios, some funds in all of them. All three groups of funds are represented (2 from each group). Under these partial facts, a composition and shares of the funds are more or less different. This case confirms a significant influence of the risk measure on the composition of the investment portfolio. In such a situation, it might be interesting to know which portfolio provides the least risk. Based on the approach described in Section 2.3, the best result is for a variance risk measurement approach.

For separate group of open unit trusts, such a portfolio diversity does not hold. The portfolio only with bond open unit trusts is the same for variance and semivariance risk measurement concept (*Trendbond* 28.63 %, *Sporobond* 30 %, *High Yield dluhopisový* 11.37 %, *Korporátní dluhopisový* 30 %). Another portfolio is formed for a standard deviation approach (*Sporoinvest* 25.48 %, *Sporobond* 30 %, *High Yield dluhopisový* 14.52 %, *Korporátní dluhopisový* 30 %). The third different portfolio is made for a risk measure as average absolute negative deviation and SRRI (*Sporoinvest* 30 %, *Sporobond* 30 %, *High Yield dluhopisový* 21.12 %, *Korporátní dluhopisový* 18.88 %). The lowest risk (according to the concept mentioned above) is for SRRI measure. Lower diversity of the results also holds for mixed open unit trust. This trend can be caused by a similar development of prices of open unit trusts of the same type or their lower number. The relative relationships among values of risk measures are not so various. Even for stock funds, all portfolios are the same. A selection of used risk measure does not influence the portfolio composition. All portfolios have the lowest possible risk. The best relative relation of this risk level to the highest one is for a variance risk measurement concept.

3.3 "January effect" time period and holiday summer time

However, it would be interesting to analyze a risk influence in another (special) investment situation from a time perspective. One of these cases is January effect. It is empirically traced pattern, where the stock returns in this month are statistically above average, often higher than in other months in the year. Then the average January return is watched in the period from 2008 to 2018. Unfortunately, the historical data (as for a holiday summer time) cannot be displayed because of lack of space. In a chosen time period, a January effect is not so significant because of the big falls at the beginning of the financial crises in January 2008. To make all portfolios, the same models (of course, with adequate specification of values of the index i) as in the previous case are formulated and solved. A similar behavioral pattern (as in the previous case) can be seen. But a diversity of the portfolios is even smaller. For all open unit trusts, there are only 3 different portfolios. Two different portfolios are made for bond and stock open unit trusts and only one for mixed funds. Lower diversity can be caused by a shorter period (some part of the year), where a development of the prices is more similar than in terms of a longer period.

Other interesting investment period is a summer. This period is related to the well-known adage *Sell in May and go away* that warns investors to sell their stock holdings in May to avoid a seasonal decline on the capital market. The analysis focuses on the summer months July and August for which a lower trade activity (traders on vacation) is typical. Then the average return for summer period from 2008 to 2018 is calculated. The mathematical models differ only in one condition representing a specification of lower limit for share of the open unit trust in the portfolio because of a lower number of stock funds. Then a lower limit of share is subjectively increased to the adequate level of 40 %. As in the previous cases, a risk measure sometimes influences the composition of the portfolio and sometimes not. In addition, we can see one quite interesting fact. Only one portfolio is computed for all open unit trusts, as well as for mixed funds. Three portfolios are made for bond and stock funds. It means that a portfolio diversity is decreased by a wider range of types of investment instruments traded by these funds.

3.4 Summarizing result discussion

According to all analyses, unambiguous influence of the risk measure on the portfolio composition has not been confirmed. In some cases (characterized by type of funds and period), all portfolios are the same. On the other side, some investment situations provide various portfolios. In these cases, it is generally not possible to say, which risk measure "provides" a portfolio with lower level of risk (measured relatively to minimum and maximum risk level). However, the variance concept provides the best results most often.

And what do the results mean for a real-life investing? At first, from my point of view, a (potential) investor (investment analytic or counsel) should use such a type of the risk measure to understand it. S/he should know to interpret its value to apply it correctly. Secondly, a selection of a risk measure is not significant namely for shorter-time investment (proportionally months) when the portfolio is made from a shortlist of funds. However, the most typical investment in the open unit trusts is for a longer-time period when a portfolio is usually made from a wider range of the offered open unit trusts. In such a situation, an influence of the risk seems to be more significant. However, any risk measure cannot ensure making a portfolio with the most attractive level of risk. This fact supports the statement from the beginning of this paragraph. Then a selection of the suitable risk measure may not be complicated.

4 CONCLUSION

To analyze an influence of the risk measure on the composition of an investment portfolio, the methodological procedure based on the reduced (modified) Markowitz model is proposed. The main advantage of this approach is its user-friendliness and applicability for a wider range of the investment instruments. It is applied on the investing in the open unit trusts. The results do not generally confirm an influence of the risk measure on the portfolio composition. However, this complex analysis shows some situations, when the portfolio composition is dependent on a type of used risk measurement approach. In all such cases, none of the selected risk measures does not generally ensure the most acceptable level of risk. Based on these results, the main criterion of a selection of the risk measure should rather be its usability for a particular investment situation, as well as its intelligibility and user-friendliness. This rather initial analysis encourages further scientific ideas. To make the analysis even more deeper and representative, other risk measures (e.g. VaR), investment instruments (e.g. bonds, stocks) or various investment time horizons (e.g. time of drops, time of rises) could be added. Further, the effect of used portfolio making technique, or setting the minimum level of risk can be also studied. This could be related to the inclusion of return as an uncertain (vague) element.

Acknowledgements: This work was supported by Grant No. IGA F4/57/2017 of the Internal Grant Agency, Faculty of Informatics and Statistics, University of Economics, Prague.

References

[1] AKAT ČR. (2018). Prezentace AKAT ke konci roku 2017. [online] Available at: https://www.akatcr.cz/download/3963-k20180226_prezentace_roku_2017_official.pdf [Accessed 3 March 2018].
[2] Alexander, G. J. and Francis, J. C. (1986). *Portfolio analysis* (3rd ed.). Englewood Cliffs: Prentice-Hall.
[3] Committee of European Securities Regulators. (2018). Methodology for the calculation of the synthetic risk and reward indicator. [online] Available at: http://ec.europa.eu/internal_market/investment/docs/legal_texts/framework/091221-methodilogies-1_en.pdf [Accessed 5 March 2018].
[4] Investiční centrum. (2018a). Produkty – fondy. [online] Available at: https://cz.products.erstegroup.com/Retail/ [Accessed 3 March 2018].
[5] Investiční centrum. (2018b). Servis a kontakt – ke stažení. [online] Available at: https://cz.products.erstegroup.com/Retail/cs/Servis_a_kontakt/Ke_stauC5uBEenuC3uAD/index.phtml [Accessed 3 March 2018].
[6] Konno, H. and Yamazaki, H. (1991). Mean-Absolute Deviation Portfolio Optimization Model and its Applications to the Tokyo Stock Market. *Management Science*, 37(5), pp. 519-531.
[7] Markowitz, H. M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons, Inc.
[8] Markowitz, H. M. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), pp. 77-91.
[9] Rényi, A. (1972). Teorie pravděpodobnosti. Praha: Academia.

Author's address

Ing. Adam Borovička. PhD.
University of Economics in Prague, Faculty of Informatics and Statistics,
Department of Econometrics
W. Churchill Sq. 4, 130 67 Prague 3
Czech Republic
email: adam.borovicka@vse.cz

IMPACT OF LABOUR-MARKET SEGMENTATION ON IDENTIFICATION OF PARAMETERS IN A DSGE FRAMEWORK

Jakub Chalmovianský,
Faculty of Economics and Administration, Masaryk University

Abstract
In this contribution, I present two small scale DSGE models of a closed economy with search and matching frictions on labor market and right-to-manage bargaining process. First model is the well-known model from (Lubik, 2009). The second model stems from Lubik's work, introducing labor market segmentation to account for different wage setting processes for two groups of workers with dissimilar level of qualification. The aim of this contribution is to examine how this modification in the second model affects the amount of information needed to properly identify its parameters. At first, I shortly introduce main aspects of both models. Based on the presented calibration, trajectories of main endogenous variables are obtained. Various subsets of these simulated trajectories are then used as observables for estimation of the model parameters to compare to what extent rich information is needed for each model to properly identify its parameters.

Keywords: *DSGE, identification issues, labor market segmentation, search and matching frictions*

JEL Classification: C32, C52, E24
AMS Classification: 91B51, 91B40

1 INTRODUCTION

In the recent years, dynamic stochastic general equilibrium (DSGE) models with search and matching frictions on labor market have become a standard tool for evaluating effects of labor market functioning on aggregate economic phenomena (e.g. economic growth), or its mutual interaction with monetary and fiscal policy. Nevertheless, there is only small body of literature on DSGE models with labor market segmentation, even though it may be of great importance. Especially, when one wants to assess impacts of various labor market policies aimed at a certain subgroup of the labor market participants, or policies with different settings for different subsets of population.

In this contribution I briefly present a novel approach to construct a DSGE model with segmented labor market and then evaluate implications of this approach for the identification of parameters when the model is estimated. The proposed model is inspired by (Porter and Vitek, 2008), and in line with their approach the model consists of two segments. First segment is populated with low-skilled household members and in the second segment high-skilled individuals reside. Porter and Vitek used their proposed model to study impacts of introducing a statutory minimum wage in Hong Kong. Separation of workers to two homogeneous subgroups based on their level of qualification, knowledge, or work experience is common in the literature on economic or social impacts of minimum wage (some examples of such works can be found in (Cahuc, Saint-Martin, and Zylberberg, 2001) or in (Neumark and Wascher, 2008)). However, this is not the case when the DSGE models are used.

I apply the before mentioned segmentation to the simple model of a closed economy with search and matching frictions on labor market and right-to-manage bargaining process by augmenting the well-known model proposed by Lubik (2009). Aim of this contribution is to evaluate the degree of information needed for proper estimation of the model parameters and to compare how this amount of information is changed when a simple model is augmented with the labor market segmentation. For this purpose, I first present some basic structure and calibration of parameters of both models used. In the first stage, based on these calibrations the trajectories of endogenous variables are simulated. These simulated trajectories are in the second stage used as inputs for the estimation of parameters. Finally, various combinations of observables and their number of observations are used in the estimation stage to see how it affects the uncertainty associated with parameter estimates for each model. Detailed discussion on the effects of various combinations of observables on DSGE estimation results can be found in (Guerrón-Quintana, 2010).

2 DESCRIPTION OF MODELS

In this section I briefly describe main characteristics of the models used for the evaluation of amount of information needed for proper estimation of model parameters. At first, I list and shortly describe main equations that form the original Lubik's model. In the second subsection, I discuss only major differences of the augmented model with dual labor market structure.

2.1 Lubik's model

As mentioned before, the model developed in (Lubik, 2009) will be augmented with segmentation on labor market and used as benchmark for comparison of the amount of information needed for proper estimation of model parameters. Lubik's model incorporates search and matching frictions on labor market and right-to-manage wage bargaining. Thanks to the search and matching process, there are frictions which make the whole procedure of creating seeker-vacancy pair for job searchers and firms, looking for workers, time-consuming and costly. I will present here only overview of its main equations. For details on the model structure, estimation techniques etc. see (Lubik, 2009).

$$\lambda_t = C_t{}^{-\sigma} \tag{1}$$

$$m(u_t, v_t) = \mu_t u_t{}^{\xi} v_t{}^{1-\xi} \tag{2}$$

$$q(\theta_t) = \frac{m(u_t, v_t)}{v_t} \tag{3}$$

$$\theta_t = \frac{v_t}{u_t} \tag{4}$$

$$n_t = 1 - u_t \tag{5}$$

$$n_t = (1-\rho)(n_{t-1} - v_{t-1} q(\theta_{t-1})) \tag{6}$$

$$y_t = A_t n_t^{\alpha} \tag{7}$$

$$y_t = \left(\frac{p_t}{P_t}\right)^{-1-\varepsilon_t} Y_t \tag{8}$$

$$\tau_t = \alpha \frac{y_t}{n_t} \frac{\varepsilon_t}{1+\varepsilon_t} p_t - w_t + (1-\rho) E_t \beta_{t+1} \tau_{t+1} \tag{9}$$

$$\kappa v_t^{(\psi-1)} = (1-\rho) q_t E_t \beta_{t+1} \tau_{t+1} \tag{10}$$

$$\beta_t = \beta \frac{\lambda_t}{\lambda_{t-1}} \tag{11}$$

$$w_t = \eta \left(\alpha \frac{y_t}{n_t} \frac{\varepsilon_t}{1+\varepsilon_t} p_t + \kappa v_t^{(\psi-1)} \theta_t \right) (1-\eta)(b + \chi_t C_t^{\sigma}) \tag{12}$$

$$Y_t = C_t + \frac{\kappa}{\psi} v_t^{\psi} \tag{13}$$

The first-order condition for households is shown in (1), where λ_t is the Lagrange multiplier on the representative household's budget constraint, C_t is aggregate consumption, $\sigma \geq 0$ is a coefficient of relative risk aversion in the household's utility function. In equation (2), $m(u_t, v_t)$ represents number of matches of unemployed job-seekers, u_t, and vacancies, v_t, matching shock μ_t measures the efficiency of the matching process, parameter $0 < \xi < 1$ stands for a match elasticity for the unemployed. Further, $q(\theta_t)$ represents aggregate probability of filling a vacancy, with $\theta_t$indicating labor market tightness, n_t is an employment rate, which development over time is given by (6), with separation rate $0 < \rho < 1$. For simplicity, this model assumes monopolistic behavior of firms, with production (and demand) y_t described by production function (7), with stochastic process A_t describing an aggregate technology shock and $0 < \alpha \leq 1$, introducing curvature in production. Demand function of a representative firm is given by (8), with aggregate output, Y_t, price p_t set by firm, and aggregate price index P_t, ε_t stands for demand elasticity assumed to be an exogenous process. The first-order conditions of the firms are (9) and (10), where τ_t is the Lagrange multiplier associated with firm's employment constraint, w_t denotes bargained wage, β_t is stochastic discount factor, E_t denotes commonly used expectations operator, parameter $\kappa > 0$ is a scaling factor on vacancy creation, and $\psi > 0$ stands for elasticity of vacancy creation. Parameters $\eta \in [0,1]$ and b in the expression for wage bargained (12) denote bargaining power of worker and unemployment benefits financed by lump-sum tax T_t, respectively.

Lubik's model is completed by specifying properties of exogeneous shocks. Logarithms of technology shock, A_t, labor shock, χ_t, demand shock, ε_t, and finally matching shock, μ_t, follow independent autoregressive processes of the first order with residuals $\epsilon_t^i \sim N(0, \sigma_i^2)$ and coefficients ρ_i, where $i \in (A, \chi, \varepsilon, \mu)$.

2.2 Model with labor market segmentation

The augmented model stems from the Lubik's model presented in the previous subsection, however I refined its original labor market sector to distinguish two sectors of economy that differ by the qualification of their labor force (education, working experience, etc.). In the first sector low-skilled labor services are supplied, and in the second one the high-skilled labor is supplied. Otherwise, the structure of the model is rather standard, consisting of households, and firms, without any explicit monetary authority. Households are populated by the before mentioned two types of individuals. Every member of a household is send out to the labor market, where he/she looks for jobs, if unemployed, or supplies labor services if employed. Job-seekers and vacancies are matched in a costly search procedure using standard search and matching process. In what follows, I briefly present the structure of households and the optimization problems of their members and discuss differences in wage determination for the two labor market sectors.

2.2.1 Households

Modelled economy is populated by a continuum of households indexed by $k \in [0,1]$. Proportion ϕ^L of household members are low-skilled workers (denoted with superscript u) and $(1 - \phi^L)$ are high-skilled workers (denoted as s), where $0 < \phi^L < 1$. Total employment, $n_{k,t} \in [0,1]$, in the representative household k can be written as follows:

$$n_{k,t} = \phi^L n_{k,t}^u + (1 - \phi^L) n_{k,t}^s \tag{14}$$

Except for the qualification, all workers are otherwise identical. Furthermore, as in (Merz, 1995), no heterogeneity in asset holdings and consumption of individual workers or households is assumed. Hence, indices of specific workers/households are dropped in all equations to come. Representative infinitely living household maximizes the following intertemporal utility function:

$$U_t = E_t \sum_{j=t}^{\infty} \beta^{j-t} \left[\frac{C_j^{(1-\sigma)} - 1}{1-\sigma} - \chi_j D n_j - D^s n_j^s \right] \tag{15}$$

with symbols having the same meaning as in section 2.1, except $D \geq 0$ is parameter of preferences over disutility from work activities, and similarly $D^s \geq 0$ is parameter of preferences over disutility from work specific only for high-skilled sector. This concept establishes additional disutility from work for high-skilled employees that may arise from their higher opportunity costs associated with gaining work experience (more years of schooling, increased work effort, etc.). Thanks to this construction, augmented model distinguishes between wages for low-skilled and high-skilled employees, allocating higher wages for the latter. It is worth noting that the proportion of employed individuals in the representative household of a given qualification is fully determined by the matching function. Hence, households cannot control for their labor supply, possibly leading to an involuntary unemployment. Finally, C_t denotes composite consumption index with constant elasticity of substitution between different consumed kinds of goods, i.e.:

$$C_t = \left[(1-\omega_u)^{\frac{1}{\eta_C}} (C_t^s)^{\frac{\eta_C - 1}{\eta_C}} + (\omega_u)^{\frac{1}{\eta_C}} (C_t^u)^{\frac{\eta_C - 1}{\eta_C}}\right]^{\frac{\eta_C}{\eta_C - 1}} \tag{16}$$

with ω_u denoting the fraction of consumed goods produced by low-skilled workers and η_C as the elasticity of substitution between the two kinds of goods. This implies the following demand functions for each type of good:

$$C_t^s = (1-\omega_u)\left[\frac{P_t^s}{P_t}\right]^{-\eta_C} C_t \tag{17}$$

$$C_t^u = \omega_u \left[\frac{P_t^u}{P_t}\right]^{-\eta_C} C_t \tag{18}$$

where P_t^s and P_t^u are aggregate price indices for the high-skilled and low-skilled production sector respectively, and P_t is the composite price index, defined as their weighted average.

$$P_t = \left[(1-\omega_u)(P_t^s)^{1-\eta_C} + \omega_u (P_t^u)^{1-\eta_C}\right]^{\frac{1}{1-\eta_C}} \tag{19}$$

Finally, the representative household's budget constraint is as follows:

$$C_t + T_t = w_t^s n_t^s + w_t^u n_t^u + (1 - n_t^u)b + (1 - n_t^s)b + \Pi_t \tag{20}$$

where Π_t are profits that representative household receives as the owner of the firms, and $w_t^l, l \in \{u, s\}$ is the real wage that each worker of given qualification receives for supplying her labor services. The only problem of household is to select the amount of consumed goods of each kind, as it cannot determine amount of labor supplied. Because there is no intertemporal aspect, the resulting first order condition is of the same form as in (1).

2.2.2 Wage bargaining

Labor market is again specified using the search and matching process with households supplying labor services to firms on a frictional market. The only difference is in the use of separate labor market for each of the two groups of workers. It is worth noting however, that the wage bargaining process reveals some differences for each kind of labor force.

Crucial assumption of this model is that labor force can be divided into two separate groups based on the qualification level of its members. Individuals in each of these two populations are entering qualification-specific labor market and participate in production of specific goods. Except for this difference all workers are otherwise identical, their qualification casts the only contrast. If this assumption holds, wage of the high-skilled worker results from the bilateral bargaining process between firms and representatives of the labor force. Both participants in the negotiations try to set the wage rate, w_t^s, to maximize the total surplus generated by the worker-firm linkage. The total surplus is split to maximize weighted average of the individual surpluses of firm and worker. The bargaining function, S^s, used in this model is of the following form

$$S_t^s \equiv \left(\frac{1}{\lambda_t}\frac{\partial\, W_t^s(n_t^s)}{\partial\, n_t^s}\right)^{\eta^s} \left(\frac{\partial\, I_t^s(n_t^s)}{\partial\, n_t^s}\right)^{1-\eta^s} \tag{21}$$

where η^s represents the high-skilled workers' power in negotiations, $\frac{\partial\, W_t^s(n_t^s)}{\partial\, n_t^s}$ is the marginal value worker contributes to the household's welfare, and from the firm's perspective $\frac{\partial\, I_t^s(n_t^s)}{\partial\, n_t^s}$ is marginal value worker contributes to the firm's welfare. The latter determined by the firm's first order condition with respect to the requested number of high-skilled workers.

Marginal contribution of the high-skilled worker for the household's welfare can be obtained making use of the options the worker has. If employed, this contribution equals the wage worker earns, w_t^s. However, at the same time she suffers from the disutility from work represented by $\chi_t D n_t$, as well as the additional disutility specific for the high-skilled work, $D^s n_t^s$. Finally, she loses the unemployment benefits, b. Additionally, the worker's marginal contribution to the welfare of her household depends also on its expected value in the next period. Note that the real payments are again valued at the marginal utility λ_t. The marginal value of a high-skilled worker for the household is:

$$\frac{\partial\, W_t^s(n_t^s)}{\partial\, n_t^s} = \lambda_t w_t^s - \lambda_t b - D\chi_t - D^s + \beta E_t \frac{\partial\, W_{t+1}^s(n_{t+1}^s)}{\partial\, n_{t+1}^s}\frac{\partial\, n_{t+1}^s}{\partial\, n_t^s} \tag{22}$$

Assuming the following evolution of employment in time: $n_t^l = (1-\rho^l)\,[n_{t-1}^l + v_{t-1}^l q(\theta_{t-1}^l)]$, the expression $\frac{\partial\, n_{t+1}^s}{\partial\, n_t^s}$ can be replaced with: $(1-\rho^s)[1-\xi^s\theta_t^s q(\theta_t^s)]$. Subsequently, the derivation of the resulting function, S_t^s, with respect to the only variable that is subject to bargaining (i.e. the wage rate of the high-skilled workers, w_t^s), applying the expressions for the worker's marginal contributions to household's and firm's value respectively, results in the following optimality condition of the bargained wage:

$$w_t^s = \eta^s \left[\alpha \frac{y_t^s}{n_t^s}\frac{\varepsilon_t}{1+\varepsilon_t} p_t^s + \kappa^s (v_t^s)^{\psi^s - 1}\theta_t^s\right] + (1-\eta^s)\left[b + \left((1-\phi^L)D\chi_t + D^s\right)(C_t^s)^\sigma\right] \tag{23}$$

The bargained wage is simply a weighted average of the payments accruing to workers and firms. Among other it includes mutual compensation of the job related costs, e.g. hiring costs, $\kappa^s (v_t^s)^{\psi^s-1}$, or costs in the form of utility loss of employed, which have two components: the loss of utility from leisure, $\left((1-\phi^L)D\chi_t + D^s\right)(C_t^s)^\sigma$, and the unpaid unemployment benefits, b.

In the case of the low-skilled workers, denoted by the superscript u, I assume slightly different scenario. The bargained wage is specified by analogous bargaining process as in previous case, except with different parameters that are typical for the low-skilled sector (particularly with different term for the workers' loss of utility from leisure). The following optimality condition then defines the resulting bargained wage:

$$w_t^u = \eta^u \left[\alpha \frac{y_t^u}{n_t^u}\frac{\varepsilon_t}{1+\varepsilon_t} p_t^u + \kappa^u (v_t^u)^{\psi^u - 1}\theta_t^u\right] + (1-\eta^u)[b + \phi^L D\chi_t (C_t^u)^\sigma] \tag{24}$$

2.2.3 Closing the model

To close the model following social resource constraints, specific for each labor market segment, need to be defined

$$Y_t^s = C_t^s + \frac{\kappa^s}{\psi^s}(v_t^s)^{\psi^s} \tag{25}$$

$$Y_t^u = C_t^u + \frac{\kappa^u}{\psi^u}(v_t^u)^{\psi^u} \tag{26}$$

Finally, I assume the logarithms of six exogeneous stochastic processes, namely the technology shock in high-skilled production, A_t^s, technology shock in low-skilled production

A_t^u, the labour shock, χ_t, the matching efficiency shock for high-skilled sector, μ_t^s, the matching efficiency shock for low-skilled sector μ_t^u, and the time-varying demand elasticity, ε_t, to follow independent $AR(1)$ processes, similarly as in the case of Lubik's model.

3 COMPARISON OF ESTIMATED PARAMETERS

Following simulation-estimation exercise has been conducted to see the effect of number and combination, or number of observations, of observables on the estimation results. Namely, posterior mean and variance of standard deviation of shocks in both models, and moments of filtered observables and exogenous variables are compared with their counterparts from stochastic simulation based on the calibrations presented in Table 1.

Table 1: Calibrated parameters

Parameter	Description	Value	
		Lubik	**Segmented labor**
β	Discount factor	0.99	0.99
α	Labor elasticity	0.67	0.67
σ	Coeff. of relative risk aversion	1	1
ξ	Match elasticity	0.7	$\xi^s = \xi^u = 0.7$
ρ	Job separation rate	0.1	$\rho^s = \rho^u = 0.1$
η	Bargaining power of workers	0.5	$\eta^s = \eta^u = 0.5$
b	Unemployment benefits	0.4	0.4
ψ	Elasticity of vacancy creation costs	1	$\psi^s = \psi^u = 1$
κ	Scaling factor on vacancy creation	0.05	$\kappa^s = \kappa^u = 0.05$
ω_u	Fraction of consumed goods from u sector	N/A	0.245
D	Scaling factor on disutility from work	N/A	0.3
D^s	Scaling factor on disutility from work in s sector	N/A	0.65
$\rho_{\{A,A^s,A^u,\chi,\mu,\mu^s,\mu^u,\varepsilon\}}$	AR coefficients of shocks	0.75	0.75
$\sigma_{\{A,A^s,A^u,\chi,\mu,\mu^s,\mu^u,\varepsilon\}}$	Standard deviations of shocks	1	1

Stochastic simulation in Dynare toolbox (version 4.4.3) for MATLAB by Adjemian et al. (2018) was applied for both models, in order to obtain simulated trajectories with 100 observations of both endogenous and exogeneous variables. Following this, I ran sequence of estimations using all 100, 75, or only 50 observations of selected combinations of simulated trajectories as observables (as shown in Table 2) to get estimates of standard deviations of exogenous processes, $\sigma_{\{A,A^s,A^u,\chi,\mu,\mu^s,\mu^u,\varepsilon\}}$, following inverse-gamma distribution with prior mean set to 0.01 and standard deviation of 1. Simulated endogenous variables used as observables during the estimation stage are u_t, v_t, w_t, Y_t for Lubik's model, and $u_t^s, u_t^u, v_t^s, w_t^s, w_t^u, Y_t^s$ for the model with labor market segmentation.

In Table 2, I present posterior means and standard deviations of the estimated standard deviations of shocks for all model specifications. For Lubik's model with full information (4 observables used) standard deviations of all 4 shocks are well estimated. This, however, does not hold for the rest of models, where only 3 variables are used, as well as for all specifications of model with segmented labor market (including the one with all 6 observables). In all those modifications standard deviation for time-varying demand elasticity seems to be poorly identified (highlighted in *italics*), even though still statistically significant. No other shock shows similar behavior as I change the inputted variables. It is worth noting that also specifications without observed one of the unemployment rates (u_t^s, u_t^u) or vacancy rate for high-skilled sector (v_t^s) were considered, however these specifications failed to successfully solve the model. All three mentioned variables therefore seem to play an important role for the model with segmented labor market.

Table 2: Estimation results

Model	Observables	Parameter	100 obs. Mean (st. d.)	75 obs. Mean (st. d.)	50 obs. Mean (st. d.)
Lubik	u_t, v_t, w_t, Y_t	σ_A	**1.09 (0.0752)**	1.10 (0.0863)	1.14 (0.1082)
		σ_μ	**0.97 (0.0662)**	1.04 (0.0818)	1.06 (0.0997)
		σ_ε	**0.97 (0.0670)**	0.90 (0.0714)	0.97 (0.0932)
		σ_χ	**1.06 (0.0729)**	1.11 (0.0872)	1.083 (0.1019)
	u_t, v_t, w_t	σ_A	1.13 (0.0889)	**1.11 (0.0877)**	1.15 (0.1102)
		σ_μ	0.97 (0.0846)	**1.04 (0.0814)**	1.06 (0.1102)
		σ_ε	*0.01 (0.0020)*	***0.01 (0.0019)***	***0.01 (0.0019)***
		σ_χ	1.23 (0.1091)	**1.23 (0.0984)**	1.22 (0.1168)
	u_t, v_t, Y_t	σ_A	**1.09 (0.0752)**	1.10 (0.0863)	1.14 (0.1082)
		σ_μ	**0.97 (0.0665)**	1.03 (0.0810)	1.06 (0.1003)
		σ_ε	***0.01 (0.0019)***	***0.01 (0.0019)***	***0.01 (0.0019)***
		σ_χ	**1.07 (0.0741)**	1.12 (0.0891)	1.09 (0.1037)
	u_t, w_t, Y_t	σ_A	**1.09 (0.0774)**	1.10 (0.0863)	1.14 (0.1082)
		σ_μ	**0.97 (0.0748)**	1.02 (0.0805)	1.07 (0.1014)
		σ_ε	***0.01 (0.0019)***	***0.01 (0.0019)***	***0.01 (0.0019)***
		σ_χ	1.05 (0.0906)	**1.11 (0.0870)**	1.07 (0.1018)
	v_t, w_t, Y_t	σ_A	**1.09 (0.0752)**	1.10 (0.0864)	1.14 (0.1083)
		σ_μ	**0.99 (0.0683)**	1.05 (0.0826)	1.08 (0.1022)
		σ_ε	***0.01 (0.0019)***	***0.01 (0.0019)***	***0.01 (0.0019)***
		σ_χ	**0.95 (0.0656)**	1.11 (0.0872)	1.07 (0.1018)
Segmented labor market	$u_t^s, u_t^u, v_t^s, w_t^s, w_t^u, Y_t^s$	σ_{A^s}	**0.92 (0.0636)**	0.96 (0.0753)	0.90 (0.0847)
		σ_{A^u}	**0.93 (0.0637)**	0.93 (0.0728)	0.96 (0.0905)
		σ_{μ^s}	**0.99 (0.0681)**	1.00 (0.0786)	1.03 (0.0987)
		σ_{μ^u}	**0.98 (0.0675)**	0.97 (0.0761)	1.10 (0.1118)
		σ_ε	***0.01 (0.0019)***	***0.01 (0.0019)***	***0.01 (0.0019)***
		σ_χ	**1.06 (0.0728)**	1.12 (0.0879)	1.07 (0.1042)
	$u_t^s, u_t^u, v_t^s, w_t^s, Y_t^s$	σ_{A^s}	**0.93 (0.0636)**	0.96 (0.0752)	0.91 (0.0852)
		σ_{A^u}	**0.93 (0.0636)**	0.93 (0.0730)	0.97 (0.0913)
		σ_{μ^s}	**0.99 (0.0682)**	1.00 (0.0786)	1.06 (0.1043)
		σ_{μ^u}	**0.98 (0.0672)**	0.97 (0.0759)	1.06 (0.1026)
		σ_ε	*0.01 (0.0019)*	***0.01 (0.0018)***	*0.01 (0.0020)*
		σ_χ	**1.06 (0.0731)**	1.12 (0.0887)	1.04 (0.0997)
	$u_t^s, u_t^u, v_t^s, w_t^u, Y_t^s$	σ_{A^s}	**0.92 (0.0626)**	0.96 (0.0751)	0.90 (0.0839)
		σ_{A^u}	**0.93 (0.0622)**	0.93 (0.0724)	0.96 (0.0908)
		σ_{μ^s}	**0.99 (0.0693)**	1.00 (0.0788)	1.07 (0.1044)
		σ_{μ^u}	**0.98 (0.0669)**	0.97 (0.0772)	1.05 (0.1008)
		σ_ε	*0.01 (0.0018)*	*0.01 (0.0018)*	***0.01 (0.0017)***
		σ_χ	**1.06 (0.0735)**	1.12 (0.0902)	1.04 (0.0979)
	$u_t^s, u_t^u, v_t^s, w_t^s, w_t^u$	σ_{A^s}	**0.92 (0.0639)**	0.96 (0.0753)	0.90 (0.0840)
		σ_{A^u}	**0.93 (0.0630)**	0.93 (0.0727)	0.96 (0.0906)
		σ_{μ^s}	**0.99 (0.0681)**	1.00 (0.0784)	1.04 (0.1016)
		σ_{μ^u}	**0.98 (0.0673)**	0.97 (0.0761)	1.04 (0.1004)
		σ_ε	***0.01 (0.0018)***	*0.01 (0.0019)*	*0.01 (0.0021)*
		σ_χ	**1.06 (0.0730)**	1.12 (0.0884)	1.04 (0.1015)

Another way to assess the amount of information needed is represented by the number of observations used. Reduced number of observations however does not seem to cause any serious adverse consequences, except only slightly greater uncertainty associated with the estimates. Estimates with the highest posterior mean to posterior standard deviation ratio in Table 2 are marked with **bold**, and as can be seen they almost unequivocally belong to the richest specification with 100 observations.

4 CONCLUSION

In this contribution I compared the requirements of two small closed economy DSGE models with search and matching frictions on labor market. One was the well-known model composed by Lubik, the second one extended the Lubik's model by introducing a novel approach for modelling labor market segmentation within DSGE framework. Based on the results of simulation-estimation exercise, number of observations of observed variables inputted to model does not play an important role for the estimation of standard deviations of shocks for presented models, however it slightly increases the uncertainty associated with the estimates. The choice of observables plays more important role, especially for the augmented model. All specifications of the augmented model and specifications of the Lubik's model with only 3 observables cannot correctly identify the standard deviation of the demand shock. Augmented model's specifications lacking any of the unemployment rates or vacancy rate for high-skilled sector fail to successfully find estimate at all. The necessity to find the right set of observables for proper estimation outbalances greater complexity of the model.

Acknowledgements: This work was supported by funding of specific research at Faculty of Economics and Administration, Masaryk University, project MUNI/A/0966/2017. This support is gratefully acknowledged.

References

[1] Adjemian, S., Bastani, H., Juillard, M., Karam, F., Mihoubi, F., Perendia, G., Pfeifer, J., Ratto, M., and Villemot, S. (2018). *Dynare: Reference Manual, Version 4*. Cepremap, Dynare Working Papers, No. 1.

[2] Cahuc, P., Saint-Martin, A., and Zylberberg, A. (2001). The consequences of the minimum wage when other wages are bargained over. *European Economic Review*, 45, pp. 337–352.

[3] Guerrón-Quintana, P. A. (2010). What you match does matter: the effects of data on DSGE estimation. *Journal of Applied Econometrics*, 25(5), pp. 774-804.

[4] Lubik, T. (2009). Estimating a Search and Matching Model of the Aggregate Labor Market. *Economic Quarterly,* 95(2), pp. 101-120.

[5] Merz, M. (1995) Search in the Labor Market and the Real Business Cycle. *Journal of Monetary Economics* 36(2), pp. 269–300.

[6] Neumark, D., and Wascher, W. T. (2008). *Minimum wages*. MIT Press, Cambridge, Massachusetts.

[7] Porter, N. and Vitek, F. (2008). *The Impact of Introducing a Minimum Wage on Business Cycle Volatility: A Structural Analysis for Hong Kong SAR*. International Monetary Fund, Working Paper WP/08/285.

Author's address
Mgr. Jakub Chalmovianský
Masaryk University, Faculty of Economics and Administration,
Lipová 41a, 602 00 Brno
Czech Republic
email: chalmoviansky@mail.muni.cz

ANALYSIS OF INCOME CONVERGENCE BASED ON DIFFERENT SPECIFICATION OF SPATIAL WEIGHTS: EVIDENCE FROM NUTS 2 EU REGIONS

Michaela Chocholatá, University of Economics in Bratislava

Abstract
This paper deals with the analysis of income convergence based on the concepts of beta and sigma-convergence for the NUTS 2 regions of the European Union during the period 2004-2014. Since the location of a region can play a key role in the analysis, also the spatial aspect was considered by the estimation of the beta-convergence models under three different specifications of spatial weights. The results proved the validity of the beta-convergence concept indicating the speed of convergence 2.54% - 3.62% per year and corresponding half-life 19.146 - 27.304 years. The sigma-convergence was not proved to be valid for the whole analysed period – tendency to elimination of disparities was proved in the pre-crisis period 2004-2008 and during the last three analysed years, i.e. 2012-2014.

Keywords: *beta-convergence, sigma-convergence, regions, European Union (EU), spatial weights*

JEL Classification: R11, C21
AMS Classification: 91B62, 91B72

1 INTRODUCTION

The analyses of regional disparities, as well as the effort to reduce or eliminate them, play nowadays the important roles at both international and national level. The European Union's (EU) regional policy, based on the Europe 2020 Strategy (European Commission, 2010), emphasizes the gradual elimination of regional disparities, particularly in the fields of economic growth, education, research and innovation, employment, social inclusion and poverty reduction. Similarly, regional policies in individual EU Member States pay considerable attention to support for economically underdeveloped regions of their own country, with a view to achieve economic, social and cultural levels comparable to the developed regions of the EU.

In the literature, the often used disparity analysis approaches are the convergence concepts – beta-convergence and sigma-convergence (Barro and Sala-i-Martin, 2007; Hančlová et al., 2010). While the beta-convergence concept is based on a negative relationship between the initial income level and the income growth rate over the analysed period, i.e., on the assumption that poorer regions will grow faster than rich regions and will "catch up" with the rich regions in the long run; the concept of sigma-convergence is based on declining income variability between regions, i.e., on the assumption that standard deviations of the logarithms of income between regions have a declining nature over time, indicating the diminishing of the differences in income levels (Barro and Sala-i-Martin, 2007; Sojková, 2001).

Although it is generally known that the economic growth of the region is influenced by the geographical location of the region, earlier works on economic growth modelling did not include the spatial aspect in the analyses. There have been published many studies dealing with the issue of regional income convergence (e.g., Szomolányi, Lukáčiková and Lukáčik, 2011; Kutscherauer et al., 2010) the similarity of which was the neglected spatial interaction of the analysed observations (for more information see e.g., Paas et al., 2007). The problem of possibly biased results and hence misleading conclusions with using of models that have ignored the influence of spatial location on the income growth is pointed out by e.g., Carrington (2003), Fingleton and López-Bazo (2006), Furková and Chocholatá (2016), Paas et al. (2007), Chocholatá and Furková (2017), Chocholatá and Furková (2016).

The aim of the paper is to examine the validity of the concepts of beta and sigma-convergence for the 249 NUTS 2 regions of the EU over the period 2004-2014 based on the development of the net disposable income of households (expressed in Euro per inhabitant). The beta income convergence hypothesis is tested based both on non-spatial and spatial models (with the different structure of the weight matrix).

The rest of the paper is organized as follows. After the introduction in section 1, it follows a brief methodical characteristic of the concepts of beta and sigma-convergence in section 2. Section 3 of the paper is devoted both to the data used and the empirical results of the analysis, the paper closes with concluding remarks in section 4.

2 METHODOLOGY – BETA AND SIGMA-CONVERGENCE

The beta-convergence analysis for *n* regions can be performed using a linear regression model in the form (Barro and Sala-i-Martin, 2007):

$$\bar{g}_i = \alpha + \beta \ln\left(income_{i,0}\right) + \varepsilon_i, \qquad \varepsilon_i \sim i.i.d\left(0, \sigma_\varepsilon^2\right) \tag{1}$$

where $\bar{g}_i = \frac{1}{T}\ln\left(\frac{income_{i,T}}{income_{i,0}}\right)$ indicates the average annual income growth in the region *i* during the period $(0, T)$, symbols α and β denote unknown parameters and ε_i is a random component. To confirm the beta-convergence hypothesis, the estimated parameter β must be statistically significant with a negative sign.

To consider the spatial linkages of analysed regions (based on values of Moran's *I* for residuals, Lagrange Multiplier tests – LM(lag), LM(err) and their robust versions)[1], the classical linear regression model (1) should be extended by inclusion of the spatial component. We will present here two kinds of econometric models that can be used to capture the spatial dependence of observations: the spatial autoregressive model (SAR) and the spatial error model (SEM).

The SAR model for analysing the beta-convergence is as follows:

$$\bar{g}_i = \alpha + \beta \ln\left(income_{i,0}\right) + \rho \sum_{j \neq i} w_{ij} \bar{g}_i + \varepsilon_i \tag{2}$$

where symbol ρ denotes the spatial lag parameter, w_{ij} are the elements of the row-standardized spatial weight matrix **W** (for more information on different specifications of spatial weights see e.g., Getis, 2010; Harris, Moffat and Kravtsova, 2011 and Furková, 2016) and all the other symbols have already been defined above. An alternative way to capture the spatial effects is to estimate the SEM model:

$$\bar{g}_i = \alpha + \beta \ln\left(income_{i,0}\right) + u_i, \quad u_i = \lambda \sum_{j \neq i} w_{ij} u_j + \varepsilon_i \tag{3}$$

where parameter λ is a spatial error coefficient indicating the extent of the spatial autocorrelation between regression residuals.

[1] For more information about testing of spatial autocorrelation and about spatial econometric models see e.g., Fischer and Getis (2010).

Sigma-convergence is usually analysed graphically based on the assessment of the development, e.g., of standard deviations of the income logarithms between regions during the analysed period. To accept the hypothesis of sigma-convergence, it is necessary to confirm the decreasing character of these standard deviations. Mathematically, the formula for the calculation of the standard deviations of income logarithms can be written in the following form (Sojková, 2001; Rey and Dev, 2006):

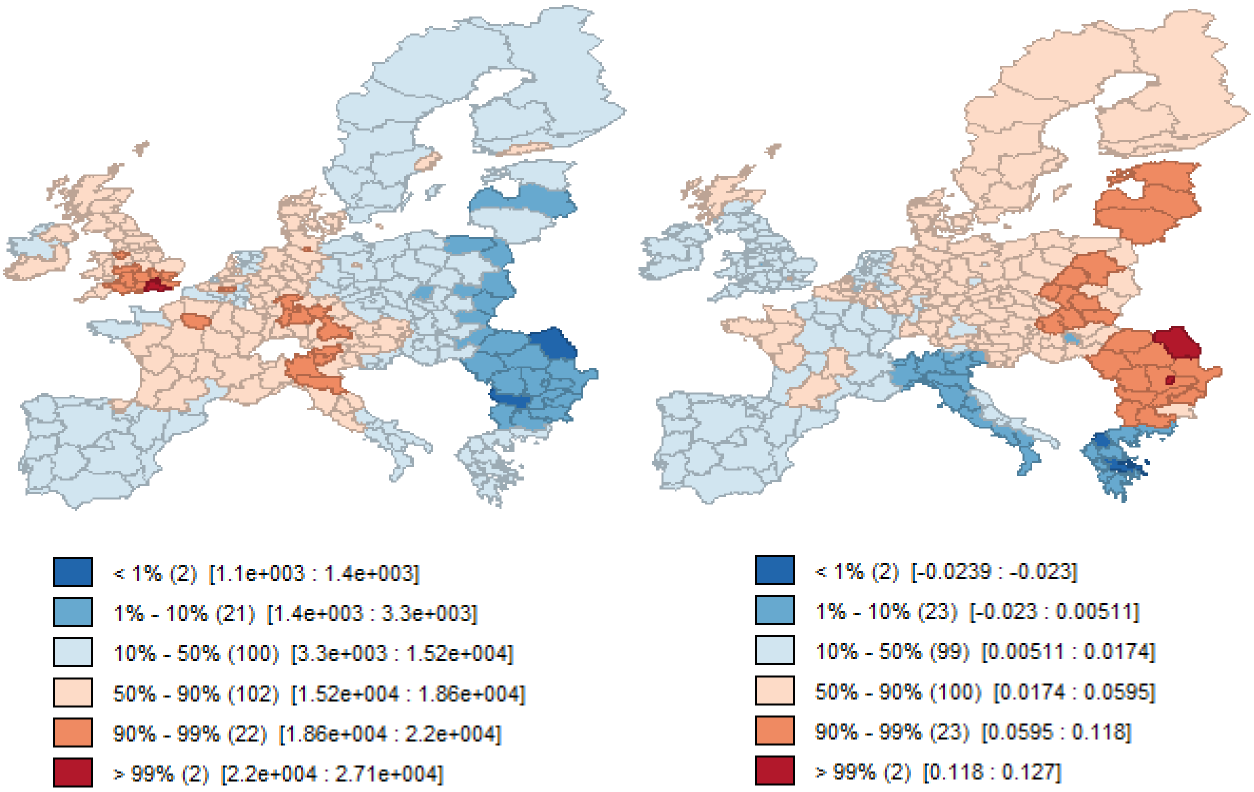
$$\sigma_t = \sqrt{\frac{\sum_{i=1}^{n}\left(\ln(income_{i,t}) - \ln(income_t)\right)^2}{n-1}} \tag{4}$$

where $\ln(income_{i,t})$ is the level of net disposable income per inhabitant (in natural logarithms) in the i - th region at time t, $\ln(income_t)$ is the average level of net disposable income per inhabitant (in natural logarithms) at time t for all analysed regions, and n represents the number of analysed regions.

3 DATA AND EMPIRICAL RESULTS

The empirical part of the paper uses the data for 249 NUTS 2 regions of the EU countries obtained from the regional REGIO database of the Eurostat (Eurostat, 2017a). The analysis of convergence is based on the development of the net disposable income of households (in Euro per inhabitant) during the period 2004-2014. Based on the values of this indicator, average annual income growth was calculated based on the logarithmic growth definition in section 2. The standard deviations for the sigma-convergence analysis were calculated using the formula (4). Analyses were carried out in the freely available software GeoDa and MS Excel. From a shapefile containing the regions of Europe (Eurostat, 2017b), a set of 249 NUTS 2 regions of the EU countries was extracted in GeoDa. In this context it is necessary to mention that the original data set contained 272 NUTS 2 regions of the 28 EU member states, but due to the possible problems with the isolated regions we had to exclude 20 island regions of Cyprus, Malta, France, Finland, Spain, Greece, Portugal and Italy and due to the missing data the further 3 regions of Luxemburg and Croatia.

To investigate whether there exists some spatial pattern in the analysed data we started with the ESDA (Exploratory Spatial Data Analysis). The levels of net disposable income in 2004 as well as the corresponding average annual income growth values for 2014/2004 are visualised via the percentile maps on Figure 1 which illustrates the unequally distribution of both analysed indicators over space, i.e. also the presence of some disparities across regions inside the analysed countries. Percentile map consists of six categories for the classification of the ranked observations: < 1%, 1%-10%, 10%-50%, 50%-90%, 90%-99% and > 99%. As we can see (Figure 1(a)), the regions with lower value of net disposable income in 2004 were detected for all the transition countries in eastern Europe, for the regions in southern Europe as well as for the majority of regions situated in the north (with exception of the UK). The lowest percentile consists of 2 regions (1 Bulgarian and 1 Romanian). On the other side, the regions with higher values of net disposable income in 2004 were situated mostly in the western part of Europe and in the UK the 2 regions of which belong to the highest percentile. To assess the issue of beta-convergence we can consider the percentile map of average annual income growth 2014/2004 (Figure 1(b)) as well as the scatter plot and the regression line for the income in 2004 and the average annual income growth rate 2014/2004 in Figure 2. The slope of the regression line (parameter β – model (1)) is negative and statistically significant which indicates that poorer regions "catch up" with the rich regions. From both the Figure 1(b) and Figure 2 it is possible to identify the huge progress of the Romanian regions (regions with the highest growth values) and Bulgarian regions with the lowest initial income values. The evidence of the beta-convergence is further given especially by the Slovak regions, some Polish regions, Latvia, Lithuania and Estonia. On the other hand, the evidence of the beta-convergence concept is not supported by the Greek (regions with extremely low growth values), Spanish, Portuguese and some Italian regions.



(a) net disposable income 2004 (b) average annual income growth 2014/2004

Figure 1 Percentile maps of net disposable income and average annual income growth[2]

Since the visualisation via percentile maps enables us to assess whether the average annual income growth is randomly distributed across the analysed regions or whether there exists spatial clustering of regions with similar percentage of growth, this approach does not provide us any information about the statistical significance or insignificance of the clustering (Mitchell, 2013). To confirm that location affects the growth variable, i.e. to analyse the existence of the spatial autocorrelation, three different weight matrices **W**, identifying the neighbourhood of regions, were used to consider the sensitivity of its specification on the results.

The first weight matrix was a contiguity weight matrix "queen" (two regions were considered as neighbours if they shared any part of a common border), secondly, we used the weight matrix based on the inverse distance function (with the threshold arc distance of 357.255 km) and finally the 8 nearest neighbours weight matrix. Testing for the spatial autocorrelation both on the global level (test for clustering) and local level (test for clusters) was based on global Moran's I statistic and local Moran's I statistic (see e.g., Fischer and Getis, 2010). Figure 3 includes both the visualisation of the global test by means of a Moran scatterplot and the local

[2] All figures presented in the paper are available in the colour version on the web-site: http://www.fhi.sk/kove/konferencie

test as a cluster map based on the local Moran's *I* statistic[3] with the use of the three types of weight matrices described above[4]. Although the used weight matrices were based on a different number of neighbours (queen – mean: 4.52, highest frequency 5 neighbours; inverse distance – mean: 24, highest frequency 12 neighbours; 8 nearest neighbours – mean and highest frequency: 8 neighbours), the values of the global Moran's *I* statistic indicated the statistically significant positive spatial autocorrelation for the analysed group of EU regions for all the three specifications. Concerning the regions specific details, the LISA (Local Indicators of Spatial Association) cluster map based on the different weights, indicated many regions with statistically significant positive autocorrelation (high-high, low-low) and only few regions with the statistically significant negative autocorrelation (high-low, low-high). Quite equivalent results were received with the queen and 8 nearest neighbours weight matrices. Based on the inverse distance weight matrix we indicated much more statistically significant regions with the positive autocorrelation especially of the low-low type – most of them located in Netherland, France, Germany and Belgium (it is interesting to point out that majority of these regions had a very high number of neighbours in comparison to the remaining two weight matrix specifications).

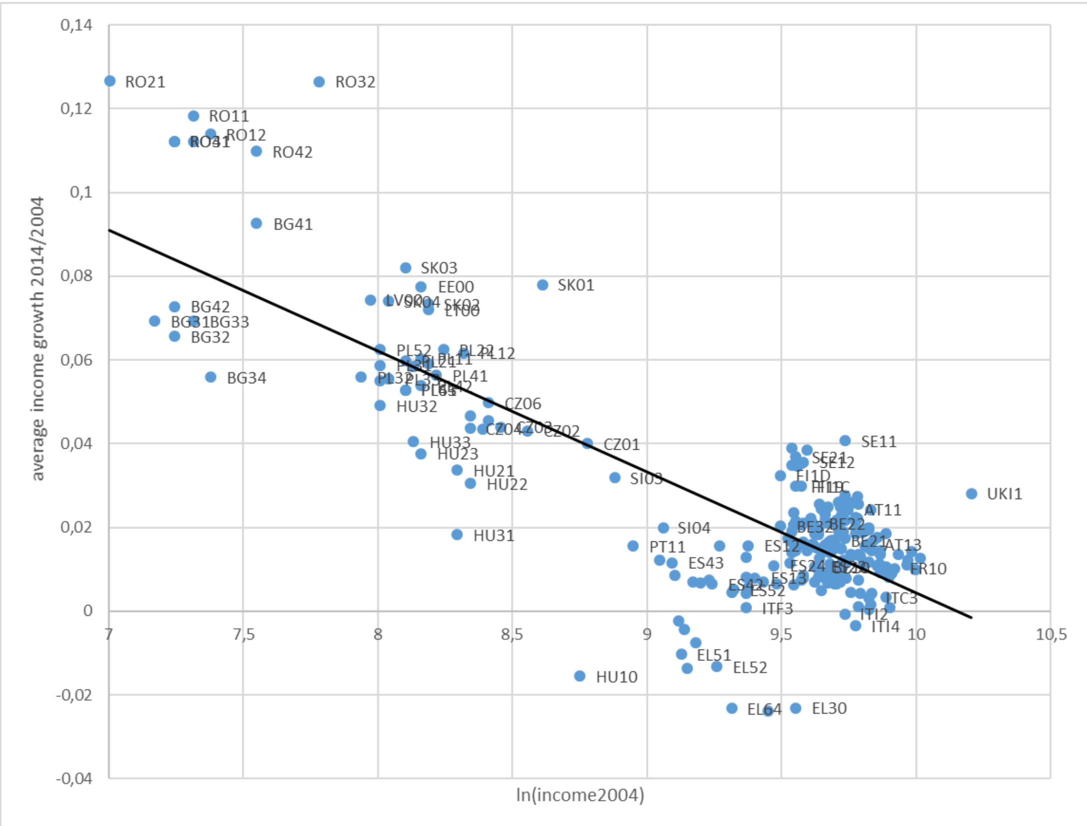


Figure 2 Scatter plot and the regression line for the income in 2004 and the average annual income growth rate 2014/2004

In the next step the beta-convergence analysis was performed by estimation of the econometric models presented in section 2. The results of estimated parameters for linear regression model (1) without spatial component as well as for the SEM specification (3) under three weights specifications are in Table 1.

[3] Randomization approach based on 999 permutations was used to prove the statistical significance of results.
[4] Weight matrices are used in a row-standardized form, i.e. each neighbour of the concrete region is given equal weight.

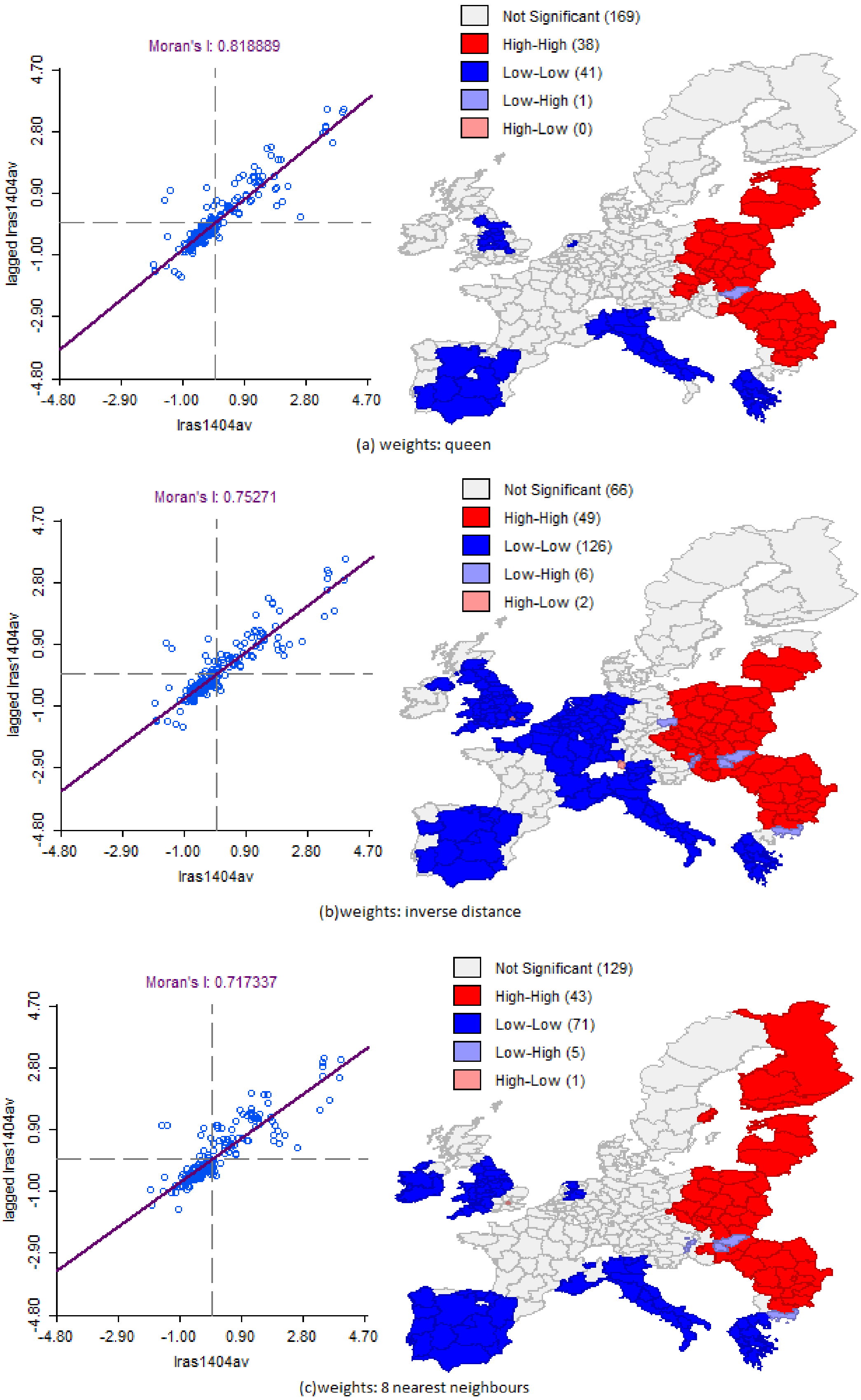


Figure 3 Moran´s *I* statistics and LISA Cluster Maps for the average annual income growth rate (*lnras1404av*) 2014/2004 under different spatial weights

All the estimated parameters are statistically significant for all models; the beta parameter is negative which is in line with the beta-convergence concept. Since the presented results showed that it is inevitable to consider the spatial dependence among the spatial units, we will concentrate on the results obtained by spatial models. Based on different weights we received slightly different convergence characteristics[5]. The speed of convergence ranging between 2.54% and 3.62% per year means that the time it takes for 50% of the initial gap to be eliminated varies between 19.146 and 27.304 years, i.e. the poorest regions are supposed to fill half the gap with the wealthiest ones in about 19.146 - 27.304 years. Chocholatá and Furková (2016) analysing the beta-convergence of 82 NUTS 2 Central European regions during 2000-2013 period based on the net disposable income per inhabitant presented similar convergence characteristics with the speed of convergence 2.91% and half-life of 23.841 years (spatial model – SEM). Paas et al. (2007) dealing with the beta-convergence presented convergence characteristics (based on spatial models) for selected EU-15 countries as well as for the "New Member States" receiving different results for the GDP per capita expressed in Euro and in PPS (Purchasing Parity Standard). The results of spatial beta-convergence for the NUTS 2 EU regions were furthermore provided e.g., by Furková and Chocholatá (2016) who confirmed the speed of convergence of 1.33% and half-life of 52.148 years based on SAR model (GDP per capita expressed in Euro) during the 2000-2011 period.

Model **Spatial weights**	**Linear model** **-**	**SEM** **queen**	**SEM** **inverse distance**	**SEM** **8 nearest neighbours**
α	0.2933(0.0000)	0.3073(0.0000)	0.2327(0.0000)	0.2842(0.0000)
β	-0.0289(0.0000)	-0.0304(0.0000)	-0.0224(0.0000)	-0.0279(0.0000)
λ	-	0.8181(0.0000)	0.8849(0.0000)	0.8324(0.0000)
Convergence characteristics				
Speed of convergence	0.0341 = 3.41%	0.0362 = 3.62%	0.0254 = 2.54%	0.0327 = 3.27%
Half-life	20.345	19.146	27.304	21.222

Table 1 Estimation results of the beta-convergence models under different spatial weights (p-values are in parentheses)

The concept of sigma-convergence refers to the development of the income disparities over the analysed period. Figure 4 presents the dynamics of both the mean values (left axis) and the standard deviation of natural logarithm (right axis) of the net disposable income for all 249 regions over the period 2004-2014. Both the mean values and standard deviations were strongly influenced by the crisis in 2008. The gradual rise of the mean income level during 2004-2008 was interrupted by a decline in 2009 with the following slightly rising tendency. The pre-crisis period, i.e. 2004-2008, with clearly downward trend of standard deviation strongly confirming the evidence of the sigma-convergence concept was followed by a period 2009-2012 with no evidence of sigma-convergence. During the last two analysed years we can again identify the declining tendency of the standard deviation giving the chance for the continual future elimination of income disparities across the EU regions. Sojková (2001) presenting the results both for the beta-convergence and sigma-convergence (GNP per capita expressed in US$) for a group of 27 European countries during 1995-1999 period does not received unambiguous

[5] Additional approaches for testing of the robustness of results can be found e.g. in Fingleton and López-Bazo (2006).

results concerning both the convergence concepts. The problems with testing for sigma-convergence in the presence of spatial effects is discussed, e.g., by Rey and Dev (2006).

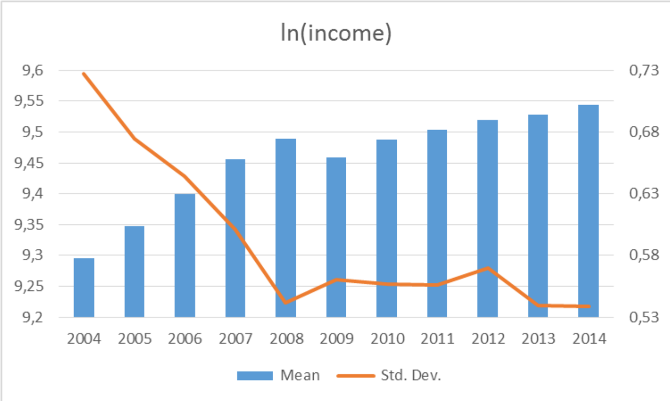


Figure 4 Mean (left axis) and the standard deviation (right axis) of the net disposable income per inhabitant (in natural logarithms)

4 CONCLUSIONS

Analysis of the income convergence for the selected set of the NUTS 2 EU regions was carried out based on testing of the beta-convergence and sigma-convergence concepts. The results proved the validity of the beta-convergence concept based both on non-spatial and spatial models. Preliminary analysis was concentrated on mapping of the analysed indicator and its growth as well as on the assessment of the scatter plot and the regression line for the income in 2004 and the average annual income growth rate 2014/2004. Further followed the consideration of the region's location supported by the results of spatial autocorrelation testing based on three differently defined spatial weights. Based on the SEM specification the speed of convergence ranged between 2.54% (spatial weights: inverse distance) and 3.62% (spatial weights: queen) per year which means that the poorest regions are supposed to fill half the gap with the wealthiest ones in about 19.146 to 27.304 years. The validity of the sigma-convergence indicating the diminishing of the disparities among the regions in time was confirmed for the periods 2004-2008 and 2012-2014.

Acknowledgements: This work was supported by the Grant Agency of Slovak Republic – VEGA grant No. 1/0248/17 "Analysis of Regional Disparities in the EU based on Spatial Econometric Approaches".

References

[1] Barro, R. and Sala-i-Martin, X. (2007). *Economic growth*. New Delhi: Prentice Hall of India Private Ltd.

[2] Carrington, A. (2003). A Divided Europe? Regional Convergence and Neighbourhood Spillover Effects. *Kyklos*, 56(3), pp.381-393.

[3] Chocholatá, M. and Furková, A. (2016). Income disparities and convergence across regions of Central Europe. *Croatian Operational Research Review*, 7(2), pp.303-318.

[4] Chocholatá, M. and Furková, A. (2017). Does the location and the institutional background matter in convergence modelling of the EU regions? *Central European Journal of Operations Research*, 25(3), pp.679-697.

[5] European Commission (2010). Europe 2020: A European strategy for smart, sustainable and inclusive growth. [online] Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:EN:PDF [Accessed 5 Jan. 2017]

[6] Eurostat, (2017a). Database – Eurostat. [online] Available at: http://ec.europa.eu/eurostat/data/database [Accessed 5 Jan. 2017]

[7] Eurostat, (2017b). GISCO – Eurostat. [online] Available at: http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units [Accessed 5 Jan. 2017]

[8] Fingleton, B. and López-Bazo, E. (2006). Empirical growth models with spatial effects. *Papers in Regional Science*, 85(2), pp.177-198.

[9] Fischer, M. and Getis, A. (2010). *Handbook of Applied Spatial Analysis*. Berlin: Springer Berlin.

[10] Furková, A. and Chocholatá, M. (2016). Spatial econometric modelling of regional club convergence in the European Union. *Ekonomický časopis*, 64(4), pp.367-386.

[11] Furková, A. (2016). Spatial pattern of innovative activity in the EU regions: exploratory spatial data analysis and spatial econometric approach. In: O. Gomes and H. Martins, ed., *Advances in applied business research: the L.A.B.S. initiative*, 1st ed. New York: Nova Science Publishers, pp.159-186.

[12] Getis, A. (2010). Spatial Autocorrelation. In: M. Fischer and A. Getis, ed., *Handbook of Applied Spatial Analysis. Software Tools, Methods and Applications*, 1 st ed. Berlin Heidelberg: Springer-Verlag, pp.255-278.

[13] Hančlová, J., Kubicová, I., Macháček, M., Melecký, A., Melecký, L., Melecký, M., Nevima, J. and Ramík, J. (2010). *Makroekonometrické modelování České ekonomiky a vybraných ekonomik EU*. Ostrava: VŠB- Technical University of Ostrava.

[14] Harris, R., Moffat, J. and Kravtsova, V. (2011). In Search of 'W'. *Spatial Economic Analysis*, 6(3), pp.249-270.

[15] Kutscherauer, A., Fachinelli, H., Sucháček, J., Skokan, K., Hučka, M., Tuleja, P. and Tománek, P. (2010). *Regionální disparity. Disparity v regionálním rozvoji země - pojetí,teorie, identifikace a hodnocení.* Ostrava: VŠB- Technical University of Ostrava.

[16] Mitchell, W. (2013). *Introduction to Spatial Econometric Modelling.* Centre of Full Employment and Equity, working paper no. 01-13, The University of Newcastle, Callaghan NSW 2308, Australia.

[17] Paas, T., Kuusk, A., Schlitte, F. and Võrk, A. (2007). Econometric Analysis of Income Convergence in Selected EU Countries and Their Nuts 3 Level Regions. *SSRN Electronic Journal*.

[18] Rey, S. and Dev, B. (2006). σ -convergence in the presence of spatial effects. *Papers in Regional Science*, 85(2), pp.217-234.

[19] Sojková, Z. (2001). *Globalizácia a konvergencia. In: Agrární perspektivy X: Zdroje trvale udržitelného ekonomického růstu ve třetím miléniu: Globalizace a regionalizace*. Praha. Available at: http://www.slpk.sk/dizertacie/agrarni_perspektivy/sojkova.pdf [Accessed 27 Jun. 2017].

[20] Szomolányi, K., Lukáčiková, A. and Lukáčik, M. (2011). The β convergence of the EU27 countries. In: *29th International Conference on Mathematical Methods in Economics 2011*, Praha: Professional Publishing, pp.287-291.

Author's address

doc. Ing. Michaela Chocholatá, PhD.
University of Economics in Bratislava, Faculty of Economic Informatics, Department of Operations Research and Econometrics
Dolnozemská cesta 1/b, 852 35 Bratislava
Slovak Republic, email: michaela.chocholata@euba.sk

QUALITY OF LIFE IN THE VISEGRAD GROUP COUNTRIES – A MULTIDIMENSIONAL APPROACH

Mariola Chrzanowska, Joanna Landmesser, Monika Zielińska-Sitkiewicz, Warsaw University of Life Sciences

Abstract
Social and economic transformations which took place after joining the EU radically changed the entire Visegrad Group. However, the rate of change is not uniform for the whole area. What is more, it can be noticed that the heterogeneity of changes is visible both on the micro and macroeconomic level. Therefore, any analysis of this topic may support the regional policy.
Regional development is defined as a transformation of factors and regional resources both in terms of goods and services, promoting the improvement of living conditions for the population. The quality of life in the regional units of a given country reflects the regional transformation and the dynamics of the changes. The main goal of the research was to group regions with a similar quality of life using selected MCA methods (*Ward's* method and *k*-means method). The analysis was based on OECD indicators and included years 2012-2016.

Keywords: *quality of life, k-means method, Ward's method, Visegrad group*

JEL Classification: C43, I15
AMS Classification: 62-07, 62P25

1 INTRODUCTION

Quality of life is one of the basic categories of social statistics research. It is an important goal of social development. It is an important goal of social policy, for which common well-being is an important task. This is the kind of policy where the realization of important human goals is in the centre. Thus, the assessment of the quality of life is strongly influenced by both economic factors, as well as the premises that enable the realization of aspirations and desires connected with existential values.
Quality of life is the subject of empirical and theoretical studies of such sciences, as: sociology, psychology and economics. The complexity of this concept is mainly due to the subjective approach to the assessment of life and objective premises (conditions and living standards) affecting this assessment. Therefore, there is no consensus in the literature as to what should be understood by the term quality of life. In social sciences, this concept primarily serves the purpose of answering the question: What lies at the heart of a valuable, satisfactory existence? (Jankowska, 2011, p. 34) Often there is more than one definition of quality of life given, because it is quite a flexible concept, definable differently depending on the researcher or the purpose of the study. Borys and Rogala (Borys and Rogala, 2008, p. 9) determined quality of life as the "degree of self-realization of a human being". In turn, Trzebiatowski (Trzebiatowski, 2011, p.28) emphasizes that the quality of life can be presented both with the use of objective indicators (e.g. material level, education, ecological safety) and subjective indicators.
Due to the remarks and the fact that the quality of life is a complex concept, the methods of its measurement are not unambiguous. One approach to measuring the quality of life is an economic approach. It uses measures to measure the quality of life, which are defined as social indicators (indices). They are most often defined as "statistical constructions based on observations, usually quantitative ones, which inform about a certain aspect of social life that

we are interested in or about changes taking place in it (Panek, 2015). However, not all data obtained from statistical and sociological surveys can be used as social indicators. Social indexes should not only be a strongly theoretically justified and empirically tested relationship with the studied phenomena, but also fulfil many other conditions. Social indicators should, among others, concern those areas that are most important for the level and quality of life of the society. First and foremost, they should reflect the effects of the activities of the society and its institutions, and to a lesser extent, the outlays for their implementation. Indicators taken into account in the measurement of quality of life should, as far as possible, be built on the basis of the existing data. And most importantly, these measures should be selected so as to enable international comparisons of living conditions, developmental processes or social regressions (Szatur-Jaworska, 2009).

This heterogeneity of criteria makes the research on the quality of life cover various aspects, as well as different research methods. And so, in the work by (Dudek and Szczesny, 2017), the fractional outcome model were used to measure the quality of life. While Somarriba, N. Pena B. (Somarriba and Pena, 2008) used the econometric analysis. In the paper (Betti, Soldi and Talev, 2015) the multidimensional analysis and fuzzy sets was used to measure the quality of life. A wide overview of approaches used in measuring the quality of life as well as the author's model of its measurement was presented by (Panek, 2015) referring to, among others, data on the quality of life gathered by Eurostat within the framework of the European Statistical System (ESS).

2 HUMAN DEVELOPMENT INDEX (HDI)

Human Development Index (HDI), or the social development index, is one of the most popular measures of the quality of life. Calculated in 1990 by the United Nations Development Programme, the measure is an attempt to determine the differences in the social and economic development of individual countries in a measurable manner. It is a synthetic measure related to three areas: living conditions, education and health. Thanks to this, it combines both economic and social criteria. The HDI index is a standardized measure that takes values from 0 to 1. This allows for linear ordering of objects in the form of rankings. It also gives the possibility to rank the countries from the most to the least developed and the dynamic analysis of changes by comparing the positions occupied in the rankings over the years. Based on the value of the indicator, the classification of objects is made, dividing the countries into four groups[1]. Countries which values do not exceed 0.535 are considered to be poorly developed countries. The averagely-developed countries fall within the range of $0.535 \leq HDI \leq 0.710$. The highly-developed countries take in the values from the range (0.710; 0.800). While the highly-developed countries obtain the HDI value above 0.8. (Jankowska, 2012, pp. 35)

HDI considerably covers a wider range than economic factors. It is widely known and used. However, its structure (i.e. the areas and representative variables included in it) is not a formula defining the standard of living and does not exhaust the topic. This fact is confirmed by the existence of related indicators, such as: Gender Inequality Index (GII), Inequality-adjusted HDI (IHDI), Multidimensional Poverty Index (MPI) and local HDI built.

3 QUALITY OF LIFE ACCORDING TO OECD METHODOLOGY

Analysing the quality of life as a subject area of the ESS, it is impossible not to mention the report: Measurement of Economic Performance and Social Progress The report developed by

[1] Before 2008, the countries were divided into three groups according to the following criteria: 1. Poorly-developed and underdeveloped (the HDI value below 0.5), 2. averagely-developed (HDI from 0.5 to 0.799), 3. highly-developed (HDI from 0.8).

Joseph E. Stiglitz, Amarty Sen and Jean-Paul Fitoussi (Stiglitz, Sen and Fitoussi, 2009) as part of the work of the French National Institute for Statistical and Economic Research (INSEE) provoked a lively discussion on the international stage. The report's conclusions recommended to extend the thematic scope of traditional indicators used to assess economic development by indicators characterizing the quality of life and inequality. Another demand was the postulate to increase the importance for sustainable development and environmental issues. Conclusions from this report and the ten-year experience of the works of the Organisation for Economic Co-operation and Development, (OECD) on the measurement of the society development became the inspiration for OECD experts to undertake the Better Life initiative. An important effect of this initiative was the publication of the first study from the series entitled How's Life? In 2011. It presents the framework for measuring the quality of life and their changes over time.

The expansion of the methodology used to compare the quality of life between the countries was the instrument for measuring the quality of life developed by OECD for the regions. It should therefore be kept in mind that in conditions of huge diversification of area, population or economic potentials of individual EU or OECD countries, it is the regional level that allows better characterization of the situation. As a consequence, the OECD project "How is the life in your region?" determined a common framework for measuring well-being at the regional level based on eleven aspects shaping the material conditions of inhabitants (income, employment and housing conditions) and their quality of life (health, education, environment, security, access to services and civic involvement) and subjective assessment of well-being (satisfaction with life, social dimension). Their measurement using a set of comparable indicators was carried out in 362 regions in 34 OECD countries.

4 THE VISEGRAD GROUP

The Visegrad Group was founded by the leaders of Poland, Hungary and Czechoslovakia on February 15, 1991. Initially functioning as a Visegrad Triangle, the group had its golden age in the beginnings of existence. During this period, the post-communist countries emerged from the allied addiction in the East and sought to gain support in the Western institutions. The supreme objective was to strengthen cooperation and mutual support as part of joining NATO and the EU. The assumptions of the V4 group was to respect the human rights, the construction of parliamentary democracy and the elimination of economic and administrative remnants of the previous system. Countries belonging to the Visegrad Group, i.e. Poland, Hungary, Slovakia and the Czech Republic were connected by similar historic experience, geographic proximity and convergent foreign policy objectives. Therefore, after joining the EU and introducing the Lisbon Treaty, they have convergent goals and views in many issues.

When analysing the quality of life of the inhabitants of the V4 group, the following conclusions can be formulated. In 2016, the Czech Republic had the highest level of HDI at the level of. Here, the Czech Republic had their best year in 2016, when it took the 28th position with the index of 0.878. Poland is on the 38th position (HDI=0.855), Slovakia is 40th with 0.845, and Hungary are on the 43rd position (0.836). In other words, the quality of life in the Visegrad Group countries is similar. So, despite the differences in approach, various behaviours and assessments, the standard of living in the V4 countries is similar. All countries recorded a slight increase in HDI as compared to 2015. It is also worth noting that Slovakia, despite entering the Euro zone, has not noted any significant improvement in social development.

The main goal of the article was to identify groups of the Polish regions with a similar quality of life based on a multidimensional comparative analysis and to determine the regional diversity of quality of life in the Visegrad Group with the use of OECD methodology. Data used to test the quality of life by OECD are the latest indicators possible to achieve –

depending on the indicator, these are data from 2012-2016. Research is continuation previous analysis (Chrzanowska, Landmesser and Dudek, 2016).

5 DESCRIPTION OF THE STUDY

The study of the quality of life was carried out using the OECD methodology (the set of indicators was specified by OECD) and the selected multidimensional comparative analysis: *Ward's* method[2] and *k*-mean method[3]. Variables for the study are presented in table 2. The study covered 35 regions of the Visegrad Group. In particular, these were:

- 8 Czech regions (Prague (CZ01); Central Bohemian Region (CZ02); Southwest (CZ03); Northwest (CZ04); Northeast (CZ05); Southeast (CZ06); Central Moravia (CZ07); Moravia-Silesia (CZ08));
- 7 Hungarian regions (Central Hungary (HU10); Central Transdanubia (HU21); Western Transdanubia (HU22); Southern Transdanubia (HU23); Northern Hungary (HU31); Northern Great Plain (HU32); Southern Great Plain (HU33);
- 16 Polish voivodeships (Łódź province (PL11); Masovia (PL12); Lesser Poland (PL21); Silesia (PL22); Lublin province (PL31); (Subcarpathia (PL32); Holy Cross province (PL33); Podlasie province (PL34); Greater Poland (PL41); Western Pomerania (PL42); Lubusz province (PL43); Lower Silesia (PL51); Opole province (PL52); Kuyavian-Pomeranian province (PL61); Warmia-Masuria province (PL62); Pomerania (PL63));
- 4 Slovak regions (Bratislava Region (SK01); West Slovakia (SK02); Central Slovakia (SK03); East Slovakia (SK04)).

The variables used in the study are presented in table 2.

Table 1. Variables included in the study

Variable symbol	Full name	Short name	Unit
Z1	Labour force with at least secondary education	Education	%
Z2	Employment rate	Employment	%
Z3	Unemployment rate	Unemployment	%
Z4	Household disposable income per capita	Income	constant USD PPP
Z5	Homicide rate	Homicide	per 100-000 people
Z6	Mortality rate	Mortality	per 1-000 people
Z7	Life expectancy	Life expectancy	Number of years
Z8	Air pollution (level of PM2.5)	Pollution	$\mu g/m^3$
Z9	Voter turnout	Frequency	%
Z10	Broadband access	Internet	% of households
Z11	Number of rooms per person	Rooms	rooms per person
Z12	Perceived social network support	Society	%
Z13	Self-assessment of life satisfaction	Self-assessment	index 0 to 10

Source: OECD.

The selected variables have been subjected to the standardization procedure. The de-stimulants were transformed into stimulants. First of all, the examined objects were divided

[2] More information on this method can be found in the paper: (Ward, 1963).

[3] The description of the method used is contained in the paper: J. Hartigan, Clustering algorithms, John Wiley& Sons, Inc 1975 (Hartigan, 1975).

into 3 groups using the *Ward's* and *k*-means methods. To check the discrimination abilities of the selected variables, an analysis of variance was additionally performed. Subsequently, (using the OECD methodology) the structure of each group was analysed. At the end, the results of the classification were compared with both methods and conclusions were formulated.

6 RESEARCH RESULTS

The first stage of the classification included grouping with the *Ward's* method. The results of this procedure are shown in fig. 1, in which three groups of regions can be clearly distinguished.

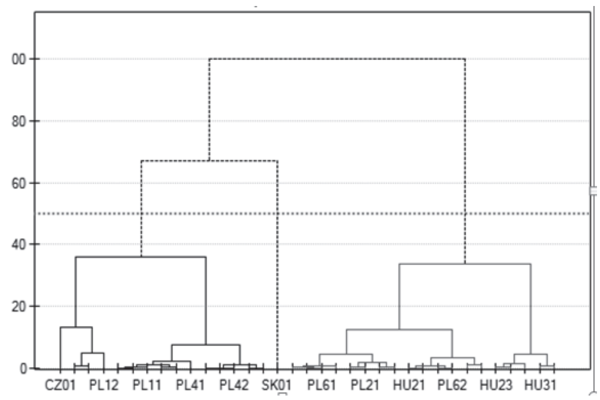


Figure 1. Grouping using *Ward's* method
Source: Own elaboration

The first of the distinguished groups includes the following regions: Prague, Central Bohemian Region, Southwest, Northeast, Southeast, Central Hungary, Łódź province, Masovia, Silesia, Greater Poland, Western Pomerania, Lower Silesia, Pomerania, West Slovakia, Central Slovakia. Four of the 13 diagnostic variables have the lowest average level in this group. These are: Z1 (education), Z7 (life expectancy), Z9 (voter turnout), Z10 (Internet). In this group, the highest level for de-stimulants was also noted – the variables: Z5 (murder), Z6 (mortality), Z8 (air pollution).

The second group consists of the regions: Northwest, Central Moravia, Moravia-Silesia, Central Transdanubia, Western Transdanubia, Southern Transdanubia, Northern Hungary, Northern Great Plain, Southern Great Plain, Lesser Poland, Lublin province, Subcarpathia, Świętokrzyskie, Podlasie, Lubusz province, Opole province, Kuyavian-Pomerania, Warmia-Masuria, East Slovakia. In this case, the highest average values were recorded for the following variables: Z4 (income), Z10 (Internet), Z11 (rooms) and the lowest ones for the Z12 variable (community).

The third region in terms of living conditions is formed by the Bratislava region. This is the group for which the highest value of income was observed and one of the highest values for the variable community (among analyzed objects). Data visualization using the OECD methodology (fig. 2) made it possible to notice that this object is characterized by high values of education, more than medium safety and community. (please note that OECD visualisation were 364 regions not only 35 like in presented analysis.

Traditionally, commuting to big, often capital, cities was a subject of research. Reverse commuting, to intermediate or rural areas, is not so popular topic of investigations (Drejerska, 2016). Undoubtedly, commuting from for example Bratislava to Vienna justifies this focus on research on the direction toward big cities.

The next stage of the research was the grouping of the regions with the *k*-mean method. The results of this grouping were as follows:

- the first group included the regions: Prague, Central Bohemian Region, Southwest, Northeast, Southeast, Central Hungary, Łódź province, Masovia, Silesia, Greater Poland, Western Pomerania, Lower Silesia, Pomerania, West Slovakia, Central Slovakia
- the second group included the regions: Northwest, Central Moravia, Moravia-Silesia, Central Transdanubia, Western Transdanubia, Southern Transdanubia, Northern Hungary, Northern Great Plain, Southern Great Plain, Lesser Poland, Lublin province, Subcarpathia, Świętokrzyskie, Podlasie, Lubusz province, Opole province, Kuyavian-Pomerania province, Warmia-Masuria, East Slovakia.
- the third group again included only Bratislava.

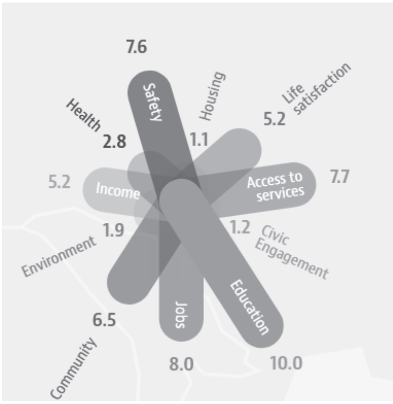


Figure 2. Visualisation of Bratislava region
Source: www.oecd.eu

Subsequently, the Kruskal-Wallis test (Kruskal and Wallis, 1952) was conducted to check the discriminatory capacity of the variables included in the grouping. (the variables analysed did not have a normal distribution and therefore the analysis of the variance could not be used). The level of significance of $p<0.05$ was assumed. Its results are presented in table 2.

Table 2. Results of the Kruskal-Wallis test

Variable symbol	Statistics H	Significance p
Z1	**11.64**	**0.00**
Z2	6.59	0.06
Z3	5.26	0.07
Z4	**25.90**	**0.00**
Z5	0.94	0.62
Z6	2.75	0.25
Z7	4.35	0.11
Z8	0.69	0.70
Z9	1.12	0.55
Z10	6.72	0.06
Z11	1.88	0.38
Z12	1.36	0.50
Z13	**9.33**	**0.01**

Source: Own calculations

Results of Kruskall-Wallis test explains why Bratislava region as a separated group. Highest level of income and education allows assume that Bratislava is the object with the best standard of living of residents. It is really an interesting case. Basis on literature (Michnak, 2016) one can notice that Bratislava is the biggest commuting centre in Slovakia. This region with the Slovak capital and the proximity of Vienna was the most intensive cross-border commuting centre in 2011 People from that area commute to Austria and its capital city Vienna supported by numerous public transport connections by trains and buses. That can explain high value of income in that region

7 SUMMARY AND CONCLUSIONS

Using two different methods, groups with the same structure were obtained. Grouping of regions using both methods allowed to identify three homogenous groups with similar quality of life. These are:

- Prague, Central Bohemian Region, Southwest, Northeast, Southeast, Central Hungary, Łódź province, Masovia, Silesia, Greater Poland, Western Pomerania, Lower Silesia, Pomerania, West Slovakia, Central Slovakia, Masovia, Podlasie, Greater Poland, Opole province, Pomerania.
- the second group included the regions: Northwest, Central Moravia, Moravia-Silesia, Central Transdanubia, Western Transdanubia, Southern Transdanubia, Northern Hungary, Northern Great Plain, Southern Great Plain, Lesser Poland, Lublin province, Subcarpathian province, Świętokrzyskie, Podlasie, Lubusz province, Opole province, Kuyavian-Pomeranian province, Warmia-Masuria, East Slovakia.
- Bratislava

The data analysis according to OECD methodology made it possible to notice that the first of the analysed groups, where the capitals of three countries from the presented groups are located, is higher. The lowest quality of life is in the eastern part of Poland and other peripheral areas.

The presented methods of measuring the standard of living of residents in territorial units can be used to monitor the level of this phenomenon, while taking into account several diagnostic variables. The institutions and organizations dealing with social policy and territorial authorities may be the recipients of such analyses. The results of the conducted analyses can be taken into account in the process of striving for achieving social, economic and territorial cohesion, e.g., during activities aimed at helping regions in which the standard of living is the lowest. They can also be used to determine the future goals of the V4 group.

References

[1] Betti, G., Soldi, R. and Talev, I. (2015). *Fuzzy multidimensional indicators of quality of life: the empirical case of Macedonia*, Social Indicators Research.

[2] Borys, T., Rogala, P. (ed), Brzozowski, T., Kusterka, M. and Rycharski T. (2008). *Quality of Life on local level. An indicators based study. United Nations Development Programme*, Warsaw: UNDP Polska. (in Polish).

[3] Chrzanowska, M., Dudek, H. and Landmesser, J. (2017). Standard of living in selected countries in Central Europe in the period 2000-2014, In: *Hradec Economic Days*, Hradec Kralove: University of Hradec Králové, 7(1), pp. 325-331.

[4] Drejerska, N. (2016). Are Polish rural areas destinations for commuting? In: *Economic Science for Rural Development 2016*, Jelgava: Latvia University of Life Science and Economics, 41, pp. 53-60.

[5] Dudek, H. and Szczesny, W. (2017). Correlates of Multidimensional Indicator of Quality of Life; Fractional Outcome Model Approach. *Statistica: Statistics and Economy Journal*, 97(4), 46-61.
[6] Hartigan, J. (1975). *Clustering algorithms,* New York: John Wiley& Sons, Inc.
[7] Jankowska, E. (2011). Concept and Tools of Measuring of the Quality of Life, *Toruńskie Studia Międzynarodowe,* 1(4), pp. 33-39 (in Polish).
[8] Jankowska, E. (2012). Zastosowanie wskaźnika rozwoju społecznego HDI do oceny sytuacji społeczno-gospodarczej krajów Unii Europejskiej w latach 2004-2011. Toruń: *Toruńskie Studia Międzynarodowe*, 1(5), pp. 33-42.
[9] Kruskal, W. and Wallis, W. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), pp. 583–621.
[10] Michnak, D. (2016). Main Trends in Commuting in Slovakia. *European Journal of Geography*, 7(2:6), pp. 6-20.
[11] How's Life in Your Region? Measuring Regional and Local Well-being for *Policy Making*. (2014). Paris: OECD.
[12] Panek, T. (2015). Hierarchiczny model pomiaru jakości życia, *Wiadomości Statystyczne,* 6, pp. 1-22.
[13] Somarriba, N. and Pena, B. (2008). Quality of life and subjective welfare in Europe: an econometric analysis. *Applied Econometrics and International Development*, 8(2), pp. 55-66.
[14] Stiglitz J., Sen, E. and Fitoussi, J.P. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress, [online]. Available at: http://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultstronaopisowa/5993/1/1/raport_stiglitza-en.pdf [Accessed 10 Jan. 2018].
[15] Szatur-Jaworska, B. (2009). Diagnoza i diagnozowanie w polityce społecznej. In: G. Firlit-Fesnak and M. Szylko-Skoczny, ed., *Polityka społeczna*, Warszawa: PWN, pp. 109-126 (in Polish).
[16] Trzebiatowski, J. (2011). Jakość życia w perspektywie nauk społecznych i medycznych – systematyzacja ujęć definicyjnych*.* *Hygeia Public Health*, 46(1), pp. 25-31 (in Polish).
[17] Ward, J.H., Jr. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), pp. 236-244.

Author's address

Mariola Chrzanowska, PhD.
Warsaw University of Life Sciences – SGGW, Faculty of Applied Informatics and Mathematics, Department of Econometrics and Statistics
Nowoursynowska 159, 02-776 Warsaw
Poland
email: mariola_chrzanowska@sggw.pl

Joanna Landmesser, dr hab.
Warsaw University of Life Sciences – SGGW, Faculty of Applied Informatics and Mathematics, Department of Econometrics and Statistics
Nowoursynowska 159, 02-776 Warsaw
Poland
email: joanna_landmesser@sggw.pl

Monika Zielińska-Sitkiewicz, PhD.
Warsaw University of Life Sciences – SGGW, Faculty of Applied Informatics and Mathematics, Department of Econometrics and Statistics
Nowoursynowska 159, 02-776 Warsaw
Poland
email: monika_zielinska_sitkiewicz@sggw.pl

JOB-FINDING AND SEPARATION RATES IN EUROPE

Ondřej Čížek,
University of Economics, Prague

Abstract
Job-finding and separation rates are calculated for many European countries and their changes are analyzed in the context of the current economic crisis. The goal is to shed some light on questions whether an increase in unemployment is caused by low job-finding rate or whether it is due to high separation rate. Steady-state unemployment rate consistent with balanced inflows and outflows is also calculated and compared to the observed unemployment rate to draw some conclusions about future development of unemployment rate. Analysis is performed for many individual European countries because of their heterogeneity. Differences between individual European countries are analyzed and discussed.

Keywords: *unemployment rate, job-finding rate, separation rate, economic crisis*

JEL Classification: E24, J64
AMS Classification: 91B40, 91B84

1 INTRODUCTION

Changes in unemployment are driven by a continuous process of job creation and job destruction. Analysis of unemployment is, therefore, in this paper based on job-finding and separation rates and is performed for many individual European economies in the context of the current economic crisis. European countries were hit by severe shocks in 2008 and 2009. Output losses were huge in all countries and expectations of economic trade declined tremendously. Nonetheless, labor market response was different across European countries. On one hand, Greece and Spain experienced sharp increase in unemployment. On the other hand, Germany is characterized by a decline in unemployment. These changes in unemployment rate are analyzed by studying job-finding and separation rates as the main forces behind this development. Similar earlier empirical analysis along these lines was performed by Arpaia and Curci (2010).

2 MEASURES OF TRANSITION RATES

Shimer's (2012) methodology is adopted in order to calculate job-finding and separation rate. The model of unemployment is in continuous time, but data are measured at discrete dates. The time interval $[t, t+1)$ is referred to as the period t. It is assumed that during period t, all unemployed workers find a job according to a Poisson process with arrival rate $f_t \equiv -\ln(1-F_t) \geq 0$, where F_t is the corresponding job finding probability.[1] Similarly, all employed workers lose their job according to a Poisson process with arrival rate $s_t \equiv -\ln(1-S_t) \geq 0$, where S_t is the corresponding employment exit probability.

The first goal is to calculate F_t and S_t from commonly available data. Short-term unemployment will be introduced in order to calculate F_t. Specifically, $U_t^s(\tau)$ will represent

[1] The difference between the rate and the corresponding probability is that the probability relates to the discrete time while the rate to the continuous time.

the number of workers who are unemployed at time $t+\tau$ but were employed at some time $t' \in [t, t+\tau]$, $\tau \in [0,1]$. Unemployment and short-term unemployment evolves according to

$$\dot{U}(t+\tau) = E(t+\tau)\cdot s_t - U(t+\tau)\cdot f_t, \tag{1}$$

$$\dot{U}_t^s(\tau) = E(t+\tau)\cdot s_t - U_t^s(\tau)\cdot f_t, \tag{2}$$

where $E(t+\tau)$, $U(t+\tau)$ represents number of employed and unemployed workers at time instant $t+\tau$. Note that t is fixed in these equations and that $\tau \in [0,1]$.

To solve for F_t, eliminate $E(t+\tau)\cdot s_t$ between these two equations, which gives $\dot{U}(t+\tau) - \dot{U}_t^s(\tau) = -\left(U(t+\tau) - U_t^s(\tau)\right)\cdot f_t$. The solution to this simple differential equation is given by $U(t+\tau) - U_t^s(\tau) = K\cdot e^{-f_t\cdot\tau}$. Because $U_t^s(0) = 0$ from definition, we have $U(t) = K$. For $\tau = 1$ we obtain $U(t+1) - U_t^s(1) = U(t)\cdot e^{-f_t}$, or[2] $U_{t+1} - U_t^s(1) = U_t\cdot(1-F_t)$. Simple algebraic manipulation leads to

$$F_t = 1 - \frac{U_{t+1} - U_t^s(1)}{U_t}. \tag{3}$$

This equation serves to calculate F_t using statistical data for unemployment U_t and short-term unemployed $U_t^s(1)$. The job-finding rate is then calculated as $f_t = -\ln(1-F_t)$.

Furthermore, solving differential equation (1) leads to

$$U_{t+1} = \frac{s_t}{s_t + f_t}\cdot\left(1 - e^{-(s_t+f_t)}\right)\cdot L_t + e^{-(s_t+f_t)}\cdot U_t, \tag{4}$$

where $L_t \equiv E_t + U_t$ represent the labor force. Job separation rate s_t is obtained from solving the equation (4) numerically and the corresponding probability is given by $S_t = 1 - e^{-s_t}$.

3 UNEMPLOYMENT RATE DYNAMICS

Under the assumption of constant labor force $L_t = L$, unemployment rate evolves according to

$$u_{t+1} = \frac{s_t}{s_t + f_t}\cdot\left(1 - e^{-(s_t+f_t)}\right) + e^{-(s_t+f_t)}\cdot u_t, \tag{5}$$

where $u_t = U_t / L$ is unemployment rate.

Steady-state unemployment rate satisfying $u_t = u_{t+1}$ is given by

$$u_t^* = \frac{s_t}{s_t + f_t}. \tag{6}$$

Equation (5) can thus be rewritten as follows

$$u_{t+1} = (1-\lambda_t)\cdot u_t^* + \lambda_t\cdot u_t, \tag{7}$$

where $\lambda_t = e^{-(s_t+f_t)}$ is a persistence of unemployment rate and $(1-\lambda_t)$ measures the speed of convergence to a steady state.

---

[2] Using the convention to denote time in the lower subscript when measured in discrete time.

4 DATA

The data for individual European countries were obtained from Eurostat (2017). Seasonal adjustment was performed in Eviews 8 (method Census X-13) in cases where the adjustment was not performed by the Eurostat. Job-finding and separation rates were calculated for Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden and United Kingdom by the method described in chapter 2. Empirical counterpart to the variable $U_t^s(1)$ is the number of workers (1000 persons) unemployed at the beginning of the quarter t, whose unemployment has not exceeded one quarter (three months).[3]

5 EMPIRICAL ANALYSIS

Empirical analysis is performed for three different periods:[4]

- pre-crisis period: before 2008,
- crisis period: from 2008 Q1 to 2009 Q4,
- post-crisis period: after 2009.

Arithmetic mean of relevant variables is calculated for each period in order to assess labor market reaction to the crisis. Firstly, unemployment rate is studied in relation to its steady state value. Changes in unemployment are then explained by movements in transition rates.

5.1 Unemployment rate and its steady state

Unemployment rate for three different time periods is shown at the following figure.

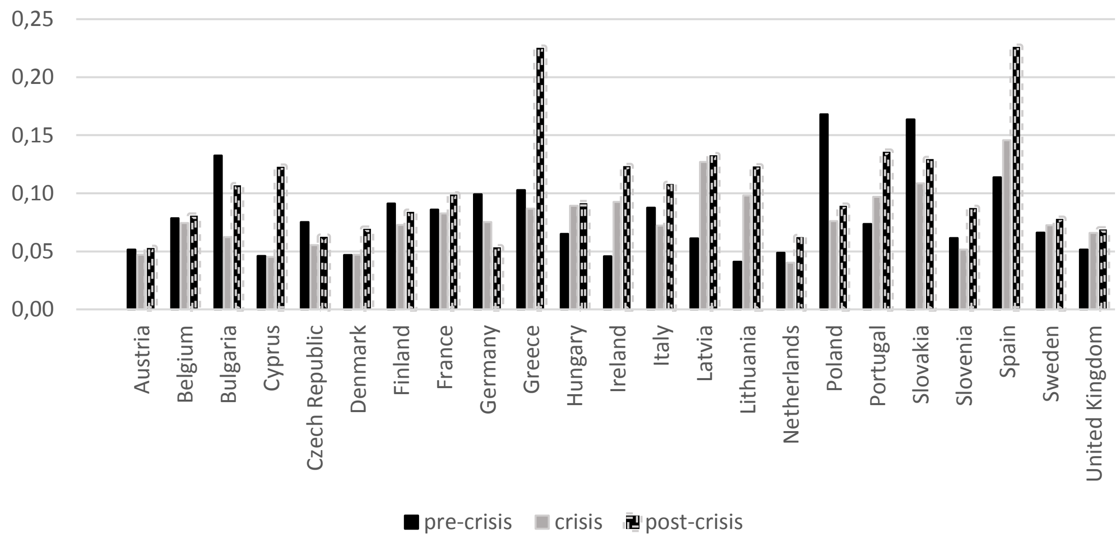


Figure 1: Unemployment rate for European countries: arithmetic mean for pre-crisis (before 2008), crisis (from 2008 Q1 to 2009 Q4) and post crisis (after 2009) periods.

---

[3] The name of this time series in the Eurostat database is 'Unemployment by sex, age and duration of unemployment [lfsq_ugad]'.

[4] Post-crisis period ends in 2016 Q3 for all countries. The beginning of the pre-crisis period differs for different countries due to data availability and is given as follows: Austria (2002 Q4), Belgium (1998 Q4), Bulgaria (2000 Q1), Cyprus (2004 Q1), Czech Republic (1997 Q4), Denmark (1998 Q4), Finland (1997 Q4), France (2002 Q4), Germany (2004 Q4), Greece (1998 Q2), Hungary (1998 Q4), Ireland (2005 Q4), Italy (1997 Q4), Latvia (2006 Q4), Lithuania (2007 Q4), Netherlands (2002 Q1), Poland (1999 Q4), Portugal (2000 Q2), Slovakia (1998 Q1), Slovenia (2001 Q2), Spain (1997 Q4), Sweden (2006 Q1), United Kingdom (1999 Q1).

Significant rise in unemployment rate in post-crisis period is detected for Greece and Spain where it is now above 20 %. An increase in unemployment rate is also detected for Cyprus, Ireland, Latvia, Lithuania and Portugal. Germany is on the other side as unemployment rate declined systematically for this country.

Steady-state unemployment rate, which is consistent with balanced inflows and outflows out of unemployment and is calculated according to the relation (6), is shown at the following figure.

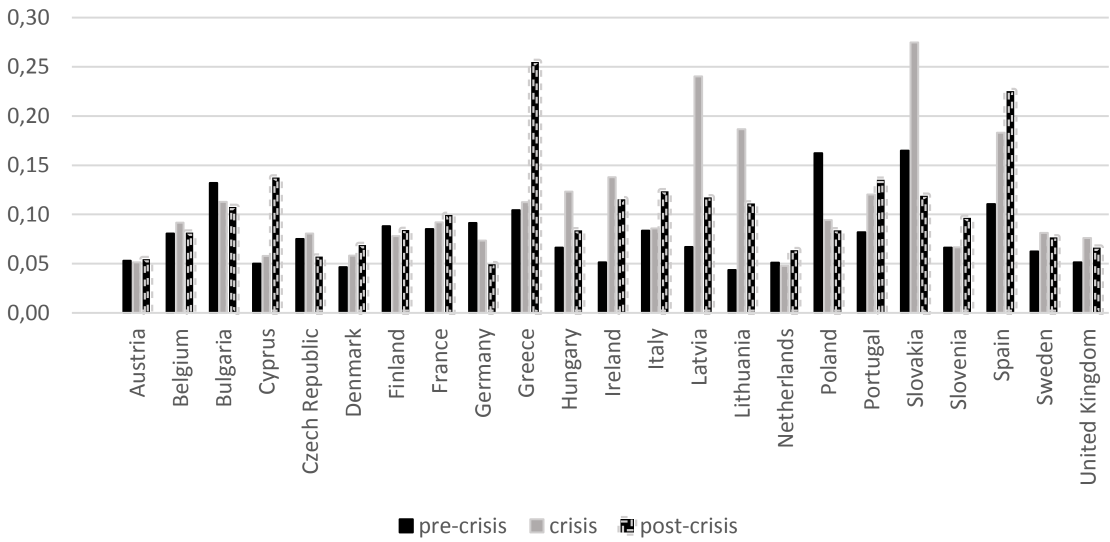


Figure 2: Steady-state unemployment rate for European countries: arithmetic mean for pre-crisis (before 2008), crisis (from 2008 Q1 to 2009 Q4) and post crisis (after 2009) periods.

Steady state unemployment rate experienced significant changes in many countries. It more than doubled in Cyprus and Greece where it is even higher than observed unemployment rate. Unemployment rate in these two countries is therefore expected to rise in the future as it will converge to its steady state value. Steady state systematically increased also in Portugal and Spain. Significant increase of the steady state was detected for many countries in the crisis period – Hungary, Ireland, Latvia, Lithuania and Slovakia. Nonetheless, the rise of the steady state in these countries was only temporary and it declined in the post-crisis period. For this reason, the corresponding rise in observed unemployment rate has not materialized in the post-crisis period in such a huge extent. Systematic decrease of the steady state is observed for Germany where it closely corresponds to the observed unemployment rate.

5.2 Transition rates

Job-finding and separation rate for selected European countries is depicted at the following figures.

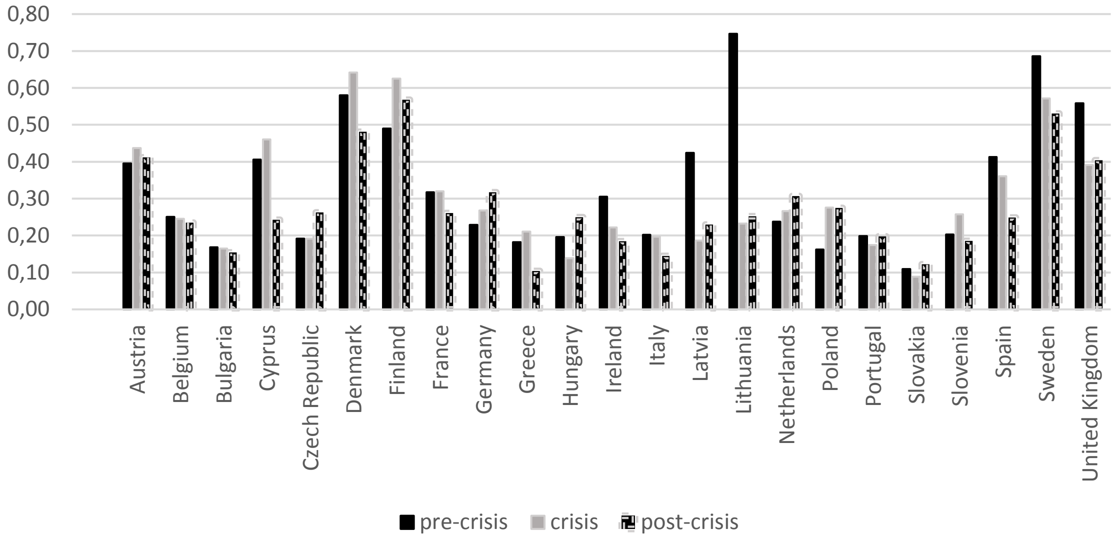


Figure 3: Job-finding rate for European countries: arithmetic mean for pre-crisis (before 2008), crisis (from 2008 Q1 to 2009 Q4) and post crisis (after 2009) periods.

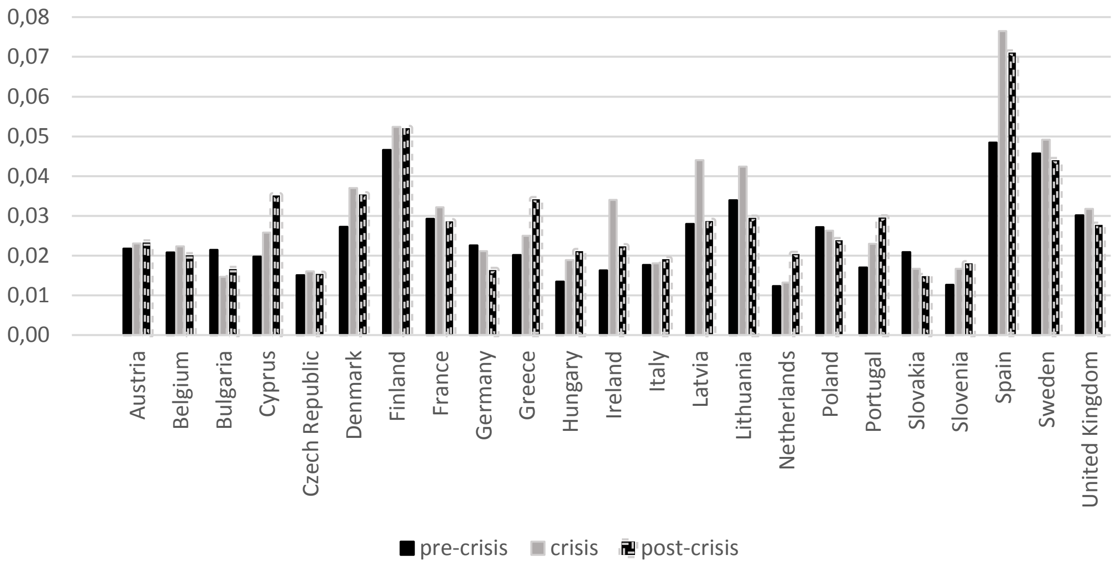


Figure 4: Separation rate for European countries: arithmetic mean for pre-crisis (before 2008), crisis (from 2008 Q1 to 2009 Q4) and post crisis (after 2009) periods.

There was a sharp decrease in job-finding rate in Latvia and Lithuania in the crisis period which caused the above mentioned contemporary increase in the steady-state value of unemployment rate. It was only temporary because separation rate decreased in the post crisis period in these two countries.

The most vulnerable labor markets are characterized by systematic decrease of the job-finding rate accompanied by a systematic increase of the separation rate. This is the case of Cyprus, Greece and Spain. It was this development of transition rates that caused the above mentioned sharp increase of the steady-state unemployment rate in these countries. Note that Cyprus and Spain have comparable job-finding rate. Much higher steady-state unemployment rate in Spain compared to Cyprus is thus caused by higher values of separation rate in Spain.

Germany again stands on the other side as this economy is characterized by a systematic increase of the job-finding rate as well as a systematic decrease of the separation rate. Other countries stand between these two extremes. An increase in unemployment rate is caused by rising separation rate in Hungary, Netherlands and Portugal and by decreasing job-finding rate in Ireland, Latvia, Lithuania, Sweden and United Kingdom.

It is clear from figures 3 and 4 that countries with high job-finding rate also have increased separation rate. Countries with high transition rates $\left(s_t + f_t\right)$ are especially Austria, Cyprus, Denmark, Finland, Spain, Sweden and United Kingdom. These countries have lower persistence of unemployment rate $\lambda_t = e^{-\left(s_t + f_t\right)}$ and increased convergence to a steady-state unemployment rate $\left(1-\lambda_t\right)$.

6 CONCLUSION

The paper examined many individual European labor markets in the context of the current economic crisis using the commonly applied measure of job-finding and separation rate. Empirical analysis revealed heterogeneous patterns between individual countries.

Latvia and Lithuania experienced dramatic decrease in job-finding rate accompanied by an increase in separation rate in the crisis period from 2008 to 2009. Nonetheless, this negative development did not last and separation rate was even improved in the post-crisis period. The initial increase of the steady-state unemployment rate was thus only contemporary and the threat of huge increase of unemployment was finally warded off.

However, this was not the case of Cyprus, Greece and Spain were job-finding rate decreased in the post-crisis period together with an increase of separation rate. Steady-state unemployment rate more than doubled in these three countries in the post-crisis period compared to the pre-crisis period. Furthermore, comparison of steady state with observed unemployment rate reveals that further increase in unemployment rate is expected for Cyprus and Greece as it will converge to its steady-state value.

Germany stands on the other side as job-finding rate increased together with a decrease of separation rate. The remaining countries stand between these two extremes. Hungary, Netherlands and Portugal is characterized by rising separation rate while for Ireland, Latvia, Lithuania, Sweden and United Kingdom is typical decreasing job-finding rate.

Empirical analysis performed in this paper could be extended in two main dimensions. Macroeconomic determinants of transition rates could be examined as by Kadeřábková and Maleček (2015). Cyclical properties of transition rates could be studied as by Fujita and Nakajima (2016), Shimer (2005) or Hairault et al. (2015).

Acknowledgements: This work was processed with contribution of long term institutional support of research activities by Faculty of Informatics and Statistics, University of Economics, Prague (IP400040) and it was supported by the Internal Grant Agency of University of Economics, Prague under Grant F4/21/2018.

References

[1] Arpaia, A., and Curci, N. (2010). EU Labour Market Behaviour during the Great Recession. [Online]. Directorate General Economic and Monetary Affairs, European Commission. *European Economy – Economic Papers*, 405. Available at: http://www.forschungsnetzwerk.at/downloadpub/ecp405_en.pdf

[2] Eurostat (2017). *Database–Eurostat*. [Online] Available at: http://ec.europa.eu/eurostat [Accessed 2 July 2017].

[3] Fujita, S., and Nakajima, M. (2016). Worker Flows and Job Flows: A Quantitative Investigation. *Review of Economic Dynamics*, 22, pp. 1-20.

[4] Hairault, J. O., Le Barbanchon, T., and Sopraseuth, T. (2015). The Cyclicality of the Separation and Job Finding Rates in France. *European Economic Review*, 76, pp. 60-84.

[5] Kadeřábková, B., and Maleček, P. (2015). Churning and Labour Market Flows in the New EU Member States. *Procedia Economics and Finance*, 30, pp. 372-378.

[6] Shimer, R. (2005). The cyclicality of hires, separations, and job-to-job transitions. *Federal Reserve Bank of St. Louis Review*, 87(4), pp. 493-508.

[7] Shimer, R. (2012). Reassessing the Ins and Outs of Unemployment. *Review of Economic Dynamics*, 15(2), pp. 127-148.

Author's address

Ing. Ondřej Čížek, Ph.D.
University of Economics in Prague, Faculty of Informatics and Statistics, Department of Econometrics
W. Churchill Sq. 4, 130 67 Prague 3
Czech Republic
email: cizeko@vse.cz

DISCRETE CHOICE WITH COMPLEMENTARY UTILITY FUNCTION AND FRAMING EFFECT

Juraj Dubovec, Department of Management Theories, Faculty of Management Science and Informatics, University of Žilina,
Lucia Pančíková, Department of Macro and Microeconomy, Faculty of Management Science and Informatics, University of Žilina
Lukáš Falát, Department of Macro and Microeconomy, Faculty of Management Science and Informatics, University of Žilina

Abstract
The paper deals with the problem of effective use of integrated transport system in Bratislava region, which presumes decline of individual transport in the city and its suburbs and to improve the transport situation in the city. At first the process of discrete choice whether to use public or private transport is discussed. Afterwards, this process is extended by knowledge from behavioral economics. Authors then suggest concrete ways how to solve some drawbacks of this system. The first one is the concept of complementarity of public and private transport. Another way authors state to improve this system is to use the framing effect in the sales process. In the end, some other concepts for making this system more efficient are outlined by authors.

Keywords: *transport, selection, behavioral economics, endowment effect, framing, complements*

JEL Classification: J6, D4

1 INTRODUCTION

The combination of changes in prices of transport, rising of incomes and changes in lifestyle cause a significant decline of interest of public transport. In the same time demand after individual transport is rising. This fact causes problems on transport infrastructure which is becoming insufficient in Slovakia. One of possible solution to this problem is so-called integrated transport system (ITS).

Integrated transport system (ITS) is the system of transport operation of a given location provided by public transport. It consists of multiple types of transports and transport lines of provided by multiple operators. Passengers are able to use unified travel ticket which is valid no matter what the transport operator and transport mean is. There is an assumption that when this is done correctly (interconnection of transport lines) all sites of the transport system can make a profit from this ITS system. This is mainly true for commuters, for which the system was created – actually, the system was made to satisfy transport needs in a sustainable way.

On June 1st, 2013 the first phase of the Integrated Transport System was initiated in Bratislava region (IDS BK). In this first phase tariff integration in public transport in Bratislava as well as in Zahorie region was performed. Three operators which were part of this phase were as follows: Dopravný podnik Bratislava, Bus operator Slovak Lines and Železničná spoločnosť Slovensko (railway operator).
ITS Bratislava used the system in which the area is divided into zones. The ticket price is dependent through how many zones passenger travels (regardless the mean of transport). Figure 1 illustrates the Bratislava ITS system.

Unfortunately, this system exhibits some drawbacks which decrease the attractivity of the system in Bratislava. The first problem is the fact that the price ticket is dependent on the zones traveled by a passenger. Integrated area is divided into tariff zones (every zone has its three-digit number) what makes the system blind, complex and difficult to understand (f. ex. in Brno, there is only one ticket for all transport means, there is only one tariff and moreover, these transport means cooperate with each other). In Bratislava, the transport means do not cooperate with each other in an efficient way. Therefore, it is no surprise a passenger has to wait for a connection. It is therefore necessary so that the system would be optimized from the site of all stakeholders.

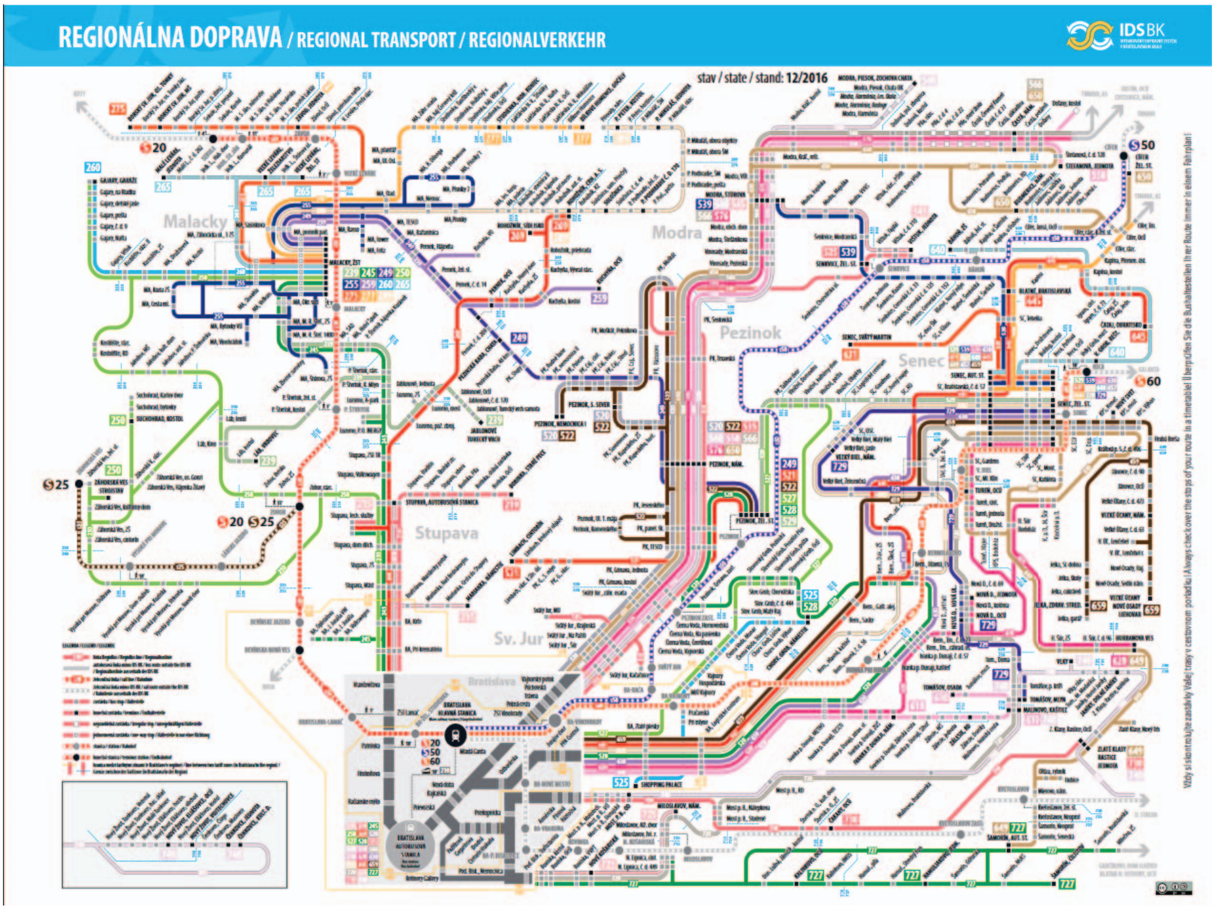


Fig.1 The network of Bratislava ITS transport lines (IDS BK)

2 DECISION-MAKING PROCESS OF THE PASSENGER

Observation, determination of the human behavior, and finding out how people decide, is one of the main objectives of social science. This is also relevant to services such as transport (Domencich and McFadden, 1975; de Dios Ortúza and Willumsen, 2001). It is therefore no surprise that to make a proper strategy for integrated transport system it is necessary to understand the human behavior in this decision-making process first. In other words, in order to make the ITS working, we need to understand the decision-making process of the commuter (economic actor), i.e. why so many people choose the private car option.

Ben-Akiva and Lerman (1985) define an election process as the series or steps with a result. This sequential decision-making process includes, according to authors, the following steps: definition of the problem of choice, generation of alternatives, evaluation of the attributes of each alternative, choice, and application of choice.

As for the commuting problem in the area of Bratislava, capital of Slovakia, there are multiple options available to the commuter (dependent on the exact location) - public transport, train (two different operators), regular bus, trolleybus, tram or private car, private bicycle. In each of these options, the economic actor (commuter) considers multiple factors to make the final decision. This decision is considered to be the best option.

The mobility of people is a complex phenomenon, due to the large number of factors that influence the decisions of individuals (Muro-Rodríguez et al., 2017). There are lots of factors which determine the selection process of the passenger such as age, sex, level of income etc. However, we consider the following three factors as the most important in the selection process (Muro-Rodríguez et al., 2017)

- travel cost – as for public transport this is the price paid for each of the means of transport; as for the private transport (car) this is the the cost of gasoline (based on the kilometers traveled), including the cost of the amortization of the vehicle (price of car, insurance, maintenance, etc.). Nowadays, parking in the centre of Bratislava, as well as in other big cities, is a big problem. Therefore as for car the cost of parking whould be also included in travel costs.
- travel time – is the total time of the trip of the commuter. The total travel time is one of the most important factors that the traveler takes into account for choosing a mode of transport (Steer Davies Gleave, 2006). It is therefore a major factor to increase the demand, as travel time is reduced (IATA, 2003).
- comfort – this factor can be understood in two ways. The first way how to understand comfort is the quality of the transport mean itself. Of course this is a qualitative indicator but also has very much to do with the selection process. Another way how to look at this indicator is through number of transfers the passenger has to make.

3 NEOCLASSICAL APPROACH

Decision-making process, regarding the selection of transport mean for commuting purpose, is an activity which enables, on base of multiple options, to choose an alternative which is an appropriate way between given locations for a passenger. This process comes out from the need to make the best selection from all possible alternatives in the space of all possible options of selection.

To solve this problem, there are multiple options. The approach used by neoclassical economics is so-called the standard to do this. „It uses consumer theory which is the economic modeling of the behavior of an economic agent that consumes goods and services. This theory relates to preferences, indifference curves and budget constraints to consumer demand curves“ (Muro-Rodríguez et al., 2017). The goal of consumer theory is to analyze individual consumer behavior using the utility function. In other words, the theory is based on the fact that the individual chooses an amount of goods, which form their basket of goods, in order to maximize their utility which translates into their level of satisfaction, subject to their income restriction (Torres, 2008). The maximization is done regarding preferences of the economic actor. Given a consumer with defined preferences, there is a deterministic utility function, which represents these preferences that will be ordered so that the individual will choose the alternative that gives him greater (maximal) satisfaction (Becker, 1965).

The selection of the transport mean is, according to knowledge of standard microeconomics (Dubovec, 2015), actually a substitution between transport means where u is the utility function which states that the utility of the individual is dependent on the amount of goods consumed as well as on their preferences (as we presented in the previous section, in this case, factors influencing preferences, i.e. the selection process or the results of the utility function, are mainly travel time, the cost or price of the mean of transport and the comfort, i.e. convenience). Formally, this utility function can be defined as

$$\mathrm{u} = \mathrm{U}(\mathrm{X}, \mathrm{Y}, \mathrm{Z}, \dots) \tag{1}$$

where *x, y* and *z* are amounts of goods

Preferences for substitutions can be expressed by the utility function

$$\mathrm{U}(\mathrm{X}, Y) = aX + bY \tag{2}$$

where *a, b*, are positive numbers, which determine the value of substitution options for given consumer.

In the consumer theory the marginal rate of technical substitution (MRTS) determined the amount of one goods (X_1) which compensates the loss of another goods (X_2). MRTS is sometimes denoted as the personal rate of substitution or the rate of commodity substitution. In this case, consumer stays on the same level of utility, on which he was originally (as both combinations are located on the same indifference curve). In the space of indifference curves the marginal rate of substitution fits to the slope of indifference curve which is equal to the ratio of marginal utilities of given goods, i.e.

$$MRTS = \frac{MU_X}{MU_Y} \tag{3}$$

where

$$MU_X = \frac{\partial U(X,Y)}{\partial X} \tag{4}$$

$$MU_Y = \frac{\partial U(X,Y)}{\partial Y} \tag{5}$$

However, it is obvious that in our problem, in case of substitution effect there must a law of exclusive OR, i.e.

$$(a = 0\ \&\ b = 1) OR\ (a = 1\ \&\ b = 0) \tag{6}$$

4 SUGGESTED APPROACH

Even though most transport choice models have been based on the approach of utility maximization with an assumption of the rational agent, other decision rules have been tried (Cantillo and Ortúzar, 2005). Sælensminde (2002) and Rouwendal and Blaeij (2004) used the concepts of declared preferences, which do not follow the principles of maximum utility. We applied. In this section we make suggestion on base of behavioral economics in order to make the Bratislava ITS working better. To do this we combine elements from neoclassical economics as well as from the behavioral economics; we use the effect of framing.

What is very important to state is the fact that in neoclassical economics, the model is based on the assumption of a rational consumer that is supposed to represent the behavior of the average consumer. However, as Kahneman and Tversky (1979) found out, people are not always rational. Decisions which people realize in the lives are not always optimal from the theory of maximalization of utility. Often, their decisions are influenced by external factors such as situation, environment as well as by internal factors (psychic, emotions). The influence of these factors in the decision-making process causes errors. However, what is more important is that these errors are not random, rather they are systematic what is a big deal. These errors are being created due to the fact that people use heuristics in the decision-making process. The heuristics are so-called mental shortcut which make this process faster. Behavioral economics which is a new field in economics tries to deal with these concepts and tries to explain human behavior on base of psychological motives and heuristics.

In our recent researches (research grants APVV-14-0658, VEGA 1/0582/16) we found out that attributes, which influence the discrete choice between public transport and individual private transport, are following ones:

- the sequence of transport lines
- unified transport and tariff terms
- quality of provided transport services
- increased mobility and the change in lifestyle
- property costs
- way how people perceive risk
- customs and habits
- values people assign to subjects and events

This knowledge widens the analytical framework for modelling the discrete choice. The use of behavioral economics is a new approach compared to traditional models of discrete choice. This behavioral approach enables to extend the standard approach.
Actually, the problem in reality is that some negative aspects of commuting are not perceived so negatively as they should be according to neoclassical economics. This is mainly true for the time loss due to morning traffic jams. In selected municipalities of the functional city region of Bratislava the morning loss due to traffic jams ranges from 12 to 35 minutes (Tóth, 2012). If we estimated the daily loss caused by traffic jams of every participant of the road transport to 20 minutes, it would be more than 6 hours a month (20 business days) at morning commuting. The question is, is this time loss for a public transport commuter and individual motorist perceived equally?

The answer to this question is no. The reason for that is the endowment effect which is known from the field of behavioral economics (Tversky and Shafir, 1992). The essence of this effect is that people tend to overestimate things they possess. The same is true in possessing a car; in this case consumer overestimates its price and advantages and he does not want to prefer the public transport. Therefore, it is important to use complementarity between car and parking, when complementary utilities are connected. In this case the endowment effect of the consumer will stay valid.

We suggest the following case: individual transport and ITS system should be in complementary ratio of utilities. This approach can significantly influence the decision-making process of the consumer (commuter). The meeting point for making this complementary effect possible can be for example to realize parking very near the bus or train stop. In this case the passenger is interested in combination not the substitution of utilities. For complements, the function of utility is defined as follow

$$U(X, Y) = \min(aX, bY) \tag{7}$$

where *a, b*, are positive numbers, which illustrated the ration in which the goods are consumed.
Let us suppose that such a transport policy focused on complementing private transport with public transport, benefits, on the one hand, users who will have shorter travel times and in general, the transport system to reduce the congestion (Li and Hensher, 2012).

Another consequence of endowment effect is the fact that passengers perceive costs for possessing and sunk costs in a different way. On base of behavioral economics, it is possible to state the economic actor (commuter in this case) does not decide optimally, i.e. he does not take into account amortization of the vehicle. Commuters, in this case, use a heuristic (fast decision-making process defined as System 1 by Daniel Kahneman) for making decision. In this case the commuter does not take into account nothing but the price of petrol. The passenger tends to evaluate costs from possessing as sunk costs and therefore does not take them into the price of commuting. Such biased decision then affects incorrect evaluation of transport jam performed by car owners.

Regarding what was stated above, we suggest offering the sale of ITS tickets as one-year tickets. Kahneman et al. (1991) showed that if payments for consumption are separated from the consumption itself, travelers perceive the spending of money as an investment which they do not consider later in the process of consumption. This advance payment causes the same effect in consumption as if someone pays in advance for electricity; he does not count every night how much electricity he spends by watching television.

To make the sale process as effective as possible we recommend using framing effect in this process. Possible implementation could be like this

- Classic sale of the one-year ITS ticket:
  - the price of one-year ITS ticket: 264.20 EUR

- The sale of one-year ITS ticket using framing effect
  - „Do you want to travel to work for only 36 cents per day? Use the ITS and buy one-year ticket! “

5 CONCLUSIONS

Nowadays, the interest in traveling by public transport is decreasing. However, the rise of private transport causes problem with transport infrastructure which is insufficient in Slovakia. Moreover, traffic jams are becoming very frequent. Bratislava region has decided to solve this problem by implementing so-called integrated transport system (ITS). This system presumes decline of individual transport in the city and its suburbs. The goal of this system is simple: to improve the transport situation in Bratislava region.

However, there are some issues of the Bratislava ITS in practice which decrease the attractiveness of this system in Bratislava. These issues include for example system complexity. Endowment effect is another problem of this concept. The consequence of this effect is the fact people do not always make rational choice which is optimal. Besides this they use a car even though it is not the optimal choice for them.

In this paper we suggested ways how to solve some drawbacks of this system. The first one is the concept of complementarity of public and private transport. This is supposed to shorter travel times and in general, the transport system to reduce the congestion. Another way how to improve this system is to use the framing effect in sales process.

Special attention should be also put to the problem shared economy which could very much decrease the intensity of vehicles by increasing the capacity and by interconnecting this system with payments for crossings route gates.

It would also be interesting to suggest the price relativity regarding the income of individual. From the field of behavioral economics, we know that people do not perceive the absolute price of goods, rather they perceive the relative price. Therefore, there is an assumption that correct setting of the ticket price for different income groups could also increase the total revenues coming from the ITS.

We believe that if this system is set up correctly (taking into account functionality, sales, marketing, interconnection of particular transport systems, parking space), all stakeholders will benefit from this integrated transport system in Bratislava region.

Acknowledgement: This work has been supported by the research grant APVV-14-0658 "Optimisation of urban and regional public personal transport".

References

[1] Becker, G. (1965). A theory of the allocation of time. *Economic Journal* 75, pp. 493-517.

[2] Ben-Akiva, M. and Lerman, S. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: MIT Press.

[3] Cantillo V. and Ortúzar J. d. D. (2005). A semi-compensatory discrete choice model with explicit attribute thresholds of perception. *Trans. Res. B* 39, pp. 641-657.

[4] de Dios Ortúza, D. and Willumsen, L. G. (2001) *Modelling Transport*, 3rd Edition. Chichester: John Wiley and Sons.

[5] Domencich, T. and AMcFadden, D. (1975) *Urban Travel Demand. A Behavioral Analysis*. Amsterdam: North-Holland.

[6] Dubovec, J. (2015) *Mikroekonómia: praktické aplikácie a prípadové štúdie*. 1. vyd. - Žilina: Žilinská univerzita.

[7] IATA (2003). Air/Rail Inter-Modality Study. Final Report, IATA International Air Transport Association Hounslow: Air Transport Consultancy Service.

[8] IDS BK. Mapy schémy. Available online at https://www.idsbk.sk/informacie/mapy-schemy

[9] Kahneman, D. (2011) *Thinking, fast and slow*. New York: Farrar, Straus and Giroux, 2011. 1st ed., 499 p.

[10] Kahneman, D., Knetsch, J. L. and Thaler, R. H. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *The Journal of Economic Perspectives*, 5(1), pp. 193-206.

[11] Kahneman, D. and Tversky, A. (1979) Prospect theory: An analysis of Decission under Risk. *Econometrica,* 47(2).

[12] Muro-Rodríguez, A. I., Perez-Jiménez, I. R. and Gutiérrez-Broncano, S. (2017). Consumer Behavior in the Choice of Mode of Transport: A Case Study in the Toledo-Madrid Corridor, *Frontier in Psychology*. 2017(8), pp. 1011.
[13] Li Z., Hensher D. A. (2012). Congestion charging and car use: a review of stated preference and opinion studies and market monitoring evidence. *Trans. Policy 20*, pp. 47-61.
[14] Rouwendal, J. and Blaeij, A. T. (2004). Inconsistent and Lexicographic Choices in Stated Preference Analysis. Tinbergen Institute Discussion Paper TI 2004-038/3.
[15] Sælensminde, K. (2002). The impact of choice inconsistencies in Stated Choice studies. *Environmental and Resource Economics* 23, pp. 403-420
[16] Steer Davies Gleave (2006). Air and Rail-Competition and Complementarity. European Commission DG TREN.
[17] Torres, O. (2008). El Problema de las Reglas de Elección Social en la Teoría de la Acción Colectiva. *Retribución, Justicia y Bienestar.* Tesis doctoral, Universidad de la Laguna.
[18] Tóth, V. (2012). Význam dopravných sektorov na príklade štúdia dennej dochádzky vo funkčnom mestskom regióne Bratislava. *Geographia Cassoviensis*, Vol. IV, 1/12. pp. 75-80.
[19] Tversky, A. and Shafir, E. (1992). Choice under conflict: The dynamics of deferred decision. *Psychological Science*, 3, 6 1992. pp. 358-361.

Authors' addresses

Ing. Juraj Dubovec, PhD.
University of Žilina, Faculty of Management Science and Informatics, Department of Management Theories
Univerzitná 8215/1, 010 26 Žilina
Slovak Republic
email: dubovec@fria.fri.uniza.sk

Ing. Lucia Pančíková, PhD.
University of Žilina, Faculty of Management Science and Informatics, Department of Macro and Microeconomy
Univerzitná 8215/1, 010 26 Žilina
Slovak Republic
email: lucia.pancikova@fri.uniza.sk

Ing. Lukáš Falát, PhD.
University of Žilina, Faculty of Management Science and Informatics, Department of Macro and Microeconomy
Univerzitná 8215/1, 010 26 Žilina
Slovak Republic
email: lukas.falat@fri.uniza.sk

SUBJECTIVE EQUIVALENCE SCALES FOR THE VISEGRAD GROUP COUNTRIES BASED ON MINIMUM INCOME QUESTION

Hanna Dudek, Warsaw University of Life Sciences – SGGW
Mariola Chrzanowska, Warsaw University of Life Sciences – SGGW

Abstract
In measurement of poverty and income inequality there is a need to compare income across different types of households. To make the well-being comparable across households of different sizes and composition, sometimes income per capita is used. However, such an approach assumes that all individuals' needs are the same in households and there are no economies of scale in consumption. Thus, to take into account these issues, equivalence scales could be applied. Equivalence scales are measures of the relative incomes required by households of different size and composition to attain a similar standard of living. A wide range of equivalence scales exist. In this study we estimate subjective equivalence scales based on MINQ (minimum income question). We choose this method mostly due to its acceptance among applied welfare researchers. The main aim of our study is to verify whether the equivalence scales in the Visegrad Group countries are the same as the OECD-modified scale. Our results based on EU-SILC data for 2015 indicate that the relative cost of a child in Poland exceeds the relative cost of an additional adult whereas in Hungary opposite relationship is observed. Moreover, in all the Visegrad Group countries stronger economies of scale than the OECD-modified scale assumes are observed.
Keywords: *equivalence scale, household, the Visegrad Group, poverty, estimation*

JEL Classification: C31, I32
AMS Classification: 62J05, 03H10, 91D10

1 INTRODUCTION

Equivalence scales are the basic tools to analyze household data for a meaningful income inequality and poverty analysis. They convert nominal incomes of heterogeneous households in comparable measures of well-being. The use of per capita income, which comes down to divide total household income by the number of household members, is not a good solution. Such an approach assumes that all individuals' needs are the same and that there are no economies of scale in consumption. This, in particular, means that children's and adults' needs are the same and two persons living together spend twice as much as a person living alone. In fact, the cost of children may differ from the cost of adults, moreover, housing and heating may be relatively fixed costs. Thus, the common approach is an employment of equivalence scales to make the adjustment for household size and composition, with household income being divided by equivalence scales to produce equivalized income. The equivalence scale, or in other words – the equivalent household size, indicates how much more money a household of a given size and composition needs, compared to a reference household, in order to have the same standard of living as a reference household. For example, if the reference household is a one-adult-member household, then it has an equivalence scale of 1.00. If the equivalence scale for a couple with one child is 1.80, then this implies that a couple with one child needs 180 percent of the income of one-adult-member household in order to enjoy the same standard of living.

In the literature there is the vast number of various methods of the estimation of equivalence scales (Deaton, 1997; Perali, 2003; Schröder, 2004; Betti et al., 2017). To the most frequently used approaches belong expert scales, scales based on consumption data and subjective scales (Schröder and Schmidt, 2005). The first ones are constructed on the basis of needs and market baskets defined by experts, the second ones are derived from household expenditures data and the last ones are estimated on the basis of households' perception of their needs.

Despite their importance and widespread use, there is no generally accepted method of the estimation of equivalence scales. Most of the research and policy recommendations rely on expert scales. Some of them are used in many countries by statistical offices as well as by academics. Among them, a primary example is the so-called OECD-modified scale (also called 50/30 scale). It assigns the first adult a value of one subsequent adult a weight of 0.5 and children below 14 a weight of 0.3. This scale was adopted by the Statistical Office of the European Union (Eurostat) in the late 1990s.

The aim of the study is to estimate the subjective equivalence scales for each of the four countries of the Visegrad Group (V4) and compare the results with the commonly used in the European Union the modified OECD scale. These V4 countries – the Czech Republic, Hungary, Poland and Slovakia – are closely linked Central European countries being once satellite states of the Soviet Union, all became EU members in 2004. An interesting issue is therefore a comparison of their equivalence scales estimated on the basis of empirical data. In the study we use the internationally harmonized European Union Statistics on Income and Living Conditions (EU-SILC) data. We estimate subjective equivalence scales based on MINQ (minimum income question). We choose this method due to its acceptance among applied welfare researchers (Bishop et al., 2014; Kalbarczyk-Stęclik, Miśta and Morawski, 2017). MINQ method explicitly focuses on households in poverty and does not impose any restrictions on the households' preferences and yields economically relevant results if the large datasets are used (Grodner and Salas, 2017).

What distinguishes our study is the application of our own model for the estimation of the subjective equivalence scales, focusing on the Visegrad Group countries and the use of the newest data from 2015.

2 DATA AND METHODOLOGY

The study uses EU-SILC cross-sectional data for the year 2015. The cross-sectional samples encompass 7914 households in the Czech Republic, 7710 households in Hungary, 5637 households in Slovakia and 12183 households in Poland. Unlike Bishop et al. (2014), Kalbarczyk-Stęclik, Miśta and Morawski (2017), Morawski, Kalbarczyk-Stęclik and Miśta (2017) we do not restrict the EU-SILC samples to the six most common household types (households with one to four adults, couples with a single child, and two-child households).

In our study we apply the approach employing minimum income question (MINQ). In this approach the data on answers to the question: "*In your opinion, what is the very lowest net monthly income that your household would have to have in order to make ends meet*?" is used. Assuming that households of various demographic types have the same utility level when their incomes are equal to the minimum needs income, one can obtain subjective equivalence scales based on MINQ (Bishop et al., 2014). Following Kapteyn, Kooreman and Willemse (1988) and Bishop et al. (2014) we apply the so-called intersection method based on the use of three steps.

1. The first step is to estimate the regression model:

$$y_{min} = f(y; x) + \varepsilon \tag{1}$$

where:

y_{min} – the answer to a question about the minimum needs income,
y – the income,
x – demographic characteristics,
ε – error term.

Rich people usually overstate minimum needs income and poor people underestimate it. Hence, the perception of the minimum needs income is distorted by the fact that actual income (y) is usually not equal the reported minimum level (y_{min}) (Goedhart et al., 1977). Thus, the intersection method postulates the existence of such an income level y^*_{min}. which is defined by

$$y^*_{min} = f(y^*_{min}; x) \tag{2}$$

Only households whose actual income (y) is equal to the minimum level (y^*_{min}) would answer MINQ correctly (Bishop et al., 2014)[1].

2. The second step is to solve the equation (2) due to y^*_{min}. This leads to achieving the subjective poverty line for households with demographic characteristics represented by x.
3. The third step is to obtain equivalence scales by dividing the subjective poverty threshold of a household of certain type by the subjective poverty threshold of a reference household, which is usually a single-member household.

Various specifications of (1) can be used. For example, minimum needs income (y_{min}) can be seen as a function of household income (y) and household size (n) (Goedhart et al., 1977):

$$ln(y_{min}) = \alpha_0 + \alpha_1 ln(y) + \alpha_2 ln(n) + \varepsilon \tag{3}$$

or it can be explained by household income and the set of household type variables (Bishop et al., 2014; Garner and Short, 2003; Kalbarczyk-Stęclik, Miśta and Morawski, 2017; Morawski, Kalbarczyk-Stęclik and Miśta, 2017):

$$\ln(y_{min}) = \alpha_0 + \alpha_1 ln(y) + \alpha_2 z_2 + \alpha_3 z_3 + \cdots + \alpha_k z_k + \varepsilon \tag{4}$$

where z_j denotes dummy variable representing j-th demographic type of household (j=1, 2,..,k).
We propose to use another model:

$$ln(y_{min}) = \alpha_0 + \alpha_y ln(y) + \theta ln(1 + \beta(adults - 1) + \gamma children) + \varepsilon \tag{5}$$

where adults denote the number of individuals who are at least 14 years old, children are the number of individuals below 14 years old, α_0, α_y, θ, β and γ are parameters to be estimated.
Taking into account a single-member household as a reference household, such a specification of regression models leads to the MINQ equivalence scale given by formula:

$$S = (1 + \beta(adults - 1) + \gamma children)^{\frac{\theta}{1-\alpha_y}} \tag{6}$$

In particular, if β=0.5, γ=0.3 and $\frac{\theta}{1-\alpha_y} = 1$, then scale (6) equals to the OECD-modified scale.
The advantages of our parametric scales are their construction, which makes it possible to establish a clear separation between adults and children needs, and the effect of the economies of scale. To get the estimates of the equivalence scales (6) we transform model (5) to the corresponding model (7):

$$y_{min} = exp\left(\alpha_0 + \alpha_y ln(y) + \theta ln(1 + \beta(adults - 1) + \gamma children)\right) + \eta \tag{7}$$

and we use nonlinear least squares method. In order to construct confidence intervals for equivalence scale (6) we apply the delta method (Cameron and Trivedi, 2009). All statistical analyses are done in Stata software.

[1] It is worth mentioning that the name of the intersection method originates from the observation that the solution of (2) can be graphically viewed as the intersection of regression line (1) and straight line $y_{min}=y$.

3 RESULTS AND DISCUSSION

Estimating the parameters of the model (7) for each country, we omitted such households for which we found missing data, households exhibiting non-positive value of income and households with at least 7 persons. The number of missed observations does not exceed 5% in each country. In table 1 we show results of the estimation.

Table 1: Results of estimation of model (7)

Parameter	The Czech Republic	Hungary	Poland	Slovakia
α_0	4.18 (0.05)	4.08 (0.05)	4.18 (0.05)	4.51 (0.10)
α_y	0.33 (0.01)	0.29 (0.01)	0.31 (0.01)	0.31 (0.02)
θ	0.26 (0.03)	0.37 (0.04)	0.12 (0.01)	0.34 (0.04)
β	1.02 (0.24)	0.81 (0.16)	6.04 (1.61)	1.19 (0.36)
γ	1.09 (0.29)	0.47 (0.10)	8.30 (2.83)	1.15 (0.37)
R^2	90.45%	89.74%	88.29%	83.06%

Source: Authors' calculations based on EU-SILC data. Standard errors in parentheses.

One can observe that all parameters are significant at least at 0.01 level and all models are well fitted to the data. Using formula (6) we report in table 2 equivalence scales for the 12 most common types of households in the V4 countries. We denote A1C0 – household type of 1 adult without children, A2C0 – 2 adults without children, A2C1 –2 adults with 1 child, A2C2 – 2 adults with 2 children and so on. For each equivalence scale we estimate 95% confidence intervals.

Table 2: Estimates of equivalence scale based on formula (6)

	Equivalence scales					Confidence intervals for equivalence scales			
HH type	CZ	HU	PL	SK	OECD	CZ	HU	PL	SK
A1C0	1	1	1	1	1	-	-	-	-
A2C0	1.32	1.36	1.41	1.47	1.5	1.28;1.37	1.32;1.41	1.36;1.46	1.36;1.57
A3C0	1.55	1.65	1.57	1.81	2	1.49;1.61	1.60;1.71	1.51;1.63	1.67;1.95
A4C0	1.74	1.90	1.68	2.10	2.5	1.67;1.81	1.84;1.97	1.61;1.74	1.94;2.26
A5C0	1.90	2.13	1.76	2.35	3	1.82;1.99	2.04;2.21	1.69;1.83	2.17;2.53
A1C1	1.34	1.22	1.48	1.45	1.3	1.27;1.40	1.18;1.27	1.40;1.56	1.33;1.58
A2C1	1.57	1.54	1.62	1.80	1.8	1.51;1.63	1.48;1.59	1.55;1.68	1.66;1.94
A2C2	1.76	1.69	1.74	2.08	2.1	1.69;1.83	1.63;1.76	1.67;1.81	1.91;2.25
A2C3	1.93	1.84	1.84	2.32	2.4	1.85;2.01	1.76;1.92	1.76;1.91	2.13;2.52
A3C1	1.75	1.80	1.71	2.09	2.3	1.69;1.82	1.74;1.86	1.65;1.78	1.93;2.24
A3C2	1.92	1.94	1.81	2.33	2.6	1.85;2.00	1.87;2.01	1.74;1.88	2.15;2.51
A4C1	1.91	2.04	1.79	2.34	2.8	1.84;1.99	1.96;2.11	1.72;1.86	2.17;2.51

Source: Authors' calculations based on EU-SILC data.

Our results indicate generally much more economies of scale than the OECD-modified scale. Taking into account the OECD-modified scale, the increased needs resulting from the additional person are independent of the size of the household the individual is being added. For example, if a single individual marries, then the increase in the household's needs, according to the

OECD-modified scale, is 50% of the required income for a single individual. Furthermore, if another adult joins the household of two adult, then the needs of the household again grow by 50% of the needs of the single person household. Thus, the absolute increase does not vary with the household's size. Taking into account our results, the addition of an adult to a single person household increases needs by 32% in the Czech Republic, by 36% in Hungary, by 41% in Poland and by 47% in Slovakia. However, if the next adult joins the household of two adult, the household's needs would increase by 23% of two-adult household in the Czech Republic, by 29% in Hungary, by 16% in Poland and by 34% in Slovakia. It means that the absolute increase in needed income falls with household's size reflecting economies of scale. The phenomenon of a strong economy of scale is even more evident when comparing three- adult households with four-adult households without children (A3C0 type with A4C0 type) and households of four adults with households of five adults (A4C0 type with A5C0 type).

The results presented in Tables 1 and 2 indicate that in the Czech Republic and in Slovakia the relative child's cost is almost the same as the relative cost of an additional adult. A completely different situation occurs in two other V4 countries. In Poland the relative cost of a child exceeds the relative cost of an additional adult, whereas in Hungary opposite relationship is observed. It is worth noting that particularly large differences in the estimated equivalence scales are observed with reference to the single-adult household with one child. These results can be partially explained by differences in the V4 countries regarding the direct state aid in respect of allowances for children and family benefits. According to the report prepared by PricewaterhouseCoopers (2015), the greatest public aid per child among the V4 countries in 2015 was in Hungary and the smallest – in Poland. In particular, allowances for children and family benefits in Hungary exceeded four times such allowances and benefits in Poland. However, it should be mentioned that in 2016 Poland introduced the child benefit 500+ program, therefore, it can be expected that the child's cost is no longer as high as in 2015 (Landmesser and Chrzanowska, 2017).

One can observe that the confidence intervals for Czech and Polish equivalence scales usually overlap whereas Slovak scales significantly differ from scales for the other V4 countries. Considering all types of households, it can be seen that Slovakia exhibits the highest scales. This result can be somewhat explained by the fact that the share of food expenditure in the Czech Republic and Poland is lower than in Hungary and Slovakia, which translates into worse economies of scale and standard of living in the latter countries (Dudek, 2014a; Dudek, 2014b). This issue, however, requires a careful analysis and should be the subject of further research. In particular, the importance of cultural dissimilarities and differences in relative prices for the determination of equivalence scales needs to be examined.

We compare our findings with results obtained by Bishop et al. (2014) and by Polish team composed of Kalbarczyk-Stęclik, Miśta and Morawski (2017). We confirm their results that in each Visegrad Group country there are stronger economies of scale than the OECD-modified scale assumes, in particular, the second child costs less than the first one. Just like in our study, analyzing data for the period of 2005-2012 by the use of fixed effect models, the mentioned Polish researchers found the lower relative cost of children in Hungary than in the Czech Republic and Poland (Morawski, Kalbarczyk-Stęclik and Miśta, 2017). Moreover, they observed that Slovakia exhibits the highest scales in the V4 (Kalbarczyk-Stęclik, Miśta and Morawski, 2017). There are also some differences in the obtained results. According to Bishop et al. (2014), the equivalence scale for two-adult households in Slovakia with the value of 1.18 was the lowest among the Euro Zone countries in the period of 2004-2007 which contradicts our finding and the result received by Kalbarczyk-Stęclik, Miśta and Morawski (2017), according to which, it equals to 1.47 and 1.49, respectively. It seems that Bishop et al. (2014) result in this respect is quite surprising, therefore, experts in the field of social statistics who know the Slovak realities should carefully examine this issue.

We limit our comparisons to the above three researches because, generally, the results of equivalence scales depend on the method of their estimation. In particular, it is worth being noted that the subjective scales are usually lower than scales based on consumption data (Schröder, 2004; Jirková and Musil, 2017). Moreover, it must be emphasized that our MINQ scales are measured at a subsistence standard of living and cannot be used for comparison at different welfare levels.

4 CONCLUSIONS

In the study we estimate subjective equivalence scales for the Visegrad Group countries using EU-SILC data. Our approach deals with self-regarding evaluations of households' own minimum income needs. The equivalence scales are calculated on the basis of the households' perception of what they consider to be the minimum income essential for survival.

Our findings give grounds for serious concern about commonly used in the European Union the OECD-modified scale. In all the Visegrad Group countries stronger economies of scale than the OECD-modified scale assumes are observed. Moreover, obtained results show that the relative cost of an additional adult in the Czech Republic, Poland and Slovakia does not exceed the relative cost of a child and Slovak scales significantly differ from scales for the other V4 countries.

Our results indicate that relative costs of children vary substantially across the V4 countries which means that the direct state aid in respect of allowances for children and family benefits, cultural differences and differences in relative prices may have a significant effect on equivalence scales.

Acknowledgement:
This study is based on data from Eurostat (research proposal 38/2017-EU-SILC). The results and their interpretation are the authors' responsibility.

References

[1] Betti, G., Karadag, M. A., Sarica, O. and Ucar, B. (2017). Regional differences in equivalence scales in Turkey. *Ekonomika regiona [Economy of Region]*, 13(1), pp. 63-69.

[2] Bishop, J. A., Grodner, A., Liu, H. and Ahamdanech-Zarco, I. (2014). Subjective poverty equivalence scales for Euro Zone countries. *Journal of Economic Inequality*, 12(2), pp. 265-278.

[3] Cameron, A.C. and Trivedi, P.K. (2009). *Microeconometrics Using Stata*, College Station, TX: Stata Press..

[4] Deaton, A. (1997). *The Analysis of Household Surveys: Microeconometric Analysis for Development Policy*, Baltimore: Johns Hopkins University Press.

[5] Dudek, H. (2014a). Economies of scales in EU household consumption – some remarks based on a country-level analysis. *Quantitative Methods in Economics*. XV(2), pp. 74-83.

[6] Dudek, H. (2014b). Do shares of food expenditure in the European Union converge? A country-level panel data analysis. *Economic Computation and Economic Cybernetics Studies and Research*, 48(4), pp. 245-260.

[7] Garner, T. and Short, K. (2003). Personal assessments of minimum income and expenses: what do they tell us about 'minimum living' thresholds and equivalence scales. In: Y.

Amiel, Y. and J. Bishop, ed., *Inequality. Welfare. and Poverty: Theory and Measurement*. London: JAI Press, pp. 191-243.

[8] Goedhart, T., Halberstadt, V., Kapteyn. A. and van Praag B. M. S. (1977). The poverty line: concept and measurement. *The Journal of Human Resources*, 12, pp. 503-520.

[9] Grodner, A. and Salas, R. (2017). Informational content of equivalence scales based on minimum needs income. EQUALITAS Working Paper No. 47.

[10] Jirková, M. and Musil, P. (2017). Analysis of economies of scale in Czech households based on different approaches. In: *Applications of Mathematics and Statistics in Economics. International Scientific Conference*, pp. 227-234, Available at: https://amse.ue.wroc.pl/papers/2017/Jirkova_Musil.pdf [Accessed 28 Feb. 2018].

[11] Kalbarczyk-Stęclik, M., Miśta, R. and Morawski, L. (2017). Subjective equivalence scale – cross-country and time differences. *International Journal of Social Economics*, 44(8), pp. 1092-1105.

[12] Kapteyn, A., Kooreman P. and Willemse, R. (1988). Some methodological issues in the implementation of subjective poverty definitions. *Journal of Human Resources*, 23(2), pp. 222-242.

[13] Landmesser, J. and Chrzanowska, M. (2017). Symulacja efektów ex ante programu "Rodzina 500+". *Research Papers of Wroclaw University of Economics*, 468, pp. 38-46.

[14] Morawski, L., Kalbarczyk-Stęclik, M. and Miśta, R. (2017). Subjective equivalence scales using EU-SILC panel data for Poland, the Czech Republic, and Hungary. *Acta Oeconomica,* 67(4), pp. 557–576.

[15] Perali, F. (2003). *The Behavioral and Welfare Analysis of Consumption*, Dordrecht: Kluwer Academic Publisher.

[16] PricewaterhouseCoopers (2015). *Family Tax Reliefs and Benefits in the EU*, https://www.pwc.pl/pl/pdf/family-tax-reliefs-and-benefits-in-the-eu.pdf, [Accessed 28 Feb. 2018].

[17] Schröder, C. and Schmidt, U. (2005). A new subjective approach to equivalence scales: an empirical investigation. In: U. Schmidt and S. Traub, ed., *Advances in Public Economics: Utility. Choice and Welfare*, Dordrecht: Springer, pp. 119-134.

[18] Schröder, C.(2004). *Variable Income Equivalence Scales*. Heidelberg: Physica-Verlag.

Authors' addresses

Hanna Dudek, PhD, Prof. SGGW, Faculty of Applied Informatics and Mathematics, Warsaw University of Life Sciences,159 Nowoursynowska Stree, 02-776 Warsaw, Poland,
hanna_dudek@sggw.pl

Mariola Chrzanowska, PhD, Faculty of Applied Informatics and Mathematics, Warsaw University of Life Sciences, 159 Nowoursynowska Street, 02-776 Warsaw, Poland,
mariola_chrzanowska@sggw.pl

THE ANALYSIS OF DIFFERENCES IN INCOME DISTRIBUTIONS FOR MEN AND WOMEN IN VISEGRAD STATES

Hanna Dudek, Joanna Małgorzata Landmesser,
Warsaw University of Life Sciences - SGGW

Abstract
In the paper, we compare income distributions in four Visegrad states – the Czech Republic, Hungary, Poland and Slovakia, taking into account gender differences. First, we examine the income inequalities of men and women in each country using the Oaxaca-Blinder decomposition procedure. The unexplained part of the gender pay gap gives us information about the wage discrimination. Then, we extend the decomposition procedure to different quantile points along the whole income distribution. In this case, to describe differences between the incomes of men and women, we utilize the residual imputation approach. The results obtained for each country are compared with each other and conclusions are drawn. The analysis is based on data from the EU-SILC study in 2014.

Keywords: *wage gap, differences in distributions, decomposition methods, Visegrad states*

JEL Classification: J31, D31
AMS Classification: 62G30

1 INTRODUCTION

The Visegrad Group (V4) is a regional form of cooperation comprising four Central-European countries – the Czech Republic, Hungary, Poland and Slovakia. The Visegrad states are united through their similar geopolitical situation, joint history and culture. Since 2004, all V4 countries have been member states of the European Union and are high-income countries with a very high Human Development Index (Chrzanowska, Dudek and Landmesser, 2017). The most developed country in the group is the Czech Republic (GDP per capita PPP based in current prices in 2015 equals to US$ 34,017), followed by Slovakia (US$ 29,209), Poland (US$ 27,654) and Hungary (US$ 26,941).

The aim of this work is to compare personal income distributions in V4 taking into account gender differences. Recently there has been an increase in interest in the studies of income inequalities. The findings of numerous empirical studies show that males earn higher wages than females. The gendered discrepancies in earnings are significant also in V4. As the *V4Revue* survey shows (Adamec, 2016) the gender wage gap between the average monthly earnings of women and men varies from Hungary's 13.2% to Slovakia's 23.1%. When using the median monthly earnings, the gap varies from the Czech Republic's 16.1% to Poland's 21.7%. But one has to be careful with these numbers, and calculating the wage gap one should consider factors such as job experience, education and sector, and the discriminatory policies.

A variety of techniques of income inequalities decomposition are becoming popular. For example, based on the Oaxaca-Blinder comparison (Oaxaca, 1973; Blinder, 1973) of average income values, Mysíková (2012) analyzed the structure of the gender wage gap in V4, Hedija and Musil (2012) in the Czech Republic, Zajkowska (2013) in Poland, Cukrowska-Torzewska and Lovasz (2016) in Poland and Hungary. New procedures go far beyond the Oaxaca-

Blinder decomposition. They allow to study differences of income distributions for various groups of people and to decompose them along the whole distributions. Examples of their applications for the analysis of income inequalities in Poland are Matysiak, Baranowska and Słoczyński (2010), Słoczyński (2012), Landmesser, Karpio and Łukasiewicz (2015), Landmesser (2016).

To decompose the differences between two distributions one uses the so-called counterfactual distribution, which is a mixture of a conditional distribution of the dependent variable and a distribution of the explanatory variables. Such a counterfactual distribution can be constructed in various ways (DiNardo, Fortin and Lemieux, 1996; Donald, Green and Paarsch, 2000; Machado and Mata, 2005; Fortin, Lemieux and Firpo, 2010). We will examine the differences in the entire range of income values by the use of the residual imputation approach (*JMP*-approach) (Juhn, Murphy and Pierce, 1993).

2 METHODS OF THE ANALYSIS

Let Y_g be the outcome variable in group g (e.g. the personal income in men's group, g=M, or in women's group, g=W) and X_g the vector of individual characteristics of the person in group g (e.g. age, education level). The expected value of y conditionally on X is a linear function $y_g = X_g \beta_g + v_g$, where β_g are the returns to the characteristics. The Oaxaca-Blinder decomposition for the average income inequality between two groups is as follows:

$$\hat{\Delta}^{\mu} = \overline{Y}_M - \overline{Y}_W = \overline{X}_M \hat{\beta}_M - \overline{X}_W \hat{\beta}_W = \overline{X}_M (\hat{\beta}_M - \hat{\beta}_W) + (\overline{X}_M - \overline{X}_W) \hat{\beta}_W . \tag{1}$$

The first component, on the right side of the equation, is the result of differences in the returns to observables. This is the result of differences in the estimated parameters, and so in the "prices" of individual characteristics of group representatives. It can be interpreted as the labor market discrimination. The second term, called the unexplained effect, gives the effect of characteristics and expresses the difference of the potentials of people in two groups (the so-called explained effect). Also the detailed decomposition may be calculated:

$$\hat{\Delta}^{\mu} = (\hat{\beta}_{0M} - \hat{\beta}_{0W}) + \sum_{j=1}^{k} \overline{X}_{jM} (\hat{\beta}_{jM} - \hat{\beta}_{jW}) + \sum_{j=1}^{k} (\overline{X}_{jM} - \overline{X}_{jW}) \hat{\beta}_{jW} . \tag{2}$$

A drawback of the approach is that it focuses only on average effects, which may lead to a misleading assessment if the effects of covariates vary across the wage distribution.

Let $F_{Y_g}(y)$ be the distribution function for the variable Y in group g. One can express it using the conditional distribution $F_{Y_g|X,D_g}(y|X = x)$ of Y and the joint distribution $F_{X|D_g}(X)$ of all elements of X ($D_g = 1$ if g=M; $D_g = 0$ if g=W):

$$F_{Y_g|D_g}(y) = \int F_{Y_g|X,D_g}(y|X = x) \cdot F_{X|D_g}(X) dx, \qquad g = M, W . \tag{3}$$

The mean decomposition analysis may be extended to the case of differences between the two distributions using the counterfactual distribution $F_{Y_W^C}(y) = \int F_{Y_W|X_W}(y|X) \cdot dF_{X_M}(X)$ (the distribution of incomes that would prevail for people in group *W* if they had the distribution of characteristics of group *M*):

$$F_{Y_M}(y) - F_{Y_W}(y) = [F_{Y_M}(y) - F_{Y_W^C}(y)] + [F_{Y_W^C}(y) - F_{Y_W}(y)]. \qquad (4)$$

The first component on the right side of the equation (4) gives the unexplained effect and the second term represents the explained effect.

The counterfactual distribution can be constructed using the residual imputation approach (*JMP*-approach) (Juhn, Murphy and Pierce, 1993). In this method, we estimate the two equations: $y_{Wi} = X_{Wi}\beta_W + v_{Wi}$ and $y_{Mi} = X_{Mi}\beta_M + v_{Mi}, i = 1,\ldots, n$. Then, the income y_M from the group *M* is replaced by a counterfactual income y_W^C, where both the returns to observables and residuals are set to be as in group *W*. The implementation of the procedure is divided into two steps:

1. The residuals are replaced by counterfactual residuals under the assumption of the rank preservation: $y_{Wi}^{C,1} = X_{Mi}\beta_M + v_{Wi}^{C,1}, i = 1,\ldots, n,$ where $v_{Wi}^{C,1} = F_{v_W|X}^{-1}(\tau_{Mi}(X_{Mi}), X_{Mi})$ and $\tau_{Mi}(X_{Mi})$ is the conditional rank of v_{Mi} in the distribution of residuals for *M*.
2. The counterfactual returns to observables are also imputed: $y_{Wi}^{C,2} = X_{Mi}\beta_W + v_{Wi}^{C,1}, i = 1,\ldots, n$.

Under the assumption that the regression residuals v_g are independent of *X*, it follows that $v_{Wi}^{C,1} = F_{v_W}^{-1}(\tau_{Mi})$. We need to compute the rank of the residual v_{Mi} in the marginal distribution of residuals for group *M*, and then pick the corresponding residuals in the marginal distribution of residuals for group *W*. For example, if v_{Mi} is at the 60$^{\text{th}}$ percentile of the distribution of residuals of group *M* ($\tau_{Mi} = 0.6$), then $v_{Wi}^{C,1}$ will be the 60$^{\text{th}}$ percentile of the distribution of residuals for group *W* (Fortin, Lemieux and Firpo, 2010). The rank preservation assumption is very strong and it means that someone with the same unobserved skills would be in exactly the same position, conditional on *X*, in either group *M* or *W*. Another limitation of the procedure is that there is no natural way of extending it to the case of the detailed decomposition.

3 DATA BASIS

The empirical data used have been collected within the European Union Statistics on Income and Living Conditions project for the Czech Republic, Hungary, Poland and Slovakia in 2014 (research proposal 234/2016-EU-SILC). The sample consists of 30,218 observations (6,501 for the Czech Republic, 8,054 for Hungary, 9,908 for Poland, 5,755 for Slovakia). Each person is described by the following characteristics: *gender* (1 – men, 0 – women), *age* (in years), *educlevel* (education level, 1 – primary, . . ., 5 – tertiary), *married* (marital status, 1 – married, 0 – unmarried), *yearswork* (number of years spent in paid work), *permanent* (type of contract, 1 – permanent job/work contract of unlimited duration, 0 – temporary contract of limited duration), *parttime* (1 – person working part-time, 0 – person working full-time), *manager* (managerial position, 1 – supervisory, 0 – non-supervisory), *big* (number of persons working at the local unit, 1 – more than 10 persons, 0 – less than 11 persons). The sample features presents Table 1.

Table 1. The selected sample features

Characteristic	The Czech Republic		Hungary		Poland		Slovakia	
	Men	Women	Men	Women	Men	Women	Men	Women
No. of obs.	3,443	3,058	4,061	3,993	5,180	4,728	2,847	2,908
Mean *grossincome*	12,414	8,865	7,566	6,209	9,619	7,948	10,293	7,981
Median *grossincome*	10,830	8,314	6,171	5,307	7,754	6,718	8,965	7,118
Std. dev. *grossincome*	9,395	5,102	8,296	5,395	7,277	5,296	7,311	6,643
Gini coef. *grossincome*	0.283	0.281	0.347	0.319	0.341	0.325	0.272	0.275
Average *age*	43.43	43.99	42.31	43.65	42.07	42.35	41.76	42.99
Average *yearswork*	22.18	20.77	20.65	20.74	20.09	18.45	19.10	19.84
Average *educlevel*	3.39	3.40	3.35	3.54	3.36	3.80	3.45	3.62
married = 1	56.26%	58.04%	55.41%	51.59%	71.54%	69.61%	62.24%	59.46%
permanent = 1	89.31%	85.84%	85.87%	87.23%	70.58%	71.64%	88.27%	86.80%
parttime = 1	1.63%	4.51%	2.66%	5.18%	4.31%	10.09%	1.51%	4.95%
manager = 1	23.93%	14.58%	10.96%	8.06%	18.69%	15.74%	14.23%	9.11%
big = 1	85.51%	78.74%	77.71%	72.80%	82.86%	80.16%	68.81%	57.43%

Source: own calculations

The annual gross employee (cash or near cash) incomes (*grossincome*) of men were compared with those obtained by women. The gross employee cash or near cash income refers to the monetary component of the compensation of employees in cash payable by an employer to an employee (e.g. wages and salaries paid for the time worked, remuneration for the time not worked, enhanced rates of pay for overtime). It includes the value of any social contributions and income taxes payable by an employee or by the employer on behalf of the employee to social insurance schemes or tax authorities. In our empirical decomposition analysis the logarithm of the annual gross employee income measured in Euro (*log_income*) constitutes the outcome variable.

4 EMPIRICAL ANALYSIS

4.1 Results of Oaxaca-Blinder decomposition for differences in mean log incomes

Table 2 presents the results of the aggregate and detailed Oaxaca-Blinder decomposition of inequalities between men's and women's log incomes for V4 countries in 2014.

The mean predicted log income is the largest in the Czech Republic: for men equals 9.280 (which corresponds to 10,721 Euro of annual income), for women equals 8.937 (7,608 Euro of annual income). The least predicted log income we noted for Hungary. There is a positive difference between the mean values of log incomes for men and women for all V4 countries. The mean log income differential is the largest in the Czech Republic (0.343) and the smallest in Hungary (0.176). The difference between the mean log income values was decomposed into two components: the first one explaining the contribution of the different values of models coefficients (the unexplained part) and the second one explaining the contribution of the attributes differences (the explained part). The unexplained effect for all states is huge and positive, but the explained is very low. It means that the inequalities examined should be assigned in the majority to the coefficients of estimated models rather than to the differentiation of individual characteristics. The unexplained part of the gender pay gap gives us information about the wage discrimination, which is the largest in the Czech Republic and in Slovakia. For the Czech Republic and Slovakia the value of the explained effect is positive and for Hungary and Poland negative. The negative value of this component means that the difference of the average log incomes between men and women is mostly reduced by the women's characteristics.

Table 2. The Oaxaca-Blinder decomposition of the average log income differences

	The Czech Republic		Hungary		Poland		Slovakia	
Mean *log_income* men	9.280		8.699		8.964		9.098	
Mean *log_income* women	8.937		8.522		8.780		8.833	
Raw differential	0.343		0.176		0.183		0.265	
Aggregate decomposition								
Unexplained effect	0.248		0.200		0.209		0.247	
Explained effect	0.095		-0.024		-0.026		0.019	
% unexplained	72.3		113.8		114.2		93.0	
% explained	27.7		-13.8		-14.2		7.0	
Detailed decomposition								
Variable	Unexpl.	Expl.	Unexpl.	Expl.	Unexpl.	Expl.	Unexpl.	Expl.
age	0.192	0.003	-0.259	0.007	-0.175	0.000	0.105	0.005
educlevel	-0.107	-0.003	0.044	-0.044	-0.091	-0.087	-0.050	-0.027
married	0.101	0.001	0.091	-0.001	0.091	0.001	0.120	-0.002
yearswork	-0.261	0.017	0.007	-0.001	-0.050	0.021	-0.172	-0.008
permanent	-0.072	0.014	-0.153	-0.010	-0.048	-0.004	-0.042	0.005
parttime	0.000	0.024	-0.006	0.012	-0.005	0.032	0.001	0.020
manager	0.003	0.030	-0.003	0.007	0.015	0.007	0.000	0.012
big	-0.008	0.009	-0.005	0.005	-0.005	0.004	-0.005	0.012
cons	0.399	0.000	0.481	0.000	0.478	0.000	0.289	0.000
Total	0.248	0.095	0.200	-0.024	0.209	-0.026	0.247	0.019

Source: own elaboration using the Stata command '`decompose`'

The detailed decomposition, which was also carried out, made it possible to isolate the factors explaining the inequality observed to a different extent (see Table 2). The strong effect of different education levels of men and women can be noticed, especially for Poland. The negative values of explained components mean that the differences of the average log incomes between men and women are reduced by the women's higher education levels. On the other hand, the values of *manager* and *big* attributes possessed by men and women increase the income inequality in all countries (see the positive explained component values). A different "evaluation" of personal characteristics (the unexplained components) allows the conclusion that in all states women are discriminated against men because of their marital status (the positive unexplained components values for variable *married*) but not because of the education levels (except Hungary), years of work (again except Hungary), the type of contract or the company size.

4.2 Results of the aggregate decomposition using the residual imputation approach

Since the Oaxaca-Blinder technique focuses only on average effects, next, we present the decomposition of inequalities along the distribution of log incomes for men and women using the *JMP*-approach. The results are shown in Table 3, where the inequalities are expressed in terms of percentiles. The symbols p5, …, p95 stand for 5^{th}, …, 95^{th} percentile (e.g. the 5^{th} percentile is the log income value below which 5% of the observations may be found). For each of the seven percentiles the total differences between the values of log incomes for men and women were computed. Then, these differences are expressed in terms of the unexplained and explained components.

Table 3. The results of aggregate decomposition using the *JMP*-approach

Percentile	The Czech Republic		Hungary		Poland		Slovakia	
	Total difference							
p5	0.616		0.325		0.324		0.393	
p10	0.401		0.239		0.304		0.200	
p25	0.370		0.122		0.125		0.304	
p50	0.264		0.151		0.143		0.231	
p75	0.279		0.157		0.162		0.266	
p90	0.285		0.226		0.205		0.274	
p95	0.349		0.286		0.254		0.307	
Percentile	Unexplained and explained components							
	Unexpl.	Expl.	Unexpl.	Expl.	Unexpl.	Expl.	Unexpl.	Expl.
p5	0.288	0.328	0.252	0.073	0.200	0.124	0.297	0.096
p10	0.236	0.164	0.281	-0.042	0.239	0.065	0.220	-0.020
p25	0.256	0.113	0.199	-0.077	0.231	-0.106	0.242	0.062
p50	0.236	0.028	0.187	-0.036	0.201	-0.057	0.237	-0.007
p75	0.264	0.014	0.191	-0.034	0.199	-0.037	0.225	0.041
p90	0.250	0.035	0.192	0.034	0.203	0.003	0.235	0.039
p95	0.236	0.113	0.204	0.082	0.198	0.056	0.227	0.080

Source: own elaboration using the Stata command '`jmpierce`'

For all V4 countries the differences between the values of log incomes for men and women along the whole log income distribution are positive. Going across the rows to compare quantile effects shows that the total differences are greater at the bottom and at the top of the log income distribution and are smaller in the middle of the distribution. Figure 1 contains the total differences between the log income distributions for men and women vs. quantile rank. The total effect for most countries is U-shaped. We note the greatest inequalities in the Czech Republic, then in Slovakia and the smallest for Poland and Hungary. The positive values indicate higher log income values for men than for women.

The unexplained effects (effects of coefficients) are bigger and the explained are lower, which indicates the importance of the "labor market value" of men's and women's attributes (see Fig. 1). The effects of coefficients are positive in the whole range of the income distribution. This is the result of differences in the "market prices" of individual characteristics of men and women (interpreted as the labor market discrimination). The size of discrimination is the highest in the Czech Republic, then in Slovakia, then in Poland and at the end in Hungary. The gender differences in characteristics increase the income inequalities at the bottom and at the top of the log income distribution (the explained effects, first, are falling and, then, are growing as we move toward the top of the distribution). It is well known that the widest gender wage gap exists among management, and mostly among those with the highest salaries. The positive values mean that the different values of characteristics of men and women increase the income inequalities in these income ranges. However, for three countries – Hungary, Poland and Slovakia – the explained effects are negative in the middle of the distribution, which means that the properties possessed by both people's groups decrease the income inequalities (probably in these income ranges the differences are reduced by the women's better characteristics).

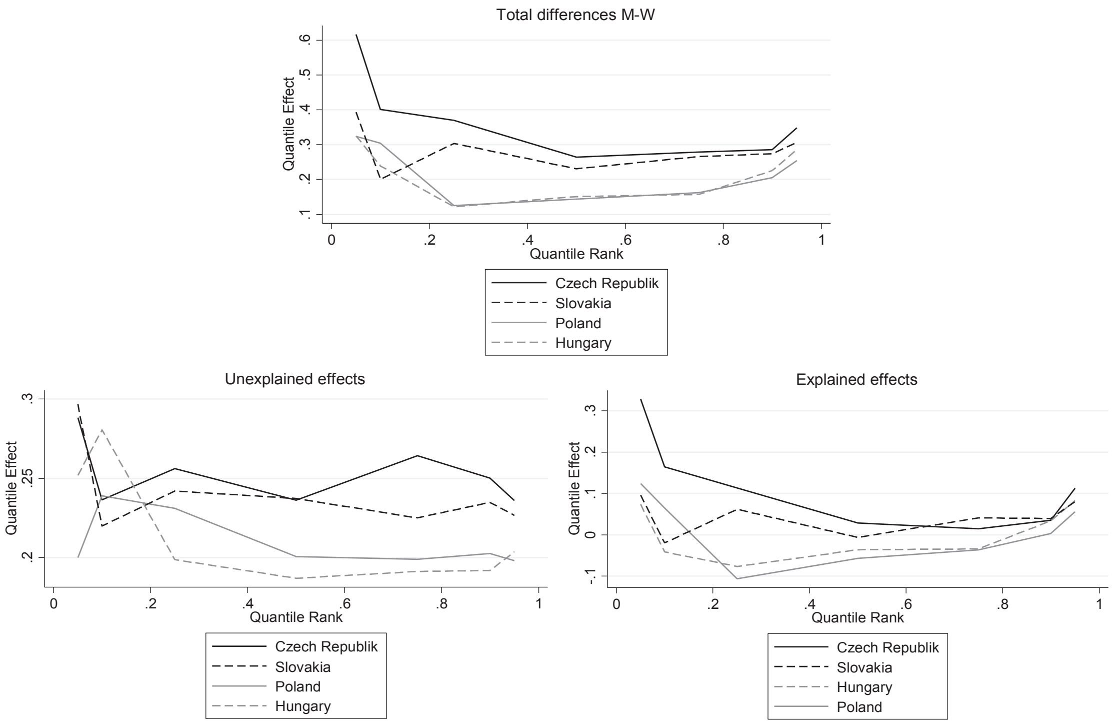


Fig. 1. The total, the unexplained and the explained differences between the log income distributions for men and women calculated using the *JMP*-approach
Source: own elaboration using `jmpierce` Stata module

5 CONCLUSIONS

The goal of this paper was to compare income distributions in V4 states taking into account gender differences. We started with the decomposition of the average values for log incomes in each country by using the Oaxaca-Blinder method. The findings are to a great extent similar for the Czech and Slovak Republics (the result as in Mysíková (2012)). For all four countries, there was a positive difference between the mean log income values for men and women. The largest mean differential was noted for the Czech Republic, then for Slovakia, Poland and the smallest for Hungary. The unexplained effect was big, but the explained low. The unexplained part of the gender pay gap gave us information about the wage discrimination, which was the largest in the Czech Republic and in Slovakia. For the Czech Republic and Slovakia the value of the explained effect was positive (increasing gender pay gap) and for Hungary and Poland negative (reducing gender pay gap). The decomposition showed also the influence of the men's and women's attributes on the average log income differences.

Then, we extended the decomposition procedure to different quantile points along the whole income distribution. In this case, to decompose the inequalities between log incomes along the whole distribution we used the *JMP*-residual imputation approach. The results obtained for each country indicate the U-shaped total effect. Our analysis confirmed that the gender wage gap in all the analyzed states can be poorly explained by gender differences in observable characteristics of people. Again, the conducted decomposition showed that the unexplained component quantitatively dominates, with the highest size of discrimination in the Czech Republic, then in Slovakia, in Poland and, at the end, in Hungary. Such a gender discrimination may lead to considerable loss in productivity and wealth, therefore inequalities induced in this way pose a serious challenge for politicians and society.

References

[1] Adamec, J. (2016). *Women earn less than men in the V4: A punishment for home and childcare?* [online] Available at: http://visegradrevue.eu/women-earn-less-than-men-in-the-v4-a-punishment-for-home-and-childcare/ [Accessed 11 Mar 2018].

[2] Blinder, A. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources*, 8(4), pp. 436-455.

[3] Chrzanowska, M., Dudek, H. and Landmesser J. (2017). Standard of living in selected countries in Central Europe in the period 2000-2014, *Double-blind peer-reviewed proceedings of the International Scientific Conference Hradec Economic Days 2017*, 7(1), pp. 325-331.

[4] Cukrowska-Torzewska, E. and Lovasz, A. (2016). Are children driving the gender wage gap? Comparative evidence from Poland and Hungary. *Economics of Transition*, 24(2), pp. 259–297.

[5] DiNardo, J., Fortin, N.M. and Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, 64(5), pp. 1001-1044.

[6] Donald, S.G., Green, D.A. and Paarsch, H.J. (2000). Differences in Wage Distributions between Canada and the United States: An Application of a Flexible Estimator of Distribution Functions in the Presence of Covariates. *Review of Economic Studies*, 67(4), pp. 609-633.

[7] Fortin, N., Lemieux, T. and Firpo, S. (2010). *Decomposition methods in economics*. Cambridge: NBER Working Paper, No. 16045.

[8] Hedija, V. and Musil, P. (2012). Gender Wage Differences in the Selected Czech Public Sector Company. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, LX(7), pp. 81-88.

[9] Juhn, Ch., Murphy, K.M. and Pierce, B. (1993). Wage Inequality and the Rise in Returns to Skill. *Journal of Political Economy*, 101(3), pp. 410-442.

[10] Landmesser, J. (2016). Decomposition of Differences in Income Distributions Using Quantile Regression. *Statistics in Transition - new series*, 17(2), pp. 331-348.

[11] Landmesser, J.M., Karpio, K. and Łukasiewicz, P. (2015). Decomposition of Differences Between Personal Incomes Distributions in Poland. *Quantitative Methods in Economics*, XVI(2), pp. 43-52.

[12] Machado, J.F. and Mata, J. (2005). Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. *Journal of Applied Econometrics*, 20, pp. 445-465.

[13] Matysiak, A., Baranowska, A. and Słoczyński T. (2010). Kobiety i mężczyźni na rynku pracy. In: Bukowski M. (ed.). *Zatrudnienie w Polsce 2008 – Praca w cyklu życia*. Human Resources Development Center, Warszawa.

[14] Mysíková, M. (2012). Gender Wage Gap In The Czech Republic And Central European Countries. *Prague Economic Papers*, 21(3), pp. 328-346.

[15] Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3), pp. 693-709.

[16] Słoczyński, T. (2012). Próba wyjaśnienia regionalnego zróżnicowania międzypłciowej luki płacowej w Polsce. *Studia Regionalne i Lokalne*, 3(49).

[17] Zajkowska O. (2013). Gender Pay Gap in Poland – Blinder-Oaxaca Decomposition. *Quantitative Methods in Economics*, XIV(2), pp. 272-278.

Author's addresses
dr hab. Hanna Dudek, prof. SGGW, dr hab. Joanna Landmesser
Warsaw University of Life Sciences - SGGW, Faculty of Applied Informatics
and Mathematics, Department of Econometrics and Statistics
Nowoursynowska 159, 02-776 Warsaw
Poland
email: hanna_dudek@sggw.pl, joanna_landmesser@sggw.pl

MILK-RUN OPTIMIZATION IN AUTOMOTIVE INDUSTRY

Jan Fábry, Škoda AUTO University, Mladá Boleslav
Jan Pelikán, University of Economics, Prague
Tomáš Václavů, Cleverlance Enterprise Solutions, Prague

Abstract
A supply chain in automotive industry is hard to coordinate. Therefore, the flexibility of logistic processes is considered as a key indicator. Companies intend to take control over all flows in production and transportation by increasing competitiveness through the optimization of supply chain. An effective connection of all material and information flows minimizes stock levels, because all inputs are available on time and in requested quantity and quality. The paper describes inbound logistics and introduces a proposal of the optimization model for planning vehicle routes. There is used the modified model for vehicle routing problem, because in the real instance special conditions must be satisfied. Milk-run routes are optimized in term of total distance using the modification of split delivery vehicle routing problem. Weight and volume attributes of delivery are considered. Proposed routes are compared with the realized routing plan.

Keywords: *Vehicle routing problem, Automotive industry, Milk-run, Split delivery problem*

JEL Classification: C44
AMS Classification: 90C90

1 INTRODUCTION

In logistic processes several principles are currently applied. Just in Time (JIT) and Just in Sequence (JIS) are the most striking concepts in the automotive industry. They are based on the interconnected supply chains, segments of which share necessary information. One of the positive consequences is the elimination of inventories at the input and partly at the output of the chain. Respecting JIT concept, each requirement is being solved exactly at the time it is prepared and able to process. This access connected with the "pull" principle is advantageous in terms of minimizing total inventory cost. Considering the basic level of this conception, the problem of delivering parts, assembling and cars expedition can be successfully solved using the operational research methodology.

The basic principles of inbound logistics, mentioned above, are applied in this paper. The main segment of the inbound logistics is a supplier delivering required inputs to the company. Deliveries can be realized using internal means of transport or outsourced transportation services (Agnetis et al., 2014, Čičková et al., 2015). A strategy of using cross-docking seems to be more advantageous approach for far-distant suppliers (Boysen and Fliedner, 2010). Wen et al. (2009) address the optimization of the vehicle routing problem where a set of homogeneous vehicles is used to transport orders from supplier to more customers via a cross-dock. In case suppliers are relatively close to the industrial company there is another strategy that is aimed at the milk-run travels of vehicles combining several suppliers on each tour (Boysen et al., 2016).

2 PROBLEM OF MILK-RUN TRAVELS

The paper is aimed at the proposal of convenient milk-run routes (Kilic et al., 2012), which are able to compete with the current situation where the transport is ensured by suppliers themselves. Although there are much more suppliers located in Liberec Region and Usti nad Labem Region, their number was reduced to 9 suppliers (close locations were grouped and their distances were negligible). The definition of the problem seems to be quite interesting due to the specific condition that must be satisfied (Václavů, 2015). Because of packaging, manipulation and storage, all delivered products are divided into 40 categories. Each supplier can deliver products from all categories, but products from maximally 4 different categories can be loaded on each vehicle. The reason is a necessity of the reduction of trucks movements between warehouse gates due to the difficulty of logistic processes and the limited unloading capacity of the warehouse. This strategy at least partially increases the consistency of the entire delivery, and the unloading of the vehicle takes shorter time than its crossing the whole area and unloading heterogeneous products.

In the real problem, requirements for a week are given in each category in term of weights. There is prepared a fleet of 7 vehicles of the identical weight capacity for products delivery. In the process of routes planning, it is necessary to consider that each supplier will be visited multiple times.

3 FORMULATION OF A MATHEMATICAL MODEL

The milk-run vehicle routing optimization model for delivering automobile parts was proposed by Hui and Jing (2013). Respecting the given conditions specified in the previous section, we propose an original mathematical model analogous to the model of split delivery vehicle routing problem (Archetti et al., 2005).

Let us suppose n locations including the depot at location 1. Distance c_{ij} between locations i and j is given ($i = 1,2,\ldots,n$; $j = 1,2,\ldots,n$). Totally K vehicles with the identical weight capacity W can be used for products delivery. All products are divided into L categories and products from maximally 4 categories can be loaded on each vehicle. Value of r_i^l is a delivery requirement of category l at location i ($l = 1,2,\ldots,L$; $i = 2,3,\ldots,n$). The following binary variables are defined:

$$x_{ij}^{kl} = \begin{cases} 1 & \text{if } k\text{-th vehicle goes from location } i \text{ to location } j \text{ and carries } l\text{-th category,} \\ 0 & \text{otherwise,} \end{cases}$$

$$y^{kl} = \begin{cases} 1 & \text{if } k\text{-th vehicle carries } l\text{-th category,} \\ 0 & \text{otherwise,} \end{cases}$$

$$h_{ij}^{k} = \begin{cases} 1 & \text{if } k\text{-th vehicle goes from location } i \text{ to location } j, \\ 0 & \text{otherwise.} \end{cases}$$

Nonnegative variable t_i^{kl} corresponds to the quantity of l-th category loaded on k-th vehicle at location i, and variable u_i^{kl} refers to the total load of l-th category on k-th vehicle at location i. The mathematical model can be formulated as follows:

$$z = \sum_{k=1}^{K}\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}h_{ij}^{k} \quad \rightarrow \ \text{min}, \tag{1}$$

$$\sum_{j=2}^{n} h_{1j}^{k} \leq 1, \qquad k = 1,2,\dots,K, \tag{2}$$

$$\frac{1}{4}\sum_{l=1}^{L} x_{ij}^{kl} \leq h_{ij}^{k}, \qquad i = 1,2,\dots,n, \quad j = 1,2,\dots,n, \quad k = 1,2,\dots,K, \tag{3}$$

$$\sum_{i=1}^{n} h_{ij}^{k} = \sum_{i=1}^{n} h_{ji}^{k}, \qquad j = 2,3,\dots,n, \quad k = 1,2,\dots,K, \tag{4}$$

$$u_i^{kl} + t_j^{kl} - W(1 - h_{ij}^{k}) \leq u_j^{kl}, \qquad \begin{array}{l} i = 1,2,\dots,n,\ j = 2,3,\dots,n,\ i \neq j, \\ k = 1,2,\dots,K,\ l = 1,2,\dots,L, \end{array} \tag{5}$$

$$\sum_{k=1}^{K} t_i^{kl} = r_i^{l}, \qquad i = 2,3,\dots,n, \quad l = 1,2,\dots,L, \tag{6}$$

$$u_1^{kl} = 0, \qquad k = 1,2,\dots,K,\ l = 1,2,\dots,L, \tag{7}$$

$$0 \leq \sum_{i=1}^{n}\sum_{l=1}^{L} t_i^{kl} \leq W, \qquad k = 1,2,\dots,K, \tag{8}$$

$$t_i^{kl} \leq W\sum_{j=1}^{n} x_{ij}^{kl}, \qquad i = 1,2,\dots,n, \quad k = 1,2,\dots,K, \quad l = 1,2,\dots,L, \tag{9}$$

$$x_{ii}^{kl}{=}0, \qquad i = 1,2,\dots,n, \quad k = 1,2,\dots,K, \quad l = 1,2,\dots,L, \tag{10}$$

$$x_{ij}^{kl}, y^{kl}, h_{ij}^{kl} \in \{0,1\}, \qquad \begin{array}{l} i = 1,2,\dots,n,\ j = 1,2,\dots,n,\ k = 1,2,\dots,K, \\ l = 1,2,\dots,L, \end{array} \tag{11}$$

$$t_i^{kl}, u_i^{kl} \in R_0^{+}, \qquad i = 1,2,\dots,n, \quad k = 1,2,\dots,K, \quad l = 1,2,\dots,L. \tag{12}$$

The objective (1) is to minimize total distance travelled by all the vehicles. Inequalities (2) assure that each vehicle goes from the depot maximally once. Inequalities (3) must be respected because the maximal number of categories on each vehicle going from location *i* to location *j* is equal to 4. In addition, if vehicle *k* does not go from *i* to *j*, no category is loaded on the vehicle. Constraints (4) are used in standard vehicle routing problem and assure that the vehicle entering the location must also leave it. Constraints (5), which are based on Miler-Tucker-Zemlin's conditions (Miller et al., 1960), are used to eliminate partial cycles and express the load balance on vehicles. All requirements of i-th location are satisfied using (6). At the depot, loads of all vehicles are set to zeros (7). Respecting (8), no vehicle weight capacity is exceeded. If in inequalities (9) the summation on the right-hand side is equal to 0 (k-th vehicle leaving location i does not contain category l), then partial requirement load t_i^{kl} must be equal to zero as well.

4 COMPUTATIONAL RESULTS

Solver Gurobi 4.6.1 was used for the optimization (Václavů, 2015). After 5 hours of calculation, the run was interrupted because there was no significant change in objective value. Total length of all routes was 1539 km, the gap value was on the level of 8 %. To be precise, total travelled distance does not include crossings between locations that were grouped. The average occupancy of a vehicle was on 90 % of the weight capacity. While the average length of the vehicle tour is approximately 220 km, there is a notable difference between the maximum and minimum lengths: 362 km and 119 km. In case of the time limitation for travels, the model can be modified as follows:

$$z = T \quad \rightarrow \ \min, \tag{13}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij} h_{ij}^{k} \leq T \quad k = 1,2,\dots,K, \tag{14}$$

where d_{ij} is the time necessary for travel from destination i to destination j. The optimal value of variable T is the longest duration of a tour.

As practical experiences show, in real distribution problems great emphasis is placed not only on the weight attributes but also on the volume character. This understandable fact can significantly influence a real decision. Therefore, model (1) – (12) must be partly modified. Let V be the volume capacity of the vehicle. Because of the specification of the problem in terms of existing 4 categories of products, s_i^l denotes total volume requirement of location i for delivery of l-th category. Because r_i^l is corresponding weight requirement, it is possible to obtain the volume per weight unit:

$$p_i^l = \frac{s_i^l}{r_i^l}, \qquad i = 1,2,\dots,n, \ \ l = 1,2,\dots,L. \tag{15}$$

Considering a vehicle volume capacity, the following constraint must be introduced:

$$0 \leq \sum_{i=1}^{n}\sum_{l=1}^{L} t_i^{kl} p_i^l \leq V, \qquad k = 1,2,\dots,K, \tag{16}$$

and several parts of the model are to be slightly modified.

Note: If necessary, the model can be modified in terms of achievement of maximal occupancy of vehicles.

After including volume restrictions to the optimization model it was necessary to increase the number of vehicles from 7 to 11. It follows the remarkable fact that the volume attributes are more important than weight attributes in terms of vehicle capacity. Unfortunately, the computational complexity of modified model is much higher and therefore, the gap for obtained solution is on 19 % after the run interruption and the total length of all routes is 2513 km. Nevertheless, this solution is close to real conditions and can be compared with real results (see Table 1).

	Reality	Proposal
Number of vehicles used	16	11
Average vehicle volume occupancy (%)	70	97
Average vehicle weight occupancy (%)	28	59
Average length of route (km)	197	228
Total distance travelled (km)	3057	2513

Table 1: Comparison of proposal with reality

5 CONCLUSIONS

In the paper, real milk-run routes in automotive industry were optimized in terms of total distance travelled. Because of logistic restrictions, there are specific conditions that must be considered. Products are divided into 40 categories and maximally 4 categories can be loaded on a vehicle. In the first model, only weight attributes were included in the model. Then, the model was modified to make the results much closer to the real distribution. Comparing the reality and the proposal, number of used vehicles was decreased from 16 to 11. This fact corresponds to increasing the occupancy of vehicles in terms of both weight and volume. While the average length of route was obviously increased, total distance (and related cost) was decreased almost by 20 %. Although achieved cost savings are remarkable, the approach of heuristic methods can be considered in future research to try to improve the solution.

Acknowledgements: This paper was supported by the grant No. 16-00408S of the Czech Grant Agency.

References

[1] Agnetis, A., Aloulou, M.A. and Fu, L.-L. (2014). Coordination of production and interstage batch delivery with outsourced distribution. *European Journal of Operational Research*, 238(1), pp. 130-142.

[2] Archetti, C., Mansini, R. and Speranza, M.G. (2005). Complexity and Reducibility of the Skip Delivery Problem. *Transportation Science*, 39(2), pp. 182-187.

[3] Boysen, N., Briskorn, D. and Emde, S. (2016). Just-in-time vehicle scheduling with capacity constraints. *IIE Transactions*, 48(2), pp. 134-145.

[4] Boysen, N. and Fliedner, M. (2010). Cross dock scheduling: classification, literature review and research agenda. *Omega*, 38(6), pp. 413-422.

[5] Čičková, Z., Brezina, I. and Pekár, J. (2015). Routes design using own and hired vehicles. In: *33rd International Conference on Mathematical methods in economics 2015*. Cheb: University of West Bohemia, pp. 105-108.

[6] Hui, J.M. and Jing, W. (2013). Milk-run Vehicle Routing Optimization Model and Algorithm of Automobile Parts. *Applied Mechanics and Materials*, Vols. 253-255, pp. 1463-1467.
[7] Kilic, H., Durmusoglu, M. and Baskak, M. (2012). Classification and modeling for in-plant milk-run distribution systems. *International Journal of Advanced Manufacturing Technology*, 62(9-12), pp. 1135-1146.
[8] Miller, C. E., Tucker, A. W. and Zemlin, R. A. (1960). Integer Programming Formulations and Traveling Salesman Problems. *Journal of the ACM*, 7(4), pp. 326-329.
[9] Václavů, T. (2015). *Milkrun optimization in automotive industry*. Diploma thesis (in Czech), University of Economics, Prague.
[10] Wen, M., Larsen, J., Clausen, J., Cordeau, J.F. and Laporte, G. (2009). Vehicle routing with cross-docking. *Journal of the Operational Research Society*, 60(12), pp. 1708-1718.

Authors' addresses

doc. Ing. Jan Fábry, Ph.D.
ŠKODA AUTO University, Mladá Boleslav,
Department of Logistics, Quality and Automotive Technology,
Na Karmeli 1457, 293 01 Mladá Boleslav
Czech Republic
email: fabry@savs.cz

prof. RNDr. Jan Pelikán, CSc.
University of Economics, Prague, Faculty of Informatics and Statistics,
Department of Econometrics,
Winston Churchill sq. 4, 130 67 Prague 3
Czech Republic
email: pelikan@vse.cz

Ing. Tomáš Václavů
Cleverlance Enterprise Solutions, Prague,
Project Manager Banking, Automotive,
Dukelských hrdinů 34, 170 00 Prague 7
Czech Republic
email: tomas.vaclavu@cleverlance.com

ECO-EFFICIENCY USING DIRECTIONAL DISTANCE FUNCTIONS WITH UNDESIRABLE OUTPUTS

Peter Fandel, Lubica Bartova, Slovak Agricultural University in Nitra

Abstract
The objective of the paper is to estimate and compare efficiency and eco-efficiency of agriculture in selected 26 EU Member States over period from 2006 to 2015. In the study we used panel of yearly aggregated data (Eurostat, 2018) of total agricultural output value (AO), labour (AWU), utilised agricultural area (UAA), fertilisers N, P, K (NPK) and greenhouse gas (GHG) emission of agriculture of EU Member States. We employed directional distance functions (DDF) approach both without and with undesirable output (GHG emission). Malmquist-Luenberger indices were employed to measure productivity change. Further they were decomposed to efficiency change and technological change. We found that countries with high eco-efficiency (Netherlands, Italy, Denmark) in agricultural production over 2006-2015, reached relatively high efficiency even disregard their environmental concerns (model without GHG). The New Member States exhibited higher total factor productivity (TFP) growth than the old MS. Significant eco-TFP growth however, was observed mainly in agriculture of the old EU Member States (MS). Bulgaria and Romania showed the worst eco-performance development.

Keywords: *Directional Distance Function, eco-efficiency, undesirable output, agriculture, EU*
JEL Classification: C61, O13, O33, Q52
AMS Classification: 90C05, 90C90

1 INTRODUCTION

To assure sustainable development of production systems, economic environmental and social factors has to be taken into account by decision making units (DMU). DMU maximize their economic aims, but at the same time should minimize negative environmental effects.
The eco-efficiency has become a tool for assessment of sustainable performance of DMU (Wang et al., 2011) By definition of WBCSD (2000) eco-efficiency is achieved by the delivery of competitively-priced goods and services that satisfy human needs and bring quality of life, while progressively reducing ecological impacts and resource intensity throughout the life cycle to the level at least in line with the Earth's estimated carrying capacity.
Sustainability assessment has been applied for different purposes, including decision making and management, at different levels: national (e.g., Coli, Nissi, and Rapposelli, 2011), regional or urban community (e.g., Hu, Sheu, and Lo, 2005; Munda and Saisana, 2011), industry sectorial (e.g. Zofío and Prieto, 2001).
Three categories of indicators and methods for evaluation of sustainability have emerged (Zhou, et al. 2018): System analysis, Flow analysis, Indicators mainly chooses from environmental, economic, social and institutional aspects.
There are many approaches dealing with multidisciplinarity in eco-efficiency assessment. One of the most used is Data envelopment analysis (DEA) (Charnes, Cooper, and Rhodes, 1978), evaluating the performance of DMUs, with multiple performance measures in terms of inputs and outputs. In general, three approaches of DEA are employed in the sustainability assessment: traditional DEA (Lovell, Pastor, and Turner, 1995), traditional DEA dealing with undesirable outcomes as input (Hu and Wang, 2006; Zhang, Bi, Fan, Yuan, and Ge, 2008) and DEA employing the concept of weak disposability technology (Färe and Grosskopf, 2004).

Among the most advanced DEA based methods frequently used in analysis of eco-efficiency are those applying on directional distance functions (Färe and Grosskopf, 2010) and intertemporal DEA models, especially using Malmquist productivity indices. Directional distance function (DDF) was introduced by Chung et al. (1997) and used as a component in Luenberger productivity indicators to model joint production of goods and bads. Later, their index method is applied in many sustainability studies (e.g., He, Zhang, Lei, Fu, and Xu, 2013). Sahoo, Luptacik and Mahleberg (2011) generate output-oriented environmental efficiency measures based on two approaches how to treat of undesirable outputs. They used data set of 22 OECD countries and propose two slacks-based formulations of environmental efficiency, one based on the range directional model, and the other on the generalized proportional distance function model. Mahleberg and Luptacik (2014) measure eco-efficiency of the Austrian economy using an augmented Leontief input–output model extended by constraints for primary inputs and a multi-objective optimisation model is applied to generate the eco-efficiency frontier. An intertemporal approach in the spirit of the Luenberger productivity indicator and its decomposition enable to examine the contributions of individual production factors, both undesirable and desirable outputs to eco-productivity change over time.
Some new DEA models with considering of undesirable outputs, or both desirable and undesirable outputs have been applied in sustainable development assessment, e.g. non-radial directional distance function (Wang et al., 2013), sequential generalized directional distance approach (Zhang, Kong and Choi, 2014), non-oriented DDF model (Chang, 2015), bounded adjusted measure (BAM) (Rashidi and Saen, 2015).
Directional distance functions approach in eco-efficiency assessment in agriculture were employ by e.g. Picazo-Tadeo, Beltrán-Esteve, and Gómez-Limón (2011, 2012) based on agricultural farm level data. Agricultural energy and environmental efficiency of the primary sectors over period 2001 – 2008 of the EU MS using non-radial DEA were examined by Vlontzos et al. (2014). They found that countries with strong environmental protection standards appear to be less energy and environmentally efficient. Low efficiency scores found for the Eastern European MS they ascribe to low level of technology in primary production process.
The objective of our paper is to estimate and assess development of eco-efficiency of agriculture in selected EU Member States from 2006 to 2015. Due to improvement of technologies after accession to the EU, we expect positive development in eco-efficiency of the new member states agriculture.

2 METHODOLOGY

In the paper we analyse eco-efficiency of agriculture of 26 EU Member States. We assume that the production technology for countries producing S outputs, $y \in R_+^S$, and U polluting by-products, $b \in R_+^U$, is represented by the output set $P(x)$, which designates the set of good and bad outputs vector (y, b) that is jointly produced from M inputs which are represented by the input vector, $x \in R_+^M$. The production technology can be mathematically expressed as follows:

$$P(x) = \{(y, b) | x \; can \; produce \; (y, b) \qquad (1)$$

Production technology when desirable and undesirable outputs are jointly produced needs to meet several assumptions in the form of axioms listed by Oh (2010) as follows:

First, the production technology set (PPS) is assumed to be compact for the input vector $x \in \mathrm{R}_+^{\mathrm{M}}$. Inputs are also assumed to be strongly (free) disposable, so that:

$$if \; x' \geq x \; then \; P(x') \supseteq P(x) \qquad (2)$$

Equation (2) suggests that PPS will not shrink when the inputs used in production are increased.
Second, the incorporation of undesirable outputs into the classical production technology requires the assumption of null-jointness. This assumption implies that the decision countries should necessarily produce the undesirable outputs when they produce the desirable outputs. The assumption of null-jointness is expressed as follows:

$$\text{if } (y, b) \in P(x) \text{ and } b = 0, \text{then } y = 0 \tag{3}$$

Equation (3) suggests that the desirable outputs cannot be produced if the undesirable outputs are not produced. This is always true when the assumption of the null-jointness is imposed on the production technology (Färe et al. 2007).
Third, a weak disposability assumption needs to be imposed onto PPS, which is stated as follows:

$$\text{if } (y, b) \in P(x) \text{ and } 0 \leq \theta \ \leq 1, \text{then } (\theta y, \theta b) \in P(x) \tag{4}$$

This assumption implies that any proportional contraction of the desirable and the undesirable outputs is also feasible if the original combination of the desirable and the undesirable outputs is in the PPS, for a given inputs x. This assumption also implies that the undesirable outputs are costly to dispose of. In other word, the cost for the abatement of the undesirable goods inevitably results in less production of the desirable outputs (Färe et al. 2007).
Fourth, the strong disposability of desirable outputs is also required, as follows:

$$if (y, b) \in P(x) and y \leq y', then\ (y', b) \in P(x) \tag{5}$$

This assumption means that some of the desirable outputs can always be disposed of without any additional cost (Färe et al. 2007).
The production technology we elaborate by employing the directional distance function. Let $g = (g_y, g_b)$ be a direction vector, where $g \in R_+^S \times R_+^U$. Then, the directional distance function is defined as follows:

$$\vec{D}(x, y, b; g_y, g_b) = \max\{\beta | (y + \beta g_y, b - \beta g_b) \in P^s(x)\}. \tag{6}$$

The direction vector, g, determines the direction of outputs, by which the desirable outputs increase and the undesirable outputs decrease. In our study the direction vector was taken as $g = (y, b)$. This function enables to find maximum increase of the desirable outputs while simultaneously reducing the undesirable (bad) outputs.
Eco-efficiency in this paper we express by the Malmquist-Luenberger productivity index (ML) (Chung, Färe, Grosskopf, 1997). It is defined between time periods t and t+1 on $P^s(x), s = t, t + 1$ as:

$$ML^s = \frac{1+ \vec{D}_0^s(x^t, y^t, b^t; y^t, b^t)}{1+ \vec{D}_0^s(x^{t+1}, y^{t+1}, b^{t+1}; y^{t+1}, b^{t+1})}, \tag{7}$$

Where the directional distance functions $\vec{D}(x, y, b; g_y, g_b) = \max\{\beta | (y + \beta g_y, b - \beta g_b) \in P^s(x)\}$, s=t, t+1, are defined on each of the contemporaneous technology set. Since without any restrictions on the two production technologies, the contemporaneous Malmquist-Luenberger productivity index is typically defined with a geometric mean form of two-period Malmquist-Luenberger productivity indexes as follows:

$$ML^{t,t+1} = \left[\frac{1+ \vec{D}_0^t(x^t, y^t, b^t; y^t, b^t)}{1+ \vec{D}_0^t(x^{t+1}, y^{t+1}, b^{t+1}; y^{t+1}, b^{t+1})} \frac{1+ \vec{D}_0^{t+1}(x^t, y^t, b^t; y^t, b^t)}{1+ \vec{D}_0^{t+1}(x^{t+1}, y^{t+1}, b^{t+1}; y^{t+1}, b^{t+1})}\right]^{1/2} \tag{8}$$

The geometric mean form of the conventional Malmquist-Luenberger productivity index can be decomposed as follows:

$$
\begin{aligned}
&\mathrm{ML}^{t,t+1} \\
&= \frac{1+\vec{D}_o^{t}(x^{t},y^{t},b^{t};\ y^{t},b^{t})}{1+\vec{D}_o^{t+1}(x^{t+1},y^{t+1},b^{t+1};\ y^{t+1},b^{t+1})} \\
&* \left[\frac{1+\vec{D}_o^{t+1}(x^{t+1},y^{t+1},b^{t+1};\ y^{t+1},b^{t+1})}{1+\vec{D}_o^{t}(x^{t+1},y^{t+1},b^{t+1};\ y^{t+1},b^{t+1})}\frac{1+\vec{D}_o^{t+1}(x^{t},y^{t},b^{t};\ y^{t},b^{t})}{1+\vec{D}_o^{t}(x^{t},y^{t},b^{t};\ y^{t},b^{t})}\right]^{\frac{1}{2}} \\
&= \mathrm{MLEC}^{t,t+1} * \mathrm{MLTC}^{t,t+1}
\end{aligned}
\tag{9}
$$

Values of $\mathrm{ML}^{t,\,t+1}$ and its components greater than 1 indicate improvements, values less than 1 indicate declines, and values equal 1 indicate no change in performance. Changes in efficiency are measured by $\mathrm{MLEC}^{t,\,t+1}$, which represents a movement of countries towards the best practice frontier. If $\mathrm{MLEC}^{t,\,t+1} > 1$, then there has been a movement toward the frontier in period *t+1*, and hence becomes more efficient. If $\mathrm{MLEC}^{t,\,t+1} < 1$, then it signifies that the country is farther away from the frontier in *t+1* than in *t*, and hence becomes less efficient. Technical (technological) changes are expressed by $\mathrm{MLTC}^{t,t+1}$ component which represents a shift in technology between t and *t*+1. Färe et al. (2001) interpret the values of the technical change component of the ML index as shifts of the production possibilities frontier in the direction of 'more goods and fewer bads', which results in the value of the $\mathrm{MLTC}^{t,t+1}$ index exceeding unity. If the $\mathrm{MLTC}^{t,t+1}$ index equals unity, this indicates that there was no shift in the production possibilities frontier. Finally, an $\mathrm{MLTC}^{t,t+1}$ index value of less than unity indicates a shift of the production possibilities frontier in the direction of 'fewer goods and more bads'. Similarly, according to Kumar (2006) $\mathrm{MLTC}^{t,t+1} > 1$ means that technical change enables more production of good and less production of bad output, whereas if $\mathrm{MLTC}^{t,t+1} <1$, there has been a shift in the frontier in the direction of fewer good outputs and more bad outputs.

The directional distance function can be calculated in several ways. In this study, the linear-programming technique data envelopment analysis (DEA) is employed. As the conventional Malmquist-Luenberger productivity index uses a geometric mean form of two contemporaneous ML indexes, ML index faces a potential LP infeasibility problem in measuring cross-period directional distance functions. To avoid potential infeasibility in DEA programs for cross-period distance function calculation we employ procedure suggested by Du et al. (2017) as a solution to the following linear programing problems:

$$
\begin{aligned}
&\overrightarrow{D_o}(x_{o,}y_{o,}b_{o,}g) = \max \beta \\
&\sum_{j=1}^{N} y_{rj}\lambda_j \geq (1+\beta)y_{ro}, r = 1,2,\dots,S \\
&\sum_{j=1}^{N} b_{pj}\lambda_j \leq (1-\beta)b_{to}, p = 1,2,\dots,U \\
&\sum_{j=1}^{N} x_{ij}\lambda_j \leq x_{io}, i = 1,2,\dots,M \\
&\lambda_j \geq 0, j = 1,2,\dots,N.
\end{aligned}
\tag{10}
$$

Constraints for bad outputs originally proposed in Chung et al. (1997) as equalities are here used as inequalities. It is assumed that good outputs would not be affected by producing less bad output when it is optimal. Constraints for undesirable outputs as $\leq$ also conform to the preferences of managers following environmental policy to reduce production of bad outputs with additional costs.

In the study we used panel of yearly aggregated data (Eurostat, 2018) of total agricultural output in million EUR (AO), labour (annual working units AWU), utilised agricultural area in thousand ha (UAA), fertilisers N, P, K in tonnes (NPK) and greenhouse gas (GHG) emission in thousand tonnes (CO_2, N_2O in CO_2 equivalent, CH_4 in CO_2 equivalent, HFC in CO_2 equivalent, PFC in CO_2 equivalent, SF_6 in CO_2 equivalent, NF_3 in CO_2 equivalent) of agriculture by selected 26 EU Member States over period from 2006 to 2015.

3 RESULTS

A growth of intensity of agricultural production was observed in almost all MS over the observed period 2006 - 2015. Agricultural output per UAA increased in all MS but Greece. Increasing agricultural output per GHG emissions in the EU MS except for Bulgaria could indicate that intensity of agricultural production in Bulgaria has been achieved without paying sufficient attention to production strategies reducing GHG emissions. (Fig. 1).

GHG emission reduction per agricultural output in the EU MS reflects positive impact of policy instruments and mitigation strategies. The increasing GHG emissions per agricultural output can been observed also in Romania and Slovakia and could be attributed e.g. to growth of inputs used (fertilisers) in primary agricultural production after the accession to the EU.

Warrying increasing trend of GHG emitted in the EU agriculture per ha of UAA has been however observed in many EU countries. Reduction of GHG emission per UAA occurred only in Greece, Ireland, Italy, Lithuania, Romania, Sweden and UK.

Growing GHG emission per labour (AWU) in the most EU MS can be attributed to declining number of AWU in the EU agriculture. GHG emission reduction per AWU was observed only in Ireland and UK.

The most efficient and eco-efficient among the EU MS in period of 2006-2015 on average were Netherlands and Denmark and Italy (Fig. 2). The most inefficient we found agriculture of Lithuania, Poland, Bulgaria, Latvia and Estonia. The highest eco-inefficiency showed agriculture of Ireland, Latvia and Lithuania.

All EU MS have been improving their eco-productivity over observed period except for Bulgaria (Fig. 3). We found that countries with highest cumulative change of eco-TFP were Cyprus, Denmark and Netherlands. These countries reached reasonably high level of TFP even we disregard ecological behaviour (model without GHG). On the other hand, Romania and Bulgaria reached the lowest level of TFP growth. Romania however payed more attention to GHG emission mitigation reaching higher level of eco-TFP compare to Bulgaria.

New MS namely Latvia, Slovenia, Estonia, Czech Republic and Slovakia showed significant improvement of TFP, while Eco TFP improvement was there negligible (Fig. 3). Technical efficiency improvement (catch-up effect) (Fig. 4) was the main contributor to the improvement of TFP in these countries since their agriculture try to reached the most productive EU MS. A gap in between of technical efficiency of Romania, Bulgaria, Hungary, Poland, some other even old MS (Austria, Greece, Spain, Portugal) and the most productive MS (Netherlands, Denmark) increased over observed years.

Agricultural eco efficiency was improved significantly in UK, Lithuania, Sweden, Cyprus, Denmark towards the most eco-efficient MS (Italy, Netherlands) (Fig. 4). Eco-efficiency fell down in Bulgaria, Romania, Greece, Portugal and Hungary.

Technological improvements contributed to the TFP growth in agriculture mainly in Latvia, Lithuania, Slovenia, Poland, Romania, Hungary and in the old MS in Austria, Greece, Portugal, Italy and Cyprus (Fig. 5).

We found eco-technological improvement in all EU MS (Fig.5). Introduction of new eco-technology measures contributed to eco-TFP growth especially in Cyprus, Netherlands and Denmark.

4 CONCLUSIONS

We found that countries with high eco-efficiency (Netherlands, Italy, Denmark) in agricultural production over 2006-2015, reached relatively high efficiency even we disregard their environmental concerns (model without GHG). The New Member States exhibited higher TFP growth than the old MS. Significant eco-TFP growth however, was observed mainly in agriculture of the old MS. Bulgaria and Romania showed the worst eco-performance development.

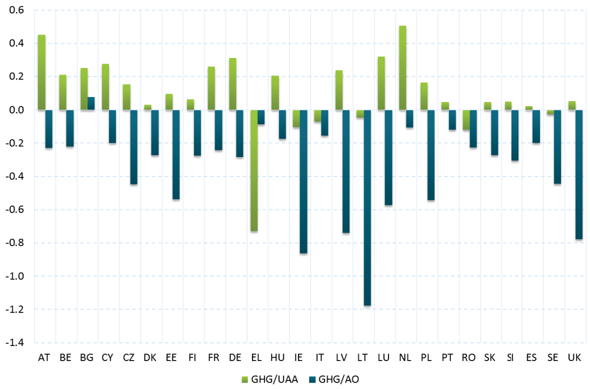


Fig. 1 Differences 2015 - 2006, GHG/UAA (tonnes per ha), GHG/AO (thousand tonnes per million EUR)

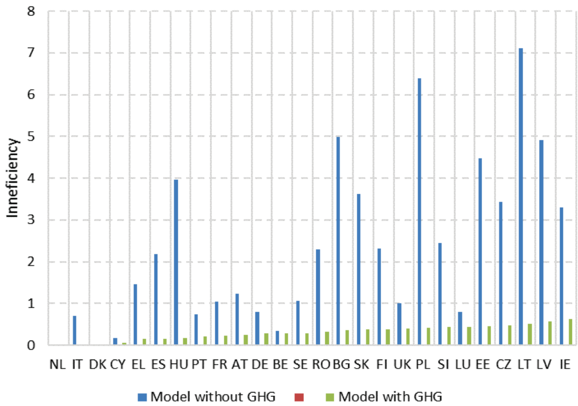


Fig. 2 Average technical inefficiency and eco-inefficiency, 2006-2015, (DDF)

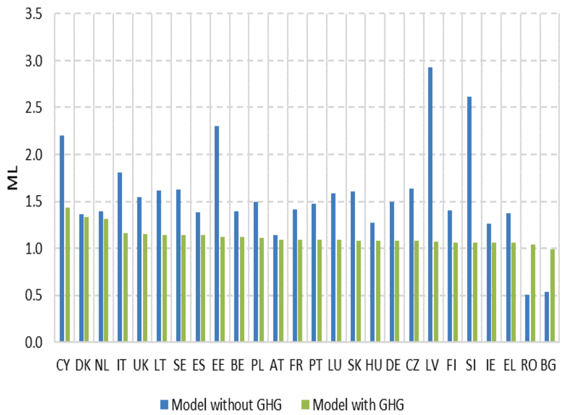


Fig. 3 Cumulative Change of TFP (ML), 2005-2016

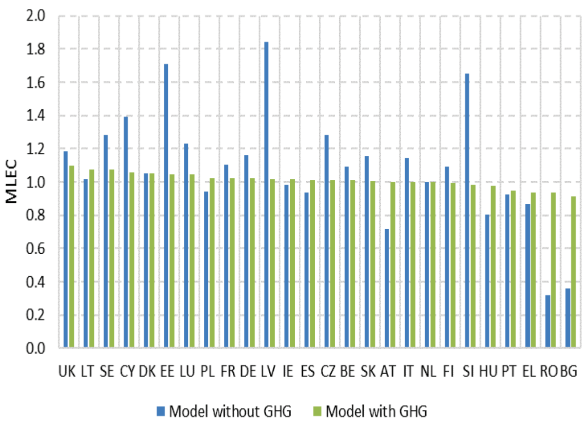


Fig. 4 Cumulative Technical efficiency and eco-efficiency changes (MLEC), 2006-2015

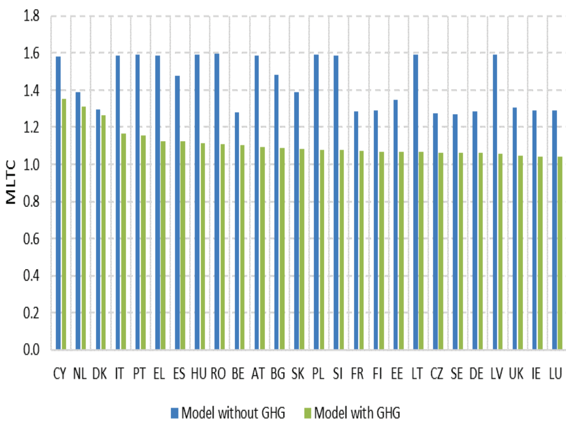


Fig. 5 Cumulative Technological Change (MLTC), 2006 - 2015

In order to estimate standard agricultural performance, we applied standard DDF without bad output. DDF approach was used as a suitable tool for eco-efficiency assessment, since in contrary to standard radial measures, it allows to model bad outputs with relatively direct approach. We employed new Du et al. (2017) model in which weak disposability axiom of Färe (2010) is redefined to avoid the most frequently occurred problem of infeasibility of linear programs used for cross-period DDF values estimation.

Acknowledgements: Authors acknowledge financial support of the Slovak Scientific Grant Agency VEGA 1/0845/17

References

[1] Chang, M. C. (2015). Room for improvement in low carbon economies of G7 and BRICS countries based on the analysis of energy efficiency and environmental Kuznets curves. *Journal of Cleaner Production*, 99, pp. 140–151.

[2] Charnes, A., Cooper, W. W. and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), pp. 429–444.

[3] Chung, Y. H., Färe, R. and Grosskopf, S. (1997). Productivity and undesirable outputs: A directional distance function approach. *J Environ Manage*, 51(3), pp. 229–240.

[4] Coli, M., Nissi, E., and Rapposelli, A. (2011). Efficiency Evaluation in an Airline Company: Some Empirical Results. *Journal of Applied Sciences*, 11(4), pp. 737–742.

[5] Du, J., Chen, Y. and Huang, Y. (2017). A Modified Malmquist-Luenberger Productivity Index: Assessing Environmental Productivity Performance in China, *European Journal of Operational Research*, http://dx.doi.org/10.1016/j.ejor.2017.01.006

[6] Eurostat (2018). Database – Eurostat. [online] Available at: http://ec.europa.eu/eurostat/data/database [Accessed 7 Jan. 2018].

[7] Färe, R., and Grosskopf, S. (2004). Modeling undesirable factors in efficiency evaluation: Comment. *European Journal of Operational Research,* 157(1), pp. 242–245.

[8] Färe, R., and Grosskopf, S. (2010). Directional distance functions and slacks-based measures of efficiency. *European Journal of Operational Research*, 200(1), pp. 320–322.

[9] Färe, R., Grosskopf, S. and Pasurka, C. A. Jr. (2001). Accounting for air pollution emissions in measures of state manufacturing productivity growth. *Journal of Reginal Science,* 41(3), pp. 381–409.

[10] Färe, R., Grosskopf, S. and Pasurka, C. A. Jr. (2007). Environmental production functions and environmental directional distance functions. *Energy*, 32(7), pp. 1055–1066.

[11] Färe, R., Grosskopf, S. and Pasurka, C. A. Jr. (2014). Potential gains from trading bad outputs: The case of U.S. electric power plants. *Resource and Energy Economics*, 36(1), pp. 99–112.

[12] He, F., Zhang, Q., Lei, J., Fu, W., and Xu, X. (2013). Energy efficiency and productivity change of China's iron and steel industry: Accounting for undesirable outputs. *Energy Policy*, 54, pp. 204–213.

[13] Hu, J. L., and Wang, S. C. (2006). Total-factor energy efficiency of regions in China. *Energy Policy,* 34(17), pp. 3206–3217.

[14] Hu, J. L., Sheu, H, J., and Lo, S. F. (2005). Under the shadow of Asian Brown Clouds: Unbalanced regional productivities in China and environmental concerns. *The*

International Journal of Sustainable Development and World Ecology, 12(4), pp. 429–442.
[15] Kumar, S. (2006). Environmentally sensitive productivity growth: A global analysis using Malmquist–Luenberger index. *Ecological Economics*, 56(2), pp. 280–293.
[16] Lovell, C. K., Pastor, J. T., and Turner, J. A. (1995). Measuring macroeconomic performance in the OECD: A comparison of European and non-European countries. *European Journal of Operational Research,* 87(3), pp. 507–518.
[17] Mahlberg, B. and Luptacik, M. (2014). Eco-efficiency and eco-productivity change over time in a multisectoral economic system. *European Journal of Operational Research*, 234(3), pp. 885–897
[18] Munda, G., and Saisana, M. (2011). Methodological considerations on regional sustain- ability assessment based on multicriteria and sensitivity analysis. *Regional Studies*, 45(2), pp. 261–276.
[19] Oh, D. (2010). A global Malmquist-Luenberger productivity index. *Journal of productivity analysis*, 34(3), pp. 183–197.
[20] Picazo-Tadeo, A. J., Beltrán-Esteve, M., and Gómez-Limón, J. A. (2012). Assessing eco–efficiency with directional distance functions. *European Journal of Operational Research,* 220(3), pp. 798–809.
[21] Picazo-Tadeo, A. J., Gómez-Limón, J. A., and Reig-Martínez, E. (2011). Assessing farming eco-efficiency: A data envelopment analysis approach. *Journal of Environmental Management*, 92(4), pp. 1154–1164.
[22] Rashidi, K., and Saen, R. F. (2015). Measuring eco-efficiency based on green indicators and potentials in energy saving and undesirable output abatement. *Energy Economics,* 50, pp. 18–26.
[23] Sahoo, B.K, Luptacik, M. and Mahlberg, B. (2011). Alternative measures of environmental technology structure in DEA: An application. *European Journal of Operational Research.* 215(3), pp. 750–762
[24] Vlontzos, G., Niavis, S. and Manos, B. (2014). A DEA approach for estimating the agricultural energy and environmental efficiency of EU countries. *Renewable and Sustainable Energy Reviews,* 40, pp. 91–96
[25] Wang, H., Zhou, P., and Zhou, D. Q. (2013). Scenario-based energy efficiency and productivity in China: A non-radial directional distance function analysis. *Energy Economics*, 40, pp. 795–803
[26] Wang, Y., Liu, J., Hansson, L., Zhang, K., Wang, R. (2011). Implementing stricter environmental regulation to enhance eco-efficiency and sustainability: a case study of Shandong province's pulp and paper industry, *China. J. Clean. Prod*. 19(4), pp. 303–310.
[27] World Business Council for Sustainable Development (2000). *Eco-efficiency, creating more value with less impact*. Available at. www.wbcsd.org
[28] Zhang, B., Bi, J., Fan, Z., Yuan, Z., and Ge, J. (2008). Eco-efficiency analysis of industrial system in China: A data envelopment analysis approach. *Ecological Economics,* 68(1), 3 pp. 06–316.
[29] Zhang, N., Kong, F., and Choi, Y. (2014). Measuring sustainability performance for China: A sequential generalized directional distance function approach. *Economic Modelling*, 41(3) pp. 92–397.
[30] Zhou, H., Yang, Y., Chen, Y. and Zhu, J. (2018). Data envelopment analysis application in sustainability: The origins, development and future directions. *European Journal of Operational Research,* 264(1), pp. 1–16

[31] Zofío, J. L., and Prieto, A. M. (2001). Environmental efficiency and regulatory standards: The case of CO 2 emissions from OECD industries. *Resource and Energy Economics*, 23(1), pp. 63–83.

Authors address
doc. Ing. Peter Fandel, CSc.
prof. Ing. Lubica Bartova, CSc.
Slovak University of Agriculture in Nitra, Faculty of Economics and Management,
Department of Statistics and Operations Research
Tr. A. Hlinku 2, 949 11 Nitra
Slovak Republic
email: peter.fandel@uniag.sk, lubica.bartova@uniag.sk

A CODE FORM GAME IN A SELECTION PROCESS

Manuel Alberto M. Ferreira,
Instituto Universitário de Lisboa (ISCTE-IUL), BRU-IUL, ISTAR-IUL, Portugal
Maria Cristina Peixoto de Matos,
Instituto Politécnico de Viseu (IPV), Escola Superior deTecnologia e Gestão de Viseu (ESTGV), CI&DETS, Portugal

Abstract
The aim of this paper is to present, in the scope of game theory, a new games representation: the code form, consisting of a table which contains the whole information, without any suppression or "adulteration "with regards to the game. After the theoretical description, its application is exemplified in solving the problem of selecting a candidate up to his/her acceptance trough game theory. The originality this paper proposes is how this problem will be approached: it will be treated as a single game which is made up of two parts, going as far as to state that the payoffs in the first part of the game will be the mediators of the second part of the game, instead of using the stochastic dynamic programming approach. The solution of the problem using stochastic dynamic programming is also presented to highlight empirically the parallelism between the two modes of resolution.

Keywords: *Game theory, two-part game, code form, stochastic dynamic programming.*

JEL Classification: C78
AMS Classification: 91A06

1 INTRODUCTION

The game here described represents a real life situation: candidate selection for a given job. It will be shown, because of the complexity of the relationships between different parties and the fact that each one of them is an essential element in the game's decision-taking, that the best candidate is not always the best option. On the other hand, through this game is also shown that knowing the fundamental elements that constitute a game - who are the players, what are the strategies for each player and the payoffs each player can receive - is not enough to know its solution, if one even exists.

This game is an example of game theory application to human resources management. On this subject see also (Andrade et al., 2012), (Ferreira et al., 2008, 2012, 2014, 2016), Ferreira (1991, 2014), (Filipe et al, 2012), (Matos and Ferreira, 2005) and (Matos et al., 2018).

2 GAME PRESENTATION

A presidential decree reduced the number of candidates to the vice-presidency to three people. Each of the three candidates are ranked on a scale from 1 (lowest) to 10 (highest).The presidential board attributed 10 points, 8 points and 5 points to the candidate classified in 1st place, 2nd place and 3rd place respectively. The probabilities of candidate i (i=1,2,3) accepting the j-th offer to run for the vice-presidency have been defined, considering that the first j-1 offers to the others have been declined, are denoted by p_{ij} and the respective values are in Table 1.

Table 1 Acceptance Probabilities

Player 1	$p_{11} = 0.5$	$p_{12} = 0.2$	$p_{13} = 0.0$
Player 2	$p_{21} = 0.9$	$p_{22} = 0.5$	$p_{23} = 0.2$
Player 3	$p_{31} = 1.0$	$p_{32} = 0.8$	$p_{33} = 0.4$

The question is:
-*What is the order in which the three potential candidates be offered the vice-presidential nomination if the presidential decree imposes the expected number of points maximization, supposing that no candidate is requested more than once and, each time a candidate rejects, another one is requested, until at least one has accepted or all have rejected.*

3 GAME DETAILS

This game which is made up of two parts – a selection process and an acceptance process – attests that the payoffs in the first part of the game (potential candidates) will be the intermediaries of the second (decision elements).Thus:

Players:
Presidential Board:**4**
Potential Candidates:
Player classified in 1st place –**1**;
Player classified in 2nd place – **2**;
Player classified in 3rd place – **3**.

Strategies[1] :
Presidential Board:
The presidential board wants to establish the order in which the potential candidates will be invited to maximize the expected number of points. In this way the strategy for the presidential board will be the order in which the three potential candidates can be offered the vice-presidential nomination until at least one has accepted or all have rejected the offer - **P**.

Potential Candidates:
The strategy of each potential candidate is to accept the offer – **A** - or to reject the offer - **R**.

Payoffs:
Presidential Board:
The presidential board payoff is (a function of the attributed points in the daily preselection and
of the potential candidates' probabilities of acceptance of the vice-presidency) the expected number of points of each possibility in the order of the proposal presented to the potential

[1] See (Bicchieri et al., 1999).

candidates. Thus, for instance: Player 1 rejects the offer, 3 rejects the offer, 2 accepts the offer,
presidential board payoff is: $\frac{1}{2} \times \frac{1}{5} \times \frac{1}{5} \times 8 = 0.16$, where the probability of player 1 rejecting the offer is $\frac{1}{2}$, $\frac{1}{5}$ is the probability of player 3 rejecting the offer and $\frac{1}{5}$ is the probability of player
2 accepting the offer.

Potential Candidates:
For these players it is possible to define: if the player accepts the offer, he/she gets the "total prize", that is, he/she gets payoff 1. On the other hand, if he/she rejects the proposal, he/she does not get anything, so his/her payoff will be 0.

4 CODE FORME REPRESENTATION FOR THE GAME

To represent the game, code form game representation, see (Matos and Ferreira, 2002, 2002a, 2003, 2005), will be used. Code form is basically a table where the strategies that are available to any player are codified.

Definition 1:
A code form game consists of a finite table, with evident extension in the case of infinite moves and infinite players, where only some cells are filled. The cells will be filled in the same order the game is played. For that it is needed: $R = \{1,2,\dots,R\}$-a set of rounds; $J = \{1,2,\dots,J\}$-a set of moves; $C = \{1,2,\dots,J+3\}$-a set of columns; $L = \{1,2,\dots L\}$-a set of rows; $N = \{1,2,\dots,N\}$-a set of players; $E_N = \{e_1, e_2, \dots, e_N\}$-a set of strategies available to each player; $E = \{E_1 \times E_2 \times \dots \times E_N\}$-a set of all such strategies profiles (space of strategies profiles); $\begin{matrix} RN: L \times \{1\} \to R \\ a_{i1} \to r \end{matrix}$, a function that indicates the round number; $\begin{matrix} PN: L \times \{2\} \to J \\ a_{i2} \to j \end{matrix}$, a function that indicates the move number; $\begin{matrix} JE: L \times C \to N \times E_N \\ a_{i,k} \to (N, E_N) \end{matrix}, c \neq 1,2,J+3$, a function that indicates who moves and what action is played; and $\begin{matrix} PJ: L \times \{J+3\} \to IR^N \\ a_{i,J+3} \to (u_1, u_2, \dots, u_N) \end{matrix}$, a function that gives the payoff of every player (for all the players) where $u_N: E \to IR$ is a von Neumann-Morgenstern utility function.■
Note:
-It can be denoted that when $RN(a_{i1}) = RN(a_{i-1,1})$, $PN(a_{i2}) = PN(a_{i-1,2})$ and $JE(a_{ic}) = JE(a_{i-1,c})$, the cells are not filled. The line is changed when the move changes. The column is changed when the player changes. ■

Code form game idea lies in the game estimated linear reading. The table is built containing the whole game information. In Table 2 is exemplified the code form game representation in the game example that is now considered. Reading from left to the right, the first column indicates the period number and the second column indicates the move[2] number. The following columns mention who moves when and in what circumstances and what action is played when somebody is called upon to move. Last column indicates the payoffs vector in

[2] See (Benoit and Krishna, 1985) and Eberwein (2000).

accordance with the strategies chosen by the players. It is easy to check that the order in which the three potential candidates can be offered the vice-presidential nomination must be:
-*To invite in the first place the candidate classified in 2nd place, 2; if he/she rejects the proposal, the candidate classified in third place, 3, should be invited and if he/she does not accept, the candidate classified in first place, 1 should be invited. The expected number of points is 7.6.*

Table 2 Code Form Game

1	1	(4,P)				
1	2		(1,A,0.5)			(5,1,0,0)
			(1,R,0.5)			
	3			(2,A,0.5)		(2,0,1,0)
				(2,R,0.5)		
	4				(3,A,0.4)	(0.5,0,0,1)
					(3,R,0.6)	(0,0,0,0)
	3			(3,A,0.8)		(2,0,0,1)
				(3,R,0.2)		
	4				(2,A,0.2)	(0.16,0,1,0)
					(2,R,0.8)	(0,0,0,0)
	2		**(2,A,0.9)**			(**7.2**,0,1,0)
			(2,R,0.1)			
	3			(1,A,0.2)		(0.2,1,0,0)
				(1,R,0.8)		
	4				(3,A,0.4)	(0.16,0,0,1)
					(3,R,0.6)	(0,0,0,0)
	3			**(3,A,0.8)**		(**0.4**,0,0,1)
				(3,R,0.2)		
	4				**(1,A,0)**	(**0**,1,0,0)
					(1,R,1)	(**0**,0,0,0)
	2		(3,A,1)			(5,0,0,1)
			(3,R,0)			
	3			(1,A,0.2)		(0,1,0,0)
				(1,R,0.8)		
	4				(2,A,0.2)	(0,0,1,0)
					(2,R,0.8)	(0,0,0,0)
	3			(2,A,0.5)		(0,0,1,0)
				(2,R,0.5)		
	4				(1,A,0)	(0,1,0,0)
					(1,R,1)	(0,0,0,0)

5 RESOLUTION THROUGH STOCHASTIC DYNAMIC PROGRAMMING

Now, a resolution using stochastic dynamic programming tools is presented. In this context this is a problem with three *stages*, representing the stage j the j^{th} position in the invitation order. Define the *states* as the list of candidates still not invited. The stage 1 has only 1 state:

$U_{11} = \{1,2,3\}$. The stage 2 has the three states: $U_{21} = \{1,2\}, U_{22} = \{1,3\}, U_{23} = \{2,3\}$ and the stage 3 has the three states: $U_{31} = \{1\}, U_{32} = \{2\}, U_{33} = \{3\}$.

Define also:
-$\boldsymbol{m_j(U_{jk})}$: Achieved points expected maximum number, beginning in stage j in state U_{jk}, supposing that there were no acceptations in the former stages,
-$\boldsymbol{d_j(U_{jk})}$: Candidate requested in stage j in order to get $m_j(U_{jk})$ and
-$\boldsymbol{V_i}$: Value in points of candidate i.

The recurrent formula, through which it is applied the so known backward induction process, is:

$$m_j(U_{jk}) = \max_{i \in U_{jk}}\left[V_i p_{ij} + (1 - p_{ij}) m_{j+1}(U_{jk} - \{i\})\right] \quad (1).$$

Follow its interpretation:
-If in state j the candidate i is requested and accepts, the value is V_i . Anyway if that candidate rejects, the best to go on is to depart from the state composed by the candidates that still were not requested. This formula is valid only for $j = 1,2,3$ if it is imposed that $m_4 = 0$.

So, going on,

Stage 3

$$m_3(U_{31}) = 10(0) = 0 \text{ with } \boldsymbol{d_3(U_{31}) = 1}$$
$$m_3(U_{32}) = 8(0.2) = 1.6 \text{ with } d_3(U_{32}) = 2$$
$$m_3(U_{33}) = 5(0.4) = 2.0 \text{ with } d_3(U_{33}) = 3$$

Stage 2

$$\begin{aligned} m_2(U_{21}) &= \max\{10(0.2) + (1 - 0.2)m_3(U_{32}), 8(0.5) + (1 - 0.5)m_3(U_{31})\} \\ &= \max\{10(0.2) + (1 - 0.2)(1.6), 8(0.5) + (1 - 0.5)(0)\} = 4 \text{ with } d_2(U_{21}) \\ &= 2 \end{aligned}$$

$$\begin{aligned} m_2(U_{22}) &= \max\{10(0.2) + (1 - 0.2)m_3(U_{33}), 5(0.8) + (1 - 0.8)m_3(U_{31})\} \\ &= \max\{10(0.2) + (1 - 0.2)(2.0), 5(0.8) + (1 - 0.8)(0)\} = 4 \text{ with } \boldsymbol{d_2(U_{22})} \\ &= \boldsymbol{3} \end{aligned}$$

$$\begin{aligned} m_2(U_{22}) &= \max\{8(0.5) + (1 - 0.5)m_3(U_{33}), 5(0.8) + (1 - 0.8)m_3(U_{32})\} \\ &= \max\{4 + (1 - 0.5)(2.0), 5(0.8) + (1 - 0.8)(1.6)\} = 5 \text{ with } d_2(U_{23}) = 2 \end{aligned}$$

Stage 1

$$\begin{aligned} m_1(U_{11}) &= \max\{10(0.5) + (1 - 0.5)m_2(U_{23}), 8(0.9) + (1 - 0.9)m_2(U_{22}), 5(1) \\ &\quad + (1 - 1)m_2(U_{21})\} = \max\{10(0.5) + (1 - 0.5)(5), 7.2 + 0.1(4), 5 + 0(4)\} \\ &= \boldsymbol{7.6} \text{ with } \boldsymbol{d_1(U_{11}) = 2} \end{aligned}$$

The the optimal policy is to request first the candidate 2. If he/she does not accept to request then the candidate 3 . And finally if this do not accept to request the candidate 1.

6 CONCLUSIONS

The analyzed game illustrates that the player who has decision power is not always the one who decides the game. In other words, the presidential board is the player who dominates the situation; therefore, they decide who the best candidate is. But, indeed, the ones who actually decide the game are the potential candidates when either they accept or they refuse the proposal. Really, the potential candidates are the ones who determine the game's outcome;

starting from a situation of weakness they gain the control of the game. They are the result of the first part of the game and become the deciding elements of the second part of the game.

On the other hand, this game shows how much Game Theory is still a subject with a long way to go for games with more than two players, see Vega-Redondo (2003). No solution concept for these kinds of games is universally accepted. One reason for this is the situation described herein since one could not find the respective equilibrium using existing solution concepts, such as the Nash equilibrium, the Shapley value, and so on.

This in no way reduces the significance of Game Theory. In fact, besides motivating the theorists and considering the already developed results, it must not be forgotten Game Theory is one of the few theories that defines rational procedures in what were previously considered irrational situations and that its concepts and ideas have already provided very important and deep knowledge in the formulation of situations to real world conflicts.It is evident to conclude how much Game Theory is an asset to humanity, see for instance Gibbons (1992), (Kennan and Wilson, 1993), (Selvarasu et al, 2009) and Weibull (1995). Finally, note how the approach through Game Theory to the considered problem gives much more relevance to the human behaviour, that appears somehow hided in the approach through Stochastic Dynamic Programming.

Acknowledgments: This work is financed by national funds through FCT - Fundação para a Ciência e Tecnologia, I.P., under the projects UID/MULTI/4466/2016 and UID/MULTI/04016/ 2016. Furthermore, the authors thank the Instituto Universitário de Lisboa and ISTAR-IUL, and Instituto Politécnico de Viseu and CI&DETS for their support.

References

[1] Andrade, M., Ferreira, M. A. M. and Filipe, J. A. (2012). Representation of reserves through a Brownian motion model. SPMCS 2012, Book Series: *Journal of Physics Conference Series*, Volume: 394, Article Number: 012036. DOI: 10.1088/1742-6596/394/1/012036

[2] Bicchieri, C., Jeffrey, R. and Skyrms, B. (1999). *The Logic of Strategy*. Oxford University Press, Inc..

[3] Benoit, J. P. and Krishna, V. (1985).Finitely Repeated Games. *Econometrica,* 53,pp. *890–904*.

[4] Eberwein, C. (2000). *The Sensitivity of Repeated Bargaining Outcomes to the Choice of First Proposer*. June.

[5] Ferreira, M. A. M.(1991). An infinite-server queuing system with an exponential busy period. *Proceedings of 15th Portuguese-Spanish Conference on Mathematics* (Braumann, C. A. ed.), 1-6, Universidade de Évora, Évora, Portugal, pp. D63-D66.

[6] Ferreira, M. A. M., Filipe, J. A., Coelho, M., et al. (2008). An infinite servers model for motor vehicles dismantling and recycling in a scarce energy ambience. In: *Quantitative Methods in Economics (Multiple Criteria Decision Making XIV).* [online] High Tatras, Slovakia: IURA EDITION SPOL SRO, pp.178-186. Available at: http://fhi.sk/files/katedry/kove/ssov/proceedings/Zbornik2008.pdf [Accessed 2 February 2018].

[7] Ferreira, M. A. M., Andrade, M., and Filipe, J. A. (2012). Studying pensions funds throug an infinite servers nodes network: a theoretical problem. SPMCS 2012, Book Series: *Journal of Physics Conference Series*, Volume: 394, Article Number: 012035.

DOI:10.1088/1742-6596/394/1/012035
[8] Ferreira, M. A. M., Filipe, J. A., Coelho, M., et al. (2014). Chaos in World Politics: A Reflection The "Drop of Honey Effect". In: *Chaos Theory in Politics* (Banerjee, S., Ercetin, S.S.,Tekin, A., eds.), Book Series: Understanding Complex Systems, Springer Complexity , pp.89-104.
[9] Ferreira, M. A. M. (2014).The pandemic period length modelled through queue systems. In: *Quantitative Methods in Economics (Multiple Criteria Decision Making XVII).*[online] Virt, Slovakia: Vydavateľstvo EKONÓM, pp. 43-47. Available at: http://fhi.sk/files/katedry/kove/ssov/proceedings/Zbornik2014.pdf [Accessed 2 February 2018].
[10] Ferreira, M. A. M., Filipe, J. A. and Coelho, M. (2016). Unemployment period quantitative approach through infinite servers queue systems. In: *Quantitative Methods in Economics (Multiple Criteria Decision Making XVIII)*. [online] Vrátna, Slovakia: Letra Interactive, s. r. o., pp. 89-94. Available at: http://fhi.sk/files/katedry/kove/ssov/proceedings/Zbornik2016.pdf [Accessed 2 February 2018].
[11] Filipe, J. A., Ferreira, M. A. M. and Andrade, M. (2012). Reserves represented by random walks. SPMCS 2012, Book Series: *Journal of Physics Conference Series*, Volume: 394, Article Number: 012034. DOI:10.1088/1742-6596/394/1/012034
[12] Gibbons, R. (1992). *Game Theory for Applied Economist*. Princeton: Princeton University Press.
[13] Kennan, J. and Wilson, R. *(1993)*. Bargaining with Private Informaiton. *Journal of Economic Literature*, 31(1), pp. 45-104.
[14] Matos, M.C. P. and Ferreira, M.A.M. (2002). Games in Code Form. *Proceedings of the Fifth Spanish Meeting on Game Theory*. Seville. July 1-3.
[15] Matos, M. C.P. and Ferreira, M. A. M. (2002a). Games in code form versus games in extensive form. In: *Quantitative Methods in Economics(Multiple Criteria Decision Making XI)*. [online] Nitra, Slovakia, pp. 185-193. Available at: http://fhi.sk/files/katedry/kove/ssov/proceedings/Zbornik2002.pdf [Accessed 2 February 2018].
[16] Matos, M. C. P. and Ferreira, M. A. M. (2003). Best payoff method. *Proceedings of XV IMGTA*. Italian Meeting on Game Theory and Applications. Urbino, Italy. 9-12 July. 2003.
[17] Matos, M. C. P.; Ferreira, M. A. M. (2005). Select and accept. *Proceedings of the 8th Joint Conference on Information Sciences*, Vols 1-3.Eds: Blair, S.; Chakraborty, U.; Chen, S.H. *et al.,* Salt Lake City, UT. July 21-26, pp. 532-534.
[18] Matos, M.C.P., Ferreira, M.A.M. and Filipe, J.A. (2018). Let the games begin and go on.,*Int. J. Business and Systems Research*, 12(1), pp.43–52. DOI: 10.1504/IJBSR.2018.10009017
[19] Selvarasu, A., Filipe, J. A., Ferreira, M. A. M., *et al.* (2009). A strategic profit model to measure Indian apparel retail performance. In: *APLIMAT 2009: 8th International Conference, Proceedings* (Kovacova, M., ed.). Bratislava, Slovakia, pp.311-331.
[20] Vega-Redondo, F. (2003) .*Economics and the Theory of Games"*. Cambridge: Cambridge University Press.
[21] Weibull, J. (1995). *Evolutionary Game theory*. Cambridge: MIT Pres. 1995.

Author's Address
Professor Catedrático Manuel Alberto M. Ferreira
Instituto Universitário de Lisboa (ISCTE-IUL), Escola de Tecnologias e Arquitectura (ISTA), Departamento de Matemática, BRU-IUL, ISTAR-IUL
Av. das Forças Armadas, 1649-026 Lisboa, Portugal
Email: manuel.ferreira@iscte.pt

Professora Adjunta Maria Cristina Peixoto de Matos
Instituto Politécnico de Viseu (IPV), Escola Superior deTecnologia e Gestão de Viseu (ESTGV), CI&DETS
Campus Politécnico, 3504-510 Viseu, Portugal
Email: cristinapeixoto@estgv.ipv.pt

APPLICATIONS OF BIFORM GAMES

Petr Fiala, University of Economics, Prague
Renata Majovská, University of Finance and Administration, Prague

Abstract
Traditional game theory is divided into non-cooperative and cooperative models. Biform games combine non-cooperative and cooperative models. There are a number of situations where the competitive and cooperative behaviors of decision-makers are combined. The paper is devoted to the analysis of some specific situations in which these behaviors occur and biform game models can be used. The sequential biform game is used in supply chains. In the non-cooperative part, a coordination mechanism based on a specific contract is applied between producers and customers. The cooperative part is merely focused on two concepts, coalition formations by resource capacity constraints and profit sharing. The second application is focused on co-opetition models. The players in the co-opetition model are the firm, customers, suppliers, competitors and complementors (competitors whose products add value). The relationship between the firm and the direct competitors is non-cooperative. The relationship between the firm and the complementors in a search for common added values is cooperative. The last example of a biform game application is an environmental subsidy negotiation where polluters behave in a cooperative way to obtain a subsidy from the authority for a joint action while competing with the distribution of this subsidy among themselves.

Keywords: *biform game, supply chain, co-opetition, environmental negotiation*

JEL Classification: C44
AMS Classification: 90B10

1 INTRODUCTION

Many economic problems can be modeled and solved by game theory. The work of John von Neumann and Oskar Morgenstern (1944) is the classic work upon which modern game theory is based. Since then, an extensive literature on game theory was published. For example, Myerson's book (1997) provides a clear and thorough examination of the models, solution concepts, results, and methodological principles of non-cooperative and cooperative game theory. Game theory models analyze situations where players make decisions to maximize their own utility, while taking into account that other players are doing the same, and that decisions, made by players, impact others utilities. Traditional game theory is divided into non-cooperative and cooperative models. The non-cooperative theory of games is strategy oriented; it studies what one may expect the players to do. The non-cooperative theory is a “micro” approach in that it focuses on precise descriptions of what happens. A biform game is a combination of non-cooperative and cooperative games (Brandenburger and Stuart, 2007). We propose to divide the biform games into so called sequential and simultaneous shapes. The concept of biform games has been refined and applied in many cases. We draw attention to the application of biform game principles in supply chains, co-opetitive models, and environmental negotiations.

Supply chain management is a philosophy that provides the tools and techniques enabling organizations to develop strategic focus and achieve sustainable competitive advantage. It presents management with a new focus and way of thinking about how their organization exists and operates within the wider business environment. We propose to use a sequential biform game for supply chain analyses.

Co-opetition combines the advantages of both competition and cooperation into new dynamic, which can be used to not only generate more profits but also to change nature of the business environment in benefit of users (Brandenburger and Nalebuff, 2011). Current business conditions are changing rapidly. New products are evolving faster. Searching for relationships with complementors brings ever new opportunities that bring added values. This is one of the basics of coopetition. This dynamics must be included in the new cooperative models.

Environmental negotiations in the search for coalition projects to reduce pollution can be modeled as a simultaneous biform game. Coalitions of polluters proposes projects to environmental authority that must simultaneously capture external and internal pressures. External pressure is determined by the limits of authority for granting subsidies. For each polluter, the internal pressure is due to the need to cover the cost of reducing pollution.

The rest of the paper is organized as follows. Section 2 presents basics of biform games and introduces our concept of dividing biform games into sequential and simultaneous shapes. Section 3 informs about using of biform games in supply chains. Biform games as basic models for modeling of co-opetition problems are analyzed in Section 4. Modeling of negotiation about forming coalition projects for environmental pollution reduction by biform games is presented in Section 5. Section 6 presents conclusions.

2 BIFORM GAMES

A biform game is a combination of non-cooperative and cooperative games (Brandenburger and Stuart, 2007). We propose to divide biform games into sequential and simultaneous shapes.

Sequential biform games

The sequential biform game is a two-stage game: in the first stage, players choose their strategies in a non-cooperative way, thus forming the second stage of the game, in which the players cooperate. First, suppliers make initial proposals and take decisions. This stage is analyzed using a non-cooperative game theory approach. The players search for Nash equilibrium by solving the next problem.

An n-player non-cooperative game in the normal form is a collection

$$\{N = \{1,2,\dots,n\};\ X_1, X_2, \dots, X_n;\ \pi_1(x_1, x_2, \dots, x_n), \pi_2(x_1, x_2, \dots, x_n), \dots, \pi_n(x_1, x_2, \dots, x_n)\}, \tag{1}$$

where N is a set of n players; X_i, i = 1, 2, …, n, is a set of strategies for player i; $\pi_i(x_1, x_2, \dots, x_n)$, i = 1, 2, …, n, is a pay-off function for player i, defined on a Cartesian product of n sets X_i, i = 1, 2, …, n.

Decisions of other players than player i are summarized by a vector

$$\mathbf{x}_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1} \ldots, x_n). \tag{2}$$

A vector of decisions $(x_1^0, x_2^0, \ldots, x_n^0)$ is Nash equilibrium of the game if

$$x_i^0(\mathbf{x}_{-i}^0) = \mathrm{argmax}_{x_i}\, \pi_i(x_i, \mathbf{x}_{-i}) \forall i = 1,2, \ldots, n. \tag{3}$$

Nash equilibrium is a set of decisions from which no player can improve the value of his pay-off function by unilaterally deviating from it.

Then, players negotiate among themselves. In this stage, a cooperative game theory is applied to characterize the outcome of negotiation among the players over how to distribute the total surplus. Each player's share of the total surplus is the product of its added value and its relative negotiation power. We propose to distribute the total surplus to players by Shapley values (8).

Cooperative game theory looks at the set of possible outcomes, studies what the players can achieve, what coalitions will form, how the coalitions that do form divide the outcome, and whether the outcomes are stable and robust.

The maximal combined output is achieved by solving the following task

$$\mathbf{x}^0 = \mathrm{argmax}_{\mathbf{x}} \sum_{i=1}^{n} \pi_i(x_i). \tag{4}$$

When modeling cooperative games is advantageous to switch from the game in normal form to the game in the characteristic function form. The characteristic function of the game with a set N of n players is such function $v(S)$ that is defined for all subsets $S \subseteq N$ (i.e. for all coalition) and assigns a value $v(S)$ with following characteristics:

$$v(\varnothing) = 0, \quad v(S_1 \cup S_2) \geq v(S_1) + v(S_2), \tag{5}$$

where S_1, S_2 are disjoint subsets of the set N. The pair (N, v) is called a cooperative game of n players in the characteristic function form.

Allocation mechanisms are based on different approaches such as Shapley values, contracts, auctions, negotiations, etc. A particular allocation policy, introduced by Shapley (1953) has been shown to possess the best properties in terms of balance and fairness. So called Shapley vector is defined as

$$\mathbf{h} = (h_1, h_2, \ldots, h_n), \tag{6}$$

where the individual components (Shapley values) indicate the mean marginal contribution of i-th player to all coalitions, which may be a member. Player contribution to the coalition S is calculated by the formula:

$$v(S) - v(S - \{i\}). \tag{7}$$

Shapley value for the i-th player is calculated as a weighted sum of marginal contributions according to the formula:

$$h_i = \sum_S \left\{ \frac{(|S|-1)!(n-|S|)!}{n!} . \left[v(S) - v(S - \{i\}) \right] \right\}, \tag{8}$$

where the number of coalition members is marked by symbol $|S|$ and the summation runs over all coalition $i \in S$.

Confidence indices $0 \leq \alpha^i \leq 1$, for all i = 1, 2, …, n, are introduced. The indices show players' anticipation of the pay-off they will receive in the cooperative stage, i.e. the proportion of the difference between the maximum and minimum core allocation achievable for players. Confidence indices of the players provide the link between the non-cooperative and cooperative stages of the biform game.

Simultaneous biform games

The simultaneous biform game is a one-stage model where combinations of concepts for cooperative and non-cooperative games are applied. The combinations will be changed according situations in problems. The first problem is a classification of situations. The situations are affected by:

- which players can cooperate,
- to what scope they can cooperate.

If all players can cooperate fully, then a standard cooperative model can be used (4) with subsequent distribution of the result according to the Shapley values (8). If no one can cooperate even in a partial content, a standard non-cooperative model is used (3). These are two extreme cases.

The pressure negotiation concept is based on the assumption that each negotiating subject decides under pressure of objective context, subject to a variety of internal and external pressures. The scope of cooperation is determined by the various constraints that result from the fact that players are under internal and external pressures. The scope of cooperation is dynamic and changes over time. The effects of pressures will be reflected in restrictive conditions.

Negotiating subject is under pressure, for example, if he wants to reach a consensus, he is aware of prices for delayed decisions; other negotiating subjects influence their behavior, etc. Pressure is a term that includes internal values and external influences and determines the decision making process. Assume that the pressure does not affect the selection of decisions directly, but through a set of conditions that have to be satisfied. Then we can consider the effects of pressure, which is reflected in changes in the set of constraints. This leads to a change of set of acceptable negotiating subjects' decisions and a change of the negotiation space and can lead to a consensus.

3 SUPPLY CHAINS

Analysis of supply chains is provided by non-cooperative (Cachon and Netessine, 2004) and cooperative (Nagarajan and Sošić, 2008) game theory. We propose to use a sequential biform game in supply chains (see Fiala, 2015, 2016)). The problem is formulated as a supply chain with layers of suppliers, producers, retailers and customers. Suppliers form a layer with *m* agents and provide *m* types of resources to producers. The layer of producers is represented by *n* agents. These agents produce one type of product. The production is characterized by consumption of *m* resources to produce one unit of the final product. Each production agent is characterized by its available production resources. The resource capacity constraints compare the total availability of resources in the production layer with total consumption of resources to produce total number of *q* units of products. Producers send the products to retailers. Retailers meet price-dependent stochastic demand of customers. This problem is solved by two-stage procedure based on combination of no-cooperative and cooperative game approaches.

The first stage solves problems by price-dependent stochastic demand of customers:

- How to get maximal profit from customers.
- How to allocate the maximal profit between retailers and producers.

The problems are solved by non-cooperative manner. A Stackelberg game is formulated between the layer of producers and the layer of retailers as a newsvendor problem with pricing. Retailers seek to maximize total profit from the sale and try to align goals with producers on a contract basis and share the total profit with them. The maximization of the profit is by the resource capacity constraints. The equilibrium point (p^0, q^0) is given by values of total number of q production units and optimal price p.

A specific buyback contract is used for coordination. The layer of producers as leader proposes the wholesale price w and the buyback price b. The layer of retailers as follower accepts the prices to coordinate the system. The allocation of the total profit between retailers and producers is given by splitting parameter λ $(0 \le \lambda \le 1)$. The value of the parameter λ is negotiated by retailers and producers.

In the second stage, producers address the following issues:

- How the determine the optimal coalition structure.
- How to allocate the profit among the members of the optimal coalition.

The problems are solved by cooperative manner. These agents compete to be members of a coalition and are willing to cooperate to produce products and sell them to customers through retailers. The optimal coalitions are determined according to the maximal profit with respect the resource capacity constraints for the coalition. The maximal profit is allocated among the members of the coalitions by Shapley values. Shapley value has been shown to possess the best properties in terms of balance and fairness.

4 CO-OPETITION MODEL

The goal of co-opetition is to move the players from a zero-sum game to a plus-sum game, a scenario in which the end result is more profitable when the competitors work together (Okura and Carfì, 2014). We will use biform games as basic models for modeling of co-opetition problems. The co-opetition business model has PARTS of a business strategy - five dimensions a company can use to identify strategies that change the game:

- Players,
- Added value,
- Rules,
- Tactics,
- Scope.

The players are the firm, customers, suppliers, competitors and complementors (competitors whose products add value). An important part of the game is to learn which variables will influence the players to either compete or cooperate. Added values are given through complementors. Rules structure negotiations between buyers and sellers. Tactics are actions taken to shape other players' perceptions. Scope means recognizing the links between games through Players, Added values, Rules, and Tactics.

The relationship between the firm and the direct competitors is non-cooperative. The relationship between the firm and the complementors in a search for common added values is cooperative. The relationship between the firm and the suppliers can be partly cooperative; for some criteria cooperative (e. g. timing of deliveries), for others non-cooperative (e. g. price).

5 ENVIRONMENTAL NEGOTIATION

An authority and polluters are looking for coalition projects to reduce pollution (Šauer et al., 2015). The relationship between the authority and the polluters is non-cooperative. The authority wants to provide as little subsidy as possible to the required reduction of pollution. Polluters are trying to get as much subsidy but are under external pressure that if their demands are too high, do not get any subsidy. The relationship among the polluters can be partly cooperative (to get subsidy) and partly non-cooperative (to maximize surplus for individual polluters). Polluters are under internal pressure to get a sufficient portion of the subsidy for themselves.

Simultaneous biform games can be used for modeling and solving environmental negotiations. Suppose n negotiating subjects (authority and polluters). Denote X as decision space for the negotiating process (set of all possible coalition projects). Elements of this space are decisions (coalition projects) $\mathbf{x} \in X$, which are vectors whose components represent the parameters of the decision (members, subsidy, conditions etc.). A consensus decision (selected coalition projects) $\mathbf{x}^*$ should be chosen from the decision space X. Each participant evaluates decisions by criterion and compares the decisions with the target value. The criteria (subsidies for members of coalition projects) are in the form of criteria functions, that all negotiating subjects want to maximize their values. Denote $f^1(\mathbf{x})$, $f^2(\mathbf{x})$, ..., $f^n(\mathbf{x})$ criteria functions that transform decision $\mathbf{x}$ and the criteria values are compared with aspiration levels. Own negotiations and exchanges of information among negotiating subjects happen in the decision space.

For each negotiating subject we can formulate a set of acceptable decisions, which is a set of decisions that are permissible and acceptable in terms of the required aspiration levels of criteria functions. The aspiration levels (required subsidies for individual polluters) $b^i(t)$, $i = 1, 2, ..., n$, $t = 1, 2, ..., T$, of criteria functions represent opportunities for added values.

Pressures acting on the aspiration levels of criteria functions that change in time points $t = 1, 2, ..., T$, and thus change the set of acceptable decisions

$$X_i(t) = \{\mathbf{x}; \mathbf{x} \in X, f^i(\mathbf{x}) \geq b^i(t)\}, i = 1, 2, ..., n. \tag{9}$$

Changes of aspiration levels (changes of required subsidies for individual polluters) are described by a vector of pressure effects $p^i(t)$ at time t ($p^i(0) = 0$)

$$b^i(t) = b^i(t\text{-}1) + \mathrm{p}^i(t). \tag{10}$$

Then we can define the negotiation space as an intersection of sets of acceptable decisions of all participants in negotiations

$$X_0(t) = \cap_{i=1}^{r} X_i(t) \tag{11}$$

If the negotiation space $X_0(t)$ is a single element set, this element is the consensus (the selected coalition projects for subsidy).

6 CONCLUSIONS

The biform game theory provides tools for different application areas. The use of tools in these areas provides feedback on more detailed classification and specification of the models and thus enriches the theory of biform games. Non-cooperative and cooperative approaches are variously combined according to asymmetric information by participants. This is the basis for our proposal for division into sequential and simultaneous shapes. Multiple criteria may be considered in our models, such as economic, environmental, technological, and others. Multicriteria evaluation approaches were tested for inclusion in the model of negotiation (Fiala, 1999). Deeper use of game theory, the theory of supply chain, co-opetition theory, negotiation models and other disciplines is required for new more precise models of biform games.

Acknowledgements

This paper was written with the support of the Czech Science Foundation project No. P402/12/G097 DYME - *Dynamic Models in Economics* and by the grant No. IGA F4/57/2017, Faculty of Informatics and Statistics, University of Economics, Prague. The paper is a result of institutional research project no. 7429/2018/08 "Analysis of ICT startups" supported by University of Finance and Administration, Prague.

References

[1] Brandenburger, A. M., Nalebuff, B. J. (2011). *Co-opetition*. Crown Business, New York.

[2] Brandenburger, A., Stuart, H. (2007). Biform games. *Management Science*, 53(4), pp. 537-549.

[3] Cachon, G., Netessine, S. (2004). Game Theory in Supply Chain Analysis. In: D. Simchi-Levi, S. D. Wu, M. Shen, eds., *Handbook of Quantitative Supply Chain Analysis: Modeling in the e-Business Era*, Boston: Kluwer, pp. 13-65.

[4] Fiala, P. (1999). Modelling of a negotiation process with multiple evaluation criteria. *Politická ekonomie,* 47(2), pp. 253–268.

[5] Fiala, P. (2015). Biform games in supply chains. In: *Mathematical Methods in Economics 2015.* Pilsen: University of West Bohemia, pp. 177–183.
[6] Fiala, P. (2016). Profit allocation games in supply chains. *Central European Journal of Operations Research*, 24(2), pp. 267-281.
[7] Myerson, R. B. (1997). *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge.
[8] Nagarajan, M., Sošić, G. (2008). Game-Theoretic Analysis of Cooperation among Supply Chain Agents: Review and Extensions. *European Journal of Operational Research*, 187(3), pp. 719-745.
[9] Okura, M., Carfì, D. (2014). Co-opetition and Game Theory. *Journal of Applied Economic Sciences,* 9(1), pp. 1-29.
[10] Shapley, L. S. (1953). A value for n-person games. In: A. W. Tucker, R. D. Luce, eds., *Contributions to the Theory of Games* II, Princeton: Princeton University Press, pp. 307-317.
[11] Šauer, P., Fiala, P., Dvořák, A., Kolínský, O., Prášek, J., Ferbar, P., Rederer, L. (2015). Improving Quality of Surface Waters with Coalition Projects and Environmental Subsidy Negotiation. *Polish Journal of Environmental Studies*, 24(3), pp. 1299–1307.
[12] von Neumann, J., Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press, Princeton.

Authors' addresses
Prof. RNDr. Ing. Petr Fiala, CSc., MBA
University of Economics, Prague, Faculty of Informatics and Statistics, Department of Econometrics
W. Churchill Sq. 4, 130 67 Prague 3
Czech Republic
email: pfiala@vse.cz

PaedDr. Renata Majovská, PhD.
University of Finance and Administration, Prague, Faculty of Economic Studies, Department of Computer Science and Mathematics
Estonska 500, 101 00 Praha 10
Czech Republic
email: renata.majovska@mail.vsfs.cz

COOPERATION IN DISTRIBUTION

Dana Figurová, Zuzana Čičková,
University of Economics in Bratislava

Abstract
The Cooperative Game Theory explores the positive effects of creating a coalition based on binding agreements. This paper is focused on the possibilities of cooperation in the Vehicle Routing Problem, which is one of the most famous routing problems with wide practical applicability. Proposed model is based on the Multi-Depot Vehicle Routing Problems, when it is assumed that the owners of individual depots are also decision makers (players). Their main interest is to cooperate with other players (owners of depots) in order to minimize their shipping costs. We will consider the benefits from merging the transport requirements of individual players. In the solution, we will accept the assumption of the mutual use of vehicles among the players in the coalition.

Keywords: *Multidepot Vehicle Routing Problem, Game Theory, Cooperative Games.*

JEL Classification: JEL C61, JEL C71
AMS Classification: AMS 91A12, AMS 91A80

1 INTRODUCTION

The Game Theory (GT) is a science that basically examines a wide range of decision situations for multiple participants. One way to explore these decisions is to accept the possibility of cooperation between the participants (players). The game is cooperative, if players can create the coalitions respecting the predetermined obligations and thus, they can profit from a common approach. In non-cooperative games the players act independently and we do not consider such commitments.

The important factor to determine the cooperative behaviour is that by its implementation, the participants are gradually increasing the achieved profit or at least there is a probability of its achievement in relation of non-cooperative strategies. Cooperative games can be characterized as "Theory, which primarily deals with a coalition of players who coordinate their activities to achieve the further benefits" (Branzei et al., 2008).

This paper is focused on the possibilities of a cooperation in the Vehicle Routing Problem (VRP), which is one of the most famous routing problems with wide practical applicability. A basic version of the VRP is the capacitated VRP (CVRP) in which each vehicle has a known capacity. Lots of new constraints are added on the route construction and made practical applications of the VRP (Toth and Vigo, 2002). Another well-known generalization of the VRP is Multi-Depot VRP (MDVRP). In the MDVRP, there are several depots instead of just one depot and every customer is visited by a vehicle based at one of the several existing depots, while every vehicle route must start and end at the same depot. For example, Messenger Problem is a relaxation of the VRP based on Multiple Travelling Salesman Problem in which the same customer can be served by different vehicles if it reduces the overall costs (Fábry, 2015).

In this paper, we will discuss the possibility of cooperation in the case of a distribution problem, where we optimize the shipment realized by vehicles that are located in the multiple depots, which belong to the individual players. The problem of distribution were solved in

(Pekár et al., 2017). Their presented mathematical model deals with synchronized distribution during time period in order to minimize the total distance. Another way to solve the real distribution problem was presented in (Čičková et al., 2013) with self-organizing migrating algorithm. In general, shipping costs represent a large part of the company's costs. One solution to reduce these costs is the cooperation between individual logistics companies. It is the theory of Cooperative Games that is used to model this type of cooperation between the companies. Lozano (2013) introduced a mathematical model to quantify the benefits from merging the transport requirements of different companies. He has shown that there is a reduction in common transport costs by using more vehicles in the distribution and increasing number of connected roads. Several authors have demonstrated that collaboration between independent owners of the logistics network can maintain a reliable maximum flow. For example, McCain (2008) focused on analyzing cooperative games among organizations to increase their profits. It is quite clear, that the cooperation between companies reduces the transport costs and therefore increases the profit of players. This issue was discussed in (Zibaei et al., 2016).

The purpose of this paper is to illustrate the cooperation in distribution, which can help the companies save their shipping costs. The contribution is structured as follows: in the first part, we will introduce a mathematical model that can be used to solve the Cooperative Distribution Problem, in the second part we will present a numerical example. In this example, we use the data reported in (Zibaei et al., 2016). However, the presented model contains two-dimensional decision variables in contrast to the three-dimensional variables used in the model in (Zibaei et al., 2016). The proposed model also involves the assumption that a vehicle can ends its distribution in different depot such as starting, which also allows to reduce the cost of the coalition.

In MDVRP, we consider that the owners of individual depots are also players, who want to cooperate with other players to minimize their shipping costs. Minimum shipping costs are obtained by solving the VRP for individual coalitions. We will also be interested in the cost savings and synergy of individual coalitions in this case. In a nutshell, we will explore how MDVRP gives the opportunity to minimize total transport costs by using cooperative games techniques.

2 COOPERATIVE DISTRIBUTION PROBLEM

The model formulation of a Cooperative Distribution Problem will be based on the formulation of MDVRP. In the MDVRP each vehicle starts its road from one of several sets of depots and its route must end at that same depot. The requirement of each customer can be satisfied by more than one means of transport. Mostly frequent case in this problem is when customer's demand exceeds vehicle capacity (collective transport). Another modification is the case, where it is more advantageous (cheaper) to visit a customer more than once. In theory, the homogeneity of the transport park is often assumed, but it is common to use different vehicles with different capacities, different age and other parameters in practice, not only to minimize overall distance, but also to minimize total costs.

In this paper, we are based on the model of distribution with multi depots mentioned in (Pekár et al., 2012), which we will subsequently modify to our requirements. In this model, we will assume that there are $k+1$ centers from which the vehicle can start its road (but the vehicle may not be used, it means that it even does not have to start its road) and the number of vehicles in the depots is not limited. We require to satisfy m customers from these depots. We will formulate the mathematical model in the full, edged and weighted graph $\bar{G} = (N, \bar{H})$. Let:

$N^{(1)} = \{0,1, \dots k\}$ is the set of vertices representing the central depots (centers) and $N^{(2)} = \{k+1, k+2, \dots m\}$ is a set of vertices representing the customers and $N = N^{(1)} \cup N^{(2)}$ is the set then the set of all the vertices of the graph. Let $\bar{H} \subset NxN$ represents the set of edges h_{ij}, $i, j \in N$ between all vertices i and j. Each edge h_{ij} being assigned to a real number called d_{ij}, also known as a price of the edge h_{ij}. This assignment is the shortest distance between nodes i and j. Each of the customers located at i, $i \in N^{(2)}$, requires the import of a certain quantity of goods, generally denominated as q_i, $i \in N^{(2)}$.

Customer's demand is accomplished by vehicles with the same capacity g units. The goal is to determine the vehicle's roads, which will satisfy the requirements of all customers. Customer requirements will only be realized in the whole (if the vehicle serves the customer, its entire delivery requirement will be realized), with no vehicle capacity exceeded. The main goal is to minimize the total traveled distance. In the model we assume implicitly, that $q_i \le g$ for all $i \in N^{(2)}$ which means, that the size of each customer's requirement will not exceed the capacity of the vehicle. On the other hand we will assume that the vehicle can be repeatedly reloaded or unloaded at another depot (not the starting one) and after that repeatedly realized the delivery or more precisely the haulage.

We will use the binary variables x_{ij} $(i, j \in N, i \neq j)$ in the model with the following meanings:

$$x_{ij} = \begin{cases} 1, & \text{if node } i \text{ immediately precedes node } j \text{ in the vehicle tour;} \\ 0, & \text{otherwise} \end{cases}$$

and also the non-negative variables u_i, $i \in N^{(2)}$ representing the cumulative load of the vehicle over a given node.

Now we accept the assumption that the owners of individual depots are different subjects (players). Players are able to cooperate with each other and create the coalitions and reduce their shipping costs. We are considering that each player owns only one depot and also has his own customers, but the player's vehicles in coalition can also be used to serve customers assigned to another player in a possible coalition. Number of possible coalitions is $2^{k+1} - 1$ (where $k+1$ represents the number of players). Let $N_S^{(1)} \subset N^{(1)}$ be a set of players' coalitions. We will also divide a set of customers based on their membership to individual players, then the coalition S will customers labeled as $N_S^{(2)}$. Thus, the set $N_S = N_S^{(1)} \cup N_S^{(2)}$ is the set representing the depots and the customers of the coalition S. Then, the optimal distributions of individual coalitions can be quantified by using this model:

$$c^S = \min f(\mathbf{X}, \mathbf{u}) = \sum_{i \in N_S} \sum_{j \in N_S} d_{ij} x_{ij} \tag{1}$$

$$\sum_{i \in N_S} x_{ij} = 1, \; j \in N_s^{(2)}, \; i \neq j \tag{2}$$

$$\sum_{j \in N_S} x_{ij} = 1, \; i \in N_s^{(2)}, \; i \neq j \tag{3}$$

$$\sum_{j \in N_S^{(2)}} x_{ij} \le 1, \; i \in N_s^{(1)} \tag{4}$$

$$u_i + q_j - g\left(1 - x_{ij}\right) \le u_j, \; i \in N_S, \; j \in N_s^{(2)}, \; i \neq j \tag{5}$$

$$\sum_{j \in N_S} x_{ij} = \sum_{j \in N_S} x_{ji} \ , \ i \in N_S \tag{6}$$

$$q_i \le u_i \le g \ , \ i \in N_s^{(2)} \tag{7}$$

$$u_i = 0 \ , \ i \in N_s^{(1)} \tag{8}$$

$$q_i = 0 \ , \ i \in N_s^{(1)} \tag{9}$$

$$x_{ij} \in \{0, \ 1\} \ , \ i, j \in N_s, \ i \ne j \tag{10}$$

Constraint set (2) and (3) guarantee that each customer will be visited exactly once. Constraint set (4) means that each route can be served by one vehicle maximally. Constraint set (5) is the sub-tour elimination condition (excluding the sub-tour eliminations, which contain one from the depots). Constraint sets (7) and (8) ensure the condition of not exceeding the capacity of vehicle and the balance of vehicle load. Conditions (9) provide a zero vehicle load in the depot. The flow conservation is expressed by constraint set (6). The scalar c^S (1) represents the value of the total transportation cost $TC(S_m)$ for the vehicles of the S. The quantification based on this model will provide the cheapest distribution of the coalition S.

3 RESULTS

Firstly, we are testing and analysing our modified model of cooperative MDVRP based on numerical example (Zibaei et al., 2016). The authors have tried to prove that transport costs can be extensively saved through the cooperation between the suppliers. The data were obtained from the website (see [8]).

So we consider the Distribution Problem with Multiple Depots, whereby we have four suppliers to serve the certain customers. Suppliers (owners of individual depots) are expressed as
$P = \{1,2,3,4\}$. Each player owns one depot with one vehicle. The maximum storage capacity in our case is $V_1 = V_2 = V_3 = V_4 = 200$, with the capacity of each vehicle given by $Q_k = 220$. Customers, who are strictly assigned to the individual depots (players), will be marked as: $D_1 = \{c_1, c_2, c_3, c_4, c_5\}, D_2 = \{c_6, c_7, c_8, c_9, c_{10}\},\ D_3 = \{c_{11}, c_{12}, c_{13}, c_{14}, c_{15}\},\ D_4 = \{c_{16},\ c_{17}, c_{18}, c_{19}, c_{20}\}$. In the case of the creation the coalition $S_m \subseteq P$ we know that there are exactly 15 possible coalitions between the players.

We solve the cooperative MDVRP by using the model (1) - (10) for each player individually and then for the created coalitions S: {1}, {2}, {3}, {4}, {1,2}}, {1, 4}, {2,3}, {2,4}, {3,4}, {1,2,3}, {1,2,4}, {1,3,4} {2,3,4 } and {1,2,3,4} by GAMS software. To obtain the optimal solution, we used the solver Cplex 12.2.0.0 on the personal computer INTEL® Core ™ 2 CPU, E8500 @ 3.16 GB RAM for Windows 10.

Our interest is to compare the optimal solutions by using the modified model of cooperative MDVRP (1) - (10) with the results introduced by authors in Zibaei et al., (2016). Our main idea is to prove that there is a reduction in total shipping costs through mutual cooperation between suppliers.

First, we summarize the results obtained by solving the model by Zibaei et al., (2016). Compared to the modified model (1) - (10), this differs in the fact that it is assumed that the vehicle returns back to its starting point (depot) after customer service. This assumption is a basic prerequisite in VRP, but in this work we modified it. We assume that the vehicle of one player can also be used to serve customers of another player in a possible coalition, which ultimately leads to saving the transport costs. It means that vehicles in a coalition do not have to return to their original depots, but they can end their road even in the depot of another supplier (player) with whom they cooperate. In the next part of this work we will see how this modification of assumption will be reflected in the achieved results.

Table 1 presents the total transport costs $TC(S_m)$ of individual coalitions. It contains results solved by Zibaei et al., (2016) and compared to our results from using our model (1) - (10). We also calculate the sum of individual players $\sum_{p \in S_m} TC(\{p\})$ for both models. We can also see the size of the cost savings $CS(S_m)$ of coalitions S_m in the case of a cooperative approach to the distribution problem. The cost savings $CS(\{S_m\})$ of coalitions S_m are expressed by the equation:

$$CS(S_m) = \sum_{p \in S_m} TC(\{p\}) - TC(S_m) \tag{11}$$

Table 1: Total costs, cost savings and synergy of individual players in coalitions.

Coalitions	**Results obtained by Zibaei S. et al**				**Results solved by model (1)-(10)**				
	$TC(S_m)$ $=c^S$	$\sum_{p \in S_m} TC(\{p\})$	CS (S_m)	Synergy	$TC(S_m)$ $=c^S$	$\sum_{p \in S_m} TC(\{p\})$	CS (S_m)	Synergy	Processing time
$S_1=\{1\}$	20,68	20,68	0	0	20,68	20,68	0	0	0,13sc
$S_2=\{2\}$	21,48	21,48	0	0	20,83	20,83	0	0	0,16sc
$S_3=\{3\}$	21,48	21,48	0	0	20,83	20,83	0	0	0,13sc
$S_4=\{4\}$	20,68	20,68	0	0	20,68	20,68	0	0	0,15sc
$S_5=\{1,2\}$	30,48	42,16	11,68	0,277	26,63	41,51	14,88	0,358	1,13sc
$S_6=\{1,3\}$	27,59	42,16	14,57	0,346	26,44	41,41	15,07	0,364	0,83sc
$S_7=\{1,4\}$	29,91	41,36	11,45	0,277	28,21	41,36	13,15	0,318	1,99sc
$S_8=\{2,3\}$	35,01	42,96	7,95	0,185	27,5	41,66	14,16	0,34	1,20sc
$S_9=\{2,4\}$	29,93	42,16	12,23	0,290	26,44	41,51	15,07	0,363	0,98sc
$S_{10}=\{3,4\}$	32,15	42,16	10,01	0,233	28,7	41,51	12,81	0,309	1,20sc
$S_{11}=\{1,2,3\}$	36,11	63,64	27,53	0,433	31,66	62,34	30,68	0,492	38,06sc
$S_{12}=\{1,2,4\}$	41,98	62,84	20,86	0,332	31,66	62,19	30,53	0,491	57,28sc
$S_{13}=\{1,3,4\}$	47,95	62,84	14,89	0,237	31,66	62,19	30,53	0,491	44,59sc
$S_{14}=\{2,3,4\}$	43,61	63,64	20,03	0,319	31,66	62,34	30,68	0,492	36,19sc
$S_{15}=\{1,2,3,4\}$	46,14	84,32	38,18	0,453	34,60	83,02	48,42	0,583	38,56sc

Source: Author

When the optimal objective function $TC(S_m)$ or c^S for any coalitional scenario is smaller than sum of the individual optimal objective function (the sum of individual players), the players have the incentive to coordinate with each other. It means that players tend to cooperate with each other when the following equation is executed in form:

$$TC(S_m) \leq \sum_{p \in S_m} TC(\{p\}) \tag{12}$$

Based on Table 1, we can confirm that in all types of coalitions is sum of the total costs of each player higher than the total cost of the coalitions. Therefore, the players tend to cooperate with each other. For this reason, we also quantify the cost savings, which players can save in case of cooperation.

Cost saving $CS(S_m)$ represents the difference between sum of the individual objective function and that of the objective function of coalition S_m. These $CS(S_m)$ can be either higher or lower, depending on the synergy between owners in different coalition. This synergy can be determined with the below function:

$$Synergy(S_m) = \frac{CS(S_m)}{\sum_{p \in S_m} TC(\{p\})} \quad (13)$$

We can see in both cases, that the highest value of cooperation (synergies) S_m is in the last coalition $S_{15}=\{1,2,3,4\}$, when all four players cooperate with each other. Therefore, we establish that this cooperation is the most efficient or the most advantageous for all players.

Considering the synergy, in the case of coalition $S_{15}=\{1,2,3,4\}$ can players save 45,3% shipment costs (using the model of the authors Zibaei et al., (2016)). On the other hand, our mathematical model of cooperative distribution reaches up to 58.3% cost savings.

We can summarize various results from Table 1. Without accepting our assumption, if player 1 cooperates with the coalition {2,3}, his costs saving is 27,53 units. On the other hand, if he decides to join the coalition {2,4}, he will save 20,86 units or to the coalition {3,4} and he will save 14.89 units. With the accepting of our assumption and using the model (1) - (10), e.g. if player 1 cooperates with the coalition {2,3}, the player will save more than in the previous case, specifically save 30,68 units.

After the implementation of the assumption into the model we can allege, that the results of the optimal paths of each player's vehicles have greatly improved. For example, if we compare the coalition $S_{15} = \{1,2,3,4\}$, we can see in both of our solved examples that in case of a cooperative distribution problem with established premise have players much lower shipping costs than without this assume (namely, 11.54 units). We can see the highest value of cost savings $CS(S_m)$ for the coalitions S_{15}. In comparison with the previous model, this will save about 10 units more on the total cost of $TC(S_m)$. In this case, we have also quantified the time for calculating individual tasks in the GAMS program, which is in the last column. We can confirm that the longest run of this program was calculated in the case of three-player coalitions.

CONCLUSION

In this paper, we focused on the Cooperative Distribution Problem assuming the cooperation between the players to minimize the total shipment costs. Our main task was to compare the results obtained by solving our model (1)-(10) and the results solved by Zibaei S. et al., (2016*)*, who proceeded on the basis from another model od cooperative MDVRP. Our main idea was to prove that there is a reduction in total transport costs through mutual cooperation between individual suppliers. By comparing our results, we have taken the decision that our modified model has produced much better results than the model of Zibaei S. et al., (2016). This is mainly because the proposed model of cooperative distribution with multiple depots has assumed the premise of using vehicles among players in individual coalitions, which was largely reflected in three-player coalitions in our case. In conclusion, we can say that after the introduction of the assumption into our model, the results of the optimal paths of each player's vehicles have greatly improved.

Acknowledgements: This work was supported by the project VEGA 1/0351/17 Application of selected models of Game Theory in economic problems of Slovakia.

References

[1] Branzei, R., Dimtrov, D., and Tijs, S. (2008). *Models in cooperative game theory*. New York: Springer Science and Business Media.
[2] Čičková, Z., Brezina, I., and Pekár J. (2013). Solving the real-life vehicle routing problem with time windows using self organizing migrating algorithm. *Ekonomický časopis,* 61(5), pp. 497-513.
[3] Fábry, J.(2015). Insertion method for multiple messenger problem with multiple depots. *Communications*, 17(2), pp. 11-14.
[4] Lozano, S., Moreno, P., Adeso-Diaz, B., and Algaba, E. (2013). Cooperative game theory approach to allocating benefits of horizontal cooperation. *European Journal of Operational Research*, 229(1), pp. 444-452.
[5] McCain R.A. (2008). Cooperative games and cooperative organizations. *The Journal of Socio-Economics*, 37(6), pp. 2155–2167.
[6] Pekár, J., Brezina, I., Čičková, Z., and Reiff, M. (2012). *Modelovanie rozmiestnenia recyklačných centier*. Bratislava: Ekonóm.
[7] Pekár, J., Brezina, I., and Čičková, Z. (2017). Synchronization of capacitated vehicle routing problem among periods. *Ekonomický časopis,* 65(1), pp. 66-78.
[8] Zibaei, S., Hafezalkotob. A., and Ghashami, S. Cooperative vehicle routing problem: an opportunity for cost saving. *Journal of Industrial Engineering International,* 12(3), pp. 271-286.

Author's Address

Ing. Dana Figurová
University of Economics
Faculty of Economic Informatics, Department of Operations Research and Econometrics
Dolnozemská cesta 1
Bratislava, 852 35
Slovak Republic
E-mail: danafigurova@gmail.com

doc. Ing. Zuzana Čičková, PhD.
University of Economics
Faculty of Economic Informatics, Department of Operations Research and Econometrics
Dolnozemská cesta 1
Bratislava, 852 35
Slovak Republic
E-mail: zuzana.cickova@euba.sk

INEFFICIENCY OF RAILWAYS IN CENTRAL AND EASTERN EUROPE – A DEA APPROACH

Hana Fitzová, Masaryk University

Abstract

Measuring the efficiency of railways is an important issue because railway transport is considered to be an important means of transport by the European Union over the long term. Many studies have dealt with the efficiency of railways in Europe. A few of them have analysed the differences between Eastern and Western Europe. However, these studies have not addressed the identification of problems in terms of a suboptimal ratio of inputs (or outputs) using data envelopment analysis (DEA). Most post-communist Central and Eastern European (CEE) countries are characterized by very low relative efficiency for a number of historical, institutional, and other reasons. The aim of this study is to explore potential factors standing behind the low efficiency of CEE countries using a DEA approach – especially to identify redundant inputs entering the railway transport production process. The study employed DEA based on four inputs (track length, number of employees, and number of cars in passenger and freight transport) and two outputs (passenger-km and tonne-km) and included 24 European countries during 1999–2016. Excessive freight fleet size and overstaffing have been identified as redundant inputs in all CEE countries with the exception of the Baltic republics.

Keywords: *Central and Eastern Europe, DEA, efficiency, railway transport*

JEL Classification: C44, C61, R40

AMS Classification: 90C05, 90B06

1 INTRODUCTION

Railway transport is considered to be an important means of transport for the long term in Europe. The main objective of the European railway reforms which started in the 1990s was to enable competition to enter the existing network in order to increase efficiency in the railway sector. There have been four reform packages in the EU so far, the last of which, concerning in particular opening the domestic passenger services market, should be implemented by December 2019, so the efficiency of European railways is still a current issue.

Many studies have already dealt with the impact of these reforms on the efficiency of railways in Europe. Most of them have directly concerned efficiency, for example Cantos (2001), Driessen (2006), Wetzel (2008), Growitsch (2009), Asmild (2009), Growitsch and Wetzel (2009), Cantos (2010), Friebel et al. (2010), and Bougna and Crozet (2016). Some studies have evaluated the impact on railway costs, for example Van de Velde et al. (2012) and Mizutani et al. (2015). Only Wetzel (2008) analysed the differences in efficiency between Eastern and Western European countries. However, these studies have not devoted attention to the identification of a suboptimal ratio of inputs (or outputs).

The aims of this study are to analyse whether the ratio of inputs used in the production process of Eastern European railways is optimal and identify the main redundant inputs using data envelopment analysis (DEA). The study utilized DEA with four inputs (track length, number of employees, and number of cars in passenger and freight transport) and two outputs (passenger-km and tonne-km). The data includes 24 European countries during 1999–2016.

The paper is structured as follows: Section 2 describes the data and the model for DEA. Section 3 presents the calculated efficiency scores. Section 4 identifies problematic inputs. Section 5 summarizes the presented issues.

2 DATA AND MODEL

This section describes the data and model used in this study.

2.1 Data

This study includes 24 European countries during 1999–2016. The data come from Eurostat (2017), UIC (2008, 2009, 2010, 2012, 2013, 2015), Trafikanalys (2016), and the European Commission (2017). Table 1 presents basic descriptive statistics for the model variables (mean, median, standard deviation, minimum, and maximum) from all years. Inputs comprise track length (in kilometres), number of employees, number of cars in passenger transport (passenger fleet), and number of wagons in freight transport (freight fleet). Outputs comprise passenger-kilometres (pkm) and tonne-kilometres (tkm), both in millions of km. The columns with the minimums and maximums mostly represent values for Luxembourg (minimum) and Germany (maximum).

Table 1: Descriptive input and output statistics

Variable	Mean	Median	Standard deviation	Minimum	Maximum
Inputs					
Track length (km)	8,076	3,633	9,243	275	34,496
Number of employees	46,675	30,934	58,170	2,989	243,455
Passenger fleet	4,076	1,810	5,556	194	21,814
Freight fleet	22,247	11,805	27,283	886	105,637
Outputs					
Passenger-km (mil.)	13,578	3,714	22,491	235	79,499
Tonne-km (mil.)	14,548	9,711	18,605	234	84,650

2.2 Model

Relative efficiency scores were calculated using DEA, which is a standard approach for such analyses. According to Agarwal, Yadav, and Singh (2010), returns to scale (RTS) do not have a clear tendency in the area of public transport. The present study employs a model (known as the CCR model) suggested by Charnes, Cooper, and Rhodes (1978) which assumes RTS to be constant. In the CCR model with constant returns to scale (CRS), a relative score in the sample of n decision making units (DMUs) with input vectors x and output vectors y was calculated from the linear programming problem

$$(LP_0) \quad \max_{u,v} u^T y_0 \tag{1}$$

$$\text{s. t.} \quad v^T v x_0 = 1 \tag{2}$$

$$u^T Y \leq v^T v X \tag{3}$$

$$u \geq 0, v \geq 0 \tag{4}$$

where v is a vector of input weights, u is a vector of output weights, and X and Y are matrices of input and output vectors for all DMUs, respectively. The lower index "0" denotes the investigated DMU. This problem was solved for all DMUs. The dual problem (DLP_0) was expressed with a real variable θ and a non-negative vector $\lambda = (\lambda_1, \dots, \lambda_n)^T$ as follows

$$(DLP_0) \quad \min_\lambda \theta \tag{5}$$

$$\text{s. t.} \quad \theta x_0 - X\lambda \geq 0 \tag{6}$$
$$Y\lambda \geq y_0 \tag{7}$$
$$\lambda \geq 0 \tag{8}$$

In accordance with Cooper, Seiford and Tone (2007), we defined the input excesses $s^- \in R^m$ and the output shortfalls $s^+ \in R^s$ and identified them as "slack" vectors by

$$s^- = \theta x_0 - X\lambda, \quad s^+ = Y\lambda - y_0, \tag{9}$$

where $s^- \geq 0$, $s^+ \geq 0$ for any feasible solution (θ, λ) of (DLP_0).

To determine potential input excesses and output shortfalls, we solved the following two-phase linear programming problem. In the first phase, we solved (DLP_0) with the optimal value θ. In the second phase, θ was included as follows:

$$\max_{\lambda, s^-, s^+} \omega = es^- + es^+ \tag{10}$$

$$\text{s. t.} \quad s^- = \theta^* x_0 - X\lambda \tag{11}$$
$$s^+ = Y\lambda - y_0 \tag{12}$$
$$\lambda \geq 0, s^- \geq 0, s^+ \geq 0 \tag{13}$$

where e is a vector of ones, so that ω is the sum of all slacks.

Only a DMU with $\theta^* = 1$ and all slacks equal to 0 is called CCR efficient. Other units are inefficient. θ is referred to as radial (technical) efficiency. This means that all inputs can be reduced simultaneously without altering the proportion in which they are utilized. The inefficiencies associated with any identified non-zero slack are referred as mix inefficiencies and mean using a different proportion of inputs (Cooper, Seiford and Tone, 2007).

3 EFFICIENCY ANALYSIS

This part summarizes the results from the input-oriented CCR model using CRS. Countries were arranged into four groups especially for historical and institutional reasons so as to display the efficiency pattern more clearly.

Western European countries comprise Austria (AT), Belgium (BE), Germany (DE), Spain (ES), Finland (FI), France (FR), Switzerland (CH), Italy (IT), Portugal (PT), and Sweden (SE). Central and Eastern European (CEE) countries comprise Bulgaria (BG), the Czech Republic (CZ), Croatia (HR), Hungary (HU), Poland (PL), Romania (RO), Slovenia (SI), and Slovakia (SK). Baltic countries (Estonia (EE), Latvia (LV), Lithuania (LT)) form a specific group mainly connected with the transport of freight from Russia to seaports. The final three countries (Luxembourg (LU), Greece (EL), and Ireland (IE)) represent very small railway systems (especially for freight transport). Denmark, the Netherlands, Norway, and the United Kingdom were not included as their data were not available.

Figure 1 depicts the calculated efficiency scores for the CCR model with CRS. Larger circles represent more recent values. The efficiency of Western European countries was clearly higher than that of Eastern European countries. The three small railway systems had efficiency scores somewhere in the middle. Estonia, Latvia, and in the final few years also Lithuania were very efficient as their railway systems served mostly as transport corridors for raw materials from Russia; their trains were usually very long and heavy, which enabled high efficiency.

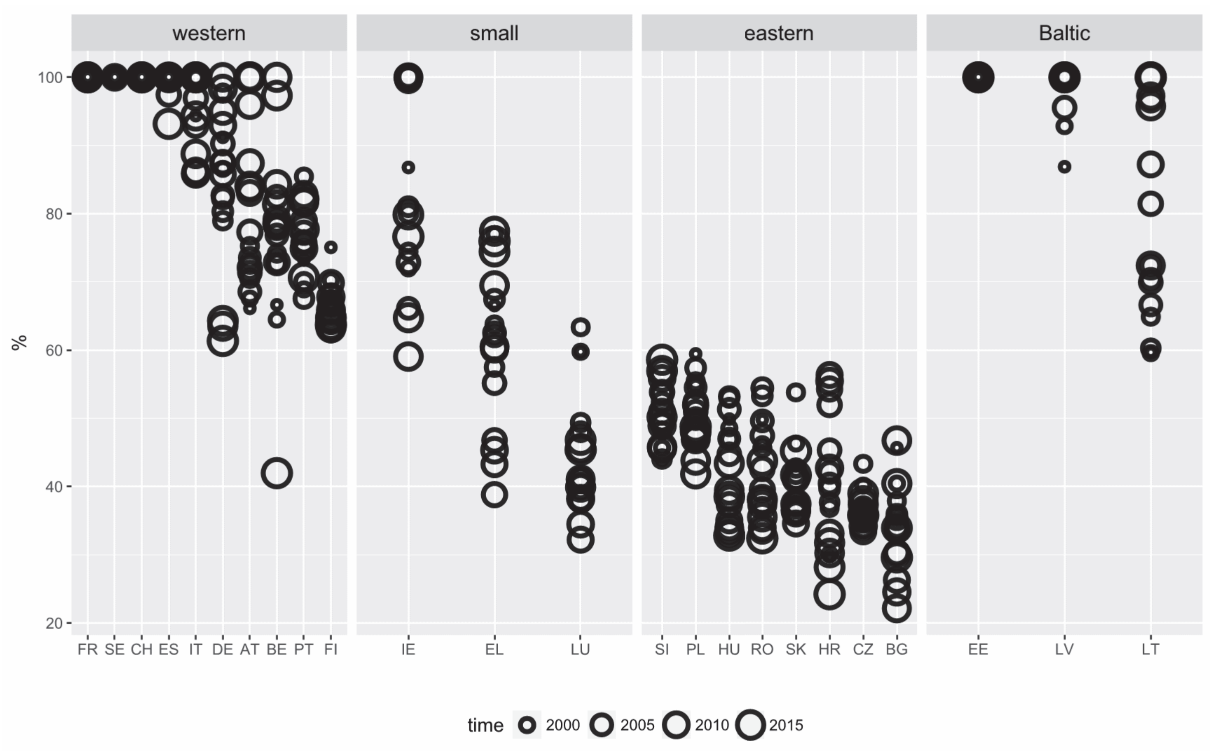


Figure 1: Efficiency scores during 1999–2016

Table 2 presents the means and medians of some variables for the four groups – Western European countries, Eastern European countries, Baltic countries, and countries with small railway systems. Network density is track length divided by area. Passenger transport share is passenger-train-km divided by total train-km. Western European countries (excluding the three smallest systems) are larger countries with higher values for operated passenger-km and train-km as well as somewhat higher passenger transport shares than Eastern European countries. Their average efficiency was about 90%. Eastern European countries are much smaller, had a median line length similar to that of Western countries, had greater network density for historical reasons, and displayed a visible slightly greater orientation to freight transport. All of these factors contributed to their lower mean efficiency of only 44%. The Baltic countries are quite small in area and had a low mean passenger transport share (35%), high values for transported tonne-km, and a high mean efficiency of around 93%. The three smallest railway systems (LU, EL, and IE), with in particular low tonne-km values and a high mean passenger transport share (89%), had efficiency scores of about 67%.

Table 2: Means and medians for characteristics of country groups

Group	**Area (1,000 km^2)**	**Track length (km)**	**Network density**	**Pkm (million)**	**Tkm (million)**	**Train-km (1000)**	**Passenger share (%)**	**Mean efficiency (%)**
Mean								

Western	**255**	12,955	61	**29,273**	22,713	**261,390**	77	**90**
Eastern	120	7,479	**67**	5,459	11,782	79,738	73	**44**
Baltic	58	1,594	**27**	**482**	11,595	11,873	**35**	**93**
Small	68	1,517	46	1,216	**386**	13,811	**89**	**67**
Median								
Western	**301**	**5,883**	55	15,808	13,559	**168,639**	80	91
Eastern	**86**	**6,034**	62	4,463	6,867	65,334	73	44
Baltic	65	1,794	28	**442**	12,166	13,836	**38**	99
Small	70	1,867	27	1,603	**392**	16,713	88	70

Table 3 presents some geographical and operational average characteristics for particular Eastern European countries to identify potential common sources of problems. Countries are ordered according to train-km. The data sample is quite variable. Poland and Romania have large areas while Slovenia, Slovakia, and Croatia are rather small. Network density varied from 37 in Bulgaria to 120 in the Czech Republic. Large differences are visible in passenger-km, tonne-km, and total train-km. Passenger transport share ranged between 65% in Slovenia and 82% in Hungary. The last column of Table 3 displays the mean efficiency score from the CRS model. Bulgaria and the Czech Republic had the lowest efficiency. Bulgaria had very few transported passengers and little freight (relative to its size) while the Czech Republic had the highest network density in Europe. It is not easy to find a unique specific problem connecting all of these countries. It could be helpful to carry out an analysis of redundant inputs using slacks (see equation (9)).

Table 3: Means of Eastern European country characteristics

Country code	**Area (1,000 km^2)**	**Lines (km)**	**Network density**	**Pkm (million)**	**Tkm (million)**	**Train-km (1000)**	**Passenger share (%)**	**Mean efficiency (%)**
PL	313	19,804	63	16,385	40,632	191,294	68	51
CZ	79	9,466	120	6,842	14,549	145,167	78	37
HU	93	7,907	85	6,414	5,045	94,361	82	42
RO	238	10,935	46	7,162	15,226	84,359	75	42
SK	49	3,632	74	2,512	8,689	46,308	70	40
BG	111	4,161	37	2,351	4,099	33,622	71	34
HR	57	2,706	47	1,251	2,586	23,757	74	40
SI	20	1,219	61	753	3,426	19,037	65	51

4 REDUNDANT INPUTS

This section contains an analysis of the suboptimal ratio of inputs for particular Eastern European countries based on the model described in equations (9)–(13). Identification of redundant inputs may be most useful as, especially for passenger transport, the outputs of railway companies are often given in the form of public service obligations. The analysis of outputs would be similar, using the "output orientation" of the model instead of its "input orientation".

Table 4 presents information about the redundancy of inputs for particular countries besides the necessity for a radial lowering of all inputs represented by the efficiency scores. For clarity, this information is presented as condensed phrases representing almost 140 values.
Neither track length nor passenger fleet represented a problem as the slacks were very small or zero. However, the slacks for number of employees (staff) and freight fleet were quite large for most countries. The trend for those problematic variables was good, although not fast enough. The positive development of the slacks for employees was probably influenced by

the economic crisis during 2008–2009. The size of the slacks for freight fleet began to decline significantly later, after 2010. Freight fleets in post-communist CEE countries seemed to be very high in some cases, but this often represented a large number of wagons standing unused somewhere in fields as a remanent from the communist era. This fact may have contributed to the low efficiency scores of CEE countries.

Table 4: Slacks for inputs

Country	Track length	Staff	Passenger fleet	Freight fleet
Bulgaria	small	large	nearly 0	large till 2013
Czech Republic	0	large till 2009	0	large
Croatia	small	large till 2010	0	large till 2010
Hungary	small	large	0	large till 2013
Poland	small	large till 2009	0	large
Romania	small	large	0	large till 2015
Slovenia	small	large	0	large till 2012
Slovakia	0	large	0	large till 2014

As this table implies, the problem of low efficiency in Eastern European countries lay not only in rather poor conversions of inputs to outputs but also in a suboptimal ratio of input mix, especially due to the large values for staff and freight fleet. Reducing the redundant inputs could help these countries to approach the efficiency of Western European countries.

5 CONCLUSIONS

This study tried to explore potential factors influencing the efficiency of Eastern European countries. An input-oriented DEA CCR model using CRS based on four inputs and two outputs was employed for 24 European countries during 1999–2016. The aims were to determine whether the ratio of inputs used in the production process of Eastern European railways was optimal and identify the main redundant inputs entering the production process. In addition to a number of common historical and geographical features which make it more difficult to achieve increased efficiency, all of the Eastern European countries were also characterized by excessive staff and freight fleet. The Baltic countries (Estonia, Latvia, and Lithuania) represented an exception. They had very high efficiency due to the different focus of their railway transport. The problem of low efficiency in Eastern European countries lay in a poor conversion of inputs to outputs represented by the relative efficiency scores from the DEA as well as in improper ratio of input mix identified by non-zero slacks in the linear programming problem. All of the redundant inputs were decreasing over time. This reduction could help these countries to become more efficient. However, it is especially necessary to improve the entire production process to increase the efficiency of Eastern European countries.

Acknowledgements: This work was supported by MUNI/A/1133/2017.

References

[1] Agarwal, S., Yadav, S. P. and Singh, S. P. (2010). DEA based estimation of the technical efficiency of state transport undertakings in India. *Opsearch*, *47*(3), 216.

[2] Asmild, M., Holvad, T., Hougaard, J. L. and Kronborg, D. (2009). Railway reforms: Do they influence operating efficiency?. *Transportation*, *36*(5), 617-638.

[3] Bougna, E. and Crozet, Y. (2016). Towards a liberalised European rail transport: Analysing and modelling the impact of competition on productive efficiency. *Research in Transportation Economics*, *59*, 358-367.
[4] Cantos, P. C. (2001). Vertical relationships for the European railway industry. *Transport Policy*, *8*(2), 77-83.
[5] Cantos, P., Pastor, J. M. and Serrano, L. (2010). Vertical and horizontal separation in the European railway sector and its effects on productivity. *Journal of Transport Economics and Policy (JTEP)*, *44*(2), 139-160.
[6] Charnes, A., Cooper, W. W. and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429-444.
[7] Cooper, W. W., Seiford. L. M. and Tone, K. (2007). Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software, 2nd edn. New York: Springer.
[8] Driessen, G., Lijesen, M. G. and Mulder, M. (2006). The impact of competition on productive efficiency in European railways (CPB Discussion Paper No. 71). The Hague: CPB Netherlands Bureau for Economic Policy Analysis.
[9] European Commission, (2017). *EU transport in Figures: Statistical Pocketbook 2017.* Brussels: European Commission. Available at: https://ec.europa.eu/transport/sites/transport/files/pocketbook2017.pdf [Accessed 10 Jan. 2018].
[10] Eurostat, (2017). Database – Eurostat. [online] Available at: http://ec.europa.eu/eurostat/data/database [Accessed 8 November 2017].
[11] Friebel, G., Ivaldi, M. and Vibes, C. (2010). Railway (de) regulation: A European efficiency comparison. *Economica*, *77*(305), 77-91.
[12] Growitsch, C. and Wetzel, H. (2009). Testing for economies of scope in European railways: An efficiency analysis. *Journal of Transport Economics and Policy (JTEP)*, *43*(1), 1-24.
[13] Mizutani, F., Smith, A., Nash, C. and Uranishi, S. (2015). Comparing the costs of vertical separation, integration, and intermediate organisational structures in European and East Asian railways. *Journal of Transport Economics and Policy (JTEP)*, *49*(3), 496-515.
[14] Trafikanalys, (2016). Bantrafik 2016. Stockholm: Trafikverket. Available at: https://www.trafa.se/globalassets/statistik/bantrafik/bantrafik/2016/bantrafik-2016_okt.pdf [Accessed 20 Feb. 2018].
[15] Van de Velde, D., Nash, C., Smith, A., Mizutani, F., Uranishi, S., Lijesen, M. and Zschoche, F. (2012). *EVES-Rail: Economic effects of vertical separation in the railway sector. Report for CER - Community of European Railway and Infrastructure Companies.,* Brussels: CER.
[16] Wetzel, H. (2008). *European railway deregulation: the influence of regulatory and environmental conditions on efficiency* (Working Paper Series in Economics No. 86). Lüneburg: University of Lüneburg.
[17] UIC. (2009). *Railway time-series data 1970-2008*. Paris: International Union of Railways.
[18] UIC. (2010). *International Railway Statistics 2009.* Paris: International Union of Railways.
[19] UIC. (2012). *International Railway Statistics 2010*. Paris: International Union of Railways.
[20] UIC. (2013). *International Railway Statistics 2011*. Paris: International Union of Railways.

[21] UIC. (2015). *International Railway Statistics 2012*. Paris: International Union of Railways.

Author's address

Mgr. Hana Fitzová
Masaryk University, Faculty of Economics and Administration, Department of Economics
Lipová 41a, 602 00, Brno
Czech Republic
email: hana.fitzova@econ.muni.cz

SPATIAL HETEROGENEITY OF REGIONAL INNOVATION PROCESSES: GEOGRAPHICALLY WEIGHTED REGRESSION APPROACH

Andrea Furková,
University of Economics in Bratislava

Abstract
The paper deals with spatial heterogeneity of innovative activity across 245 NUTS 2 European Union regions in 2011 and 2012 period. The geographically weighted regression (GWR) estimation of the regional knowledge production function model and resulting statistical tests confirmed that GWR local regressions describe the data set significantly better than the global model. Estimated parameters of both innovation inputs, namely research and development (R&D) expenditure and human resources in science and technology vary significantly across the EU area. Spatial heterogeneity of the innovation activities invoke local R&D policy implication not region wide policy implication.

Keywords: *Research and Development, Spatial Heterogeneity, Geographically Weighted regression*

JEL Classification: O21, R12
AMS Classification: 91B72

1 INTRODUCTION

In a regression context, spatial effects concern two categories of specifications. The first deals with spatial dependence (spatial autocorrelation) and the second with the spatial heterogeneity. Spatial dependence can be perceived as a special case of cross-sectional dependence and spatial heterogeneity as a special case of cross-sectional heterogeneity. The difference is that the structure of the dependence is related to the geographical location and distance as well as in a more general economic or social network space. Spatial effects are subject of interest of spatial econometrics which is a subfield of econometrics that deals with the incorporation of spatial effects in econometric methods.

Spatial econometric tools have been used in various fields of applied economics and also in the analysis of regional innovation processes. One motivation for considering spatial aspects in the modelling of the innovation processes comes from very strong spatial polarization of economic activities in space (Autant-Bernard, 2011). This uneven distribution means that there is a high spatial heterogeneity in the innovation processes and it should be taken into account. This spatial heterogeneity is very likely to lead to spatial dependence. In general, ignoring spatial heterogeneity leads to biased parameter estimates and consequently to misleading significance tests. In recent year, spatial regression techniques have been developed to deal with spatial heterogeneity in relationships between variables. In particular, geographically weighted regression (GWR) is considered as an effective technique in dealing with spatial heterogeneity. Empirical works aimed at regional innovation processes primarily focus on problems related to spatial autocorrelation and research and development (R&D) spatial spillover effects. The problem of spatial heterogeneity is neglected issue, not only as far as it is concerned the innovation processes. This fact was the main motivation for analysis presented in this paper. The purpose of this study is to perform a spatial heterogeneity analysis of innovative activity of the European Union (EU) regions based on the GWR estimation of the regional knowledge production function (RKPF) model. We will concentrate on the question of spatial differentiation of the RKPF parameters. Our proxy for innovative activity refers to patent application at the EPO (European Patent Office) and the R&D expenditure and human recourses in science and technology were opted as innovative input indicators of the RKPF model.

The study is organized as follows: section 2 deals with main theoretical issues concerning the GWR methodology, section 3 presents empirical results and the paper closes with concluding remarks.

2 METHODOLOGY

The concept of spatial heterogeneity is associated with the situation where it is assumed that the relationship between the dependent variable and the explanatory variables may vary spatially. Instead of fixed values of regression parameters for all spatial units, it is assumed that their values may be different for spatial unit groups, which we refer to as spatial regimes (for more details see Anselin and Rey, 2014). In the most general case we might expect a different relationship to hold for every point in space. Formally, we write a linear relationship depicting this as:

$$y_i = \mathbf{x}_i^{\mathrm{T}} \boldsymbol{\beta}_i + u_i \qquad (1)$$

where i indexes observations collected at $i = 1, 2, \dots, N$ points in space (spatial units), $\mathbf{x}_i^{\mathrm{T}}$ represents a $1 \times k$ vector of explanatory variables with an associated set of $k \times 1$ parameters $\boldsymbol{\beta}_i$, y_i is the dependent variable at observation i and u_i denotes a stochastic disturbance in the linear relationship. A set of estimated local regression parameters can be obtained by applying the Geographically Weighted Regression (GWR) method. The methodological issues of this method will be briefly discussed in next section. With respect to the goal of our analysis, the methodology related to the spatial regimes, i. e., situation when we assume different regression parameter values for groups of spatial units, will be not discussed here.

2.1 A brief description of the GWR method

The GWR method is part of spatial econometric techniques belonging to geographically weighted (GW) models (see e.g., Brunsdon et al., 1999 or Wheeler, 2007). GW models are appropriate when the spatial data analysed can not be adequately captured by a single global model, but a properly localized model calibration provides a better data description.

The goal of GWR methodology is to obtain local linear regression estimates for each point in the space (estimation of model (1)). The GWR method uses the so called window weighing technique, i.e., local linear regressions are performed on geographically weighted data sets. The search window systematically moves within a data file from one point to another when the search window stops at a point, the other points in the neighbourhood are identified, and the local regression is performed on that data set. Larger weights are assigned to the points closest to the central point and local linear regression for N observation is performed N times. The neighbouring points (location) are weighted based on the selected spatial kernel function. The search window range is controlled by the bandwidth of spatial kernel function.

It is evident that the initial step of the GWR method is the spatial kernel function and its bandwidth selection. Next, based on this information, a geographical weight matrix $\mathbf{W}_i$ for each location i is constructed. This N dimensional diagonal matrix is constructed such that $\mathbf{W}_i = f(\mathbf{d}_i, h)$, where $f(\ )$ is a spatial kernel function, $\mathbf{d}_i$ is a distance vector between the central point and all neighbours, and h is a bandwidth or decay parameter. Distance can be calculated in various ways (for more details see e. g., Anselin and Rey, 2014). The geographical weight matrix $\mathbf{W}_i$ has following form:

$$\mathbf{W}_i = \begin{bmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & w_{iN} \end{bmatrix} \qquad i = 1, 2, \dots, N \tag{2}$$

Commonly used continuous or discontinuous kernel functions are e. g., Gaussian, Exponential, Box-car, Bi-square or Tri-cube (for more details see e.g., Gollini et al., 2015). The key parameter in all kernel functions is the bandwidth *h*. For the discontinuous functions (Box-car, Bi-square or Tri-cube), bandwidths can be specified either as a fixed distance or as a fixed number of local data (i.e., an adaptive distance). For the continuous functions (Gaussian or Exponential) bandwidths can be specified either as a fixed distance or as a decay parameter - fixed quantity that reflects local sample size (i.e., still an "adaptive" distance, but the actual local sample size will be the sample size, as functions are continuous). Bandwidths for GWR procedure can be user specified or some optimization procedures (e.g. cross–validation method (CV) approach of Fotheringham et al., 2002, 2010; the generalized cross-validation criterion (GCV) of Loader, 1999; Akaike Information Criterion (AIC) of Hurvich et al., 1998) can be applied. The basic form of the GWR model can be written as follows:

$$\mathbf{W}_i\mathbf{y} = \mathbf{W}_i\mathbf{X}\boldsymbol{\beta}_i + \mathbf{u}_i \tag{3}$$

where $\mathbf{y}$ is $N \times 1$ vector of the observed dependent variable for all N locations, $\mathbf{X}$ is a $N \times k$ matrix of independent variables, $\mathbf{u}_i \sim N\left(0, \sigma^2\mathbf{I}\right)$ is $N \times 1$ vector of random errors, $\boldsymbol{\beta}_i$ is $k \times 1$ vector of parameters corresponding to location i and $\mathbf{W}_i$ is N dimensional weighting matrix at location i. It is necessary mention that $\mathbf{W}_i\mathbf{y}$ is N dimensional vector of geographically weighted dependent variable used to produce estimates for location i. Similarly, $\mathbf{W}_i\mathbf{X}$ is a geographically weighted matrix of independent variables and does not represent individual observation. The parameters in the GWR model are estimated by the weighted least squares approach. The parameter vector at location i is estimated as (LeSage, 1999):

$$\hat{\boldsymbol{\beta}}_i = \left(\mathbf{X}^{\mathrm{T}}\mathbf{W}_i\mathbf{X}\right)^{-1}\left(\mathbf{X}^{\mathrm{T}}\mathbf{W}_i\mathbf{y}\right) \tag{4}$$

The GWR model produces N such vectors of parameter estimates, one for each location. Exploration commonly consists of mapping obtained local regression parameter estimates and associated *t*–values to determine evidence of spatial heterogeneity. Testing procedures and corresponding statistics related to GWR model can be found in Leung et al. (2000).

3 EMPIRICAL RESULTS

The spatial heterogeneity analysis of the innovation activities was performed for 245 NUTS 2 (Nomenclature of Units for Territorial Statistics) EU regions[1] observed over the 2011 – 2012[2] period. The geographic characteristics of the regions - spatial units, in this case polygons, contain the .shp file obtained from the web page of Eurostat which was subsequently corrected in the GeoDa software. The basis of our analysis was RKPF model (see e.g., Moreno, 2005). Based on the RKPF concept, we have chosen patent applications (*PAT*) at the

---

[1] We considered NUTS 2 regions of the EU corresponding to actual state in 2012. At the beginning of the empirical analysis, we had to exclude 20 island regions of Cyprus, Malta, France, Finland, Spain, Greece, Portugal and Italy from our sample of data in order to avoid possible problems with isolated regions. Another data file reduction had to be done due to missing data. We excluded 7 regions of Bulgaria, Germany and Greece.
[2] 2012 was the last year of published statistics by Eurostat.

EPO (number per million of inhabitants) as a proxy for innovative output. Despite some limitations associated with patent applications as an indicator of innovation output, patent applications are considered to be an adequate "representative" of innovation activities as the patent process is a demanding procedure, and only innovations that are potentially of high value are considered. The *PAT* variable represents the averaged value of patent applications for years 2011 and 2012. The first innovative input is represented by the *RDE* variable, expressed as total R&D expenditure in 2011 (in % of GDP) and the second one is the *HRST* variable representing human resources (university graduates) in science and technology in 2011 (% of active population). The *HRST* variable captures the ability to generate new knowledge as well as the ability to absorb external knowledge in the form of knowledge spillover effects. We assumed time lags between innovation inputs and patents and in order to approximate long-run values we averaged two year's data for patents applications. All data comes from Eurostat regional statistics (Eurostat, 2017). GWR estimation of the RKPF model was carried out in R studio software.

The model under the consideration following a Cobb–Douglas technology in logarithmic form is given by:

$$\ln PAT_i = \beta_0 + \beta_1 \ln RDE_i + \beta_2 \ln HRST_i + u_i, \qquad i = 1, 2, ..., N \tag{5}$$

where β_0, β_1 and β_2 are regression parameters to be estimated, all remaining terms have been already defined before. The model defined in (5) will be estimated by ordinary least squares method – OLS (see e.g. Szomolányi et al.), also called as global model and next geographically weighted least squares method (GWR or local model) will be applied in order to analyse the spatial differentiation of the innovation inputs parameters. The GWR technique is considered to be a useful for detecting spatial heterogeneity. The purpose of this paper is to answer two main questions associated with the GWR estimation. The first, we want to know if the GWR model describes the data set significantly better than an OLS model. And second one, if the given GWR model is valid, does each set of its parameters β_{ik} $(i = 1, 2, ..., N)$ vary significantly across the study area? As for the first question, it is, in fact, a goodness of fit test for a GWR model. For this purpose, we applied two different tests for testing goodness of fit, namely F_1 test, which is based on the residual sum of squares and its approximated distribution and F_2 test, which uses analysis of variance (for more details see Leung et al., 2000 and Brunsdon et al., 1999). After a final GWR model has been selected, we can further test whether or not each set of parameters in the model drifts significantly over the study area. To answer this question, test for variation of each set of parameters – F_3 test (for more details see Leung et al., 2000) was used.

In the process of estimating a GWR model, the weighting matrix should first be decided. We decided for Gaussian weighting scheme and the bandwidth was specified as a decay parameter. The value of this parameter was calibrated by cross–validation optimization procedure. The selected results of OLS and GWR estimation are presented in tab. 1. Local GWR parameters and local R^2 coefficients are presented in the form of natural breaks maps in Fig. 1 and Fig. 2 respectively.

Tab.1 Summary of GWR and OLS estimations

Parameter	GWR					Global (OLS)
	Minimum	First Quartile	Median	Third Quartile	Maximum	
Intercept	-14.0341	-2.5953	1.4234	4.1133	14.8900	-3.8969***
ln*RDE*	0.0386	0.4572	0.5905	0.7536	1.5358	0.9362***
ln*HRST*	-2.6838	0.0918	0.7827	1.6694	5.1943	2.1411***
Quasi-global R^2	0.8867					–
AIC_C	471.1504					–
AIC	–					656,3260
Adjusted R^2	–					0.5975

Source: own calculations in R studio.

Note: Symbol *** indicates statistical significance at 1% level of significance.

The evidence for spatial heterogenity is already supported by basic statistics listed in tab. 1. GWR parameter estimate e.g., for ln*RDE* is varying from 0.0386 to 1.5358 with median value equal to 0.5905 while its global OLS parameter estimate is 0.9362. Also, following the results of the F_1 and F_2 tests (see *p*–values in tab. 2) we can conclude that the GWR estimation describes the data set significantly better than the OLS model. The global adjusted R^2 is 0.5975 and the adjusted R^2 (quasi-global R^2) corresponding to GWR estimation is 0.8867 which suggests that there has been some improvement in the model performance. Local R^2 values indicate how well the local regression model fits observed dependent variable values. Mapping local R^2 values (see Fig. 2) we can see that the most of local models have achieved this value higher than 0.6 what could be perceived as a good GWR model performance. A large difference between the AIC and AIC_C (for more details see e.g. Charlton and Fotheringham, 2009) as another measure of fit can be seen as a strong evidence of an improvement in the fit of the model to the data as well.

Tab.2 The results of the *F* – tests

	F – statistic	NDF	DFF	*p*–value
F_1 test	0.3894	203.01	242	0.0000
F_2 test	2.6412	101.36	242	0.0000
F_3 test				
Intercept	2.7493	67.6068	203.01	0.0000
ln*RDE*	1.4379	86.9929	203.01	0.0193
ln*HRST*	2.3571	69.1526	203.01	0.0000

Source: own calculations in R studio.

Note: NDF and DFF refer to the degrees of freedom of the numerator and the denominator, respectively, of the corresponding *F* – distribution.

Based on F_3 test, all three sets of parameters were tested (see *p*–value in tab. 2). The test confirmed the hypotheses that the relation between both innovation inputs and the innovation output is not uniform over the EU area when the other variable is taken to be fixed. Fig. 1 (on the left) shows spatial distribution of local R&D expenditure parameters. The highest values (0.8162 – 1.5358) are evident especially for regions of north Europe, the Baltic countries and e.g. some regions of Czech Republic, Croatia, Hungary, Austria or Slovakia. In the rest of the regions, a different response to the change in R&D expenditure on a patent applications were detected (from 0.0386 to 0.795). Fig. 1 also shows (on the right) spatial distribution of local parameters for human resources in science and technology. Compared to the local R&D expenditure parameters, spatial distribution of *lnHRST* parameters seems to be rather different. The highest values (1.1182 – 5.1943) are also evident for regions of north Europe and the Baltic countries as in previou case, but this group of regions was extented by the regions of e.g. Spain, Portugal, Italy, Greece, Poland, Bulgaria or Romania. Again, the rest of the regions performed a different response to the change in human resources on a patent applications (from -2.6838 to 1.0853). Surprisingly we detected some regions with negative relationship between human resources and patent applications.

In the OLS model, it is conventional to test statistical significance of estimated parameters. Carrying out such tests in GWR model seems to be a little more contentious (see Charlton and Fotheringham, 2009) and raises the problem of multiple significance testing. The problem of individual parameter significance testing is out of the scope of this paper.

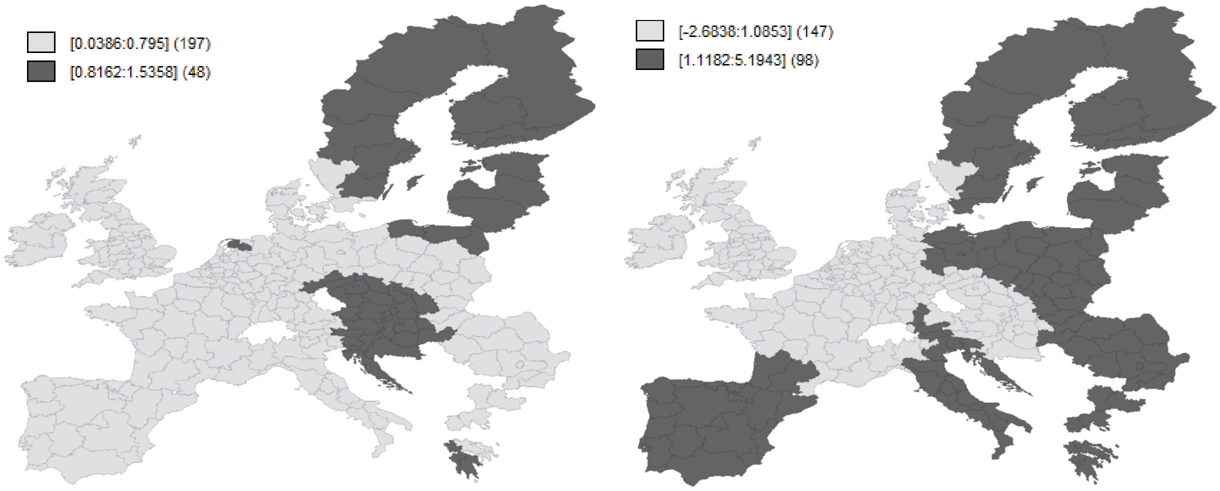


Figure 1: Natural breaks maps for lnRDE (on the right) and lnHRST (on the left) GWR parameters

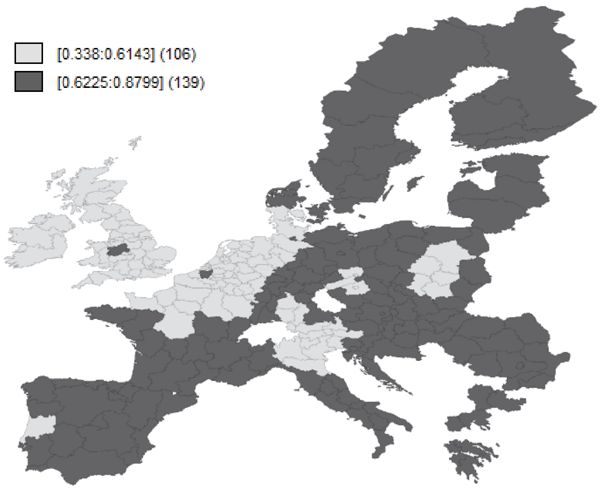


Figure 2: Natural breaks map for local GWR R^2

Next, we have examined the GWR residuals with respect to spatial autocorrelation. GWR residuals have undergone spatial autocorrelation tests (for more details see e.g., Chocholatá, 2017 or Anselin and Rey, 2014). Moran´s I statistic for the residuals is 0.0674 (pseudo p–value = 0.09), so there is little evidence of spatial autocorrealation in them (see Fig. 3 on the left). Fig. 3 (on the right) also provides $G_i(d)$ cluster map for the GWR residuals and according to graphical visualisation of $G_i(d)$ statistics only 20 regions have been found to have positive spatial autocorrelation. Spatial dependencies which might have been present in the residuals for the global – OLS model have been removed with the geographical weighting in the local – GWR model.

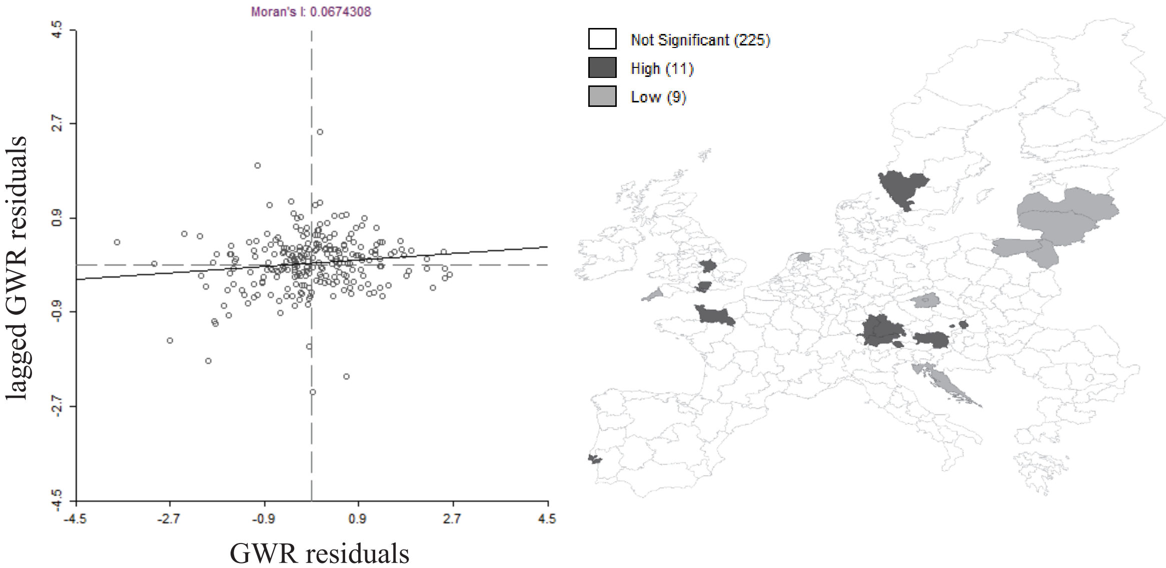


Figure 3: Moran´s I and $G_i(d)$ cluster map for the GWR residuals

Note: calculation based on contiguity queen's case weight matrix (the first order)

4 CONCLUSION

In this study GWR approach has been exploited as a tool for analysing spatial heterogeneity of regional innovative activity of the EU. Based on the GWR estimation of the RKPF model and resulting statistical tests we found out that GWR local regressions describes the data set significantly better than the global OLS model. The second finding was that the both innovation input parameters vary significantly across the EU area. If we look more closely e.g., at the spatial distribution of R&D expenditure parameter, we can conclude that financial support of R&D activities will have significantly different effect on regional patenting activities. The R&D expenditure appear to be the most effective mainly in regions of north Europe, the Baltic countries and some central European regions. This uneven distribution in other words spatial heterogeneity of the innovation activities should be taken into account especially in the EU context. Both innovation input parameters were statistically significant at global level and both exhibited strong regional variation. This situation invoke local R&D policy implication not region wide policy implication.

Acknowledgements: This work was supported by the Grant Agency of Slovak Republic – VEGA grant no. 1/0248/17 „Analysis of regional disparities in the EU based on spatial econometric approaches".

References

[1] Anselin, L. and Rey, S. J. (2014). *Modern spatial econometrics in practice: a guide to GeoDa, GeoDaSpace and PySAL*. Chicago, IL: GeoDa Press LLC, p.369.

[2] Autant-Bernard, C. (2011). *Spatial Econometrics of Innovation: Recent Contributions and Research Perspectives*. Working paper GATE no. 1120. Available at: https://halshs.archives-ouvertes.fr/halshs-00605056/document [Accessed 20 Feb. 2018].

[3] Brunsdon, C., Fotheringham, A. S. and Charlton, M. (1999). Some Notes on Parametric Significance Tests for Geographically Weighted Regression. *Journal of Regional Science*, 39(3), pp. 497–524.
[4] Charlton, M. and Fotheringham, A. S. (2009) *Geographically Weighted Regression. A Tutorial on using GWR in ArcGIS 9.3*. Available at https://www.geos.ed.ac.uk/~gisteac/fspat/gwr/gwr_arcgis/GWR_Tutorial.pdf [Accessed 20 Feb. 2018].
[5] Chocholatá, M. (2017). Analýza mzdových disparít v okresoch Slovenska v období 2001-2015. In: *AIESA – Building of Society based on Knowledge (17th International Scientific conference*). [online] Bratislava: Letra Edu, pp. 316–326. Available at: https://fhi.euba.sk/www_write/files/veda-vyskum/konferencie/aiesa/AIESA_2017/AIESA_2017__cely_zbornik.pdf [Accessed 20 Feb. 2018].
[6] Fotheringham, A. S., Brunsdon, C. and Charlton, M. (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester: Wiley, p.269.
[7] Fotheringham, A. S., Brunsdon, C. and Charlton, M. (2010). *Geographically weighted regression the analysis of spatially varying relationships*. Chichester: Wiley, p.282.
[8] Gollini, I., Lu, B., Charlton, M., Brunsdon, C. and Harris, P. (2015). GW model: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. *Journal of Statistical Software*, 63(17), pp. 1–50.
[9] Hurvich, C. M., Simonoff, J. S. and Tsai, C. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology*), 60(2), pp. 271–293.
[10] LeSage, J. P. (1999). *The theory and practice of spatial econometrics*. Available at: http://www.spatial-econometrics.com/html/sbook.pdf. [Accessed 20 Feb. 2018].
[11] Leung, Y., Mei, C. and Zhang, W. (2000). Statistical Tests for Spatial Nonstationarity Based on the Geographically Weighted Regression Model. *Environment and Planning A*, 32(1), pp. 9–32.
[12] Loader, C. (1999). *Local regression and likelihood.* New York: Springer, p.290.
[13] Moreno, R., Paci, R. and Usai, S. (2005). Spatial spillovers and innovation activity in European regions. *Environment and Planning A*, 37(10), p. 1793–1812.
[14] Wheeler, D. C. (2007). Diagnostic Tools and a Remedial Method for Collinearity in Geographically Weighted Regression. *Environment and Planning A*, 39(10), pp. 2464–2481.
[15] Szomolányi, K., Lukáčiková, A. and Lukáčik, M. (2011). The β-convergence of the EU27 countries. In: *29th International Conference on Mathematical Methods in Economics 2011*. Jánska Dolina, Slovakia, Praha: Professional Publishing, pp. 287–291.
[16] Eurostat (2017). Database – Eurostat. [online] Available at: http://ec.europa.eu/eurostat/web/regions/data/database [Accessed 20 Jun. 2017].

Author's address
doc. Andrea Furková PhD.
University of Economics in Bratislava, Faculty of Economic Informatics, Department of Operations research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
Email: andrea.furkova@euba.sk

CONTRIBUTION TO CREW ROSTERING IN AIR TRANSPORT

Vojtěch Graf, VSB - Technical University of Ostrava; Dušan Teichmann VSB - Technical University of Ostrava; Michal Dorda VSB - Technical University of Ostrava

Abstract
Air transport offers a broad spectrum of problems that can be successfully solved by using mathematical programming methods. We can mention for example flight schedule planning, assigning airplanes or crew members to individual flights. Matching the crew members can bring many real constraints such as different experience, training skills, language skills or more specific constraints resulting in the fact that some pilots can be assigned to the same flight or not. The article is focused on solving the problem of crew rostering based on graph theory and linear programming

Keywords: *Crew rostering, Air transport, Linear programming, Graph theory*

JEL Classification: C61
AMS Classification: 90C08

1 INTRODUCTION

One of the main targets of each air transport company is to provide services of corresponding quality, especially with respect to fluency of the transportation process and delay minimisation, to its customers. On the other hand, because of strong competition each air transport company wants to minimise its costs. One of possible ways how to reach requested quality standards of the offered services with minimal costs is optimal planning of flight schedules and all the processes connected with them. The planning process of the air transport company is a chain consisting of several partial planning steps.

Within the first step it is necessary to create the flight schedule (a plan of the individual flights operated by the air transport company) on the basis of destinations which the company wants to offer. Consequently, the company must define arrival and departure times of all the flights which have been defined within the first step of the planning process. The arrival and departure times are influenced by operational conditions of airports above all. We can mention for example service hours of the airport, free time slots for take-off and landing or available technical equipment for aircraft ground handling.

The next step of the planning process is assigning suitable airplanes to the individual flights. The assignment is also influenced by some factors, the most important are operational and technical parameters of the airplanes such as their range of flight, capacity and so on. Another important criterion which should be considered is an expected demand for transportation on the air route or technical parameters of the airports. If the air transport company operates a homogenous fleet of airplanes, then this part of the planning process is simplified essentially.

After matching the available airplanes with the planned flights, it is necessary to plan crews (pairs of pilots) for the flights. The crew planning process can be divided into two partial steps. Within the first step the pilots are matched, that means we create the pairs of them. After it we can assign the created crews to the individual flights. Both planning processes can be successfully solved by means of mathematical programming methods. Using the methods may simplify the decision process and we obtain an optimal solution when applying the methods correctly.

2 CREW PLANNING

Crew planning (in our case pairing the available pilots) is influenced by many factors. In other words, many reasons why two pilots can be or cannot be matched may be given in reality. One of the most important factors is the fact that each crew has to be formed by a captain (a pilot in command) and a first officer. That means we cannot create any crew consisting of the pilots who do not have the rank of captain – the pilots are not qualified to be the pilot in command. Another influencing factor is seniority of the pilots. For our purposes we can divide the pilots into the experienced pilots and inexperienced ones on the basis of their level of experience. When pairing the pilots, we should prefer the crews consisting of an experienced pilot and an inexperienced one.

Another factor which can influence the process of matching the pilots is dividing the pilots into instructors and cadet pilots (under training). In such cases the instructor pilot has to be matched with the cadet pilot. Another factor which should be mentioned is dividing airports into three classes according to difficulty of take-off and landing – three categories A, B and C are defined, where the class C is the most difficult category of the airports. Each pilot has to be qualified in order to land at the airports of the individual categories. That means a pilot who is qualified only for the category A must not land at any airport of the category B or C.

Other reasons for matching or not matching the pilots can be represented by different personal or family reasons. Two pilots must not be matched it they do not like one another because such aversion may bring potentially dangerous situations during the whole flight. Analogously, in case of married couples we try not to match them. From the psychological point of view it is usual that one person is dominant in the married couples and this fact may negatively influence decisions, especially in case of emergency, of the pilot in command if he/she is the person who is not dominant in the marriage.

However, with no respect to the created plan some unexpected disruptions of the flying staff, especially of the pilots, may happen in practice due to many reasons. In such cases the original plan has to be modified in order to react to the disruptions. The modifications can represent an optimisation problem – the air transport company has a set of the pilots and on the basis of their qualification and experience the company has to create new aircrews. The pilots can be found in different localities that are different from the airports of the planned flight departures. That means the newly planned aircrews are assigned to the individual flights with respect to costs associated with transportation of the aircrews to the airports of their departure.

3 STATE OF THE ART

Many authors have addressed the problem of crew planning. The problem is still topical because almost all means of transport cannot be operated automatically with any human operator. Many constraints – for example work shifts or ordered rest times which are typical for all the modes of transport – arising in reality complicate the problem of crew planning. Therefore, we can draw inspiration also from publications which are devoted to other modes of transport – see for example (Palúch, 2013) or (Majer, Palúch and Peško, 2017).

There are different approaches to planning crews which are applied in available publications. The main difference lies in methods which are employed in the individual publications and in complexity of the applied approaches. Some authors focus only on solving a single planning problem. In (Doi, Nishi and Voß, 2018) the authors create the crews with respect to uniformly distributed shifts of all the pilots. The authors present a two-step algorithm which matches the pilots to create the crews and generates their rest times. On the other hand, in (Erdogan et al., 2015) the authors focus on assigning the crews to the individual flights. The authors developed a heuristic method which can be applied for planning up to 27 000 flights. An optimization criterion which is used in the heuristic corresponds to the total costs of assigning the crews to the flights which bear the air transport company. The developed method was applied on real examples. Results obtained by the method were compared with the planning process of the air transport company Air France (Desoulniers et al., 1997). The comparison revealed that applying the heuristic brings potential savings in planning the crews.

An interesting approach to crew planning was used in (Yildiz,Gzara and Elhedhli, 2017). The authors of the article defined constraints which result from natural human body cycles; they incorporated the constraints in a mathematical model. They point out that even when fatigue of the pilots from previous flights is considered in the crew planning process, so often ordered breaks are not enough to eliminate their fatigue.

In (Mercier and Soumis, 2017) the crew planning problem with time coordination of the flights is integrated. To solve the integrated problem a combination of several methods – such as column generation method or generating dynamic constraints – was applied. An application of the proposed method brought a decrease in need for the airplanes and their crews. In (Deveci ane Demirel, 2018) and (Quesnel, Desaulniers and Soumis, 2017) the authors focus on integrating the crew planning problem with assigning the crews to the individual airplanes.

4 PROBLEM FORMULATION AND PROPOSED HEURISTIC METHOD

In the article we present a heuristic that can be used for crew rostering in air transport. In our case we create as many pairs of the pilots as possible and the pairs are consequently assigned to the individual flights. The problem is solved as a two-step optimisation task – a linear mathematical model creates the pairs of the available pilots in the first step and another linear model assigns the pairs to the individual flights in the second step. Thus, the presented models can be employed for this part of the planning process of the air transport companies.

The first model, intended for matching the pilots, proceeded from graph theory, more specifically from an algorithm for matching in general graphs (Palúch and Peško, 2006). Vertices of a graph represent the individual pilots who are available. Edges of the graph model admissible matching – if the pilots can be matched then the vertices corresponding to the pilots are joined by an edge. Weights of all the vertices and edges of the graph are equal to 1. An example of such graph is shown in Figure 1. We can see that vertex 1 is incident with 3 edges. Therefore, vertex 1 is adjacent to 3 vertices – vertex 2, 4 and 6. This means pilot 1 can be matched with pilots 2, 4 and 6.

Matching in the graphs is based on a simple algorithm, the algorithm searches for a path in the graph which visits as many vertices as possible. The path begins in any vertex (for example vertex 1) and any incident edge is included in the path – let us choose the edge connecting vertex 1 with vertex 2. The edge which has been chosen is highlighted by a bold line – see Figure 2. From vertex 2 we can visit any adjacent, not visited yet, vertex – in our case we visit vertex 3. However, now the edge is not highlighted. In the same manner we visit other, not visited yet, vertices until the maximum possible number of the vertices is visited. We also alternate the highlighted edges with the unhighlighted ones. A resulting path is shown in

Figure 2 – the path visits vertices 1→2→3→5→6→4. The highlighted edges indicate matching in the graph. In our example we have three pairs of the pilots – pilot 1 with pilot 2, pilot 3 with pilot 5 and pilot 4 with pilot 6. If there are any vertices that have not been matched (excluding a vertex in a graph with an odd number of the vertices), it is necessary to test whether a better solution cannot be found. The test of optimality is based on searching for an alternative path which begins and ends in the vertices that have not been matched. But the rule that the highlighted edges alternate with the unhighlighted ones must be kept in the alternative path. The algorithm is more precisely described in (Demel, 2002).

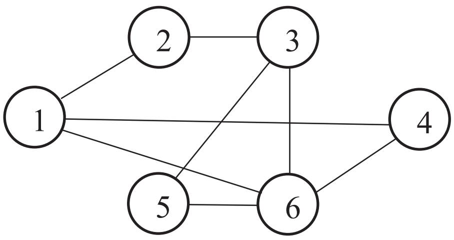


Figure 1: Example of the graph

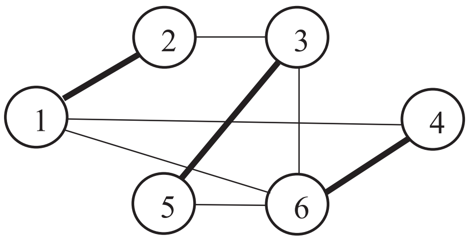


Figure 2: Matching in the graph

Let a set I of the pilots that can be assigned to the planned flights be given. A set $J(I = J)$ contains the same elements as the set I. The set J which is equal to the set I is defined to depict the possible links between the individual pilots easily. Possible matching of the pilots is modelled by a matrix d_{ij}. If it holds that $d_{ij} = 1$, then the pilot *i* can be matched with the pilot *j*. If $d_{ij} = 0$, then the pilot *i* cannot be matched with the pilot *j*. A group of variables x_{ij} modelling matching the corresponding pilots is used in the model. The mathematical model can be defined as follows:

$$\max f(x) = 0,5 \cdot \sum_{i \in I} \sum_{j \in J} x_{ij} \tag{1}$$

subject to:

$$\sum_{j \in J} x_{ij} \le 1 \text{ for } i \in I, \tag{2}$$

$$\sum_{i \in I} x_{ij} \le 1 \text{ for } j \in J, \tag{3}$$

$$x_{ij} = x_{ji} \text{ for } i \in I, j \in J, \tag{4}$$

$$x_{ij} \le d_{ij} \text{ for } i \in I, j \in J, \tag{5}$$

$$x_{ij} \in \{0,1\} \text{ for } i \in I, j \in J. \tag{6}$$

Function (1) corresponds to the objective function of the model; the function models the number of the crews (the pairs of the pilots). The value of the optimisation criterion must be multiplied by 0.5 in order to get the real number of the crews and not the number of the matched pilots. The modification of the objective function relates to constraints (4). Let us consider that the pilots *i* and *j* are matched, that means $x_{ij} = 1$. However, matching the pilots *i* and *j* means that also the pilots *j* and *i* are matched – that means $x_{ji} = 1$. This fact causes that the value of the optimisation criterion is increased by 2, but only a single crew was created.
Constraints (2) ensure that each pilot i is assigned to a single or no crew. Constraints (3) have the same meaning, but for the pilot j. Constraints (5) ensure that the pilots who cannot create the crew are not matched by the model. Constraints (6) define a domain of definition of the

variables x_{ij} – the variables are binary. If $x_{ij} = 1$, then the pilots i and j are matched. If $x_{ij} = 0$, then the pilots are not matched.

The second linear mathematical model continues in results of the first model. In this case we deal with an assignment task; the optimisation criterion minimises the costs of assigning the crews to the individual flights. Let a set K of the crews be given. The individual crews $k \in K$ must be assigned to a set of the planned flights L. We consider that cardinalities of both sets are equal – the number of the crews is equal to the number of the flights. Moreover, a matrix c_{kl} must be defined. Elements of the matrix correspond to the costs that are associated with assigning the crew k to the flight l. A group of binary variables y_{kl} is defined in the model; the variables model our decision about assigning or not assigning the crew k to the flight l.

The mathematical model of the assignment problem is defined in the following form:

$$\min f(y) = \sum_{k \in K} \sum_{l \in L} c_{kl} \cdot y_{kl} \tag{7}$$

subject to:

$$\sum_{l \in L} y_{kl} = 1 \text{ for } k \in K \tag{8}$$

$$\sum_{k \in K} y_{kl} = 1 \text{ for } l \in L \tag{9}$$

$$y_{kl} \in \{0,1\} \text{ for } k \in K, l \in L \tag{10}$$

Function (7) expresses the optimisation criterion, its value corresponds to the total costs of assigning the crews to the flights. Constraints (8) assure that each crew is assigned to a flight. On the other hand, constraints (9) secure that only single crew is assigned to each flight. Constraints (10) define a domain of definition of the variables y_{kl}, the variables are binary as well. If $y_{kl} = 1$, then the crew k is assigned to the flight l. If $y_{kl} = 0$, then the crew k is not assigned to the flight l.

In the presented heuristic method, we make some assumptions. All the airplanes are found at the airports where the individual flights begin – that means it is not necessary to deal with time coordination of the individual flights. Analogously, for the crews we assume the pilots are going on duty and, therefore, we do not consider limitations regarding their required rest periods between the consecutive flights. In addition, we assume that no flight is longer than the pilot can be in duty. And finally, the number of the available crews is equal to the number of the planned flights.

5 OPTIMISATION EXPERIMENT

Functionality of the presented method will be tested on an example. An air transport company employs 21 pilots who can form the individual crews. The matrix d_{ij} gives information whether the concrete two pilots can be matched or not.

$$d_{ij} = \begin{bmatrix}
0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}$$

The first model (1) – (6) was solved in optimisation software Xpress-IVE for this example. We found an optimal solution. The maximum number of the crews is equal to 10. In Table 1 the variables x_{ij} which are equal to 1 are listed – the variables give information about the crews that should be created. In the table one can see 20 values of the variables x_{ij} but as discussed in the previous text the final number of the crews corresponds to one half of this number. We can see that pilot 2 was not matched.

x_{ij}	**Value**	x_{ij}	**Value**
x_{13}	1	x_{31}	1
x_{45}	1	x_{54}	1
x_{67}	1	x_{76}	1
x_{89}	1	x_{98}	1
x_{1013}	1	x_{1310}	1
x_{1112}	1	x_{1211}	1
x_{1416}	1	x_{1614}	1
x_{1517}	1	x_{1715}	1
x_{1821}	1	x_{2118}	1
x_{1920}	1	x_{2019}	1

Table 1: Matching the pilots

The results of the first model correspond to input data for the second model, which assigns the crews to the individual airplanes that have been chosen for the corresponding flights. In our example the air transport company have 10 crews which have to be assigned to 10 planned flights. Each crew that have been created is numbered – see Table 2.

Crew x_{ij}	Crew number k	Crew x_{ij}	Crew number k
x_{13}	1	x_{1112}	6
x_{45}	2	x_{1416}	7
x_{67}	3	x_{1517}	8
x_{89}	4	x_{1821}	9
x_{1013}	5	x_{1920}	10

Table 2: Numbering of the crews

A matrix c_{kl} contains the costs that the air transport company has to bear for assigning the crew k to the place where the flight l starts; the costs represent for example travel costs of the pilots. One can see that values of T are in the matrix. The value T corresponds to a so-called prohibitive constant; the constant takes a big value which has to be greater than the other values in the matrix (in our case $T = 100000$). The constant is used in cases when a certain crew cannot be assigned to a certain flight for some reasons, for example due to insufficient qualification of the pilots.

$$c_{kl} = \begin{bmatrix} 2 & 3 & 5 & 1 & 6 & 2 & 4 & 5 & 1 & 2 \\ 3 & 1 & 5 & 4 & 2 & 6 & 3 & 4 & T & 4 \\ 2 & 5 & 6 & 4 & 4 & 2 & T & 1 & 3 & T \\ 1 & 1 & 2 & 3 & 5 & 1 & 1 & 2 & 4 & 2 \\ 5 & 6 & 2 & 4 & 3 & 1 & 5 & 4 & 2 & 4 \\ 1 & 2 & T & 4 & 2 & 5 & 1 & 3 & 1 & 3 \\ T & 1 & 2 & 4 & 6 & 2 & 1 & 1 & 2 & 3 \\ 5 & 1 & 3 & 8 & 4 & 6 & 3 & 1 & 5 & 4 \\ 4 & 2 & 5 & 6 & 3 & 5 & 1 & 4 & 2 & 6 \\ 3 & 4 & T & 4 & 6 & 4 & 2 & 3 & 2 & 1 \end{bmatrix}$$

The second model (7) – (10) was also implemented in optimisation software Xpress-IVE and resolved. An optimal solution was found as well. The total costs of assigning the crews to the flights are equal to 12 monetary units. The proposed assignment is shown in Table 3. In the last column of the table the costs c_{kl} are listed in order to check the value of the optimisation criterion.

y_{kl}		$c_{kl}[mu]$
Crew number k	**Airplane number** l	
1	4	1
2	5	2
3	8	1
4	1	1
5	6	1
6	9	1
7	3	2
8	2	1
9	7	1
10	10	1
		$\sum c_{kl} = 12$

Table 3: Resulting assignment

6 CONCLUSIONS

Air transport cannot be operated without precise planning. In order to minimise economical losses incurred by delays of the flights, the air transport companies are forced to plan the flight schedules and all the processes connected with them conscientiously. Planning the flight plan consists of partial subsequent phases. In the first phase the air routes are chosen on the basis of the passengers' demand and offered destinations. Based on the air routes the individual flights have to be planned. For each flight a suitable airplane must be chosen.

The last part of the planning process is crew planning which comprises two tasks – matching the pilots (each crew consists of two pilots) and assigning the crews to the individual flights. The presented article dealt with the crew planning process. The heuristic method was presented; the method consists of two consequent steps. The linear mathematical model for the first step matches the pilots in order to create the crews. The optimisation criterion corresponds to the number of the created crews which we want to maximise. The results of the first phase enter to the second phase of the heuristic method; the second phase assigns the crews to the planned flights under minimisation of the total travel costs of the crews.

Functionality of the proposed method was tested on the example. Both linear models can be solved in optimisation software Xpress-IVE. Because both parts of the crew planning process are solved separately, we cannot state that the final solution is optimal. Therefore, it is desirable to develop an approach which would cover both partial problems. One of the possible ways is to combine the presented linear models.

References

[1] Demel, J. (2002). *Grafy a jejich aplikace.* Praha: Academia.

[2] Desoulniers, G., Desrosiers, J., Dumas, Y., Marc, S., Rioux, B., Solomon, M. M., Soumis, F. (1997). Crew pairing at Air France. *European Journal of Operational Research, 97(2)*, pp. 245-259.

[3] Deveci, M., Demirel, N. C. (2018). Evolutionary algorithm for solving the airline crew pairing problem. *Computers & Industrial Engineering, 115*, pp. 389-406.

[4] Doi, T., Nishi, T., Voß, S. (2018). Two-level decomposition-based matheuristic for airline crew rostering problems with fair working time. *European Journal of Operational Research, 267(2)*, pp. 428-438.

[5] Erdogan, G., Haouari, M., Matoglu, M. O., Ozener, O. O. (2015). Solving a large-scale crew pairing problem. *Journal of the Operational Research Society, 66(10)*, pp. 1742-1754.

[6] Majer, T., Palúch, S., Peško, Š. (2017). Algorithm for vehicle crew scheduling in regular bus transport. In: *18th International Carpathian Control Conference ICCC 2017*. Sinaia, pp. 300-305.

[7] Mercier, A., Soumis, F. (2017). An integrated aircraft routing, crew scheduling and flight retiming model. *Computers &Operations Research, 34(8)*, pp.2251-2265.

[8] Palúch, S. (2013). Vehicle and crew scheduling problem in regular personal bus transport. In: *Proceeding of 16th International Conference on Transport Science ICTS 2013*, Portoroz, pp. 297-304.

[9] Palúch, S., Peško, Š. (2006). *Kvantitatívne metódy v logistike.* Žilina: Žilinská univerzita v Žiline.

[10] Quesnel, F., Desaulniers, G., Soumis, F. (2017). A new heuristic branching scheme for the crew pairing problem with base constraints. *Computers and Operational Research, 80*. pp. 159-172.

[11] Yildiz, B. C., Gzara, F., Elhedhli, S. (2017). Airline crew pairing with fatigue: Modeling and analysis. *Transportation Research Part C, 74*, pp. 99-112.

Author's address
Ing. Vojtěch Graf
VŠB - Technical University of Ostrava, Faculty of Mechanical Engineering, Institute of
17. listopadu 15/2172, 708 33, Ostrava - Poruba
Czech Republic
email: vojtech.graf@vsb.cz

doc. Ing. Dušan Teichmann, Ph.D.
VŠB - Technical University of Ostrava, Faculty of Mechanical Engineering, Institute of
17. listopadu 15/2172, 708 33, Ostrava - Poruba
Czech Republic
email: dusan.teichmann@vsb.cz

doc. Ing. Michal Dorda, Ph.D.
VŠB - Technical University of Ostrava, Faculty of Mechanical Engineering, Institute of
17. listopadu 15/2172, 708 33, Ostrava - Poruba
Czech Republic
email: michal.dorda@vsb.cz

EFFICIENCY EVALUATION IN MULTI-STAGE DEA MODELS: A MINIMAX GOAL PROGRAMMING APPROACH

Josef Jablonský,
University of Economics, Prague

Abstract
Traditional data envelopment analysis (DEA) models evaluate efficiency of transformation of multiple inputs into multiple outputs in one stage. The main information given by this evaluation is splitting the units according to their efficiency scores into two disjunctive sets – efficient and inefficient ones. Multi-stage models assume that all or at least a certain part of outputs of a stage *s* serve as the inputs of the next stage *s*+1. The outputs of the last stage of the model are final outputs of the production process. It is possible to derive the efficiency scores in all particular periods but there are proposed several models attempting to compute aggregated efficiency characteristics over all periods considered. Our aim is to propose a goal programming based model that minimizes maximum deviation of the particular efficiency scores from their ideal values. The model is verified on a real data set and its results compared with the results of traditional models.

Keywords: *Data envelopment analysis, Multi-stage models, Minimax, Multiple criteria decision making*

JEL Classification: C44
AMS Classification: 90C29, 90C05

1 INTRODUCTION

Data envelopment analysis (DEA) models allow evaluation of relative efficiency and performance of decision making units (DMUs) that spent multiple inputs and produce multiple outputs. In general, DEA models measure efficiency of transformation of multiple inputs into multiple outputs relatively to other members of the set of DMUs under evaluation. Let us suppose that the set of DMUs contains *n* elements. The DMUs are evaluated by *m* inputs and *r* outputs with input and output values x_{ij} ($i = 1,\dots,n$; $j = 1,\dots,m$) and y_{ik}, ($i = 1,\dots,n$; $k = 1,\dots,r$), respectively. The efficiency of DMU$_q$ is defined as the weighted sum of outputs divided by the weighted sum of inputs with weights expressing the importance of single inputs/outputs v_i ($i = 1,\dots,m$) and u_k ($k = 1,\dots,r$):

$$\theta_q = \frac{\sum_{k=1}^{r} u_k y_{qk}}{\sum_{j=1}^{m} v_j x_{qj}} \tag{1}$$

where θ_q is the efficiency score of the unit DMU$_q$.

Traditional DEA models have been introduced by Charnes et al. (1978) and further extended by Banker et al. (1984). They deal with efficiency analysis of single-stage processes. However, reality is usually more complex and production process under consideration can be defined as a multi-stage process. A simplest case is the process that consist of two serial sub-processes. These two sub-processes can be evaluated independently using traditional DEA models but several attempts have been published how to analyze efficiency and performance of the whole two-stage system. (Kao and Hwang, 2008) and (Chen et al., 2009a) are two first DEA models of this nature. They have been further studied and extended by (Chen et al., 2009b). A slack-based model (SBM) for two-stage processes is proposed in (Jablonsky, 2017). These models can be easily extended for a more general case of multi-stage processes. Multi-stage systems and their efficiency evaluation by DEA models is still subject of research among many researchers. This is illustrated by the work (Li et al., 2018) that discusses latest developments in this field.

The aim of this paper is to propose a multi-stage DEA model that is based on the methodology of goal programming. This model works in two stages. The first stage consists in computation of efficiency scores in particular stages independently using one of the traditional DEA models. The second stage minimizes maximum deviation of the efficiency scores given by an aggregated multi-stage model from their ideal values.

The paper is organized as follows. Section 2 contains formulation of (Kao and Hwang, 2008) model and its extensions. Section 3 introduces an original multi-stage DEA model and illustrates its use on a numerical example taken from (Kao and Hwang, 2008). Final section discusses given results and possibilities of a future research in this field.

2 DEA MULTI-STAGE SERIAL MODELS

A Multi-stage production process is graphically illustrated on Figure 1. The first stage of this process spends m inputs and produces p_1 outputs. For simplicity, we suppose that all outputs of the first stage serve as inputs of the second stage (in general, some of them may leave production process). The second stage produces r final outputs of the production process.

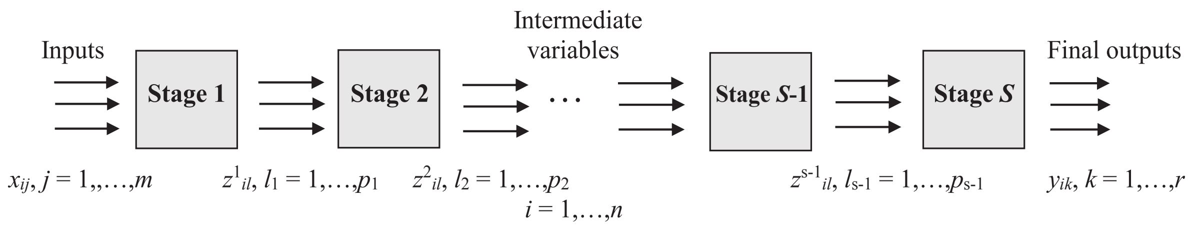


Figure 1: *Multi-stage serial model*

Let us denote for the DMU$_i$: x_{ij}, $i = 1,\ldots,n$; $j = 1,\ldots,m$, the value of the j-th input of the first stage, y_{ik}, $i = 1,\ldots,n$; $k = 1,\ldots,r$, the value of the k-th final output of the last stage, and $z^s_{il_s}$, $i = 1,\ldots,n$; $l_s = 1,\ldots,p_s$, $s = 1\ldots,S{-}1$ the value of the l_s-th output of the s-th stage and at the same time the value of the l_s-th input of the $(s{+}1)$-th stage. Individual efficiency scores of the first and the last stages θ^1_q and θ^S_q are defined as follows:

$$\theta^1_q = \frac{\sum_{l_1=1}^{p_1} w^1_{l_1} z^1_{ql_1}}{\sum_{j=1}^{m} v_j x_{qj}}, \qquad \theta^S_q = \frac{\sum_{k=1}^{r} u_k y_{qk}}{\sum_{l_{S-1}=1}^{p_{S-1}} w^{S-1}_{l_{S-1}} z^{S-1}_{ql_{S-1}}}, \tag{2}$$

where $w^s_{il_s}$ $l_s = 1,\ldots,p_s$, $s = 1\ldots,S{-}1$, are the weights of the intermediate variables in the s-th stage. The overall efficiency score θ^0_q for the DMU$_q$ is the weighted sum of final outputs divided by the weighted sum of the inputs of the first stage, i.e.

$$\theta_q^0 = \frac{\sum_{k=1}^{r} u_k y_{qk}}{\sum_{j=1}^{m} v_j x_{qj}}. \tag{3}$$

Obviously the relation $\theta_q^0 = \theta_q^1 \theta_q^2 ... \theta_q^S$ holds.

(Kao and Hwang, 2008) was originally proposed for two-stage serial models but can easily be generalized for a multi-stage case. This generalized model maximizes overall efficiency score θ_q^0 of the unit under evaluation under the assumption that efficiency scores in all other stages for all other units of the set do not exceed 1. This model is not linear in its objective function but can be easily transformed into a linear program using Charnes and Cooper transformation. The linearized version of this model is below (the model assumes $S \geq 3$; for two-stage models the second set of constraints is removed):

$$\begin{aligned}
&\text{Maximize} && \theta_q^{*KH} = \sum_{k=1}^{r} u_k y_{qk} && \\
&\text{subject to} && \sum_{l_1=1}^{p_1} w_{l_1}^1 z_{il_1}^1 \leq \sum_{j=1}^{m} v_j x_{ij}\,, && i = 1,\dots,n, \\
& && \sum_{l_s=1}^{p_s} w_{l_s}^s z_{il_s}^s \leq \sum_{l_{s-1}=1}^{p_{s-1}} w_{l_{s-1}}^{s-1} z_{il_{s-1}}^{s-1}, && i = 1,\dots,n,\ s = 2,\dots,S-1, \\
& && \sum_{k=1}^{r} u_k y_{ik} \leq \sum_{l_{S-1}=1}^{p_{S-1}} w_{l_{S-1}}^{S-1} z_{il_{S-1}}^{S-1}, && i = 1,\dots,n, \\
& && \sum_{j=1}^{m} v_j x_{qj} = 1, && \\
& && v_j > 0, w_l > 0, u_k > 0. &&
\end{aligned} \tag{4}$$

Maximum efficiency score of the unit DMU$_q$ in all stages can be derived using conventional input-oriented DEA models with constant returns to scale technology. These models for the first and the last (S-th) stages are following:

$$\begin{aligned}
&\text{Maximize} && \theta_q^{*1} = \sum_{l=1}^{p_1} w_l^1 z_{ql}^1 && \theta_q^{*S} = \sum_{k=1}^{r} u_k y_{qk} && \\
&\text{subject to} && \sum_{l=1}^{p_1} w_l^1 z_{il}^1 \leq \sum_{j=1}^{m} v_j x_{ij}\,, && \sum_{k=1}^{r} u_k y_{ik} \leq \sum_{l=1}^{p_{S-1}} w_l^{S-1} z_{il}^{S-1}, && i = 1,\dots,n, \\
& && \sum_{j=1}^{m} v_j x_{qj} = 1, && \sum_{l=1}^{p_{S-1}} w_l^{S-1} z_{ql}^{S-1} = 1, && \\
& && v_j > 0, w_l^1 > 0, w_l^{S-1} > 0, u_k > 0. && &&
\end{aligned} \tag{5}$$

The optimal efficiency scores for other stages of the model can be computed in a similar way. Note that optimal weights given by model (4) and by models (5) are not identical.

3 A MINIMAX MULTI-STAGE MODEL

Efficiency scores (2) computed using optimal weights of models (4) cannot be higher (better) than those computed using weights obtained by models (5). The idea of the proposed model is to derive weights of all variables (inputs, intermediate variables in all stages and final outputs) in order the efficiency scores (2) computed using these weights differ from their ideal values (maximum efficiency scores computed by models (5)) as little as possible. The proposed model minimizes maximum deviation of ideal efficiency scores and efficiency scores obtained by common weights using model (4). Its formulation follows:

$$
\begin{aligned}
\text{Minimize}\quad & D \\
\text{subject to}\quad & \theta_q^s + d_s = \theta_q^{*s}, \quad s = 1,\ldots,S, \qquad (6)\\
& d_s \le D, \qquad s = 1,\ldots,S,\\
& d_s \ge 0,\\
& \text{+ positivity constraints for all other variables (weights of inputs, intermediate}\\
& \text{variables and final outputs),}
\end{aligned}
$$

Where θ_q^s is the efficiency score of DMU$_q$ in the s-th stage, θ_q^{*s} is the ideal efficiency score computed by model (5), and d_s are deviational variables that must be lower or equal than the maximum deviation D.

Model (6) is not linear and not easily solvable. That is why we propose the following modification (like in the case of model (4) this model assumes $S \ge 3$; in case of two-stage model the second set of constrains is removed):

$$
\begin{aligned}
\text{Minimize}\quad & D \\
\text{subject to}\quad & \sum_{l_1=1}^{p_1} w_{l_1}^1 z_{il_1}^1 \le \sum_{j=1}^{m} v_j x_{ij}, \qquad i = 1,\ldots,n, \qquad (7)\\
& \sum_{l_s=1}^{p_s} w_{l_s}^s z_{il_s}^s \le \sum_{l_{s-1}=1}^{p_{s-1}} w_{l_{s-1}}^{s-1} z_{il_{s-1}}^{s-1}, \qquad i = 1,\ldots,n,\ s = 2,\ldots,S-1,\\
& \sum_{k=1}^{r} u_k y_{ik} \le \sum_{l_{S-1}=1}^{p_{S-1}} w_{l_{S-1}}^{S-1} z_{il_{S-1}}^{S-1}, \qquad i = 1,\ldots,n,\\
& \sum_{l_1=1}^{p_1} w_{l_1}^1 z_{ql_1}^1 + d_1 = \theta_1^* \sum_{j=1}^{m} v_j x_{qj},\\
& \sum_{l_s=1}^{p_s} w_{l_s}^s z_{ql_s}^s + d_{s-1} = \theta_s^* \sum_{l_{s-1}=1}^{p_{s-1}} w_{l_{s-1}}^{s-1} z_{ql_{s-1}}^{s-1}, \qquad s = 2,\ldots,S-1,\\
& \sum_{k=1}^{r} u_k y_{qk} + d_S = \theta_S^* \sum_{l_{S-1}=1}^{p_{S-1}} w_{l_{S-1}}^{S-1} z_{ql_{S-1}}^{S-1},\\
& d_s \le D, \qquad s = 1,\ldots,S,\\
& d_s \ge 0,\\
& v_j > 0, w_l > 0, u_k > 0.
\end{aligned}
$$

Table 1: *Data set for 24 non-life insurance companies in Taiwan*

	X1	X2	Z1	Z2	Y1	Y2
Taiwan Fire	1178744	673512	7451757	856735	984143	681687
Chung Kuo	1381822	1352755	10020274	1812894	1228502	834754
TaiPing	1177494	592790	4776548	560244	293613	658428
China Mariners	601320	594259	3174851	371863	248709	177331
Fubon	6699063	3531614	37392862	1753794	7851229	3925272
Zurich	2627707	668363	9747908	952326	1713598	415058
Taian	1942833	1443100	10685457	643412	2239593	439039
MingTai	3789001	1873530	17267266	1134600	3899530	622868
Central	1567746	950432	11473162	546337	1043778	264098
TheFirst	1303249	1298470	8210389	504528	1697941	554806
KuoHua	1962448	672414	7222378	643178	1486014	18259
Union	2592790	650952	9434406	1118489	1574191	909295
Shingkong	2609941	1368802	13921464	811343	3609236	223047
South China	1396002	988888	7396396	465509	1401200	332283
Cathay Century	2184944	651063	10422297	749893	3355197	555482
Allianz	1211716	415071	5606013	402881	854054	197947
Newa	1453797	1085019	7695461	342489	3144484	371984
AIU	757515	547997	3631484	995620	692731	163927
North America	159422	182338	1141950	483291	519121	46857
Federal	145442	53518	316829	131920	355624	26537
R.Sunalliance	84171	26224	225888	40542	51950	6491
Aisa	15993	10502	52063	14574	82141	4181
AXA	54693	28408	245910	49864	0.1	18980
Mitsui Sumito	163297	235094	476419	644816	142370	16976

Let us denote (**v***, **w**^{1}*, …, **w**^{S}*, **u***) vectors of optimal weights computed by model (7). The efficiency scores of the first and the last stages and overall efficiency score using these optimal weights are as follows (efficiency scores of intermediate stages are computed in the same way):

$$\theta_q^1 = \frac{\sum_{l_1=1}^{p_1} w_{l_1}^* z_{ql_1}^1}{\sum_{j=1}^{m} v_j^* x_{qj}}, \qquad \theta_q^S = \frac{\sum_{k=1}^{r} u_k^* y_{qk}}{\sum_{l_{S-1}=1}^{p_{S-1}} w_{l_{S-1}}^* z_{ql_{S-1}}^{S-1}}, \quad \theta_q^0 = \theta_q^1 \dots \theta_q^S. \qquad (8)$$

Results (8) allow ranking of the DMUs easily according to their overall efficiency scores θ_i^0, $i = 1,\dots,n$.

The results of our proposed model (7) will be illustrated on a case taken from (Kao and Hwang, 2008) which is two-stage model. The original data set is presented in Table 1. In this study, 24 non-life insurance companies in Taiwan are described by:

- Two inputs in the first stage (X1 – operation expenses, X2 – insurance expenses),
- Two intermediate variables (Z1 – direct written premiums, Z2 – reinsurance premiums), and
- Two final outputs (Y1 – underwriting profit, Y2 – investment profit).

The results of numerical experiments are presented in Table 2. The following models have been applied:

- Conventional input-oriented DEA model with constant returns to scale technology (1st stage).
- Conventional input-oriented DEA model with constant returns to scale technology (2nd stage).
- Conventional input-oriented DEA model with constant returns to scale technology (overall efficiency - inputs of the first stage and final outputs of the second stage).
- Kao and Hwang model (3).
- Minimax Model (7).

The efficiency scores calculated using the mentioned models are completed by ranking of DMUs (in parentheses).

Table 2: *Results of numerical experiments*

	Stage 1	Stage 2	Overall	Model (3)	Model (7)
Taiwan Fire	0.993 (7)	0.713 (7)	0.984 (6)	0.699 (3)	0.695 (3)
Chung Kuo	0.998 (6)	0.627 (10)	1.000 (1)	0.625 (5)	0.624 (5)
TaiPing	0.69 (23)	1.000 (1)	0.988 (5)	0.690 (4)	0.690 (4)
China Mariners	0.724 (21)	0.432 (16)	0.488 (14)	0.304 (15)	0.301 (15)
Fubon	0.838 (13)	1.000 (1)	1.000 (1)	0.767 (1)	0.756 (2)
Zurich	0.964 (8)	0.406 (18)	0.594 (13)	0.390 (12)	0.377 (11)
Taian	0.752 (16)	0.538 (13)	0.470 (16)	0.277 (17)	0.276 (17)
MingTai	0.726 (19)	0.511 (15)	0.415 (19)	0.275 (18)	0.275 (18)
Central	1.000 (1)	0.292 (23)	0.327 (22)	0.223 (20)	0.219 (20)
TheFirst	0.862 (11)	0.674 (9)	0.781 (10)	0.466 (9)	0.436 (9)
KuoHua	0.741 (18)	0.327 (22)	0.283 (23)	0.164 (23)	0.139 (23)
Union	1.000 (1)	0.760 (6)	1.000 (1)	0.760 (2)	0.760 (1)
Shingkong	0.811 (14)	0.543 (12)	0.353 (20)	0.208 (21)	0.205 (21)
South China	0.725 (20)	0.518 (14)	0.470 (16)	0.289 (16)	0.288 (16)
Cathay Century	1.000 (1)	0.705 (8)	0.979 (7)	0.614 (6)	0.585 (7)
Allianz	0.907 (10)	0.385 (19)	0.472 (15)	0.320 (14)	0.320 (13)
Newa	0.723 (22)	1.000 (1)	0.635 (11)	0.360 (13)	0.345 (12)
AIU	0.794 (15)	0.374 (20)	0.427 (18)	0.259 (19)	0.244 (19)
North America	1.000 (1)	0.416 (17)	0.822 (9)	0.411 (11)	0.310 (14)
Federal	0.933 (9)	0.901 (5)	0.935 (8)	0.547 (8)	0.513 (8)
R.Sunalliance	0.751 (17)	0.280 (24)	0.333 (21)	0.201 (22)	0.200 (22)
Aisa	0.590 (24)	1.000 (1)	1.000 (1)	0.590 (7)	0.590 (6)
AXA	0.850 (12)	0.560 (11)	0.599 (12)	0.420 (10)	0.386 (10)
Mitsui Sumito	1.000 (1)	0.335 (21)	0.257 (24)	0.135 (24)	0.100 (24)

4 CONCLUSIONS

Results of numerical example show a very close similarity in ranking of DMUs between Kao and Hwang model (3) and the model proposed in this paper. It is possible easily prove that the final efficiency scores obtained by our model (7) are always lower or equal to those obtained by Kao and Hwang model. There are several rank reversals – e.g. Fubon and Union companies on the first and second rank. In detail, the reason for this rank reversal between these two companies may be illustrated by efficiency scores of two stages obtained by both models. The efficiency scores in the first and second stages given by Kao and Hwang model are 0.831 and 0.923 respectively. Their product is the final efficiency score obtained by this model, i.e. θ^{*KH}_{Fubon} = 0.767. The same values for Union company are 1.000, 0.767 and θ^{*KH}_{Union} = 0.767. Maximum difference from ideal efficiency scores (see Table 2) are for Fubon D = 1 - 0.923 = 0.077 and for Union this difference is zero because the model returns for this company maximum efficiency scores. Our model (7) returns for Fubon the following values in the first and second stages – 0.797 and 0.949. The maximum difference in this case is D = 0.051 which is lower than in the previous case. Overall efficiency score by our model for Fubon company is the product of two particular efficiencies, i.e. 0.756. For Union company our model returns the same results as Kao and Hwang model and overall efficiency score is 0.760.

Measuring efficiency in multi-stage models is an important task and this paper is a contribution in research in this field. Goal programming or more general multiple criteria decision making methodology allows a more flexible approach to solving this problem – except minimax approach a lexicographic approach or weighted sum of deviations may be used for measuring the distance from ideal values. This can be a good starting point for a future research.

Acknowledgements: The research is supported by the Czech Science Foundation, project no. 16-01821S.

References

[1] Banker, R.D., Charnes, A. and Cooper, W.W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), pp. 1078–1092.

[2] Charnes, A., Cooper, W.W. and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), pp. 429–444.

[3] Chen, Y., Liang, L. and Zhu, J. (2009a). Equivalence in two-stage DEA approaches. *European Journal of Operational Research*, 193(2), pp. 600-604.

[4] Chen, Y., Cook, W.D., Li, N. and Zhu, J. (2009b). Additive efficiency decomposition in two-stage DEA models. *European Journal of Operational Research*, 196(3), pp. 1170-1176.

[5] Jablonský, J. (2017). Benchmarking of countries at Summer Olympic Games 2016 using two-stage DEA models. In: *Mathematical Methods in Economics 2017 - Conference Proceedings*. University of Hradec Králové, pp. 277-282.

[6] Kao, C. and Hwang, S.N. (2008). Efficiency decomposition in two-stage data envelopment analysis: An application to non-life insurance companies in Taiwan. *European Journal of Operational Research*, 185(1), pp. 418-429.

[7] Li, H., Chen, C., Cook, W.D., Zhang, J. and Zhu, J. (2018). Two-stage network DEA: Who is the leader? *Omega*, 74(1), pp. 15-19.

Author's address

Prof. Ing. Josef Jablonský, CSc.
University of Economics, Prague, Faculty of Informatics and Statistics, Department of Econometrics
W. Churchill Sq. 4, 13067 Praha 3
Czech Republic
email: jablon@vse.cz

FAST ALGORITHM FOR DETERMINATION OF FULL NETWORK ARC CHARACTERISTIC IN A SERVICE SYSTEM

Jaroslav Janáček, Marek Kvet,
University of Žilina, Faculty of Management Science and Informatics

Abstract
Reliability of a transportation network links influences importantly function of public service systems constituted on the network. To be able to study impact of an individual network arc on proper functionality of the service system, we suggested characteristic function of the arc transit time elongation and completed a fast algorithm, which enables to determine progress of the characteristic and provided a reader with computational study performed with real road networks.

Keywords: *Detrimental event, Robust service system design, Critical arc characteristic, Full characteristic computation*

JEL Classification: C61
AMS Classification: 90C27

1 INTRODUCTION

Reliability of any service system is strongly influenced by vulnerability of the underlying transportation network used for the service delivery. An expected traversing time of individual arcs of the transportation network can be elongated by many traffic density, weather conditions, accidents etc. The traversing time elongation of the individual arc may or may not impact performance of the service system depending on mutual position of the considered arc and the network paths, along which the service is delivered. Within this paper, we concentrate our effort on revealing the arcs, which are critical from the point of proper functioning of the service system. The effect of individual arc collapse was originally studied in [4, 6, 7] and the associated research issued establishing measures of the arc criticality called „Network robustness index“ and „Network trip robustness“. The index values were computed from completely exclusion of the arc from the network graph. Other approach [2] was based on a definition of so-called arc characteristics an on suggestion of the fast algorithm, which is capable to determine upper and lower bound of the characteristics progress.
Within this paper, we focus on suggesting such an algorithm, which can compute complete arc characteristics by one run of the distance matrix re-computation.
The remainder of the paper is organized as follows. In the next section, we will define the full arc characteristics and perform analysis of possible shapes of the characteristics in a symmetrical transportation network. The third section will be devoted to an explanation of the suggested algorithm. The fourth section will contain computational study performed with real service systems operating on real transportation networks. The fifth section will conclude the computational study and give some perspectives of the future research.

2 NETWORK ARC CHARACTERISTICS

Most of public service systems provide a finite set of system users with the kind of service from a finite set of service centers. Service is delivered along the shortest (fastest or cheapest) paths in the transportation network, which node set contains the service center locations and the user locations [1, 3, 5].The system is represented by a set I_1 of p service center locations, by a set J of user locations, where volume b_j of demand is associated with the user location j. Symbol t_{ij} denotes the traversing time of the fastest path from i to j in the underlying network. The set of associated arcs forming the path is denoted P_{ij}. Let us denote $ass(j) \in I_1$ the nearest center location to user j. Then the transport performance TP^* of the service system can be computed as follows.

$$TP^* = \sum_{j \in J} b_j t_{ass(j),j} \tag{1}$$

Let symbol a denote an arc with traversing time t_a. We will deal with the impact of t_a elongation by Δ on the total transport performance of the system. We realize that arc a may or may not be element of some of the paths $P_{ass(j),j}$ for $j \in J$. If the arc does not belong to any of the paths, it cannot influence the transport performance and we denote it irrelevant. A relevant arc can belong to several paths We denote $J^a \subseteq J$ the set of $j \in J$, for which $a \in P_{ass(j),j}$. Let us analyze the progress of transport performance $TP^a(\Delta)$ with increasing value of Δ representing elongation of the traversing time t_a. If the value of Δ is near to zero, the value $TP^a(\Delta)$ will grow linearly with a slope B_1, because each path $P_{ass(j),j}$ for $j \in J^a$ stays the shortest one. If Δ reaches the value, for which first one of the paths becomes not to be the shortest one, then an associated user j will be serviced by another shortest path to some service center from I_1 and this path will avoid the arc a. Then the slope of $TP^a(\Delta)$ will decrease by b_j because the shortest path from j to the nearest service center will be not influenced by elongation of t_a. This situation can repeat several times, while elongation reaches subsequently values $e_1, \ldots, e_q$, what is accompanied by subsequently decreasing series of slopes $B_2, \ldots, B_{q+1}$. If the elongation Δ moves across the range $[e_i, e_{i+1}]$, then the slope of $TP^a(\Delta)$ equals to B_{i+1} (see Fig. 1). The function $TP^a(\Delta)$ is piecewise linear concave increasing function. If the network graph is symmetrical, i.e. $t_{ij} = t_{ji}$ for every node pair (i, j), then either $TP^a(\Delta)$ is linear function with slope B_1 or B_{q+1}=0.

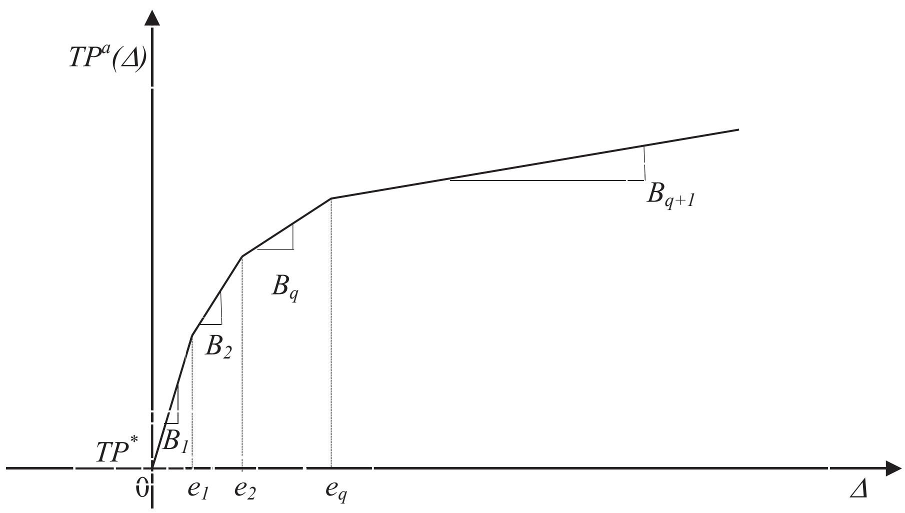


Fig. 1 Network arc full characteristic

3 COMPUTATION OF ARC CHARACTERISTICS

The following algorithm uses Pappe's approach for computation of the shortest paths from all users to the nearest service center in the adjusted network. The resulting nearest center to a user j will be denoted by $\underline{ass}(j)$ and the time length of the associated path will be denoted by $t'_{\underline{ass}(j),j}$. Let T be a big constant, e.g. sum of all arc traversing times is satisfactorily big. Then, the algorithm for determination of arc a characteristic parameters can be stated as follows.

0. Increase traversing time t_a of the arc a by T. Define $B = \sum_{j \in J^a} b_j$
1. Re-compute the shortest paths form each $j \in J^a$ to the nearest service center $\underline{ass(j)}$ and denote the corresponding traversing time by $t'_{\underline{ass(j)},j}$.
2. Initialize set L of difference values by $L=\emptyset$.
   Process each user $j \in J^a$ in the following way:
   Define $Dif= t'_{\underline{ass(j)},j} - t_{ass(j),j}$.
   If $Dif \notin L$, then update $L=L\cup\{Dif\}$ and initialize set $J(Dif)=\{j\}$,
   else update $J(Dif)= J(Dif)= \cup\{j\}$.
3. Order the set L of noL elements increasingly so that $L(1)$ is the minimal value of computed differences and $L(noL)$ is the biggest one.
   If $L(1) = 0$, then update $B = B - \sum_{j \in J(L(0))} b_j$ and exclude $L(1)$ from the ordered set L.
4. If $L(noL) \geq T$, then the progress of $TP^a(\Delta)$ is linear function with slope $B_1 = B$.
   Otherwise apply the following process to determine characteristic parameters q, $B_1, \ldots, B_{q+1}$, and $e_1, \ldots, e_q$.
   $q=noL$,
   $B_1 = B$,
   for i=1, …, noL perform
   $e_i=L(i)$,
   $B_{i+1} = B_i - \sum_{j \in J(L(i))} b_j$
5. Terminate.

4 COMPUTATIONAL STUDY

Within this section, we present a computational study focused on real network arc characteristics regarding a given public service system. The used benchmarks were derived from the real emergency health care system, which was originally implemented in three selected regions of Slovakia. For each self-governing region, i.e. Banská Bystrica (BB), Prešov (PO) and Žilina (ZA), all cities and villages with corresponding number b_j of inhabitants were taken into account as aggregate users with associated demands. The coefficients b_j were rounded to hundreds. These sub-systems cover demands of all communities - towns and villages spread over the particular regions by given number of ambulance vehicles. In the benchmarks, the set of communities represents both the set J of users' locations and also the set I of possible service center locations. The basic characteristics of used benchmarks are summarized in Table 1. Each row of the table corresponds to one of the studied networks. The individual benchmark is described by the cardinality of set I and by the value of p, which expresses the number of located service centers. The finite set of located service centers corresponds to current centre locations. The right part of the table is used to report the results of edges classification. The total number of all edges of the network is given in the column denoted by "Total". The column "Relevant" contains the number of relevant arcs obtained by suggested algorithm. The relevant arcs are either cut-links, number of which is denoted by "CL" or non-cut-links, number of which is reported as "nCL".

Table 1: Benchmark characteristics

Region	$\|I\|$	p	Edges classification			
			Total	Relevant	CL	nCL
BB	515	36	843	615	168	447
PO	664	32	1018	771	256	515
ZA	315	29	494	370	130	240

An individual experiment was organized so that the set of relevant network arcs was identified first. Then, the list of relevant arcs was partitioned into the set of cut-links, which was excluded from further process of their characteristics evaluation, and the set of non-cut-links. The set of non-cut-links was reduced by the arcs, which lie on the shortest paths from users to the nearest service center locations, but for each such user exists detour, which avoids the arc and has the same length as the shortest path. To perform the algorithm described in the previous sections, the programming environment NetBeans 8.1 was used and the associated algorithms were programmed in Java language and the experiments were run on a PC equipped with the Intel® Core™ i7 5500U processor with the parameters: 2.4 GHz and 16 GB RAM.
The analysis of network non-cut-links characteristics obtained for the individual self-governing regions is presented in pair of figure and table. The figure depicts the distribution of arcs according to the number q of characteristics segments. The table contains statistics of arc characteristics belonging to the individual classes specified by the value of q. Each table is organized so that the individual rows correspond to mentioned classes of network arcs specified by q. For each class we present the following statistics, which correspond to the individual columns:

- min e_1 – minimal value of the first segment length
- max e_1 – maximal value of the first segment length
- min e_q– minimal length of characteristics range (i.e. interval, where transportation performance increases with increasing traversing time elongation)
- max e_q– maximal length of characteristics range (i.e. interval, where transportation performance increases with increasing traversing time elongation)
- min B_1 – minimal value of slope of characteristics in the first segment
- max B_1 – maximal value of slope of characteristics in the first segment
- avg B_1 – average value of slope of characteristics in the first segment
- avg *meanP* – average value of mean transportation performance increment in the characteristics range

The following pair of Fig. 2 and Table 2 corresponds to the results obtained for the region of Banská Bystrica (BB). The pair of Fig. 3 and Table 3 corresponds to the results obtained for the region of Prešov (PO) and the pair of Fig. 4 and Table 4 corresponds to the results for the region of Žilina (ZA)

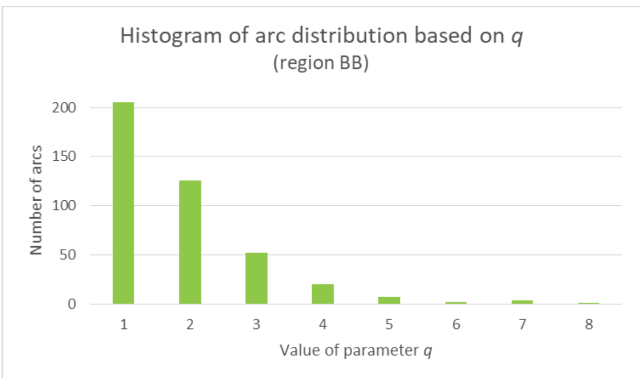


Fig. 2 Non-cut-links distribution based on the value of q for the region of Banská Bystrica

Table 2: Results of the network analysis for the self-governing region of Banská Bystrica

q	min e_1	max e_1	min e_q	max e_q	min B_1	max B_1	avg B_1	avg *meanP*
1	1	29	1	29	1	55	8.7	21.3
2	1	15	2	23	2	71	18.1	53.9
3	1	8	4	28	5	86	27.4	118.9
4	1	5	7	23	8	87	31.3	140.4
5	1	2	8	19	17	56	33.1	174.7
6	1	1	18	23	21	48	34.5	238.8
7	1	1	14	20	21	46	30.0	171.7
8	2	2	27	27	38	38	38.0	434.4

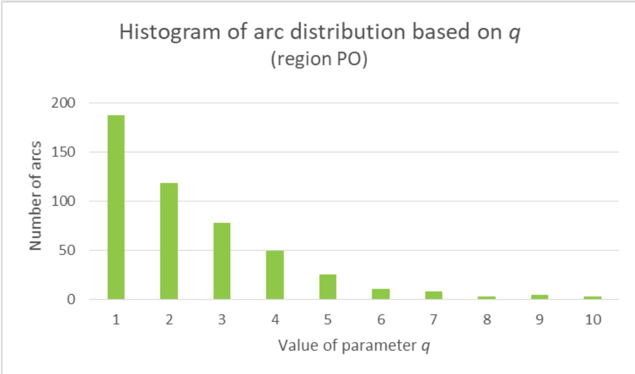


Fig. 3 Non-cut-links distribution based on the value of q for the region of Prešov

Table 3: Results of the network analysis for the self-governing region of Prešov

q	min e_1	max e_1	min e_q	max e_q	min B_1	max B_1	avg B_1	avg *meanP*
1	1	21	1	21	1	61	10.6	20.3
2	1	21	2	27	2	220	23.6	75.3
3	1	10	3	29	4	137	29.4	124.7
4	1	8	4	35	6	140	39.4	186.6
5	1	8	9	31	8	176	44.6	272.3
6	1	3	13	24	17	78	41.2	209.2
7	1	5	12	26	21	142	73.3	437.4
8	1	2	22	26	54	83	68.5	648.5
9	1	2	13	24	28	145	80.3	581.7
10	1	1	27	30	37	99	68.0	763.3

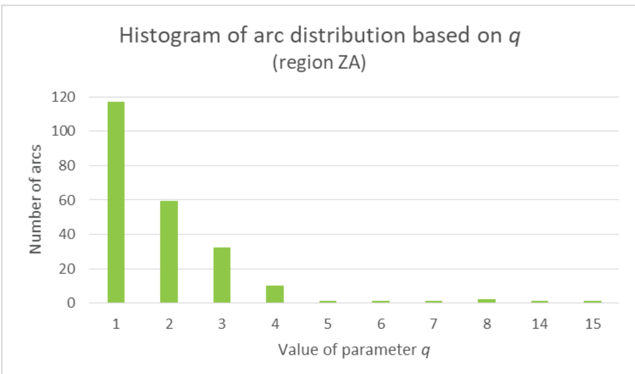


Fig. 4 Non-cut-links distribution based on the value of q for the region of Žilina

Table 4: Results of the network analysis for the self-governing region of Žilina

q	min e_1	max e_1	min e_q	max e_q	min B_1	max B_1	avg B_1	avg *meanP*
1	1	25	1	25	1	107	17.2	39.6
2	1	8	2	18	3	113	32.4	95.2
3	1	4	3	24	6	147	45.4	178.7
4	1	3	4	16	13	84	50.1	225.6
5	1	1	9	9	25	25	25.0	98.9
6	1	1	6	6	67	67	67.0	169.4
7	1	1	13	13	32	32	32.0	117.6
8	1	1	13	15	40	103	71.5	365.0
14	3	3	20	20	70	70	70.0	554.7
15	3	3	25	25	107	107	107.0	1086.7

5 CONCLUSIONS

This paper was focused on means of network arc importance investigation, where importance is considered to be relative regarding a given public service system. The suggested algorithm produced full characteristics of each relevant arc. The obtained characteristics enable to partition the set of relevant arcs into several classes and separate so-called cut-links from the other relevant arcs, which have an alternative if the arc collapses. This way, we presented a tool for transportation network analysis from the point of performance of a service system. The attached computational study demonstrates usage of the algorithm for real road network analysis. The future research in this field will be focused on exploitation of the suggested algorithm and associate arc evaluations for construction of an effective set of detrimental scenarios destined for robust service system designing.

Acknowledgements: This work was supported by the research grants VEGA 1/0342/18 "Optimal dimensioning of service systems", VEGA 1/0463/16 "Economically efficient charging infrastructure deployment for electric vehicles in smart cities and communities" and APVV-15-0179 "Reliability of emergency systems on infrastructure with uncertain functionality of critical elements".

References

[1] Janáček, J. and Kvet, M. (2012). Relevant Network Distances for Approximate Approach to Large p-Median Problems. In: *Operations Research Proceedings 2012*, Hannover, Germany, Springer, pp. 123-128.

[2] Janáček J. and Kvet M. (2017). Characteristic of a critical network arc in a service system. *Transport Problems*, 12, pp. 141-146.

[3] Jánošíková, Ľ. (2007). Emergency Medical Service Planning. *Communications - Scientific Letters of the University of Žilina*, 9(2), pp. 64-68.

[4] Jenelius, E. (2009). Network structure and travel patterns: explaining the geographical disparities of road network vulnerability. *Journal of Transport Geography*, 17, pp. 234-244.

[5] Kvet, M. (2014). Computational Study of Radial Approach to Public Service System Design with Generalized Utility. In: *Proceedings of International Conference Digital Technologies 2014*, Žilina, Slovakia, pp. 198-208.

[6] Scott, D. M., Novak, D. C., Aultman-Hall, L. and Guo, F. (2006). Network robustness index: a new method for identifying critical links and evaluating the performance of transportation networks. *Journal of Transport Geography*, 14(3), pp. 215-227.

[7] Sullivan, J. L., Novak, D. C., Aultman-Hall, L. and Scott, D. M. (2010). Identifying critical road segments and measuring system-wide robustness in transportation networks with isolating links: A link-based capacity-reduction problem. *Transportation Research Part A*, 44, pp. 323-336.

Authors' addresses
Prof. RNDr. Jaroslav Janáček, CSc.
University of Žilina - Faculty of Management Science and Informatics
Univerzitná 8215/1, 010 26 Žilina, Slovakia
Email: jaroslav.janacek@fri.uniza.sk

Ing. Marek Kvet, PhD.
University of Žilina - Faculty of Management Science and Informatics
Univerzitná 8215/1, 010 26 Žilina, Slovakia
Email: marek.kvet@fri.uniza.sk

KERNEL SEARCH FOR THE CAPACITATED P-MEDIAN PROBLEM

Ľudmila Jánošíková,
University of Žilina, Faculty of Management Science and Informatics

Abstract
The capacitated p-median problem is an NP-hard combinatorial optimization problem. Its solution is extremely time demanding, thus the researchers are looking for approximate methods that are able to provide near-optimal solutions in a reasonable computation time. The kernel search is one of the recently developed matheuristics. It is a general and simple heuristic framework based on the idea of intensification of the search process in promising regions of the solution space. Promising variables are identified by the linear programming relaxation of the problem. The kernel search has been successively applied to a variety of mixed integer programming problems. In the paper we propose its implementation for the capacitated p-median problem and evaluate the efficiency of the algorithm using benchmark instances.

Keywords: *Capacitated p-median problem, Matheuristic, Kernel search*

JEL Classification: JEL C61
AMS Classification: AMS 90C09, AMS 90C27, AMS 90C59

1 THE CAPACITATED *P*-MEDIAN PROBLEM

The capacitated p-median problem (CPMP) is a classical problem in location analysis (ReVelle and Eiselt, 2005). It seeks a location of a given number of facilities in the region under consideration. The facilities are supposed to serve customers in such a way that a customer must be served from exactly one facility. Moreover, the facilities have restricted capacities which means that every facility can serve only a limited volume of demands. Problems of this type arise in the industrial sphere as well as in the public sector. An example is ambulance location where one must take into account that an ambulance is able to answer only a limited number of emergency calls.

Let us specify the problem in a more formal way.

We are given a set I of candidate locations, where exactly p facilities are to be placed. These facilities serve customers from the set J. Each customer j has an associated demand b_j and must be assigned to exactly one facility. Each facility has a given capacity Q which must not be exceeded by the total demand of customers located in its service area. Further the distance d_{ij} between location i and customer j is given. The objective is to find location of p facilities and the assignment of customers to them so that the total distance between facilities and their relevant customers could be as small as possible.

To formulate the problem in terms of mathematical programming, decision variables have to be introduced first. The decision on opening a facility must be done for each candidate location $i \in I$. To model this decision we need a binary variable y_i, which takes the value 1 if a facility is located in node i, otherwise it takes the value 0. The assignment of customer j to the facility located in node i is modelled by a binary variable x_{ij}. Variable x_{ij} takes the value 1, if customer j will be served by a facility located in node i, otherwise $x_{ij} = 0$.

After these preliminaries, the model of the capacitated *p*-median problem can be written as follows:

$$\text{minimize} \quad \sum_{i \in I} \sum_{j \in J} d_{ij} x_{ij} \tag{1}$$

$$\text{subject to} \quad \sum_{i \in I} x_{ij} = 1 \quad \text{for } j \in J \tag{2}$$

$$x_{ij} \le y_i \quad \text{for } i \in I,\ j \in J \tag{3}$$

$$\sum_{j \in J} b_j x_{ij} \le Q y_i \quad \text{for } i \in I \tag{4}$$

$$\sum_{i \in I} y_i = p \tag{5}$$

$$x_{ij}, y_i \in \{0,1\} \quad \text{for } i \in I,\ j \in J \tag{6}$$

The objective function (1) minimizes the total distance between facilities and their relevant customers. Constraints (2) ensure that every customer *j* will be assigned to exactly one facility *i*. Constraints (3) ensure that if a customer *j* is assigned to a node *i*, then a facility will be open in the node i. Constraints (4) limit the total demand of customers assigned to each open facility. Constraint (5) limits the total number of facilities. The remaining obligatory constraints (6) specify the definition domains of the variables.

The CPMP is known to be an *NP*-hard combinatorial optimization problem. It means that it cannot be solved to optimality in a polynomial time. That is why approximate methods have been developed that are able to provide near-optimal solutions in a reasonable computation time. The state-of-the-art approximate method was published by Stefanello et al. (2015) under the name IRMA (Iterated Reduction Matheuristic Algorithm). IRMA is an efficient heuristic based on a domain decomposition. It applies the local optimization method as a master method and the integer programming solver as a slave solving a subproblem. In addition, the solution space is reduced by eliminating variables that probably will not be in the optimal solution. All popular metaheuristics were applied for solving the CPMP as well, e.g. a genetic algorithm (Correa et al., 2004), a grouping genetic algorithm (Jánošíková and Vasilovský, 2017) or a scatter search method (Scheuerer and Wendolsky, 2006). A plenty of hybrid solution methods that combine concepts of various algorithmic approaches have also been developed. Several hybrid approaches were proposed by the research group around Masoud Yaghini from the Iran University of Science and Technology. In the paper (Yaghini et al., 2010) they propose a hybrid metaheuristic called GACO that combines two metaheuristics: a genetic algorithm (GA) and ant colony optimization metaheuristic (ACO). The GA guides the search. It acts with an incomplete representation of candidate solutions, specifically with sets of medians. The ACO is used to obtain corresponding actual solutions, it means to allocate demand nodes to the medians. The paper (Yaghini et al., 2013) introduces a hybrid metaheuristic based on a tabu search metaheuristic that applies mathematical programming to explore the neighbourhood of the current solution. Another hybrid metaheuristic that combines path relinking and scatter search principles was proposed by Díaz and Fernández (2006). Path relinking is used within the scatter search metaheuristic to combine two solutions from the reference set.

Matheuristics represent another promising research direction. These approximate methods are based on a mathematical programming formulation of the problem. In principle, the approach consists in successive modifications of the original problem and their solution by a solver, which is used as a black-box. The kernel search is one of the recently developed matheuristics.

2 KERNEL SEARCH

The kernel search was proposed by Professor Grazia Speranza and her students from the University of Brescia, Italy as an efficient heuristic for the solution of mixed integer linear programs (MILPs) with binary variables (Guastaroba et al., 2017). The main idea of the heuristic is to solve a series of restricted problems that contain only a subset of promising decision variables (a kernel) and a small subset of the remaining variables. The problems are solved using a general-purpose MILP solver as a black-box.

The initial kernel K consists of promising binary variables that are identified by solving the linear programming (LP) relaxation of the original MILP. The variables that are positive in the optimal solution to the LP relaxation are regarded as promising because it is likely that they take the value 1 in an optimal (or near-optimal) solution to the original MILP. All of them or a part of them form the initial kernel. The remaining variables are sorted in non-decreasing order of their reduced costs and then partitioned into buckets. In the most simple setting, all buckets (possibly except the last one) have the same number of variables l. If we denote the number of all binary variables as n and the number of variables in the initial kernel as m, then $N_b = \left\lceil \frac{n-m}{l} \right\rceil$ buckets will be created. The initial feasible integer solution is calculated by solving the MILP(K), which is the original MILP restricted to the variables in the kernel only. Then the improving phase follows. It consists of the loop that repeats two steps: (i) solving a restricted MILP($K \cup B_k$) problem that contains only variables from the kernel and one bucket B_k; (ii) updating kernel K so that new promising variables are added and variables that are not perspective any more are removed from the kernel. The loop terminates if all buckets or the first $\overline{N_b}$ buckets have been explored. A general scheme of the procedure is in Algorithm 1.

Algorithm 1 The KS algorithmic framework (Guastaroba et al. 2017)

(*Initialization*)

1. Solve the linear programming relaxation of the 0–1 MILP.
2. Determine an initial kernel K and a sequence of buckets $\{B_k\}$ for $k = 1, ..., N_b$ containing the variables that are not in the kernel.
3. Solve MILP(K).

(*Improvement*)

4. **while** $k \leq \min \{\overline{N_b}, N_b\}$ **do**
    Solve MILP($K \cup B_k$).
    Analyze the solution and adjust kernel K.
   end while

3 THE KERNEL SEARCH FOR THE CPMP

The general scheme of the kernel search has to be adapted to a particular problem. In this section we describe the implementation of the kernel search for the CPMP. We focus on the practical-sized problems with hundreds or thousands of customers and candidate locations.

The first issue is the construction of the initial kernel. We found inspiration in the application of the kernel search for the single source capacitated facility location problem (Guastaroba and Speranza, 2014). Promising location variables are specified by the optimal solution to the LP relaxation of the original problem (1)-(6). The open facilities (for which $y_i > 0$ in the optimal solution) are sorted in non-increasing order of the total demand they serve. The remaining candidate locations, where a facility was not open (for which $y_i = 0$) are sorted in non-decreasing order of their reduced costs. The variables y_i that correspond to the first m

locations in the ordered list are selected to the initial kernel, the remaining location variables create the buckets. The size of the initial kernel m and the size of the buckets l are algorithm's parameters that significantly influence its behavior. The smaller these values, the smaller the solution space is, i.e. the shorter computation time is needed to obtain the optimal solution to the restricted CPMP. On the other hand, a small solution space decreases the quality of the solution. So the size of the kernel and the buckets must be a compromise between speed and solution quality. In the CPMP setting there is another factor as for the parameter m. To get a feasible problem with respect to constraints (4) and (5), m must be at least p and large enough to ensure that the aggregated capacity of all facilities in the kernel will meet all customers' demands. That is why we set $m = \max\{\lceil \sum_{j \in J} b_j / Q \rceil, p\} + l$. The size of the bucket l depends on the size of the problem instance. It will be specified in the following section devoted to computational experiments.

For each location variable y_i the promising assignment variables x_{ij} are selected and they become a part of the initial kernel or a bucket, respectively. A variable x_{ij} is promising if its reduced cost in the optimal solution to the LP relaxation is below a pre-specified threshold γ. In (Guastaroba and Speranza, 2014) the authors recommend to set γ to the median of the reduced costs of the variables x_{ij}. However, in large-scale instances of the CPMP with thousands of customers to find the median would be extremely memory and time demanding task. That is why we set γ to a fraction of the maximal reduced cost. The precise fraction will again be specified with the computational experiments. Before solving a restricted CPMP(K) or CPMP($K \cup B_k$) the feasibility of the problem is checked with regard to constraints (2). If there is a customer j that is not linked to at least one candidate location in K or $K \cup B_k$ respectively, then the variables x_{ij} are added connecting the customer j with all candidate locations in K or $K \cup B_k$ respectively.

In the final step the kernel is updated so that new promising location variables are added and those variables that are nor perspective any more are deleted. Promising variables are those variables y_i from the current bucket B_k that take the value 1 in the optimal solution to the restricted CPMP($K \cup B_k$). The set including these variables is denoted as B_k^+. If the value of a variable y_i is 0 in the optimal solution to the CPMP($K \cup B_k$) and also in r restricted problems solved since it has been added to the kernel, where r is a given parameter, then that variable is assumed to be no longer promising and is removed from the kernel. Let B_k^- be the set including these variables. Thus at the end of iteration k the kernel is updated so that variables in B_k^+ are added and variables in B_k^- are removed form it. The associated variables x_{ij} are added or removed as well.

The general kernel search supposes that two new constraints are added to each consecutive restricted MILP solved. One constraint sets a cut-off value to the objective function and the other constrains at least one binary variable from the current bucket to be set to 1 (Guastaroba et al., 2017). However, our experiments with large-scale CPMP instances revealed that the former constraint permits the solver to find any feasible integer solution in an acceptable time and the latter does not improve the final solution. That is why we omit these additional constraints and allow the solver to finish with a feasible integer solution that may be worse than the best solution found so far. As a consequence, the search is more diversified in several first iterations and the kernel is broader giving the algorithm a chance to find better solution in final iterations.

4 COMPUTATIONAL EXPERIMENTS

The computational experiments were performed on a personal computer equipped with the Intel Core i7 processor with 1.60 GHz and 8 GB RAM. The application was implemented in FICO Xpress-IVE using Mosel language. Mosel is a programming and modelling language that enables easily to modify the model. We appreciated its features in the improving phase of the kernel search where it is necessary to add or delete some variables and to re-define the constraints so that they can include only existing variables. The solver Xpress Optimizer 8.3 (64-bit, release 2017) was used to solve the CPMP models.

We tested the kernel search algorithm on two sets of benchmark instances. The first set was proposed by Lorena and Senne (2004). It consists of five large-scale instances named p3038_600 to p3038_1000 containing 3038 nodes. The number of medians ranges from 600 to 1000. The second set comprises four Spain instances by Díaz and Fernández (2006). Both sets are available on the site http://www-usr.inf.ufsm.br/~stefanello/instances/cpmp/. The sets of candidate locations and customers are identical in all instances, i.e. every customer can be a median.

The instances in the first benchmark set have thousands of candidates and customers. In such a case the number of assignment variables $\boldsymbol{x}$ in the mathematical programming model is huge. In order to decrease computational complexity and speed up the solver, we propose a reduction of the solution space. The model is reduced by heuristic elimination of those variables $\boldsymbol{x}$ which are less likely to belong to a good or optimal solution (see also Stefanello et al., 2015). The elimination is based on the assumption that customers will not be served by those facilities that are too far away. That is why only those variables $\boldsymbol{x}$ are included in the model for which coefficient d_{ij} is less than a predefined threshold. The threshold is defined by the value $\alpha \cdot d^{max}/\sqrt{p}$, where α is a parameter (α = 1.5 in our experiments) and $d^{max} = max\{d_{ij}: i \in I, j \in J\}$.

The setting of the kernel search parameters for these two sets of instances is summarized in Table 1. The values of parameter l was chosen so that the number of variables after presolve does not exceed 20,000 in any restricted problem. In such a case the solver was able to find at least one integer solution in several minutes and to improve it in the specified time limit. In order to reduce the total running time we set a time limit equal to 600 seconds for the solution of each restricted problem. Moreover, the solver was forced to stop with a sufficiently good solution. This was achieved by setting the relative gap parameter. The relative gap starts with the value 0.005 and is reduced by 0.0004 in each iteration. The improvement loop of the kernel search algorithm terminates when all buckets have been explored, it means that $\overline{N_b} = N_b$.

Table 1. Parameter setting for test instances

Parameter	p3038 instances (Lorena and Senne, 2004)	Spain instances (Díaz and Fernández, 2006)
l	200	50
γ	0.2	0.1
r	3	3

Table 2. Computational results for the 3038-node benchmark set

Instance	$\lvert I \rvert$	p	Best solution	Kernel search	Gap (%)	CPU (second)
p3038_600	3038	600	122711.17	123279.28	0.46	3191.86
p3038_700	3038	700	109677.30	110345.12	0.61	3241.64
p3038_800	3038	800	100064.94	100744.14	0.68	3581.53
p3038_900	3038	900	92310.09	93099.41	0.86	2797.66
p3038_1000	3038	1000	85854.05	86406.51	0.64	3048.32

Table 3. Computational results for the Spain benchmark set

Instance	$\lvert I \rvert$	p	Best solution	Kernel search	Gap (%)	CPU (second)
spain737_74_1	737	74	8845	8890	0.51	4111.39
spain737_74_2	737	74	8870	8880	0.11	2849.70
spain737_148_1	737	148	5901	5926	0.42	1573.75
spain737_148_2	737	148	5912	5931	0.32	1782.47

Table 2 presents the results of the kernel search algorithm for the 3038-node benchmark set. The first column gives the instance name followed by the instance parameters (the number of customers and the number of medians) and the best solution found so far. The optimal solutions are unknown so the table reports the best known solutions achieved by the IRMA algorithm (Stefanello et al., 2015).The other columns give the solution of the KS, the relative gap between the KS and IRMA solutions, and the computation time of the KS. Table 3 presents analogous results for the Spain data set. The Spain instances with 148 medians can be solved by a general-purpose solver to optimality in a couple of hours. So the best solution for these two instances in Table 3 is the optimal solution.

We can observe that the kernel search achieves very good results in an acceptable amount of time. All solutions are within 1% of the best known solution. The computation time does not exceed 1 hour (with the exception of the spain737_74_1 instance). It should be noted that we did not pay attention to solver tuning. Also the parameters of the kernel search algorithm were set according to preliminary experiments but without a serious statistical evaluation. It may happen that with a more careful algorithm tuning the results will be better in terms of the quality of the solution and the computation time as well.

5 CONCLUSIONS

The paper describes the implementation of the recently developed matheuristic kernel search to the capacitated p-median problem. Compared with the state-of-the-art decomposition heuristic it achieved slightly worse results. However, the great advantage of the kernel search is that it is a general concept that is invariable with respect to changes in the model. It can be applied without modifications in more practical problems that differ from the basic formulation (1)-(6) in such a way that the decomposition heuristic like IRMA is not able to cope with. For example, one may require that the location of medians does not differ from the reference solution too much. Another alternative formulation might relax integrality constraints on variables $\boldsymbol{x}$ allowing a customer to be served from multiple facilities. Such modification of the basic model is relevant in the case of the emergency medical system, where patients living in one municipality are served by ambulances from multiple stations.

Application of the kernel search for such practical problems will be the topic of our future research.

Acknowledgements: This work was supported by the research grants VEGA 1/0463/16 „Economically efficient charging infrastructure deployment for electric vehicles in smart cities and communities", VEGA 1/0342/18 "Optimal dimensioning of service systems", and APVV-15-0179 "Reliability of emergency systems on infrastructure with uncertain functionality of critical elements".

References

[1] Correa, E. A., Steiner, M. T. A., Freitas, A. A. and Carnieri, C. (2004). A genetic algorithm for solving a capacitated p-median problem. *Numerical Algorithms*, 35(2-4), pp. 373-388.

[2] Díaz, J. A. and Fernández, E. (2006). Hybrid scatter search and path relinking for the capacitated p-median problem. *European Journal of Operational Research*, 169(2), pp. 570-585.

[3] Guastaroba, G. and Speranza, M. G. (2014). A heuristic for BILP problems: The Single Source Capacitated Facility Location Problem. *European Journal of Operational Research*, 238(2), pp. 438-450.

[4] Guastaroba, G., Savelsbergh, M. and Speranza, M. G. (2017). Adaptive Kernel Search: A heuristic for solving Mixed Integer linear Programs. *European Journal of Operational Research*, 263(3), pp. 789-804.

[5] Jánošíková, Ľ. and Vasilovský, P. (2017). Grouping genetic algorithm for the capacitated p-median problem. In: *Proceedings of the International Conference on Information and Digital Technologies 2017*, Žilina, Slovakia: IEEE, 5 – 7 July 2017, pp. 152-159.

[6] Lorena, L. A. N. and Senne, E. L. F. (2004). A column generation approach to capacitated *p*-median problems. *Computers & Operations Research*, 31(6), pp. 863-876.

[7] ReVelle, C. S. and Eiselt, H. A. (2005). Location analysis: A synthesis and survey. *European Journal of Operational Research*, 165(1), pp. 1-19.

[8] Scheuerer, S. and Wendolsky, R. (2006). A scatter search heuristic for the capacitated clustering problem. *European Journal of Operational Research*, 169(2), pp. 533-547.

[9] Stefanello, F., Araújo, O. and Müller, F. (2015) Matheuristics for the capacitated p-median problem. *International Transactions in Operational Research*, 22(1), pp. 149-167.

[10] Yaghini, M., Lessan, J. and Gholami Mazinan, H. (2010). An efficient hybrid metaheuristic for capacitated p-median problem. *International Journal of Industrial Engineering & Production Research*, 21(1), pp. 11-15.

[11] Yaghini, M., Karimi, M. and Rahbar, M. (2013). A hybrid metaheuristic approach for the capacitated p-median problem. *Applied Soft Computing*, 13(9), pp. 3922-3930.

Author's address

prof. Ing. Ľudmila Jánošíková, PhD.
University of Žilina, Faculty of Management Science and Informatics, Department of Mathematical Methods and Operations Research
Univerzitná 1, 010 26 Žilina
Slovak Republic
e-mail: Ludmila.Janosikova@fri.uniza.sk

SECOND ORDER STOCHASTIC DOMINANCE CONSTRAINTS IN MULTI–OBJECTIVE STOCHASTIC PROGRAMMING PROBLEMS

Vlasta Kaňková,
Institute of Information Theory and Automation of the AS CR

ABSTRACT Many economic and financial applications lead to deterministic optimization problems depending on a probability measure. These problems can be either static (one stage) or dynamic with finite (multistage) or infinite horizon, single– objective or multi–objective. Constraints sets can be "deterministic", given by probability constraints or stochastic dominance constraints. We focus on multi–objective problems and second order stochastic dominance constraints. To this end we employ the former results obtained for stochastic (mostly strongly) convex multi–objective problems and results obtained for one objective problems with second order stochastic dominance constraints. The relaxation approach will be included in the case of second order stochastic dominance constraints.

Keywords Stochastic multi–objective optimization problems, efficient solution, Wasserstein metric and $\mathcal{L}_1$ norm, Lipschitz property, second order stochastic dominance constraints, relaxation
AMS classification: 90 C 15
JEL classification: C 44

1 Introduction

Let $(\Omega, \mathcal{S}, P)$ be a probability space, $\xi := \xi(\omega) = (\xi_1(\omega), \ldots, \xi_s(\omega))$ an s–dimensional random vector defined on $(\Omega, \mathcal{S}, P)$, $F := F_\xi$ the distribution function of ξ, P_F, and Z_F the probability measure and the support corresponding to F, respectively; E_F denote the operator of mathematical expectation corresponding to F. Let, moreover, $g_i := g_i(x, z)$, $i = 1, \ldots, l$, $l \geq 1$ be real–valued (say, continuous) functions defined on $\mathbb{R}^n \times \mathbb{R}^s$; $X_F \subset X \subset \mathbb{R}^n$ be a nonempty set generally depending on F, and $X \subset \mathbb{R}^n$ be a nonempty deterministic set. If for every $x \in X$ there exist finite $\mathsf{E}_F g_i(x, \xi)$, $i = 1, \ldots, l$, then a rather general (often employed) type of "multi–objective" one–stage optimization problem depending on a probability measure can be introduced in the form:

$$\text{Find} \quad \min \mathsf{E}_F g_i(x, \xi), \ i = 1, \ldots, l \quad \text{subject to} \quad x \in X_F. \tag{1}$$

There are known (from the literature) mainly X_F: "deterministic", given by probability constrains, by mathematical expectation and recently often appear stochastic dominance constraints. To define second order stochastic dominance constraints let $g : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^1$ be a real-valued function, $Y : \mathbb{R}^s \to \mathbb{R}^1$ random value. If for $x \in X$ there exists finite $\mathsf{E}_F g(x, \xi)$, $\mathsf{E}_F Y(\xi)$ and if

$$F^2_{g(x,\,\xi)}(u) = \int\limits_{-\infty}^{u} F_{g(x,\,\xi)}(y)dy, \quad F^2_Y(u) = \int\limits_{-\infty}^{u} F_Y(y)dy, \quad u \in \mathbb{R}^1,$$

then we can define the second order stochastic dominance constraints X_F by

$$X_F \ = \{x \in X : F^2_{g(x,\,\xi)}(u) \leq F^2_Y(u) \quad \text{for every} \quad u \in \mathbb{R}^1\}. \tag{2}$$

Consequently multi–objective stochastic programming problems with second order stochastic dominance constraints can be defined by the relations (1), (2). Employing the results of [9](see also Lemma 1), the multi–objective stochastic programming problem with second order stochastic dominance constraints can be written in a more friendly form:

$$\text{Find} \quad \min \mathsf{E}_F g_i(x, \xi), \ i = 1, \ldots, l \quad \text{subject to} \quad x \in X_F, \tag{3}$$

where

$$X_F = \{x \in X : \mathsf{E}_F(u - g(x, \xi))^+ \le \mathsf{E}_F(u - Y(\xi))^+ \quad \text{for every} \quad u \in \mathbb{R}^1\}. \tag{4}$$

Remark 1. Second order stochastic dominance corresponds to an order in the space of non negative concave utility functions.

The paper [6] is focus on the investigation of stability (obtained on the base of Wasserstein metric) and empirical estimates for the multi–objective stochastic problems. However replacing there general convex X_F by second order stochastic dominance constraints (4) we obtain an infentisimal optimization problem for which the Slater's condition is not generally fulfilled. The aim of this contribution is to relax constraints set to obtain problems for which Slater's condition is already fulfilled and to estimate the error of approximation. To this end the stability based on the Wasserstein metric is employed.

The stochastic multi–objective problem defined by (3),(4) is a deterministic multi–objective problem depending on the probability measure; consequently to analyze this problem it is possible to employ classical well known results (see, e.g., [2], [6] and [7]).

2 Some Definitions and Auxiliary Assertions

2.1 Deterministic Problems

First, we recall some results from the deterministic multi–objective optimization problems. To this end let $f_i(x)$, $i = 1, \ldots, l$ be real–valued functions defined on $\mathbb{R}^n$, $\mathcal{K} \subset \mathbb{R}^n$ be a nonempty set. We consider a multi–objective deterministic optimization problem in the form:

$$\text{Find} \quad \min f_i(x),\ i = 1, \ldots, l \quad \text{subject to} \quad x \in \mathcal{K}. \tag{5}$$

Definition 1. The vector x^* is an efficient solution of the problem (5) if and only if there exists no $x \in \mathcal{K}$ such that $f_i(x) \le f_i(x^*)$ for $i = 1, \ldots, l$ and such that for at least one i_0 one has $f_{i_0}(x) < f_{i_0}(x^*)$.

Definition 2. The vector x^* is a properly efficient solution of the multi–objective optimization problem (5) if and only if it is efficient and if there exists a scalar $M > 0$ such that for each i and each $x \in \mathcal{K}$ satisfying $f_i(x) < f_i(x^*)$ there exists at least one j such that $f_j(x^*) < f_j(x)$ and

$$\frac{f_i(x^*) - f_i(x)}{f_j(x) - f_j(x^*)} \le M. \tag{6}$$

Proposition 1. [4] Let $\mathcal{K} \subset R^n$ be a nonempty convex set and let $f_i(x)$, $i = 1, \ldots, l$ be convex functions on $\mathcal{K}$. Then $x^0 \in \mathcal{K}$ is a properly efficient solution of the problem (5) if and only if x^0 is optimal in

$$\min_{x \in \mathcal{K}} \sum_{i=1}^{l} \lambda_i f_i(x) \quad \text{for some} \quad \lambda_1, \ldots, \lambda_l > 0, \quad \sum_{i=1}^{l} \lambda_i = 1.$$

A relationship between efficient and properly efficient points is introduced, e.g., in [3] or [4]. We summarize it in the following Remark.

Remark 2. Let $f(x) = (f_1(x), \ldots, f_l(x))$, $x \in \mathcal{K}$; $\mathcal{K}^{eff}$, $\mathcal{K}^{peff}$ be sets of efficient and properly efficient points of the problem (5). If $\mathcal{K}$ is a convex set, $f_i(x)$, $i = 1, \ldots, l$ are convex functions on $\mathcal{K}$, then

$$f(\mathcal{K}^{peff}) \subset f(\mathcal{K}^{eff}) \subset \bar{f}(\mathcal{K}^{peff}), \tag{7}$$

where $\bar{f}(\mathcal{K}^{peff})$ denotes the closure set of $f(\mathcal{K}^{peff})$.

2.2 Wasserstein Metric in Stochastic Optimization Problems

To recall the Wasserstein metric and its application to single–objective stochastic optimization problem we consider the case $l = 1$. To this end let $\mathcal{P}(\mathbb{R}^s)$ denote the set of all (Borel) probability measures on $\mathbb{R}^s$ and let the system $\mathcal{M}_1^1(\mathbb{R}^s)$ be defined by the relation:

$$\mathcal{M}_1^1(\mathbb{R}^s) := \left\{ \nu \in \mathcal{P}(\mathbb{R}^s) : \int_{\mathbb{R}^s} \|z\|_1 d\nu(z) < \infty \right\}, \quad \| \cdot \|_1^s := \| \cdot \|_1 \text{ denotes } \mathcal{L}_1 \text{ norm in } \mathbb{R}^s. \quad (8)$$

If the assumption A.0, A.1 are defined by

A.0 $g_1(x, z)$ is for $x \in X$ a Lipschitz function of $z \in \mathbb{R}^s$ with the Lipschitz constant L (corresponding to the $\mathcal{L}_1$ norm) not depending on x,

A.1 $g_1(x, z)$ is either a uniformly continuous function on $X \times \mathbb{R}^s$ or there exists $\varepsilon > 0$ such that $g_1(x,\, z)$ is a convex bounded function on $X(\varepsilon)$ ($X(\varepsilon)$ denotes the ε–neighborhood of the set$X$),

and if $P_F,\, P_G \in \mathcal{M}_1^1(\mathbb{R}^s)$; $F_i,\, G_i,\, i = 1,\, \ldots,\, s$ denote one-dimensional marginal distribution functions corresponding to F and G, then

Proposition 2. [5] Let $P_F, P_G \in \mathcal{M}_1^1(\mathbb{R}^s)$. If Assumption A.0 is fulfilled, then

$$|\mathsf{E}_F g_1(x,\, \xi) - \mathsf{E}_G g_1(x,\, \xi)| \le L \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| dz_i \quad \text{for} \quad x \in X. \quad (9)$$

If, moreover, X is a compact set and Assumption A.1 is fulfilled, then also

$$|\inf_{x \in X} \mathsf{E}_F g_1(x,\, \xi) - \inf_{x \in X} \mathsf{E}_G g_1(x,\, \xi)| \le L \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| dz_i. \quad (10)$$

To study the constraints set defined by (2) we recall the next Lemma.

Lemma 1. [7] Let $g(x, z)$, $Y(z)$ be for every $x \in X$ Lipschitz functions of $z \in \mathbb{R}^s$ with the Lipschitz constant L_g not depending on $x \in X$. Let, moreover, $P_F \in \mathcal{M}_1^1(\mathbb{R}^s)$. If X_F is defined by the relation (2), then

1. $X_F = \{x \in X : \mathsf{E}_F(u - g(x, \xi))^+ \le \mathsf{E}_F(u - Y(\xi))^+ \quad \text{for every} \quad u \in \mathbb{R}^1\}$,
2. $(u - g(x, z))^+,\ (u - Y(z))^+, \quad u \in \mathbb{R}^1, x \in \mathbb{R}^n$ are Lipschitz functions of $z \in \mathbb{R}^s$ with the Lipschitz constant L_g not depending on $u \in \mathbb{R}^1$, $x \in \mathbb{R}^n$. (See before employed the relation (4.))

If the assumptions of Lemma 1 are fulfilled, $P_F,\, P_G \in \mathcal{M}_1^1(\mathbb{R}^s)$, $u \in \mathbb{R}^1$, $x \in X$, then it follows from Proposition 2 that

$$\begin{aligned} |\mathsf{E}_F(u - g(x, \xi))^+ - \mathsf{E}_G(u - g(x, \xi))^+| &\le L_g \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| dz_i, \\ |\mathsf{E}_F(u - Y(\xi))^+ - \mathsf{E}_G(u - Y(\xi))^+| &\le L_g \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| dz_i. \end{aligned} \quad (11)$$

Further defining the sets X^ε by

$$X_F^\varepsilon = \{x \in X : \mathsf{E}_F(u - g(x,\, \xi))^+ - \mathsf{E}_F(u - Y(\xi))^+ \le \varepsilon \quad \text{for every} \quad u \in \mathbb{R}^1\}, \quad \varepsilon \in \mathbb{R}^1, \quad (12)$$

we can obtain

$$x \in X_F \Longrightarrow x \in X_G^{\varepsilon}, \quad x \in X_G \Longrightarrow x \in X_F^{\varepsilon} \quad \text{with} \quad \varepsilon = 2L_g \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - G_i(z_i)| dz_i,$$

$$\text{and generally} \qquad X_G^{\delta-\varepsilon} \subset X_F^{\delta} \subset X_G^{\delta+\varepsilon} \quad \text{for} \quad \delta \in \mathbb{R}^1. \tag{13}$$

2.3 Relaxation

Till now we have considered stochastic multi–objective problems with constraints set (4). According to the well known fact from the infitisimal programming theory, there can be problem with Slater's condition. Consequently to this fact Dentcheva and Ruszczynski [1] suggested to relax one objective problems by the modification of the constraints set; replacing X_F by $X_F^{a,\,b}$:

$$X_F^{a,\,b} = \{x \in X : \mathsf{E}_F(u - g(x, \xi))^+ \le \mathsf{E}_F(u - Y(\xi))^+ \quad \text{for every} \quad u \in \langle a,\, b\rangle\},\ a,\, b\, \in \mathbb{R}^1. \tag{14}$$

However they did not specified how to choice a, b. Surely, it is desirable to determine a, b to be a difference between X_F and $X_F^{a,b}$ small. More precisely, it is desirable to be small difference between the corresponding optimal solutions and optimal values. On the other side they have proven the following assertion.

Proposition 3. [1]. Let $\bar{Y} := \bar{Y}(\xi)$ be a random value defined on $(\Omega, \mathcal{S}, P)$. Let, moreover, $Y(\xi)$ has a discrete distribution with realizations y_i, $i = 1, \ldots, m$, where $a \le y_i \le b$, $a,\, b \in \mathbb{R}^1$ for all i. Then the inequality

$$\mathsf{E}_{F_{\bar{Y}}}(u - \bar{Y}(\xi))^+ \le \mathsf{E}_{F_Y}(u - Y(\xi))^+ \quad \text{for all} \quad u \in \langle a,\, b\rangle$$

is equivalent to

$$\mathsf{E}_{F_{\bar{Y}}}(y_i - \bar{Y})^+ \le \mathsf{E}_{F_Y}(y_i - Y)^+, \quad i = 1, \ldots, m.$$

If $Y(\xi)$ is a random value defined on $(\Omega, \mathcal{S}, P)$ $(\xi = \xi(\omega))$ and if $a < b,\ a,\, b \in \mathbb{R}^1$, then we can define a random value $Y^{a,\,b} := Y^{a,\,b}(\xi)$ by

$$\begin{aligned} Y^{a,\,b}(\xi) = \quad & Y(a) \quad \text{if} \quad \xi \le a, \\ & Y(\xi) \quad \text{if} \quad \xi \in (a,\, b), \\ & Y(b) \quad \text{if} \quad \xi \ge b \end{aligned} \tag{15}$$

and to note the corresponding distribution function by $F_Y^{a,\,b}$.

Employing (15), under rather generalize assumptions, we can for constants $a,\, b,\, a < b; a_1\, b_1,\, a_1 < b_1$ define

i. the random value $Y^{a,\,b}(\xi)$ $(a \le Y^{a,\,b}(\xi) \le b)$ with the distribution function $F_Y^{a,\,b}$,

ii. for every $x \in X$ the random value $g^{a,\,b}(x,\,\xi)$ $(a \le g^{a,\,b}(x,\,\xi) \le b)$ with the distribution function $F_{g(x,\,\xi)}^{a,\,b}$.

Evidently, $Y(\xi)$, $g(x,\,\xi)$, for every $x \in X$, are functions of the random vector ξ and, simultaneously, they are functions of the components $\xi_1, \ldots, \xi_s$ of the random vector ξ. Consequently, under rather general assumptions it is possible to choose $a,\, b,\ a_1,\, b_1$ such that

iii. $a_1 < b_1,\, i = 1, \ldots, s \Longrightarrow a < Y(\xi) < b, \quad a < g(x, \xi) < b \quad \text{for every} \quad x \in X.$

The constants a_1, b_1 determine a distribution function $F^{a_1, b_1} := F_{\xi}^{a_1, b_1}$ with a support $\prod_{i=1}^{s} \langle a_1, b_1 \rangle$.

iv there exists (for a natural number m) points $y_1, \ldots, y_m \in \prod_{i=1}^{s} \langle a_1, b_1 \rangle$ those define discrete distribution function $\bar{F}^{a, b} := \bar{F}_Y^{a, b}$ with atoms $y_1., \ldots, y_m$ ($\bar{F}_Y^{a, b}$ approximates $F_Y^{a, b}$).

According to the above recalled assertions we can see that now it is possible to define two constraints sets $X_F^{a, b} := X_F^{a_1, b_1}$ and $\bar{X}_F^{a, b} = \bar{X}^{a_1, b_1}$ by

$$X_{F^{a,b}}^{a, b} = (X_F^{a_1, b_1}) = \{x \in X : \mathsf{E}_{F^{a_1, b_1}}(u - g(x, \xi))^+ \leq \mathsf{E}_{F^{a_1, b_1}}(u - Y(\xi))^+ \quad \text{for every} \quad u \in \langle a, b \rangle\}, \tag{16}$$

$$\bar{X}_{\bar{F}}^{a_1, b_1} = \{x \in X : \mathsf{E}_{F^{a_1, b_1}}(y_i - g(x, \xi))^+ \leq \mathsf{E}_{\bar{F}^{a, b}}(y_i - Y(\xi))^+ \quad \text{for every} \quad i = i, \ldots, m\}, \tag{17}$$

and to define optimization problems

$$\text{To find} \quad \varphi^{a_1, b_1}(F^{a_1, b_1}, X_{F^{a_1, b_1}}^{a_1, b_1}) = \inf\{\mathsf{E}_{F^{a_1, b_1}} g_0(x, \xi) | x \in X_F^{a_1, b_1}\}, \tag{18}$$

$$\text{To find} \quad \bar{\varphi}^{a_1, b_1}(F^{a_1, b_1}\, \bar{X}_{\bar{F}}^{a_1, b_1}) = \inf\{\mathsf{E}_{F^{a_1, b_1}} g_0(x, \xi) | x \in \bar{X}_{\bar{F}}^{a_1, b_1})\}. \tag{19}$$

It has been proven in [8] that the optimization problems (18), (19) already fulfil the Slater's condition. The next assertion follows also from [8].

Proposition 4. Let X_F, $X_F^{a_1, b_1}$, $\bar{X}_F^{a_1, b_1}$ be compact sets, $P_F \in \mathcal{M}_1^1(\mathbb{R}^s)$, Assumptions A.1, i., ii., iii., iv be fulfilled. If

1. $g(x, z)$ is for every $z \in Z_F$ a Lipschitz function of $x \in X$ with the Lipschitz constant $\bar{L}$ not depending on $z \in Z_F$,
2. there exists a constant $D > 0$ such that

$$\Delta[X_F^{\varepsilon'}, X_F^{\varepsilon''}] \leq D\varepsilon \quad \text{for every} \quad \varepsilon', \varepsilon'' \in \langle -3\varepsilon, 3\varepsilon \rangle,$$

with $\varepsilon = 2L_g \max[\sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i^{a_1, b_1}(z_i) - \bar{F}_i^{a_1, b_1}(z_i)| dz_i,\ \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - F_i^{a_1, b_1}(z_i)| dz_i]$,

then

1. $$\begin{aligned} |\varphi(F, X_F) - \varphi(F, X_F^{a_1, b_1})| &\leq 2D\bar{L}L_g \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - F_i^{a_1, b_1}(z_i)| dz_i, \\ |\varphi(F, X_F) - \varphi(F, \bar{X}_{\bar{F}}^{a_1, b_1})| &\leq 2D\bar{L}L_g \sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - \bar{F}_i^{a_1, b_1}(z_i)| dz_i, \end{aligned}$$

2. If, moreover, Assumptions A.0 is fulfilled, then

$$\begin{aligned} &|\varphi(F, X_F) - \varphi^{a_1, b_1}(F^{a_1, b_1}, \bar{X}_{\bar{F}}^{a_1, b_1})| \leq (2L + 2D\bar{L}L_g)[\sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i(z_i) - \bar{F}_i^{a_1, b_1}(z_i)| dz_i, \\ &|\varphi(F^{a_1, b_1}, X_F^{a_1, b_1}) - \bar{\varphi}^{a_1, b_1}(\bar{F}^{a_1, b_1}, \bar{X}_{\bar{F}}^{a_1, b_1})| \leq \\ &\qquad (2L + 2D\bar{L}L_g)[\sum_{i=1}^{s} \int_{-\infty}^{+\infty} |F_i^{a_1, b_1}(z_i) - \bar{F}_i^{a_1, b_1}(z_i)| dz_i. \end{aligned} \tag{20}$$

($\Delta[\cdot, \cdot] = \Delta_n[\cdot, \cdot]$ denotes the Hausdorff distance in the space of subsets of n-dimensional Euclidean space; for more details see, e.g., [10].)

Remark 3. Evidently, it is reasonable to choose a_1, b_1, m to be "small" the values

$$|\varphi(F, X_F) - \varphi(F, X_F^{a_1, b_1})|, \quad |\varphi(F, X_F) - \varphi(F, \bar{X}_F^{a_1, b_1})|.$$

To this end the relation (11) (with $G := F_\xi^{a_1,\, b_1}$), $G := \bar{F}_Y^{a,\, b}$ and Proposition 4 can be employed. The constant a, b, has to be chosen with respect to the function g. (The assumption i. ii. iii. iv. has to be fulfilled.)

3 Application to Multi-Objective Stochastic Programming Problems

First, we recall one very simple assertion.

Lemma 2. Let $X \subset \mathbb{R}^n$ be a compact convex set, $a, b \in \mathbb{R}^1$, $a < b$. Let, moreover, $g(x, z)$ be for every $z \in \mathbb{R}^s$ a concave function of $x \in X$, then

$$\begin{aligned}
X_F &= \{x \in X : \mathsf{E}_F(u - g(x, \xi))^+ \le \mathsf{E}_F(u - Y(\xi))^+ \quad \text{for every} \quad u \in \mathbb{R}^1\},\\
X_F^{a,\, b} &= \{x \in X : \mathsf{E}_{F^{a.,\, b.}}(u - g(x, \xi))^+ \le \mathsf{E}_{F^{a_1,\, b_1}}(u - Y(\xi))^+ \quad \text{for every} \quad u \in \langle a, b\rangle\},\\
\bar{X}_{\bar{F}}^{a,\, b} &= \{x \in X : \mathsf{E}_{F^{a_1,\, b_1}}(y_i - g(x, \xi))^+ \le \mathsf{E}_{\bar{F}^{a_1,\, b_1}}(y_i - Y(\xi))^+ \quad \text{for every} \quad i = 1, \ldots, m\},
\end{aligned}$$

are convex sets.

Proof. The assertion follows immediately from the properties of convex functions and convex sets.

Further setting successively

$$\begin{aligned}
f_i(x) &= \mathsf{E}_F g_i(x, \xi),\ i = 1, \ldots, l & \mathcal{K} &= X_F,\\
f_i(x) &= \mathsf{E}_{F^{a_1,\, b_1}} g_i(x, \xi),\ i = 1, \ldots, l, & \mathcal{K} &= X_F^{a,\, b},\\
f_i(x) &= \mathsf{E}_{\bar{F}^{a_1,\, b_1}} g_i(x, \xi),\ i = 1, \ldots, l, & \mathcal{K} &= \bar{X}_F^{a,\, b},
\end{aligned} \tag{21}$$

we obtain three "deterministic" (depending on the probability measures) multi–objective problems. If the "original" functions $g_i(x, z)$, $i = 1, \ldots, l$ are for every $z \in \mathbb{R}^s$ convex functions, $g(x, z)$ for every $z \in \mathbb{R}^s$ concave function, then the corresponding problems (5) are convex multi–objective problems. According to Proposition 1 we can obtain the sets of properly efficient points and further according to Remark 2 these sets approximate the corresponding sets of efficient points. Moreover the Slater's condition is fulfilled for the second and the third optimization problem in (21). Of course, the second and the third problem "approximate" the original one. Employing the results of Subsections 2.2, 2.3 we can estimate the errors of approximations.

Furthermore, employing the results of the paper [6] we can study the stability of these problems and their empirical estimates. It means, the case when F is replaced by empirical distribution function or the case when $F^{a,\, b}$, $\bar{F}$ are determined by random sample. However the corresponding results have been obtained under the assumptions of strongly convex $g_i(x, z)$, $i = 1, \ldots, l$. (For the definition of strongly convex functions see, e.g., [6]). The investigation in this direction is surely very interesting and important but it is beyond of the scope of this paper. Moreover it can be obtained on the base of the paper [6].

4 Conclusion

The contribution deals with multi–objective stochastic programming problems with second order stochastic dominance constraints. In particular, the aim of the paper is to show a possibility to generalize the paper [6] with rather general convex constraints set to the special case of second order stochastic dominance constraints. To this end a relaxation approach has been employed.

Acknowledgements: This work was supported by the Czech Science Foundation under Grant No. 18–02739S.

References

[1] Dentcheva, D. and Ruszczynski, A. (2003). Optimization with Stochastic Dominance Constraints. *SIAM Journal on Optimization,* 14(2), pp. 548-566.

[2] Dupačová, J., Hurt, J. and Štěpán, J. (2002). *Stochastic modeling in economic and finance.* Dordrecht: Kluwer Academic Publishers.

[3] Ehrgott, M. (2005). *Multicriteria optimization.* Berlin: Springer.

[4] Geoffrion, A. M. (1968). Proper Efficiency and the Theory of Vector Maximization. *J. Math. Anal. Appl.* 22(3), pp. 618–630.

[5] Kaňková, V. and Houda, M. (2006). Empirical Estimates in stochastic Programming. In:*Prague Stochastics 2006.* Prague: MATFYZPRESS, pp. 26–436.

[6] Kaňková, V. (2016). A Remark on Multiobjective Stochastic Optimization with Strongly Convex Functions. *CEJOR,* 24(2), pp. 309–333.

[7] Kaňková, V. (2017). Stability, Empirical Estimates and Scenario Generation in Stochastic Optimization–Applications in Finance. *Kybernetika,* 53(6), pp. 1026–1046.

[8] Kaňková, V. and Omelčenko, V. (2018). Stochastic Optimization Problems with Second Order Stochastic Dominance Constraints wia Wassertein Metric. *Kybernetika* (submitted).

[9] Ogryczak, W. and Ruszczynski, A. (1999). From Stochastic Dominance to Mean–Risk Models: Semideviations as Risk Measures. *European J. Oper. Res.,* 116, pp. 33–50.

[10] Rockafellar, R. and Wets, R. (1983) *Variational analysis.* Berlin: Springer.

Author's address
RNDr. Vlasta Kaňková, CSc.
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Department of Econometrics
Pod Vodárenskou věží 4, 182 08 Prague 8
Czech Republic
e–mail: kankova@utia.cas.cz

AN APPLICATION OF HYBRID FUZZY MULTI-CRITERIA DECISION-MAKING METHOD ON A LOCATION-SELECTION PROBLEM

Michal Koláček, Markéta Matulová, Masaryk University Brno

Abstract
In this paper we present a useful tool for multi-criteria decision making under uncertainty. The proposed approach involves several steps: identification of potential alternatives, determination of decision criteria, application of fuzzy theory to quantify criteria values under uncertainty and evaluation of alternatives with a hybrid method combining fuzzy TOPSIS and fuzzy AHP. We apply the proposed method in the case study of the construction of a decision making framework for the selection of the best location for organizing a conference.

Keywords: *AHP, fuzzy numbers, TOPSIS, multi-criteria decision-making under uncertainty.*

JEL Classification: C61, D81
AMS Classification: 62C86, 90B50

1 INTRODUCTION

Decision making is an integral part of everyone's life. The problems we face are often of a very complex nature and the alternatives are judged by numerous criteria at a time. Over the last decades, decision support systems have been extended to use mathematical tools to facilitate the decision-making process. If we desire to truly represent reality, it becomes necessary to also incorporate uncertainty. Thus the development of fuzzy logic made the fundamental contribution to mathematical decision models. It allows for working with information that is inaccurate and often specified by natural language. In the paper, we will describe one of the ways of extending mathematical models of decision making for the use under uncertainty and apply the presented method to a specific decision-making problem.

2 SELECTED METHODS OF MULTI-CRITERIA DECISION MAKING

For multi-criteria evaluation of alternatives, we can use various methods that differ in various aspects, e.g. in the way of constructing criteria functions and matrices, basis for ranking alternatives, computational demands, versatility with respect to the goal of the decision problem, etc.

Let us further consider the existence of k criteria functions $K_j,\ j = 1,\dots,k$. Their values in n alternatives $a_i, i = 1,\dots,n$, are denoted by y_{ij} and we call them criteria values. The criteria are typically arranged in a criteria matrix $\mathbf{Y} = (y_{ij})_{n \times k}$. Many MCDM methods require that the criteria are compared and their importance is specified. Most commonly, the so-called relative significance of the criterion K_j, denoted by $v_j, j = 1,\dots,k$, is used to describe the relevance of the criterion to the decision problem. In order to ensure comparability across different methods, these relative measures of criteria are normalized and the resulting weights of the criteria $w_j, j = 1,\dots,k$, are obtained. Let us describe methods that use the weights of the criteria.

2.1 Method TOPSIS

Method TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) was first introduced in the 1980s by Ching-Lai Hwang and Kwangsun Yoon [12] and is based on the evaluation of the alternatives in terms of their distance from extreme solutions. The procedure for using the TOPSIS method is given below (further we assume that all criteria are maximized and all the elements in the criterion matrix are positive).

2.1.1 Construction of normalized criteria matrix

First, we create a standardized matrix $\mathbf{R} = (r_{ij})_{n \times k}$ by normalization of the criteria matrix **Y**. This can be achieved by the original formula proposed by Hwang and Yoon [12] or, more simply, according to Garcia-Cascalez [4], where column maxima are used for normalization of the respective vectors. The matrix **R** is then calculated as follows: $r_{ij} = \frac{y_{ij}}{y_j^{max}}, i = 1,\ldots,n,$ $j = 1,\ldots,k,$ where y_j^{max} is determined for every criteria K_j, as $y_j^{max} = \max_{i = 1,\ldots,n} y_{ij}$, $j = 1,\ldots,k.$

2.1.2 Construction of weighted criteria matrix

Regardless of the chosen way of normalization, the next step of the TOPSIS method is creation of a weighted criteria matrix $\mathbf{Z} = (z_{ij})_{n \times k}$, using the matrix **R** and the weights of criteria $w_j, j = 1,\ldots,k,$ in a way that: $z_{ij} = r_{ij} \cdot w_j,\ i = 1,\ldots,n,\ j = 1,\ldots,k.$

2.1.3 Distance of alternatives from ideal and basal solution

The resulting matrix **Z** allows the identification of ideal solution $A^+ = (A_1^+,\ldots,A_k^+)$ and basal solution $A^- = (A_1^-,\ldots,A_k^-)$, where $A_j^+ = \max_{i = 1,\ldots,n} z_{ij}$, $A_j^- = \min_{i = 1,\ldots,n} z_{ij}$, for every $j = 1,\ldots,k,$ that is for all criteria. The next step is determining distance of the individual alternatives a_i, $i = 1,\ldots,n,$ from ideal and basal solution respectively:

$$d_i^+ = \sqrt{\sum_{j=1}^{k} (z_{ij} - A_j^+)^2}\ , \qquad d_i^- = \sqrt{\sum_{j=1}^{k} (z_{ij} - A_j^-)^2}\ , \qquad i = 1,\ldots,n. \tag{1}$$

2.1.4 Computation of similarity of alternatives to the ideal solution

The last computational step consists of the determination of similarity to the ideal solution which is defined as $c_i = \frac{d_i^-}{d_i^+ + d_i^-},\ i = 1,\ldots,n.$ It is clear that $c_i \in \langle 0,1 \rangle$, where the extreme value of 0 is achieved in the basal $A^-$and the extreme value of 1 in the ideal alternative A^+.

2.1.5 Selection of the compromise alternative

Alternatives $a_i, i = 1,\ldots,n,$ are ranked according to the values c_i. The alternative a_i with a maximal value of c_i is selected as the compromise alternative a^*. It is the one that is "as far from the basal alternative and, at the same time, as close to the ideal alternative as possible".

2.2 Analytic Hierarchy Process

Method AHP (Analytic Hierarchy Process) suggested by T. Saaty [10] breaks down the decision-making problem into a hierarchical system. The simplest three-level hierarchy puts the objective of the decision-making problem on the top level, criteria on the middle level and the alternatives on the bottom level. If there are more criteria, it is desirable to split the middle level into two parts: the specific criteria become the lower-level sub-criteria that are grouped into more general criteria put on the upper middle level.

The essence of the AHP method is to construct Saaty's matrices for pairwise comparisons of elements from the lower level of the hierarchy with respect to the element from the higher level of the hierarchy. Subsequently, the significance of individual elements is computed from these matrices. The originally proposed computation procedure was based on a dominant eigenvector of Saaty's matrix. However, a simplified procedure using geometric mean is often used in practice. This approach is applied in this paper as well as it is more suitable for fuzzification. The relative significance of i -th criterion v_i, and its weight w_i are computed from Saaty's matrix $\boldsymbol{S} = (s_{ij})_{k \times k}$ according to the formula

$$v_i = \sqrt[k]{\prod_{j=1}^{k} s_{ij}} \text{ , } w_i = \frac{v_i}{\sum_{p=1}^{k} v_p} \text{ , } i = 1, \ldots, k. \tag{2}$$

The overall assessment of the alternative, with respect to the decision goal, is calculated as the weighted average of its significance to individual sub-criteria (the weight of a sub-criterion is the product of its relevance to the higher criterion and the contribution of that higher criterion to the decision target). The compromise alternative a^* is the one with highest overall rating.

3 FUZZY EXTENSION OF MCDM METHODS

Until now, we have assumed that all the quantitative information entering the model is provided with certainty. However, uncertainty, which is present in real decision-making situations (mainly in terms of expressing preferences), greatly reduces credibility of the results. To overcome this issue, L. A. Zadeh introduced the so-called fuzzy set theory [13] in 1965. Fuzzy theory has become very popular and has been widely used, primarily because of its ability to work relatively simply with the imprecise information expressed by natural language (the adjective crisp is used for precise values in contrast with fuzzy). The core of the whole theory is the concept of a fuzzy set; defined below.

Definition: Let U be the set called universe of discourse. Then the fuzzy set $\tilde{A}$ on universe U is defined by a mapping $\mu_{\tilde{A}}: U \to \langle 0, 1 \rangle$, called membership function of the fuzzy set $\tilde{A}$. For every $x \in U$, the value $\mu_{\tilde{A}}(x)$ determines the grade of membership of x to the fuzzy set $\tilde{A}$. The set of all fuzzy sets on universe U is denoted by $\mathcal{F}(U)$.

Fuzzy numbers are defined as a special case of fuzzy sets on universe of real numbers. A very simple class of fuzzy numbers is represented by triangular fuzzy numbers. Every triangular fuzzy number can be written as $\widetilde{C} = (c_l, c_m, c_u)$, where its membership function is:

$$\mu_{\tilde{C}}(x) = \begin{cases} \dfrac{x - c_l}{c_m - c_l}, & x \in \langle c_l, c_m) \\ \dfrac{c_u - x}{c_u - c_m}, & x \in \langle c_m, c_u \rangle \\ 0, & \text{elsewhere} \end{cases} \tag{3}$$

Real numbers c_l, c_m, c_u represent the smallest likely value, the most probable value and the largest possible value of a corresponding fuzzy event respectively. Due to its simplicity, this class of fuzzy numbers is widely used. Specifically in multi-criteria models of decision making, positive triangular fuzzy numbers are used (i.e. numbers satisfying $c_l > 0$). They very well represent the language expressions of the decision maker or the expert without significantly increasing the complexity of the calculations. Basic arithmetic operations with fuzzy numbers can be introduced using standard fuzzy arithmetic based on Zadeh's extension principle. However, the set of triangular fuzzy numbers is not closed under multiplication and division, so approximations to the actual results are used (see e.g. [8]). We define for positive triangular fuzzy numbers $\widetilde{C} = (c_l, c_m, c_u)$, $\widetilde{D} = (d_l, d_m, d_u)$ following operations:

$$\begin{aligned}
\widetilde{C} + \widetilde{D} &= (c_l + d_l, c_m + d_m, c_u + d_u),\\
\widetilde{C} \cdot \widetilde{D} &= (c_l \cdot d_l, c_m \cdot d_m, c_u \cdot d_u),\\
\frac{\widetilde{C}}{\widetilde{D}} &= \left(\frac{c_l}{d_u}, \frac{c_m}{d_m}, \frac{c_u}{d_l}\right).
\end{aligned} \tag{4}$$

For the interpretation of results of fuzzy MCDM method a defuzzification procedure is needed to transform the fuzzy numbers into a specific decision or real value. Let us describe the method used in this paper. It is based on calculating the distance between triangular fuzzy numbers, namely the vertex method introduced by Chen-Tung Chen [1], which defines the distance between fuzzy numbers $\widetilde{C} = (c_l, c_m, c_u)$, $\widetilde{D} = (d_l, d_m, d_u)$ by the formula

$$d(\widetilde{C}, \widetilde{D}) = \sqrt{\frac{1}{3}[(c_l - d_l)^2 + (c_m - d_m)^2 + (c_u - d_u)^2]}\,. \tag{5}$$

3.1 Fuzzification of the TOPSIS method

There are multiple approaches to fuzzy extension of the TOPSIS method. We have chosen procedures described in [1]. The input data consists of fuzzy weights of criteria $\widetilde{w_j} = (w_{jl}, w_{jm}, w_{ju}), j = 1, \ldots, k$, and the fuzzy criteria matrix $\widetilde{\mathbf{Y}} = (\tilde{y}_{ij})_{n \times k}$. Its elements are triangular fuzzy numbers $\tilde{y}_{ij}$ with significant values $(y_{ijl}, y_{ijm}, y_{iju})$, which are generated by the decision-maker's (or expert's) evaluation of the alternatives with regard to the individual criteria.

3.1.1 Construction of normalized fuzzy criteria matrix

To create a normalized fuzzy matrix $\widetilde{\mathbf{R}} = (\tilde{r}_{ij})_{n \times k}$, the linear transformation is used for computational simplicity. Criteria values are transformed using numbers y_j^{max}, which are defined for each criterion as $y_j^{max} = \max_{i = 1,\ldots,n} y_{iju}$. We then obtain:

$$\tilde{r}_{ij} = \left(\frac{y_{ijl}}{y_j^{max}}, \frac{y_{ijm}}{y_j^{max}}, \frac{y_{iju}}{y_j^{max}}\right), \qquad i = 1, \ldots, n, \qquad j = 1, \ldots, k. \tag{6}$$

3.1.2 Construction of weighted fuzzy criteria matrix

After normalization we create a weighted fuzzy criteria matrix $\widetilde{\mathbf{Z}} = (\tilde{z}_{ij})_{n \times k}$ using the formula $\tilde{z}_{ij} = \tilde{r}_{ij} \cdot \tilde{w}_j$, $i = 1, \ldots, n$, $j = 1, \ldots, k$. Then the elements of matrix $\widetilde{\mathbf{Z}}$ are normalized positive triangular fuzzy numbers $(z_{ijl}, z_{ijm}, z_{iju})$.

3.1.3 Fuzzy ideal and fuzzy basal solution and their distance from other alternatives

Next, it is necessary to determine the fuzzy ideal solution $\widetilde{A}^+ = (\widetilde{A}_1^+, \ldots, \widetilde{A}_k^+)$ and the fuzzy basal solution $\widetilde{A}^- = (\widetilde{A}_1^-, \ldots, \widetilde{A}_k^-)$. As all elements of matrix $\widetilde{\mathbf{Z}}$ have significant values within the range of $\langle 0,1 \rangle$, the determination of these extreme solutions can be simplified to the choice of $\widetilde{A}_j^+ = (1,1,1)$, $\widetilde{A}_j^- = (0,0,0)$, $j = 1, \ldots, k$, (see e.g. [10] and [11]). The distance of individual alternatives from the fuzzy ideal and from the fuzzy basal solution is then calculated from the weighted fuzzy criteria matrix $\widetilde{\mathbf{Z}}$ using (5) according to formulas

$$d_i^+ = \sum_{j=1}^{k} d(\tilde{z}_{ij}, \widetilde{A}_j^+)\,, d_i^- = \sum_{j=1}^{k} d(\tilde{z}_{ij}, \widetilde{A}_j^-)\,, \qquad i = 1, \ldots, n. \tag{7}$$

3.1.4 Computation of similarity to ideal solution

The next step is to determine, for each alternative, the index $c_i = \frac{d_i^-}{d_i^+ + d_i^-}$ expressing its relative distance from the fuzzy basal alternative $\widetilde{A}^-$.

3.1.5 Selection of compromise alternative

Finally, we conclude the procedure as in terms of certainty – we rank individual alternatives according to the corresponding values c_i so that we can select the one with the maximum value to be the compromise alternative a^*.

3.2 Fuzzification of Saaty's method for weights of criteria

Weights of criteria can be determined using fuzzy extension of analytical hierarchical process described in [6]. Pairwise comparisons of criteria $K_j,\ j = 1, \dots, k$, are used to construct a fuzzy matrix $\widetilde{\boldsymbol{S}} = (\tilde{s}_{ij})_{k\times k}$. Its entries express the ratio of the significance of the i-th criterion to the significance of the j-th criterion. The diagonal of the matrix consists of crisp values, $\tilde{s}_{ii} = 1,\ i = 1, \dots, k$. Non-diagonal elements $\tilde{s}_{ij}, i \neq j$, must satisfy the reciprocity condition $\tilde{s}_{ij} = 1/\tilde{s}_{ji}$, where the reciprocals are understood within the meaning of the definition (4). Different fuzzy scales can be prescribed for the elements $\tilde{s}_{ij}$; we used the scale given by the Table 1 (either $\tilde{s}_{ij}$ or $\tilde{s}_{ji}$ must belong to this scale for every $i \neq j$). Fuzzy version of formula (2) is used for the computation of weights where the conditional arithmetic described in [6] is applied instead of the standard fuzzy arithmetic.

4 APPLICATION OF FUZZY MCDM METHOD ON THE PROBLEM OF EVENT-LOCATION SELECTION

The theory described in the previous sections was applied to a specific decision-making problem of selecting a suitable space for the opening ceremony of Brno 2017 – 85th International Session of the European Youth Parliament. Alternatives were represented by four locations in Brno:

- Besední dům (a_1),
- Lužánky Park (a_2),
- Academic Cinema Scala (a_3),
- Masaryk University Hall (a_4).

The organizers of the event formed an expert team that identified the 21 most important characteristics for evaluation of the individual alternatives. These characteristics were grouped, according to similarity, into five criteria on the upper level of the hierarchy, see Figure 1. The criteria and sub-criteria weights were determined by pair comparisons, as proposed in subchapter 3.2. The alternatives were evaluated by the experts with respect to all sub-criteria to get the criteria matrices needed for fuzzy TOPSIS (fuzzy scales for alternatives and criteria are described in Table 1).

Value	Alternatives ratings	Value	Criteria Preference
$(1,1,3)$	Poor	$\left(\frac{1}{3},1,3\right)$	Criteria are of the same importance.
$(1,3,5)$	Medium poor	$(1,3,5)$	First criterion is slightly preferred to the other.
$(3,5,7)$	Fair	$(3,5,7)$	First criterion is preferred to the other.
$(5,7,9)$	Medium good	$(5,7,9)$	First criterion is strongly preferred to the other.
$(7,9,9)$	Good	$(7,9,9)$	First criterion is absolutely preferred to the other.

Tab. 1. Fuzzy scale for evaluation of alternatives and criteria

Saaty's fuzzy matrices $\widetilde{\mathbf{S}}$ and fuzzy criteria matrices $\widetilde{\mathbf{Y}}$ were aggregated using geometrical mean (applied componentwise on respective significant values as in [7]) of the fuzzy values given by the experts. Calculations were performed using programming language R [9] and its packages FuzzyAHP [5], MCDM [3] and FuzzyMCDM [2]. Input matrices and intermediate results are too extensive to be included; Table 2 gives an overview of the resulting evaluation.

Alternatives	d_i^+	d_i^-	c_i	Ranking
a_1	1,372841	1,2127350	0,4690387	2
a_2	1,501998	0,9879444	0,3967740	4
a_3	1,360780	1,2210642	0,4729426	1
a_4	1,398845	1,1658255	0,4545713	3

Tab. 2. Final evaluation of alternatives by fuzzy MCDM method

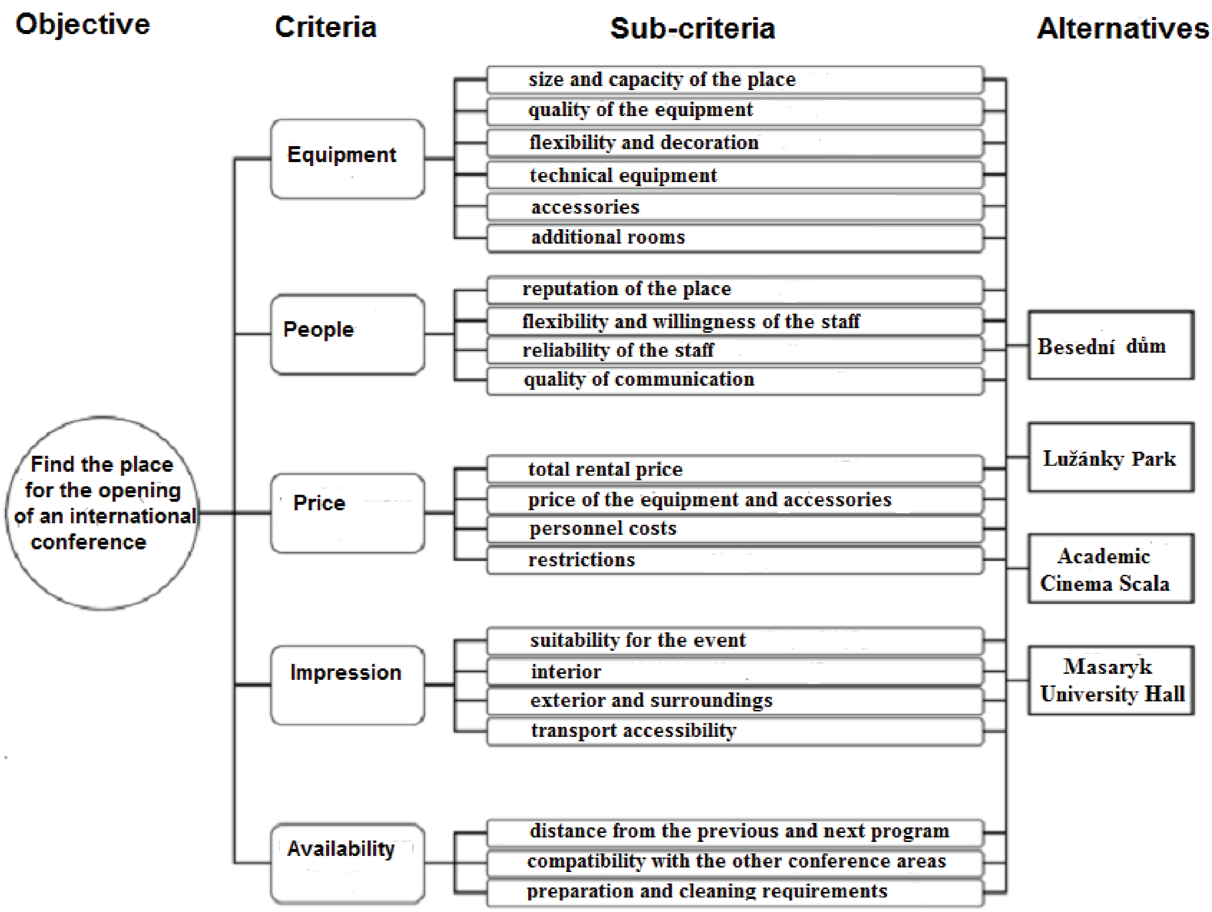


Fig. 1. Hierarchy of considered decision-making problem

The alternative a_3 Academic Cinema Scala was evaluated as the most appropriate. The decision framework including the R procedure is versatile and can be used repeatedly for organizing other events in the future.

5 CONCLUSION

The presented fuzzy multi-criteria decision-making method appears to be highly suited to working with inaccurate input information based on natural language. Compared to the standard approaches, fuzzy MCDM allows working with a degree of uncertainty to describe the real world more accurately. The results of the analyses are, therefore, more plausible and relevant for decision makers. The method can be further extended to allow for different types of fuzzy numbers.

Ackowledgements: This paper was supported by project MUNI/A/1018/2017.

References

[1] Chen, C. (2000). Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy Sets and Systems*, 114(1), pp.1-9.
[2] Cran.r-project.org. (2018). CRAN - Package FuzzyMCDM. [online] Available at: https://CRAN.R-project.org/package=FuzzyMCDM [Accessed 21 Apr. 2018].
[3] Cran.r-project.org. (2018). CRAN - Package MCDM. [online] Available at: https://CRAN.R-project.org/package=MCDM [Accessed 21 Apr. 2018].
[4] García-Cascales, M. and Lamata, M. (2012). On rank reversal and TOPSIS method. *Mathematical and Computer Modelling*, 56(5-6), pp.123-132.
[5] GitHub. (2018). JanCaha/FuzzyAHP. [online] Available at: http://github.com/JanCaha/FuzzyAHP/ [Accessed 21 Apr. 2018].
[6] Krejčí, J., Pavlačka, O. and Talašová, J. (2016). A fuzzy extension of Analytic Hierarchy Process based on the constrained fuzzy arithmetic. *Fuzzy Optimization and Decision Making*, 16(1), pp.89-110.
[7] Meixner, O. (2009). Fuzzy AHP group decision analysis and its application for the evaluation of energy sources, In: *Proceedings of the 10th International Symposium on the Analytic Hierarchy/Network Process*. Pittsburgh, pp 1-14.
[8] Novák, V. (2000). *Základy fuzzy modelování*. Praha: BEN - technická literatura.
[9] R-project.org. (2018). R: The R Project for Statistical Computing. [online] Available at: https://www.R-project.org/ [Accessed 21 Apr. 2018].
[10] Saaty, T. L. (1980). *The Analytic Hierarchy Process*. New York: McGraw-Hill.
[11] Santos, F. J. J. and Camargo, H. A. (2010). Fuzzy systems for multicriteria decision making. *CLEI Electronic Journal*, 13(3), pp. 1-8.
[12] White, D., Hwang, C. and Yoon, K. (1982). Multiple Attribute Decision Making - A State-of-the-Art Survey. *The Journal of the Operational Research Society*, 33(3), pp.289-298.
[13] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), pp. 338-353.

Author's address

Mgr. Michal Koláček
Masaryk University, Faculty of Science
Kotlářská 7, 60200 Brno
Czech Republic
email: 116502@mail.muni.cz

Ing. Mgr. Markéta Matulová, PhD.
Masaryk University, Faculty of economics and administration, Department of applied mathematics and informatics
Lipová 41a, 60200 Brno
Czech Republic
email: marketa.matulova@econ.muni.cz

DETERMINANTS OF SLOVAK FOREIGN TRADE – COMPARISON AMONG V4 COUNTRIES BASED ON GRAVITY MODEL

Brian König - University of Economics in Bratislava,
Gabriela Dováľová - Institute of Economic Research SAS

Abstract
The article deals with the analysis of the factors influencing the foreign trade of Slovak Republic. The comparative analysis among V4 countries is based on the gravity model of international trade with applying Poisson pseudo-maximum likelihood estimator. Importance of this topic bases on the fact, that the high openness of the Slovak economy while having the relatively low absorption capacity of the domestic market at the same time, still creates a need for orientation on foreign markets. The paper main findings may be summarized as follows: there is no common border effect in SK, CZ and HU; lower effect of distance for SK compared to the rest of V4 countries; distinctly higher effect of EU membership for SK international trade partners compared to CZ, HU and PL.

Keywords: *Gravity model, International trade, Export, V4 countries*

JEL Classification: C460, F140
AMS Classification: 91B60

1 INTRODUCTION

Improving the labour market situation in Slovakia encourages domestic consumption, but the wage level is still at a relatively low level compared to the western European Union (EU) countries which creates the need also for orientation on foreign markets. In the future from the export perspective, the more pressure from other countries with even cheaper labour and lower production costs can be expected. Slovakia's export possibilities will largely depend, among other things, on the development of major determinants of consumer market developments such as incomes, savings and price level. From the point of view of future consumer demand, the area of indebtedness will also play an important role.

In 2018, it is expected that the contribution of Slovakia's foreign trade to economic growth will slightly increase, which should be supported in particular by the increase in demand in the European countries where Slovakia exports the most. The growth in exports should also be supported by the launch of production in the new Jaguar Land Rover´s plant near Nitra. Concerning the foreign trade of EU countries, including Slovakia, the impact of Brexit on the European Union is increasingly being discussed. The impacts of Brexit have not yet been unambiguous, but several analytical studies have been carried out describing possible scenarios of further developments and subsequently assessing possible impacts on further economic developments in the EU countries. Most of the studies so far conducted to assess the impact of Brexit on the EU countries have concluded that the United Kingdom's exit from the EU will mean economic losses for both parties. Many authors, in estimating the economic impacts of Brexit, use gravitational models in their analyses (e.g. Gudgin et al, 2017). As an example of different way how to use gravity model, the authors Grančay et al (2015) could be mentioned, who used the gravity model for identifying changes in foreign trade determinants of Slovakia and the Czech Republic. These authors showed, that importance of common border for the countries' bilateral trade with third countries has been gradually declining, which was according to the authors in line with accelerating globalization and lower shipping prices. The authors also emphasized the importance of the EU membership as an important determinant of bilateral trade.

The great advantage of the gravity model is its simplicity and relatively good approximation of available data. Shepherd (2016) shows, that gravity model has a comparative advantage in using of data to assess the sensitivity of trade to particular trade cost factors. This kind of model is very useful in testing the response of trade flows to intended reforms. For example nowadays, the gravity model is very often used for estimating the economic impact of Brexit (e.g. Gudgin et al, 2017). According this author, one of the main disadvantages of gravity model is that it is not an appropriate tool for describing the behaviour of economic welfare. Authors Ivus and Strong (2007) state, that the main advantage of gravity models in comparison with CGE models is, that gravity models measure the effect on trade flows of a past trade policy. By contrast, CGE models take most of the time ex-ante approach, which means quantifying the future effects of a new policy.

In the following part of the paper, we focused on the analysis of Slovakia's export territories and on identifying the significance of selected determinants of Slovak foreign trade using the gravity model. The aim of the analysis is to identify the importance of EU countries as the export territories of Slovakia and to define the significance of selected determinants, including the effect of EU accession on the size of Slovak foreign trade in comparison with the rest of V4 (Czech Republic, Hungary, Poland, Slovakia) countries, as there are many analysis pointing to the common specifics of V4 countries (e.g. Szomolányi, Lukáčik and Lukáčiková, 2013).

1.1 Exporting territories of SR

Slovakia is one of the most open economies of the European Union. Using a long term view, the main trading partners of Slovakia in terms of value are European countries. In 2017 Europe accounted for 90.8 % of total Slovak exports (Figure 1b), with total imports from Europe at the level of 65.3 % (Figure 1a).

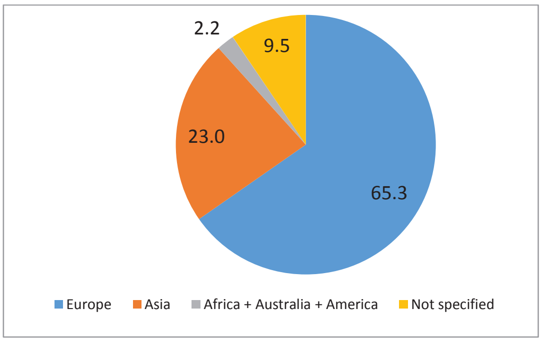


Figure 1a Total import

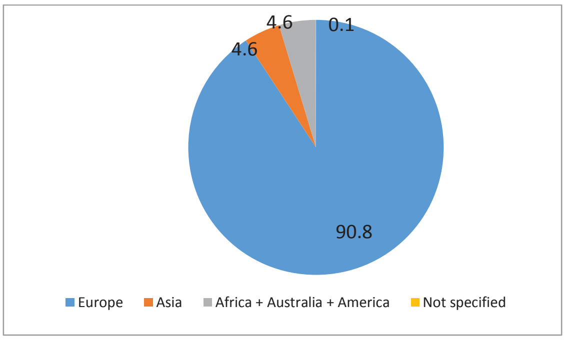


Figure 1b Total export

Figure 1: Total imports and exports of the Slovak Republic by continents, 2017 (in %)

Source: Authors' calculation on the basis of Statistical Office data of SR

Slovakia´s export activities are only to a very limited extent oriented towards the so called „third markets". Slovakia´s dominant export territory is the EU28. In January-December 2017, the total import value of Slovakia reached 71,817.2 million euro, out of which approximately 66.9 % came from EU28. The total value of Slovakia export reached 74,813.3 million euro in 2017, again highlighting the importance of EU28 with the share of 85.4 %. In 2017, the most important export partners of Slovakia were countries like Germany, V3 countries (Czech Republic, Hungary, Poland), France, Italy and Austria with a share of 64 percent of total export (Figure 2b). Countries V3 are also important from the Slovak´s import point of view representing 20.2 % of total import (Figure 2a).

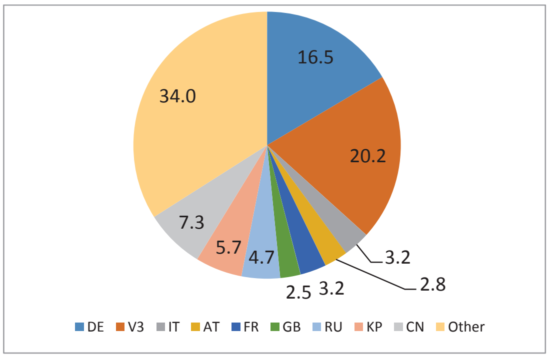


Figure 2a Total import

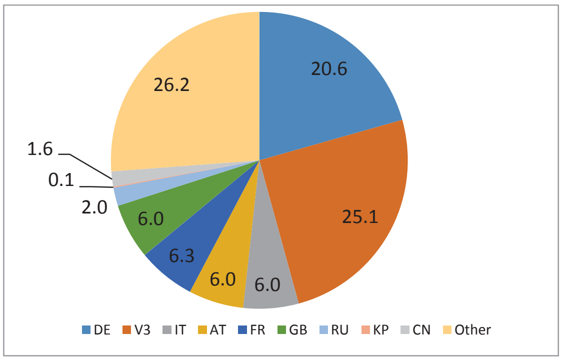


Figure 2b Total export

Figure 2: Total import and export of Slovak Republic from the selected countries, 2017 (in %)

Source: Authors' calculation on the basis of Statistical Office data of SR

Out of Slovakia´s total export to the EU28, around 56.6 % went to the countries with a common Euro currency. The rest went to 9 countries that have not yet adopted the euro, including the countries V3. From the Slovak´s import from EU28 point of view, approximately 62.2 % was delivered from countries with the common currency.

2 METHODOLOGY

The gravity model is a very often used instrument of international trade theory which serves to empirical analyses of international trade. It is called "gravity" as it is based on the Newton's gravitation law, which described relation of two mutually attracted planets with the forces depending on the size of the planets and the distance between them. Similarly, countries' trade depends on their economic size and their mutual distance. The gravity models of international trade were for the first time introduced by Dutch economist, Tinbergen (1962). Basic gravity model can be written as follows:

$$X_{ij} = \alpha \left(\frac{Y_i^{\beta_1} . Y_j^{\beta_2}}{D_{ij}^{\beta_3}} \right) \tag{1}$$

where X_{ij} represents value of trade between countries i and j; α, β_1, β_2, β_3 are the unknown parameters; Y_i and Y_j express the size of the economy i and j respectively; D_{ij} is a distance between country i and j. There are more approaches how to express a value of trade between two countries, but in general, it is represented by the size of export or import between two countries. The size of economy is usually expressed by the level of GDP or by the combination of per-capita GDP and country population which in addition, allows for controlling the effect of economic development of particular country. Very often gravity model is expressed in a log-log form as follows:

$$\ln X_{ij} = \beta_0 + \beta_1 \ln Y_i + \beta_2 \ln Y_j + \beta_3 \ln D_{ij} + u_{ij} \tag{2}$$

where u_{ij} is error term. Later on, the gravity models have been exposed to a criticism for being just an econometric tool without sufficient theoretical foundation. However, subsequent research shown that the gravity models can be supported by the range of trade theories (e.g. Bergstrand (1985), Eaton and Kortum (2002), Chaney (2008), etc.) and many others variables were added to the gravity equation (common borders, import duties, common currency, international trade agreements, common language, island country, seaside country etc.)[1].

Since the data of the bilateral trade contain lot of zero values, as each individual country doesn't trade with all the other countries, ordinary least square method (OLS) would drop such observations as logarithm of zero is not defined, what might lead to a sample selection bias as the zero values are not normally distributed. For that reason we apply Poisson pseudo-maximum likelihood estimator (PPML) which has more advantage compared to the OLS estimator (Shepherd, 2016). PPML allows for estimation of nonlinear models (what some of the gravity models are) and in addition, it naturally includes zero values. Moreover, even though the dependent variable for the PPML estimator is expressed in levels, interpretation of the parameters is the same than in case of OLS (elasticities). For these reasons the PPML estimator is very often used estimation method in the gravity models.

[1] See Bacchetta et al (2012).

3 DATA

For the purpose of presented analysis we used data from 2000-2015 for 194 countries. Dependent variables of all four models are represented by the bilateral nominal exports of V4 countries in US dollars downloaded from the database UNCTAD. The rest of the variables are used from the CEPII gravity database which gathers data for the purpose of gravity models. Per-capita GDP is expressed as nominal GDP in US dollars divided by the country population. Distance represents a bee-line between capitals of origin and destination country. Population is measured as total population in millions. We also used standard dummy variables for common borders, common currency and the membership in the European Union.[2]

4 RESULTS

For the purpose of our analysis we assume gravity model in the following form:

$$EXPORT_{ij} = \beta_0 + \beta_1 lnGDPCAP\_D + \beta_2 lnPOP\_D + \beta_3 lnDISTCAP_{ij} + \beta_4 CONTIG_{ij} + \beta_5 EU\_O + \beta_6 EU\_D + \beta_7 YEAR\_D_{2001} + \cdots + \beta_{21} YEAR\_D_{2015} + u_{ij} \quad (3)$$

where $EXPORT_{ij}$ is a nominal value of export from country *i* (origin country) to country *j* (destination country); *GDPCAP_D* is per-capita GDP of importer (D - destination country); *POP_D* expresses the population of importer; $DISTCAP_{ij}$ is distance between the capitals of *i* and *j* countries; $CONTIG_{ij}$ is a dummy variable for common border between countries *i* and *j*; *EU_O* is a dummy variable for membership of the origin country (exporter) in the EU; *EU_D* is a dummy variable for membership of the destination country (importer) in the EU; $YEAR\_D_{year*}$ are dummy variables for the particular years. These variables are often used in current gravity models. *EU_O* is also included as the analysed period contains period of EU accession of all four countries.

Table 1 – Gravity model of bilateral export of V4 countries (PPML estimates)

variable/country	SK	CZ	HU	PL
ln_gdpcap_d	**0.53*** **	**0.4875*** **	**0.6878*** **	**0.6394*** **
	(0.0997)	(0.1031)	(0.1094)	(0.0639)
ln_pop_d	**0.8992*** **	**0.6879*** **	**0.9512*** **	**0.7072*** **
	(0.1013)	(0.1006)	(0.0877)	(0.0409)
ln_distcap	**-0.649** **	**-1.0871*** **	**-1.2274*** **	**-1.1164*** **
	(0.266)	(0.1064)	(0.1855)	(0.1285)
contig	**0.7751**	**0.4575**	**0.0864**	**0.4457** **
	(0.6187)	(0.2792)	(0.4764)	(0.1881)
eu_o	**1.2415*** **	**1.1853*** **	**0.6661*** **	**1.2217*** **

[2] We also used time dummy variables according to the authors Grančay et al (2015).

	(0.1666)	(0.0707)	(0.087)	(0.0534)
eu_d	**1.4959*** **	**0.4181*** **	**0.4833** **	**0.5392*** **
	(0.4521)	(0.1438)	(0.2361)	(0.1653)
_cons	**14.7828*** **	**20.5939*** **	**18.9987*** **	**19.6133*** **
	(1.7628)	(1.5172)	(1.0247)	(1.2075)
R-squared	0.6787	0.955	0.8386	0.9699
Num. of obs.	2,959	2,959	2,959	2,959

Remark: Values in parenthesis are standard deviations of the estimated parameters.

*** - denotes p-value less than 0.01; ** - denotes p-value less than 0.05; * - denotes p-value less than 0.1. Dummy for the years are omitted from the table.
Source: Authors' calculations

Based on the results in Table 1 we can conclude that the expected effects of individual variables on export are in line with our expectations as well as with the international trade theory, as the coefficients $\beta_1, \beta_2, \beta_4, \beta_5, \beta_6$ are positive and parameter β_3 belonging to the distance variable, which represents a proxy for travel costs is negative. Surprisingly, coefficient β_4 representing the effect of common border is in all countries statistically insignificant except the case of Poland. This might be contributed to the high economic openness of these countries and to the growing effect of globalization. Concerning the EU membership, there is a positive effect of EU accession on the size of trade for all four countries as the β_5 is positive. Furthermore, all V4 countries in average have higher export to the EU countries compared to the non-EU members, while the strongest effect of the importer membership is recorded by the SK, pointing out that the export among SK and EU countries is approximately 1.5 % higher compared to the non-EU countries. The importance of distance is for the HU, PL, and CZ almost twice as important as in the case of SK.

5 CONCLUSIONS

The aim of the paper is to analyse the main determinants of SK export flows and its comparison with the rest of the V4 countries. For this purpose we adopted gravity model which represents very often used tool in international trade analysis. As the main findings we may mention following points: there are no common border effects in SK, CZ and HU; lower effect of distance for SK compared to the rest of V4 countries; distinctly higher effect of EU membership for SK export partners compared to CZ, HU and PL. This research may serve as a starting point for the analysis of the effects of EU members exit on the bilateral trade of V4 countries. Such research would be vital especially in context of SR, which economy is to high extent dependent on the foreign trade. Slovakia´s export activity is still relatively low reliant on the so-called "third markets", accounting for almost 86% of total exports to the EU28. Therefore, for Slovakia, the shape of future trade relations within the EU countries will be important, including the post-Brexit trade deal between the UK and EU.

Acknowledgements: This work was supported by VEGA 1/0248/17 and VEGA 2/0109/16.

References

[1] Bacchetta, M., Beverelli, C., Cadot, O., Fugazza, M., Grether, J.M., Helble, M., Nicita, A. and Piermartini, R. (2012). *A practical guide to trade policy analysis.* Geneva: World Trade Organization.

[2] Bergstrand, J. H. (1985). The gravity equation in international trade: some microeconomic foundations and empirical evidence. *The Review of Economics and Statistics,* 67(3): pp. 474–81.

[3] Chaney, T. (2008). Distorted gravity: the intensive and extensive margins of international trade. *American Economic Review,* 98(4): pp. 1707–21.

[4] Eaton, J. and Kortum, S. (2002). Technology, geography and trade. *Econometrica,* 70(5): 1741–79.

[5] Grančay, M., Grančay, N., Drutarovská, J. and Mura, L. (2015). Gravity Model of Trade of the Czech and Slovak Republics 1995-2012: How the determinants of Trade Changed. *Political Econonomy*, 63(6), pp. 759-777.

[6] Gudgin, G., Coutts, K., Gibson, N. and Buchanan, J. (2017). The Role of Gravity Models in Estimating the Economic Impact of Brexit. Centre for Business Research, University of Cambridge, working paper no. 490. Available at: *https://ideas.repec.org/p/mde/wpaper/0077.html* [Accessed 27 March 2018].

[7] Ivus, O. and Strong, A. (2007). Modelling approaches to the analysis of trade policy: computable general equilibrium and gravity models. In: *Handbook on International Trade Policy*. Edited by W.A. Kerr and J.D. Gaisford. Great Britain: Edward Elgar Publishing, pp. 44-54.

[8] Shepherd, B. (2016). *The Gravity Model of International Trade: A User Guide.* 1st edn. Bangkok: United Nation Publication [pdf]. Available at: http://www.unescap.org/resources/gravity-model-international-trade-user-guide-updated-version [Accessed 27 March 2018].

[9] Szomolányi, K., Lukáčik, M. and Lukáčiková, A. (2013). Estimation of the Production Function Parameters in V4 Economies. In: *31st International Conference on Mathematical Methods in Economics 2013*. Jihlava: College of Polytechnics Jihlava, pp. 898-902.

[10] Tunbergen, J. (1962). *Shaping the World Economy: Suggestions for an International Economic Policy*. 1st edn. New York: The Twentieth Century Fund.

Author's address

Ing. Brian König, PhD.
University of Economics in Bratislava
Faculty of business informatics
Department of Operations research and Econometrics,
Dolnozemská cesta 1/b, 852 35 Bratislava
Slovak Republic

Institute of Economic Research SAS
Šancová 56, 811 05 Bratislava
Slovak Republic
email: brian.konig@euba.sk

Ing. Gabriela Dováľová, Phd.
Institute of Economic Research SAS
Šancová 56, 811 05 Bratislava
Slovak Republic
Email: gabriela.dovalova@savba.sk

WORLD TAX INDEX: MEASURING TAX BURDEN IN THE CZECH REPUBLIC

Zlatica Konôpková, Masaryk University, Brno,
Jakub Buček, Masaryk University, Brno

Abstract
World Tax Index (WTI) is an overall multi-criteria indicator of the tax burden in OECD countries that can be considered as an alternative to the tax quota or implicit tax rates. This index is a combination of hard data weighted by soft data (QEO) gathered through an online questionnaire survey among tax specialists. The new methodology of WTI is based on factor analysis because we consider the weights of tax categories in WTI to be latent variables that cannot be measured directly (as it is common in social sciences). Initial group of tax specialists from the academic field has been extended to the private and public sphere. Using the answers from the pilot project held in the Czech Republic in 2016 we obtained the weights of tax categories represented by the tax rates (most used or average). The indicator defined in this way is actually unique weighted average tax rate for the country. WTI for the Czech Republic was calculated for the period 2000-2017 and we can conclude it is quite stable over examined period.

Keywords: *Tax Burden, World Tax Index (WTI), Tax Quota, Factor Analysis, Czech Republic*

JEL Classification: H21, C38
AMS Classification: 62H25, 91B64

1 INTRODUCTION

After years of prosperity, poor performance of the world economy during global financial crisis has increased the indebtedness of the countries. Consolidation of the public sector was one of the key tasks for the governments. Unfortunately, it is still a task to be done and governments are struggling with the consequences of the latest global financial crisis for almost a decade now.

Taxes are one of the two main instruments of fiscal policy, they are collected by government to cover its second instrument, spending. Politicians must find equilibrium, the ideal setting of policy in which they collect enough funds to cover their activities, but do not burden private sector too much. Taxes may have negative impact on economy, they burden part of economic activity, create distortions and can slow down the economic growth. The impact of taxation on economy is the subject of numerous studies and since there is a requirement to incorporate it into models of economic growth we need to be able measure the size of tax burden correspondingly. Tax burden is also used to compare the countries.

The goal of this article is to present the World Tax Index calculated for the Czech Republic based on new methodology. Firstly, we will briefly present the different options to measure tax burden, then we will explain the details of calculating WTI for the Czech Republic and discuss the results.

2 MEASURING TAX BURDEN

Tax quota and implicit tax rates are the mostly used indices to measure tax burden. The first mentioned, tax quota simply expresses the ratio of tax revenues to nominal GDP, or in compound form ratio of tax revenues with social security contributions to nominal GDP. The indicator is often used for comparison and in models of economic growth because is simple and available for many countries. However, according to Kotlán and Machová (2012), the tax quota reflects only the percentage of GDP redistributed through the public budgets rather than the level of the tax burden. A higher tax quota can indicate only more efficient tax collection and not the higher tax burden. Another disadvantage is a usage of GDP that has different methodology in countries.

Implicit tax rate is the second popular indicator of tax burden. It again relates the tax revenues but only to the activities or commodities in GDP connected to the tax instead of the whole nominal GDP. However, this change does not eliminate the disadvantages mentioned above. The implicit tax rates are available for consumption, labour and capital but only for European Union and therefore international comparison is not possible outside European Union. More about disadvantages of using these two indicators can be found in Kotlán, Machová and Janíčková (2011).

In the literature we can find alternative indicators of tax burden. Raimondos-Møller and Woodland (2006) present a non-money metric index of optimality that measures the distance of some current tax structure from the optimal tax structure in the presence of public goods. De Laet and Wöhlbier (2008) define a tax as a compulsory unrequired payment to the government, differentiate the types of taxes according to the national account and assigned them various economic functions. Kiss, Jedrzejowicz and Jirsáková (2009) eliminates the influence of government in tax quota.

3 WORLD TAX INDEX

World Tax Index (WTI) is an overall multi-criteria indicator of the tax burden designed for OECD countries. It was firstly presented in Kotlán and Machová (2012)[1] as an alternative for standardly used indicators (tax quota, implicit tax rates). WTI is not limited only to the tax revenues and nominal GDP, but it includes information from different areas that can possibly affect the tax burden, e.g. progressivity of the tax system, administration, incentives, deductions etc. The results presented in Kotlán and Machová (2013) show that the WTI generally allows to prove more significantly and clearly the negative effect of taxation in case of particular types of taxes, which is not possible using the tax quota.

WTI is structured according to OECD classification into 5 main groups of taxes - Personal Income Tax (PIT), Corporate Income Tax (CIT), Value Added Tax (VAT), Individual Property Taxes (PRO) and Other Taxes on Consumption/Excises (OTC). Available hard tax data are combined with soft data (Qualified Expert Opinion – QEO) gained from questionnaire survey conducted in all OECD countries. Initially, the questionnaire was distributed among tax specialists in academic sphere who expressed their opinion on how individual components of WTI contribute to the tax burden in their country by distributing the 100% in every major category. The Saaty method[2] of pair-wise comparisons was used to obtain the weights of the main categories in WTI and the experts were also asked to compare all pairs of the categories according to their importance.

In Konôpková and Buček (2016) we introduced new methodology of WTI. As it is common in social sciences[3], we have chosen factor analysis to obtain the weights of tax categories in

[1] The current values of WTI are avalaible at http://www.eaco.eu/about-eaco/research/.
[2] See e.g. Saaty (2008).
[3] See e.g. Bartholomew, Steele, Galbraith and Moustaki (2008), Fabrigar, Wegener, MacCallum and Strahan (1999).

WTI and have opened the questionnaire to public and private sector due to the following reasons. Firstly, the initial method was built on assumption that experts can evaluate the tax burden in their country directly. According to Borsboom, Mellenbergh and van Heerden (2003) human behavior cannot be observed directly and the variables expressing the behavior are latent, their values are hidden for us, but we can use a mathematical method to estimate them. Secondly, the questionnaire was distributed only to experts in academic sphere, but the weights should reflect proportionally the opinion of the whole diverse population in country. Academic researchers usually do not have real experience in other sectors of economy and therefore cannot assign values objectively.

Factor analysis[4] is a method built on asking different questions that indirectly refers to the same thing. In line with this method we created online questionnaire[5] to measure the tax burden. As we haven't found any similar research the questions were prepared based on theory or our own opinion. The questionnaire is built on 32 "*Do you think…?*" questions with 5-scale answer from *Strongly disagree* to *Strongly agree.* There are seven general questions on tax system and then 5 questions per each major category. The respondents also have to order the basic components of WTI in order to compare the values of previous index. Finally, in order to compute the WTI over time, they are asked to express their opinion how the importance of category changed in period 2010-2015 on 5-scale from *High decrease* to *High increase.*

Applying the factor analysis we receive the factors and calculate the weights of tax categories for each individual respondent, the average value of all respondents from the country is the weight for the country. WTI is then calculated using the standardized weights and representing tax rates for every category. The indicator defined in this way is actually unique weighted average tax rate for the country. Because the questions were prepared without any prior testing we made the pilot in the Czech Republic in 2016.

4 PILOT PROJECT – THE CZECH REPUBLIC

The pilot project run from October 2016 till April 2017. Questionnaire was sent to academic researches chosen from SCOPUS database, public national institutions that are part of the tax system - Finance Ministry, Financial Administration. In private sector it was sent in newsletter of Chamber of Tax Consultants of the Czech Republic and to "Big Four" group. The group is created by the four largest companies in the world offering tax, audit, assurance, consulting, corporate finance and legal services operating in all OECD countries and therefore we consider them as good sample for our research. This enlargement of possible respondents together with questionnaire available online should increase the response rate, the number of respondents, their heterogeneity and therefore the credibility of the index.

Together 57 respondents filled the questionnaire, most of them from the private sphere (82%). The public sphere did not respond too much (only 3 responses) and did not even respond to our repeating emails. More than 75% respondents strongly agree or agree with the statement that they consider themselves as experts on taxes, less than 9% strongly disagree or disagree with it and remaining 16% are neutral. We can consider respondents suitable for our research. Applying exploratory factor analysis (EFA) with varimax rotation on data we extracted the values of immeasurable variables, the latent factors (details in Konôpková and Buček (2016)). Then we used standardization to receive the weights of tax categories stated in Table 1.

[4] See e.g. Fruchter (1954), Brown (2009).
[5] The questionnaire is available at http://goo.gl/forms/kWCsMvKO2tWUirIk2.

Table 1 Calculated weights in WTI for the Czech Republic

PIT	CIT	VAT	PRO	OTC
0.224	0.242	0.163	0.158	0.213

Source: Author's construction.

First of all, we have to remind that new WTI represents now a single weighted average tax rate reflecting the tax burden in the country. As we can see the most important in WTI is Corporate Income Tax. This could be expected because the most of the respondents are experts from private sector, working in companies that take care of finance of the other companies. The second most important is Personal Income Tax. Both categories are direct taxes and are followed by Other Taxes on Consumption/Excises. Excises are the highest taxes in the Czech Republic and are imposed on goods that are desired. Their contribution to WTI will be the highest in absolute values. The last categories are Value Added Tax and Property Taxes. The results are in line with OECD (2011) stating that consumption taxes and property taxes have the smallest negative effects on economy and the corporate and personal income taxes are on the opposite side. It supports the main trend in fiscal policy, to make a move from direct taxes to indirect ones

We will consider the weights of categories to be the same over time because the experts do not consider the tax system very dynamic (neutral changes in importance of categories) as it was the same in previous version of WTI.

Table 2 Representing tax rates for the Czech Republic and their source

PIT	Average personal income tax and social security contribution rates on gross labour income (100% average wage)	OECD (2017) Table I.5
CIT	Statutory corporate income tax rate	OECD (2017) Table II.1
VAT	Basic tax rate	OECD (2017) Table 2.A2.1
PRO	Tax on the acquisition of immovable property/property transfer tax + land tax	Finance.cz (2017)
OTC	Excises – average rate weighted by consumption	OECD (2017) Table 4.A4.1-8

Source: Author's construction.

To calculate WTI for the Czech Republic we have chosen the representing tax rate/rates for each category. Table 2 shows a list of tax rates and their source. PIT, CIT and VAT categories are represented by single tax rate – the main tax rate. PRO category is represented by national taxes, local taxes are not taken into account because they are different in each city. PRO is hence represented by the sum of the tax on the accusation of immovable property (previously called property transfer tax) and the land tax. It was the most difficult to choose representing tax rate for category OTC. There are excises on different goods (wine, beer, alcohol, oil, gas) with different definition (e.g. tax amount per l, tax rate, etc.) therefore we calculated tax rate per unit for every commodity and then calculated average tax rate weighted by the consumption ratio of the commodity in GDP. Changes in tax rates over time are shown in the Figure 1. The most significant change is in Excises caused by the entry of the Czech Republic into the European Union in 2004 when the tax harmonization has started. It has also affected CIT that has gradually decreased in order to improve business environment.

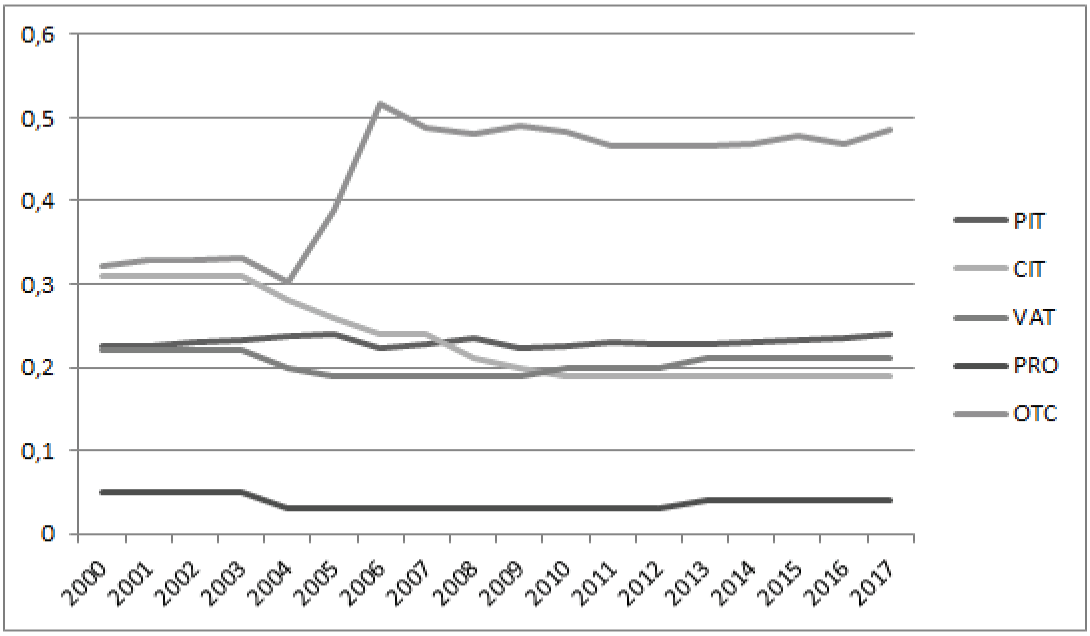


Figure 1 Representing tax rates in time. Source: Author's construction.

Final results are shown in Table 3, detailed results for components of WTI can be found in Table 4. WTI for the Czech Republic was calculated for the period 2000-2017 and we can conclude it is quite stable over examined period, the only significant deviation was in 2004 when the Czech Republic joined the European Union and tax harmonization started.

Table 3 World Tax Index for the Czech Republic

2000	2001	2002	2003	2004	2005	2006	2007	2008
0.238	0.239	0.240	0.241	0.222	0.235	0.254	0.249	0.242
2009	2010	2011	2012	2013	2014	2015	2016	2017
0.238	0.236	0.235	0.234	0.237	0.238	0.241	0.239	0.243

Source: Author's construction.

Table 4 Components of WTI for the Czech Republic

	2000	2001	2002	2003	2004	2005	2006	2007	2008
PIT	0.050	0.050	0.052	0.052	0.053	0.054	0.050	0.051	0.053
CIT	0.075	0.075	0.075	0.075	0.068	0.063	0.058	0.058	0.051
VAT	0.036	0.036	0.036	0.036	0.033	0.031	0.031	0.031	0.031
PRO	0.008	0.008	0.008	0.008	0.005	0.005	0.005	0.005	0.005
OTC	0.068	0.070	0.070	0.070	0.064	0.083	0.110	0.104	0.102
	2009	2010	2011	2012	2013	2014	2015	2016	2017
PIT	0.050	0.050	0.052	0.051	0.051	0.052	0.052	0.053	0.054
CIT	0.048	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046
VAT	0.031	0.033	0.033	0.033	0.034	0.034	0.034	0.034	0.034
PRO	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.006	0.006
OTC	0.104	0.103	0.099	0.099	0.099	0.100	0.102	0.100	0.103

Source: Author's construction

5 CONCLUSIONS

In this paper we presented new values of World Tax Index to measure tax burden in the Czech Republic for period 2000-2017. New methodology of WTI is built on factor analysis in order to estimate the weighs of main categories in index which we consider to be latent variables. The indicator of tax burden defined in this way is actually a single weighted average tax rate for the country.

The most important categories of WTI are direct taxes, Corporate Income Tax and Personal Income Tax, and are followed by Other Taxes on Consumption/Excises. Excises are the highest taxes in the Czech Republic and are imposed on goods that are desired. The last categories are Value Added Tax and Property Taxes. The results are in line with OECD (2011) stating that consumption taxes and property taxes have the smallest negative effects on economy and the corporate and personal income taxes are on the opposite side. It supports the main trend in fiscal policy, to make a move from direct taxes to indirect ones

We can conclude tax burden is quite stable with no significant shock over examined period, the only deviation was in 2004 when the Czech Republic joined the European Union and tax harmonization started.

Acknowledgements: This work was supported by funding of specific research at Faculty of Economics and Administration, project MUNI/A/1139/2017. This support is gratefully acknowledged.

References

[1] Bartholomew, D.J., Steele, F., Galbraith, J. and Moustaki, I. (2008). *Analysis of Multivariate Social Science Data. Statistics in the Social and Behavioral Sciences Series.* (2nd ed.). Abingdon: Taylor & Francis.

[2] Borsboom, D., Mellenbergh, G.J. and van Heerden, J. (2003). The Theoretical Status of Latent Variables. *Psychological Review*, 110 (2), pp. 203–219.

[3] Brown, J. D. (2009). Principal components analysis and exploratory factor analysis – Definitions, differences and choices. [online]. Shiken: JALT Testing & Evaluation SIG Newsletter. Available at http://jalt.org/test/PDF/Brown29.pdf [Accessed 28 Aug. 2017].

[4] De Laet, J.-P. and Wöhlbier, F. (2008). Tax Burden by Economic Function: A Comparison for the EU Member States. *MPRA Paper No. 14761*.

[5] Fabrigar, L.R., Wegener, D.T., MacCallum, R.C. and Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), pp. 272-299.

[6] Finance.cz, (2017) [oline]. Available at: https://www.finance.cz/dane-a-mzda/majetkove-dane/dan-z-prevodu-nemovitosti/ [Accessed 5 Oct. 2017].

[7] Fruchter, B. (1954). *Introduction to Factor Analysis*. New York: Van Nostrand.

[8] Kiss, G. P., Jedrzejowicz, T. and Jirsáková, J. (2009). How to Measure Tax Burden in an Internationally Comparable Way? *National Bank of Poland Working Paper No. 56*.

[9] Konôpková Z. and Buček J. (2016). World Tax Index: Methodology for 2017. In: *6th International Scientific Conference TAXES IN THE WORLD*. Krakow: Ostrava: European Association Comenius – EACO and VSEO, pp. 77-87.

[10] Kotlán, I., Machová, Z. and Janíčková, L. (2011). Taxation Influence on the Economic Growth. *Politická ekonomie,* 59(5), pp. 638–658.

[11] Kotlán, I. and Machová, Z. (2012). World Tax Index: Methodology and Data. *DANUBE: Law and Economics Review*, 3(2), pp. 19–33.

[12] Kotlán, I. and Machová, Z. (2013). The Impact of the Tax Burden on the Living Standard in OECD Countries. *Ekonomický časopis*, 61(9), pp. 951–962.
[13] OECD, (2011). *Tax Policy Reform and Economic Growth*. Washington: OECD.
[14] OECD, (2017). Database – OECD. [online]. Available at: http://www.oecd.org/tax/tax-policy/tax-database.htm [Accessed 5 Oct. 2017].
[15] Raimondos-Møller, P. and Woodland, A. D., (2006). Measuring Tax Efficiency: A tax optimality index. *Journal of Public Economics*, 90(12), pp. 1903-1922.
[16] Saaty, T. L. (2008). Decision Making with the Analytic Hierarchy Process. *International Journal of Services Sciences*, 1(1), pp. 83–98.

Author's address

Mgr. Zlatica Konôpková.
Masaryk University, Faculty of Economics and Administration,
Department of Economics
Lipová 41a, 602 00 Brno
Czech Republic
email: konopkova.zlatica@mail.muni.cz

Mgr. Bc. Jakub Buček.
Masaryk University, Faculty of Economics and Administration,
Department of Economics
Lipová 41a, 602 00 Brno
Czech Republic
email: jakubbucek@mail.muni.cz

MULTICRITERIA EVALUATION OF THE CZECH REGIONS FROM THE SELECTED ECONOMIC ACTIVITY ASPECTS POINT OF VIEW

Martina Kuncová, University of Economics, Prague
Jana Sekničková, University of Economics, Prague

Abstract
The Czech Republic is divided into 14 regions that can be compared with respect to a large amount of not only economic criteria. The data for such economic comparisons are published by the Czech Statistical Office. In this paper based on the data from the years 2012 and 2016 we compare all Czech regions regarding such economic indices as unemployment rate, economic activity, average age, wages, free workplaces, income, consumption and investments. We evaluate these regions from the multi-criteria point of view and as tools for analyses we use data envelopment analysis (DEA) and multi-criteria evaluation of alternative (MCEA). We compare results with ranks gained five years ago by similar analysis. In the end of paper we try to interpret obtained results and explain main changes.

Keywords: *Czech Regions, economic indicators, budgets, DEA, multi-criteria evaluation*

JEL Classification: C67
AMS Classification: 90B50, 90C29, 91B06

1 INTRODUCTION

The Czech Republic is divided into fourteen regions – Prague, the capital city, Central Bohemian Region, South Bohemian Region, Plzeň Region, Karlovy Vary Region, Ústí nad Labem Region, Liberec Region, Hradec Králové Region, Pardubice Region, Vysočina Region, South Moravian Region, Olomouc Region, Moravian-Silesian Region and Zlín Region. Each region has its own budget. In this paper we would like to analyze selected data from regional budgets and compare all 14 Czech regions with each other on the basis of several economic activity characteristics. For our analysis we used official budgets of regions from years 2012 and 2016 available from official web pages of Czech Statistical Office (www.czso.cz). Relevant data from these budgets are income of region (in thousands of CZK per capita), common expenditures (mainly transferred grants and subsidies, in thousands of CZK per capita) and capital expenditures (mainly investments, in thousands of CZK per capita). We also use other regional data, especially unemployment rate (in percentage), ratio of economic activity (in percentage), average wage (in CZK), average age (in years) and the number of free workplaces per capita.

We assume that region with higher income is more effective than the one with lower budget. Higher income indicates higher amount of grants and subsidies transferred to the economic subjects. Also the region with higher amount of investments is evaluated better as well as region with lower unemployment rate and higher economic activity. Of course we would like region with young people so the lower average age is desirable and higher average wage is also wanted.

For comparison of regions it is possible to use various kinds of techniques and methods such as statistical methods, econometric models, data envelopment analysis (DEA) models, or multicriteria decision making (MCDM) models, especially multicriteria evaluation of alternatives (MCEA). In this paper we use DEA models and selected MCEA methods.

DEA models are widely used in comparison of countries, regions or districts from various points of view. Melecky and Stanickova (2012) compared four Visegrad countries by DEA models, Friebel and Friebelova (2012) measured life quality in 14 southwest Czech districts by DEA using 4 inputs and 1 output and Furkova (2014) evaluated by these models the efficiency of Slovak economic faculties. MCEA models are also used very often for multi-criteria analysis of countries, regions and districts. Koutroumanidis, Papathanasiou and Manos (2002) compared regions of Greece using multi-criteria methods, especially PROMETHEE II., Dincer (2011) used TOPSIS and WSA methods for evaluation of European Union member states and candidate countries. The TOPSIS method was also used by Hashemabadi and Razmi (2014) for the comparison of countries from the tourism point of view or by Latuszynska (2014) and Kuncova and Doucek (2011) for the EU countries comparison from the ICT development point of view. In (Kuncova and Seknickova, 2013) we evaluated Czech regions by MCEA methods, especially by WSA, TOPSIS and ELECTREE III and we used also DEA and econometric models. In this paper we follow this analysis with newer data about the Czech regions.

The aim of this article is to compare the Czech regions from the economic point of view via DEA models (chapter 2) and MCEA (chapter 3) and also compare the results of these methods applied on data from 2012 and 2016 (chapter 4).

2 DATA ENVELOPMENT ANALYSIS MODEL

The basic idea of DEA models consists in estimation of an efficient frontier that defines production possibility set of the problem. Based on the set of available decision making units (DMUs) DEA estimates so-called efficient frontier, and projects all DMUs onto this frontier. If a DMU lies on the frontier, it is referred to as an efficient unit, otherwise inefficient. DEA models can be oriented to inputs or outputs. In the case of input oriented models we assume fixed level of all outputs (CCR-I) and the output oriented model assumes fixed level of all inputs (CCR-O) (Cooper, Lawrence and Zhu, 2004). These models are used if we assume constant return to scale. In the case of variable return to scale, we work with BCC (Banker, Charnes, Cooper) models that can be similarly input or output oriented. The review and detailed information about DEA models were described by Cooper, Lawrence and Zhu (2004) and by Cooper, Seiford and Tone (2006). The basic idea for the efficiency calculation is to maximize the rate of weighted sum of outputs divided by weighted sum of inputs (Charnes, Coper and Rhodes, 1978).

The mathematical model of DEA considers r DMUs $U_1, U_2, \dots, U_r$ with m inputs ($i = 1,2,\dots,m$) and n outputs ($j = 1,2,\dots,n$). The vector of input values of DMU k ($k = 1,2,\dots,r$) is denoted as $\mathbf{x}_k = (x_{1k}, x_{2k}, \dots, x_{mk})^{\mathrm{T}}$ and matrix of all input values for all DMUs is denoted as $\mathbf{X} = \{x_{ik}, i = 1,2,\dots,m, k = 1,2,\dots,r\}$. Similarly, the vector of output values of DMU k is denoted as $\mathbf{y}_k = (y_{1k}, y_{2k}, \dots, y_{nk})^{\mathrm{T}}$ and matrix of all output values for all DMUs is denoted as $\mathbf{Y} = \{y_{jk}, j = 1,2,\dots,n, k = 1,2,\dots,r\}$. The relative technical efficiency of given DMU q can be generally expressed as ratio of weighted sum of outputs and weighted sum of inputs

$$TE_q = \frac{\sum_{j=1}^{n} u_j y_{jq}}{\sum_{i=1}^{m} v_i x_{iq}}, \tag{1}$$

where v_i, $i = 1, 2, \dots, m$ is a weight for i-th input and u_j, $j = 1, 2, \dots, n$ is a weight for j-th output.

CCR input oriented model has the following form (the other models are determined from this model and are described by Cooper, Lawrence and Zhu (2004)):

maximize

$$z = \mathbf{u}^{\mathrm{T}}\mathbf{y}_q,$$

subject to

$$\begin{aligned} \mathbf{v}^{\mathrm{T}}\mathbf{x}_q &= 1, \\ \mathbf{u}^{\mathrm{T}}\mathbf{Y} - \mathbf{v}^{\mathrm{T}}\mathbf{X} &\leq 0, \\ \mathbf{u} &\geq \varepsilon, \\ \mathbf{v} &\geq \varepsilon. \end{aligned} \tag{2}$$

The efficient unit U_q lies on the efficient frontier in case that the optimal efficiency (calculated by the model) $z = 1$. The inefficient units have z lower than 1 (in CCR-I model, higher than 1 in CCR-O models).

3 MULTI-CRITERIA EVALUATION OF ALTERNATIVES MODEL

In the second step of our analysis we used the multi-criteria evaluation of alternatives (MCEA). MCEA belongs to the category of discrete multi-criteria decision making models where all the alternatives (a_1, a_2, …, a_p) and criteria (f_1, f_2, …, f_k) are known. To solve this kind of model it is necessary to know the preferences of the decision maker. These preferences can be described by aspiration levels (or requirements), criteria order or by the weights of the criteria. For our analysis we used three methods – WSA, TOPSIS and ELECTREE III that need only decision matrix **Y** and weights of criteria as inputs – for other ELECTRE methods more additional parameters are necessary (Figueira, Greco and Ehrgott, 2005; Zopounidis and Pardalos, 2010).

WSA (Weighted Sum Approach) sorts the alternatives based on the values of their utility functions which in this case are assumed to be linear. It requires the information about the weights of the criteria. Higher value of utility means better alternative. The basic concept of the TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) method is that the best alternative should have the shortest distance from the ideal alternative and the farthest from the basal alternative. The method is also able to rank the alternatives using the relative ratio of distance (*R.R.D.*). of the alternatives from the basal alternative. Higher relative ratio of distance means better alternative. ELECTREE III method uses pairwise comparison of the alternatives and summarizes the weights of the criteria where the alternative is better than the other. According to the strength of the preference the final indifference classes are made. It is necessary that all the alternatives are nondominated (so there is no alternative that is better in at least one criterion and no worse in any other) (Figueira, Greco and Ehrgott, 2005).

4 RESULTS

In our case we first compared 14 regions with respect to available data from years 2012 and 2016 using DEA models. The models have 5 inputs (unemployment rate, economic activity, average wage, average age and free workplaces per capita) and 3 outputs (income per capita, consumption per capita and investments per capita). The DEA model assumes inputs are minimized and outputs are maximized for efficient units. In this case we had to transform input data to minimize them (especially economic activity and average wage). The results (objective values) obtained after transformation from DEA models based on data of the year 2016 are in Table 1. Note that the last column (*AAD*) displays average absolute deviation from efficient frontier. Zero values denote regions those are efficient in all four models.

From the data in Table 1 we can easily conclude that only two regions are efficient with respect to all four DEA models. There are Prague, the capital city and South Moravian Region. This result could be expected as these are regions with the two largest cities in the Czech Republic. Prague has more than two times higher the consumption per capita than other regions, the highest income per capita and eight times higher investments than other regions. It is also the best in the economic activity, has the highest average wage and the lowest unemployment rate. On the other hand the South Moravian Region has the lowest consumption per capita, but it has the second highest income per capita and it is above the average in other criteria. The absolutely inefficient are South Bohemian and Plzeň Region with average absolute deviation more than 21% and Pardubice Region with even more than 26%. These regions loose mainly is the criterion free working places per capita.

Region	*CCR-I*	*CCR-O*	*BCC-I*	*BCC-O*	***AAD***
Prague, the Capital City	1	1	1	1	**0**
Central Bohemian	0,8862	1,1283	1	1	**0,0605**
South Bohemian	0,7641	1,3085	0,9839	1,3042	**0,2162**
Plzeň	0,7611	1,3136	0,9728	1,2977	**0,2193**
Karlovy Vary	0,8184	1,2218	0,9917	1,2168	**0,1571**
Ústí nad Labem	0,9744	1,0263	1	1	**0,0130**
Liberec	0,8417	1,1880	0,9865	1,147	**0,1267**
Hradec Králové	0,8987	1,1127	0,9773	1,1041	**0,0852**
Pardubice	0,7265	1,3760	0,9763	1,3709	**0,2610**
Vysočina	0,9609	1,0407	1	1	**0,0200**
South Moravian	1	1	1	1	**0**
Olomouc	0,9317	1,0733	1	1	**0,0354**
Zlín	0,8524	1,1731	0,975	1,1642	**0,1275**
Moravian-Silesian	0,9797	1,0208	1	1	**0,0103**

Table 1 Data envelopment analysis – results (year 2016)

For the second analysis we use the same criteria as in DEA model and the weight of each criterion was equal ($v_j = 1/8, j = 1, \ldots, 8$). Note that all alternatives for data of the year 2012 are nondominated, but in the case of year 2016 Pardubice and Plzeň Regions are dominated by Prague, the capital city. This fact does not influence the order of the regions obtained via ELECTRE III method.

We obtained the results displayed in Table 2. Column *Utility* displays weighted sum gained by WSA and *Rank* presents order with respect to maximal utility. Similarly, *R.R.D.* column displays the relative ratio of distance to the basal alternative by TOPSIS method and *Rank* presents order with respect to maximal relative ratio of this distance. Column *ELECTRE III Rank* presents order with respect to ELECTRE indifference classes. Regions with respect to data of the year 2016 were divided into ten indifference classes. The column *Average rank* is the mean of three previous ranks. *Total rank* sorts regions with respect to average rank.

MCEA with equal weights places Prague, the capital city at the first place before Central Bohemian Region (that has the lowest average age and very good level of income per capita, unemployment rate and economic activity) and Vysočina Region as third (especially because of the second highest investment). South Moravian Region is the fourth and Liberec Region has the fifth place. The last one is Zlín Region (although it has no worst value in any criterion, all criteria values are below average). Note that dominated Pardubice Region has finished as eleventh and dominated Plzeň Region was the seventh together with South Bohemian Region.

Region	*WSA*		*TOPSIS*		*ELECTRE III.*	*Average rank*	***Total rank***
	Utility	*Rank*	*R. R. D.*	*Rank*	*Rank*		
Prague, the Capital City	0,8340	1	0,6720	1	1	1,0	**1**
Central Bohemian	0,4653	2	0,4779	2	5	3,0	**2**
South Bohemian	0,2812	8	0,3128	8	7	7,7	**7**
Plzeň	0,2867	6	0,2499	14	3	7,7	**7**
Karlovy Vary	0,2730	10	0,2607	13	10	11,0	**12**
Ústí nad Labem	0,2829	7	0,3991	3	10	6,7	**6**
Liberec	0,2991	5	0,3570	6	8	6,3	**5**
Hradec Králové	0,2735	9	0,2914	9	6	8,0	**9**
Pardubice	0,2462	12	0,2762	11	9	10,7	**11**
Vysočina	0,3192	4	0,3756	4	2	3,3	**3**
South Moravian	0,3890	3	0,3654	5	4	4,0	**4**
Olomouc	0,2630	11	0,3355	7	10	9,3	**10**
Zlín	0,2148	14	0,2707	12	10	12,0	**14**
Moravian-Silesian	0,2161	13	0,2911	10	10	11,0	**13**

Table 2 Multicriteria evaluation of alternatives – results (year 2016)

Now we can compare results of both parts. Table 3 summarizes the previous ranks for data of the year 2016. Remember that only two regions are efficient by all four DEA models – Prague, the capital city (the winner also by MCEA models) and South Moravian Region (it was 4th by MCEA models).We can see that both DEA and MCEA models placed Prague as the first and the total rank for this region is so the first. The other efficient region – South Moravian – has the second total rank before Vysočina Region and Central Bohemian Region. The fifth is Ústí nad Labem Region.

In the year 2016 two regions were dominated – Plzeň Region and Pardubice Region. In Table 3 we see that both regions are placed on the bottom by DEA models (the thirteen and fourteen position). But from the Table 2 we can see that Plzeň region is placed in the third indifference class and the difference between DEA and MCEA results is very large in this case. The difference between results is given by the fact that Pilsen Region has better values in 7 out of 8 criteria than Zlín Region (which was the last in MCEA) but it is the second worst in the last criterion (free working places per capita) and as DEA calculates the ratios of outputs and inputs, this very bad input leads to the second worst position of the region in DEA models. The same influence we see in the situation of the Central Bohemian or South Bohemian Regions. Similar it is in Moravian-Silesian Region which was the worst in the unemployment rate and the investment per capita and that is why the MCEA methods put it in the worst position than DEA models for which the main fact is that this region has the best value in the criterion free working places per capita.

This order is very different from ranks based on data of the year 2012. In 2012 eight regions were marked as efficient by all four DEA models (see zero values in Table 4, column AAD(2012)). Also MCEA methods placed these efficient regions at the top except of Olomouc and Hradec Králové Regions that are the ninth and tenth by MCEA. And vice-versa inefficient South Moravian Region has the sixth place by MCEA methods in 2012. Table 4 summarized these results. The columns *AAD* display the average absolute deviations gain by DEA models based on data of 2012 and 2016. The column *M.I.* displays measure of improvement as a difference of AAD between years 2012 and 2016. Positive values mean approximation to the efficient frontier, negative values denotes removing regions. Similarly, columns *Rank* display total ranks of regions set by MCEA methods based on data of the year 2012 and 2016 and *M.I.* displays measure of improvement as a difference of ranks. Again positive values mean improvement and negative values marked regions that became worse.

Region	*DEA*	*MCEA*	*Average Rank*	*Total Rank*
Prague, the Capital City	1	1	1	1
Central Bohemian	7	2	4,5	4
South Bohemian	12	7	9,5	10
Plzeň	13	7	10	11
Karlovy Vary	11	12	11,5	12
Ústí nad Labem	4	6	5	5
Liberec	9	5	7	6
Hradec Králové	8	9	8,5	9
Pardubice	14	11	12,5	14
Vysočina	5	3	4	3
South Moravian	1	4	2,5	2
Olomouc	6	10	8	7
Zlín	10	14	12	13
Moravian-Silesian	3	13	8	7

Table 3 Recapitulation of results (year 2016)

From the Table 4 we can see that Prague is the absolute winner. It was efficient in 2012 and also in 2016 and in both years MCEA methods placed it at the first position. In 2016 Prague is the first by each of all three methods of MCEA, in 2012 it is winner by WSA and TOPSIS and it is placed in the second indifference class by ELECTRE III. This fact is clear from budget point of view. Prague, the capital city, has large income that divides between grants, investments and also expenditures of regional assembly. These expenditures are much higher than in the case of others regions. In DEA and also MCEA models the other economic indicators play also the role and so average age and free working places per capita are not so important with respect to income and investments where Prague places the first rank.

Table 4 also shows that only three regions have a positive measure of improvement by DEA models – Moravian-Silesian Region (with improvement about more than 6,7%), Ústí nad Labem Region (improvement about almost 4%) and South Moravian Region (that is newly efficient region). All these three regions has the positive measure of improvements also by MCEA methods. Together with Prague these three regions can be denoted as progressive. The improvement is evident in almost all criteria except of the investment per capita, the average age and the free working places per capita which were better in all regions in the year 2012.

On the other hand 8 regions became worse by DEA and MCEA in 2016 compared with 2012. Six of them – South Bohemian Region, Plzeň and Karlovy Vary Regions, Pardubice, Olomouc and Zlín Regions – can be denoted as degressive by both DEA and MCEA models. The other two regions – Central Bohemian and Vysočina – did not change their rank by

MCEA but slightly moved away the efficient frontier. Liberec and Hradec Králové Regions are worse by DEA models but their ranks by MCEA are higher in 2016 than in 2012. The rank of Liberec Region even increased about five places. The reason was mentioned above – in the MCEA methods (especially when equal weights are used) the differences between good and bad indicators are wiped out, whereas DEA can highlight this ratio.

Region	DEA			MCEA		
	AAD(2012)	AAD(2016)	***M.I.***	Rank(2012)	Rank(2016)	***M.I.***
Prague, the Capital City	0,0000	0,0000	**0,0000**	1	1	**0**
Central Bohemian	0,0000	0,0605	**-0,0605**	2	2	**0**
South Bohemian	0,0000	0,2162	**-0,2162**	5	7	**-2**
Plzeň	0,0000	0,2193	**-0,2193**	4	7	**-3**
Karlovy Vary	0,0000	0,1571	**-0,1571**	7	12	**-5**
Ústí nad Labem	0,0517	0,0130	**0,0387**	8	6	**2**
Liberec	0,0301	0,1267	**-0,0966**	10	5	**5**
Hradec Králové	0,0000	0,0852	**-0,0852**	10	9	**1**
Pardubice	0,0093	0,2610	**-0,2517**	10	11	**-1**
Vysočina	0,0000	0,0200	**-0,0200**	3	3	**0**
South Moravian	0,0233	0,0000	**0,0233**	6	4	**2**
Olomouc	0,0000	0,0354	**-0,0354**	9	10	**-1**
Zlín	0,0030	0,1275	**-0,1245**	13	14	**-1**
Moravian-Silesian	0,0775	0,0103	**0,0672**	14	13	**1**

Table 4 Measure of improvement of average absolute deviation (DEA) and rank (MCEA) between 2012 and 2016

5 CONCLUSIONS

Similarly as in (Kuncova and Seknickova, 2013) the aim of this contribution was to analyze the economic position of the 14 regions of the Czech Republic by two different approaches. At first DEA models were used and then multi-criteria evaluation of alternatives using 3 methods was applied. It is clear that DEA models are not able always to create the order of the alternatives but the results in this paper are close to the multi-criteria models (Prague, the capital city, was placed as the first by DEA and MCEA methods in 2012 and also in 2016). From the results of the year 2016 we can see that Prague do not need to be the best region, as South Moravian Region as well as Vysočina Region and Central Bohemian Region are placed on the top. From comparison of years 2012 and 2016 we can conclude that South Moravian Region starts to be better together with Ústí nad Labem and Moravian-Silesian Regions.

Acknowledgements

The research was supported by the grant No. 16-01821S of the Czech Science Foundation and also by the Internal Grant Agency of the University of Economics in Prague IGA F4/57/2017.

References

[1] Charnes A., Cooper W.W. and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research,* 2(6), pp. 429–444.

[2] Cooper, W. W., Lawrence, M.S. and Zhu, J. (2004). *Handbook on Data Envelopment Analysis*. Kluwer Academic Publishers, Norwell.
[3] Cooper, W.W., Seiford, L.M. and Tone, K. (2006). *Introduction to Data Envelopment Analysis and Its Uses*. Springer Science + Business Media Inc., New York.
[4] Czech Statistical Office, (2018). [online] Available at: https://www.czso.cz/csu/czso/ [Accessed 17 Jan 2018]
[5] Dincer, S. E. (2011). Multi-criteria Analysis of Economic Activity for European Union Member States and Candidate Countries: TOPSIS and WSA Applications. *European Journal of Social Sciences*, 21(4), pp. 563-572.
[6] Figueira, J., Greco, S. and Ehrgott M. (2005). *Multiple Criteria Decision Analysis – State of the Art Surveys*. New York : Springer Science + Business Media Inc.
[7] Friebel, L. and Friebelova, J. (2012). Quantitative evaluation of life quality of Czech districts. In: *30th International Conference Mathematical Methods in Economics*. Karvina: Silesian University, School of Business Administration, pp. 190-195.
[8] Furkova, A. (2014). Are Slovak Economic Faculties Efficient in Their Scientific Activities? DEA Aproach. In: *Quantitative Methods in Economics (Multiple Criteria Decision Making XVII)*. Bratislava: Ekonóm, pp. 53–57.
[9] Hashemabadi, A. G. and Razmi, M. J. (2014). Studying Iran's Tourism Industry Position in Middle-East using Tourism Development Indicators and TOPSIS Method. *Advances in Management & Applied Economics*, 4(5), pp. 85-97.
[10] Koutroumanidis, T., Papathanasiou, J. and Manos, B. (2002). A multicriteria analysis of productivity of agricultural regions of Greece. *Operational Research*, 2(3), pp. 339–346.
[11] Kuncova, M. and Doucek, P. (2011). Comparison of the Cluster Analysis and the Methods of the Multi-criteria Evaluation of Alternatives Used to Create a Groups of Countries Similar in the Take up of the Internet Services. In: *29th Conference Mathematical Methods in Economics* 2011, Praha: Professional Publishing, pp. 419–424.
[12] Kuncova, M. and Seknickova, J. (2013). Economic Analysis of Czech Regions by Econometric Models, MCDM and DEA. In: *31th Conference Mathematical Methods in Economics 2013*. Jihlava: College of Polytechnics, pp. 500–505.
[13] Latuszynska, A. (2014). Multiple-criteria Decision Analysis Using TOPSIS Method for Interval Data in Research Into the Level of Information Society Development. *Folia Oeconomica Stetinensia*, 13(2), pp. 63-76.
[14] Melecky, L. and Stanickova, M. (2012). Assessment of efficiency in Visegrad countries and regions using DEA models. *Central European Review of Economic Issues,* 15, pp. 145-156.
[15] Zopounidis, C. and Pardalos, P. (2010). *Handbook of Multicriteria Analysis*. Springer, Berlin.

Author's address
Ing. Martina Kuncová, Ph.D. and Mgr. Jana Sekničková, Ph.D.
University of Economics Prague, Faculty of Informatics and Statistics, Dpt. of Econometrics
W. Churchill Sq. 4, 130 67 Prague 3
Czech Republic
email: kuncovam@vse.cz; jana.seknickova@vse.cz

EVALUATION OF THE CZECH REGIONS FROM THE TOURISM INFRASTRUCTURE POINT OF VIEW USING TOPSIS METHOD

Martina Kuncová, College of Polytechnics Jihlava,
Zuzana Tučková, College of Polytechnics Jihlava

Abstract
The development of regions inside a country can be compared and measured via several mathematical, statistical or econometric models. Reasons for these analysis might be different – the comparison of economic activity, ICT development, natural resources, cultural heritage etc. The aim of this article is to compare the 14 Czech regions from the tourism infrastructure point of view. In the analysis 22 criteria are used and the TOPSIS method, as one of the multi-criteria evaluation methods, is applied so as to find out the order of the regions in the years 2015 and 2016. Afterwards the main differences among the regions and between the selected years are mentioned.

Keywords: *Tourism infrastructure, Czech regions, multi-criteria comparison*

JEL Classification: C44, Z32
AMS Classification: 90B50, 91B06

1 INTRODUCTION

Tourism industry belongs to the one of the largest economic drivers. World Travel and Tourism Council's latest annual research (WTTC, 2017) shows that the travel and tourism direct, indirect and induced impact accounted for more than 10% of global GDP and the sector grew much faster in 2017 than economy as a whole. In the Czech Republic the contribution of this sector to total GDP was 2.6% reflecting mainly the economic activity generated by industries such as hotels, travel agents, transportation services or the activities of the restaurant and leisure industries directly supported by tourists. According to the WTTC economic impact research the Czech Republic ranks 52nd in total direct contribution to GDP (out of 185 countries), 69th in the total contribution to employment, and 53rd in investment. Although it might seem to be a good positions still it is below of the World and European average so there is a possibility of improvement.

The development of infrastructure inside the countries and regions is one of the main factors influencing the growth of tourism and the overall economic growth of the region. According to Bookman and Bookman (2007) the efficient infrastructure is a vital factor that ensures the effective functioning of a country's economy. Tourism infrastructure is that part of the infrastructure that includes facilities and services that are used not only by the inhabitants of the region but also by tourists (Goeldner and Ritchie 2009; Bagheri, Shojaei and Khorami, 2018). This infrastructure has the potential not only to increase the tourism itself but also to increase the competitiveness and economic development of the region.

2 METHODOLOGY AND DATA

For the comparison of countries, regions or cities from the tourism point of view various methods can be used such as statistical methods, econometric models, data envelopment analysis (DEA) models, or multi-criteria decision making (MCDM) models, especially multi-criteria evaluation of alternatives (MCEA). The theory of multi-criteria evaluation of alternatives offers many different methods for this kind of problems, such as WSA, TOPSIS, ELECTRE, PROMETHEE, VIKOR, AHP etc. A systematic review of the methods and techniques used in the tourism-connected papers and articles was described by Mardani et. al. (2016). Selection of the method is influenced by the aim of the decision as some methods are able only to find good (or acceptable) and bad (no-acceptable) alternatives. According to other similar papers and articles and according to the similar methodology used by Bagheri, Shojaei and Khorami (2018) we use the TOPSIS method in this paper either.

The model of multi-criteria evaluation of alternatives (MCEA) contains a list of alternatives $A = \{a_1, a_2, \mathrm{K}\ , a_p\}$, a list of criteria $F = \{f_1, f_2, \mathrm{K}\ , f_k\}$ and an evaluation of the alternatives by each criterion in the criteria matrix (Lai and Liu, 1994; Figueira, Greco and Ehrgott, 2005).

TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) method is able to rank the alternatives using the relative index of distance of the alternatives from the ideal and non-ideal alternative. Higher relative index of distance means better alternative. The user must supply only the information about the weights of criteria (Lai and Liu, 1994). This method was used in various research topics. The TOPSIS method was first used in 1981 (Hwang and Yoon, 1981). In literature, it is mentioned in a variety of comparisons and evaluations, for example for the evaluation of companies (Kuncová, Hedija and Fiala, 2016; Yalcin, Bayrakdaroglu and Kahraman, 2012; Wang and Hsu, 2004), for the selection of a suitable material (Jee and Kang, 2000)¨or for the comparison of the electricity production methods (Sarkar, 2014). However, it is also possible to compare the impact of the economic crisis on the selected economic indicators in 2002-2009 in selected EU countries and in Turkey (Mangir and Erdogan, 2011), to compare countries from the perspective of the development of the information society between 2005 and 2010 (Latuszynska, 2014) or from the ICT development point of view (Kuncová and Doucek, 2011) or to compare the Czech regions on the basis of economic indicators (Kuncová and Sekničková, 2013). Except of these articles the TOPSIS method was widely used in tourism-related analyzes, such as finding the position of countries with regard to tourism (Hashemabadi and Razmi, 2014), rating of the suitable areas for tourism in Iran (Anvari, 2016), evaluation of online travel agencies (Soleymaninejad, Shadifar and Karimi, 2016) or in the article which we follow and which was aimed at the comparison of the condition of tourism infrastructure in Iranian provinces (Bagheri, Shojaei and Khorami, 2018). The output provided by TOPSIS is a complete arrangement of possible alternatives with respect to the distance to both the ideal and the basal (non-ideal) hypothetical alternatives incorporating relative weights of criterion importance. The required input information includes decision matrix Y with the information about all selected alternatives $a_1, .., a_p$ according to all criteria $f_1, .., f_k$ and weight vector of these criteria. Finally the relative ratio of distance (*R.R.D.*) of the alternatives from the basal alternative is calculated and on this basis the order of the alternatives is created - the higher the ration the better position the alternative has (detailed description of steps and notation in Hwang and Yoon, 1981; Lai and Liu, 1994).

Analysis and comparison of the tourism infrastructure can be based on hard and soft types of criteria, sometimes called social and economic sub-sectors (Ghosal, 2013). According to the analysis made by Bagheri, Shojaei and Khorami (2018) and Pearce and Wu (2015) we used similar criteria to compare 14 Czech regions. Data was taken from the Czech Statistical Office reports (CSO, 2017). Bagheri, Shojaei and Khorami (2018) suggested 20 criteria for the tourism infrastructure comparison. As some of them are not available or suitable for the Czech Republic regions, we have selected similar 22 criteria and we have used similar weights (Table 1) like in the analysis of the Iran's provinces (Bagheri, Shojaei and Khorami, 2018). The criteria are aimed at:

- healthcare situation (criterion 1-5),
- cultural and natural heritage of the region (criterion 6-10 and 19),

- environmental characteristics (criterion 11-13),
- transportation infrastructure (criterion 14-16),
- development of ICT and energy (criterion 17-18),
- number of enterprises (criterion 20-21),
- housing situation (criterion 22).

All these criteria describe somehow the tourism infrastructure – only purely economic indicators are not included (however, these are the basis of most of the above-mentioned indicators). For the comparison data from the years 2015 and 2016 was used (CSO, 2017).

Table 1: Criteria for the comparison and their weights (Bagheri, Shojaei and Khorami, 2018, own criteria customization)

	Criterion name	**weight**
1	Hospitals	0.039962
2	Other independent health establishments	0.027570
3	Specialized therapeutic institutions	0.027570
4	Pharmacies, incl. Detached workplaces	0.028934
5	Beds total	0.054192
6	Museums, galleries, and monuments in operation, total	0.131626
7	Historical and other monuments accessible to visitors for admission	0.117396
8	National parks	0.010672
9	Protected landscape areas	0.010672
10	Small-size protected areas, total	0.010672
11	Share of population supplied with water from public water supply systems (%)	0.020159
12	Share of population living in houses connected to public sewerage systems (%)	0.041504
13	Public wastewater treatment plants (WWTPs)	0.033203
14	Length of operated railway lines	0.022531
15	Length of roads and motor-ways	0.030831
16	Motor vehicles (total) by region	0.079450
17	Fixed broadband subscriptions	0.073521
18	Installed system power (MW)	0.022531
19	Afforestation total	0.046247
20	Number of enterprises-industry	0.040318
21	Number of enterprises-construction	0.040318
22	Dwellings completed	0.090122

3 RESULTS AND DISCUSSION

Based on the data of the Czech Statistical Office (CSO, 2017), the 14 Czech regions were compared via selected 22 criteria in two years 2015 and 2016 by TOPSIS method. We obtained the results displayed in Table 2. Column *R.R.D.* displays the relative ratio of distance to the basal alternative by TOPSIS method and *Rank* presents order with respect to maximal relative ratio of this distance.

In both years the Central Bohemian Region is on the top followed by Prague, the capital city, which is usually the best in comparison of regions (especially from the economic point of view). Although Prague is famous through its cultural monuments, the Central Bohemian Region is better in the criteria connected with culture and natural heritage (in criteria 6,7,9 and 10). As it has larger area than the capital city, it is also better in the length of railways and motorways. But behind the selected criteria the influence of the position of the Central Bohemian Region close to the capital city is evident. South Moravian Region was the third one. It is very close in most criteria to the Central Bohemian Region but it is worse in dwellings completed, transportation infrastructure or cultural monuments. Also here the influence of the second biggest city Brno, which is located in this region, is evident.

Table 2: Results of the TOPSIS method

Region	2015		2016	
	Rank	*R.R.D.*	*Rank*	*R.R.D.*
Central Bohemian	1	0.73702	1	0.74245
Prague, the Capital City	2	0.57090	2	0.58458
South Moravian	3	0.53357	3	0.53819
South Bohemian	4	0.40277	4	0.42546
Moravian-Silesian	5	0.38511	5	0.40686
Ústí nad Labem	6	0.38136	6	0.38576
Hradec Králové	7	0.32492	7	0.32764
Plzeň	8	0.30023	8	0.32128
Vysočina	9	0.26868	9	0.27302
Pardubice	10	0.22480	10	0.21636
Olomouc	11	0.19003	11	0.21087
Zlín	13	0.17259	12	0.19496
Liberec	12	0.18711	13	0.16751
Karlovy Vary	14	0.07845	14	0.08410

All these three regions have the final score very high and all were better in 2016. The improvement of other regions between the years 2015 and 2016 is visible from the final score in 2016 compared to the score in 2015. Only two regions were worse in 2016: Pardubice and Liberec Regions and only Liberec Region lost its position in 2015. The worst region from the tourism infrastructure point of view seems to be Karlovy Vary Region although it has famous spa but it belongs to the smaller regions. The same trends we see when we take all regions and both years together (so we have 28 alternatives, all are non-dominated). The best region was the Central region in 2016 followed by the same region in 2015 (Table 3). The third and fourth positions were for Prague in 2016 and 2015 followed by the South Moravian region in 2016 and 2015. We see the first differences in the 13th place, from which it is clear that the situation in the Hradec Králové region deteriorated in comparison with 2015 in 2016 (because of the worse afforestation and fewer museums). It was similar in other regions (Vysočina, Pardubice, Liberec) where the tourism situation was better in 2015 than in 2016. The biggest shift in the *R.R.D.* score was recorded by the Zlín region (mainly because of the higher number of the small–size protected areas, more motor vehicles in region, higher installed electricity network power and more dwellings completed in 2016 compared to 2015).

Table 3: Results of the TOPSIS method for both years data

Region	*Rank*	*R.R.D.*	Region	*Rank*	*R.R.D.*
Central Bohemian_2016	1	0.69695	Hradec Králové_2016	15	0.30377
Central Bohemian_2015	2	0.68712	Plzeň_2015	16	0.29023
Prague, the Capital City_2016	3	0.56203	Vysočina_2015	17	0.26435
Prague, the Capital City_2015	4	0.54907	Vysočina_2016	18	0.25577
South Moravian_2016	5	0.50900	Pardubice_2015	19	0.22227
South Moravian_2015	6	0.50044	Olomouc_2016	20	0.21136
South Bohemian_2016	7	0.40350	Pardubice_2016	21	0.20833
South Bohemian_2015	8	0.38876	Zlín_2016	22	0.19999
Moravian-Silesian_2016	9	0.38651	Olomouc_2015	23	0.19755
Moravian-Silesian_2015	10	0.37873	Liberec_2015	24	0.18466
Ústí nad Labem_2016	11	0.36594	Zlín_2015	25	0.17808
Ústí nad Labem_2015	12	0.36126	Liberec_2016	26	0.16548
Hradec Králové_2015	13	0.31973	Karlovy Vary_2016	27	0.09153
Plzen_2016	14	0.30673	Karlovy Vary_2015	28	0.08754

The results are positively correlated with the population but the highest positive correlation has been detected between the results of the TOPSIS method and the employment rate. On the other hand the correlation between the final score from the TOPSIS method and the area of the region is not too strong as it might be expected – only 0.44. The results indicate that, from the point of view of tourism infrastructure, the regions with higher population and higher employment rate are better developed. However, the large increase in tourism may have a negative effect on frequently visited places in the form of a decline in indigenous population.

The same results we obtain when other MCEA methods were used (WSA, PROMETHEE II) – in the pairwise comparison with usual criterion type in PROMETHEE II method only the South Moravian Region and Prague change its places in the order (Prague was the 3rd region). So we can conclude that the results are not influenced by he method.

Although the situation in the Czech Republic and Iran in Bagheri, Shojaei and Khorami study (2018) is not similar the results of the TOPSIS method show that in Iran only the capital city province has the final score higher than 0.5 but in the Czech Republic 3 regions have the score higher than 0.5 that means 21% of the regions (compared to 3% in Iran) and in Iran 87% of provinces has the score lower than 0.2 compared to 21% in the Czech Republic in 2016. From this point of view the regions in the Czech Republic are more similar to each other from the tourism infrastructure point of view than the provinces in Iran (but of course the area plays an important role).

4 CONCLUSIONS

Tourism industry is currently an important area for country development. Several techniques could be used to measure the impact of tourism or the interactions between tourism and other industries. The tourism infrastructure of the region is given by hard and soft factor. Except of economic ones there are factors connected with the cultural and natural heritage, environment, energy, ICT, number of enterprises or households, etc. that have also an important influence on the tourism development. According to the analysis comparing 31 Iran's provinces (Bagheri, Shojaei and Khorami, 2018) and using TOPSIS method, all 14 regions of the Czech Republic were compared via the same method and with similar criteria (22 criteria were used) with data from the year 2016. The results showed that the Central Bohemian Region was the best one from the tourism infrastructure point of view followed by Prague, the capital city, and the South Moravian Regions. According to the results we see that tourism infrastructure is highly dependent on the region's population and employment, which affects the region's amenities and makes it attractive for tourism as well. The results are, of course, dependent on the choice of criteria, the choice of weights and the choice of method. However, when using equal weights, there have been only slight changes in the order of the regions.

Acknowledgements: This work was supported by the grant "Analysis of the relationship between sustainable tourism and the performance of tourism enterprises or services" of the College of Polytechnics Jihlava.

References

[1] Anvari, M., R. (2016). Rating Suitable Areas of Tourism in Iran, Case Study: Kerman Province. *The IIOAB Journal*, 7(4), pp. 401-412.

[2] Bagheri, M., Shojaei, P. and Khorami, M., T. (2018). A comparative survey of the condition of tourism infrastructure in Iranian provinces using VIKOR and TOPSIS. *Decision Science Letters*, 7 (2018), pp. 87-102.

[3] Bookman, M. Z. & Bookman, K. R. (2007). *Medical tourism in developing countries.* New York, NY:Palgrave Macmillan.

[4] CSO (2017). Czech Statistical Office – comparison of the Czech Republic regions [online] Available at: https://www.czso.cz/csu/czso/srovnani-kraju-v-ceske-republice-2016 [Accessed 1 Feb. 2018].

[5] Figueira, J., Greco, S. and Ehrgott M. (2005). *Multiple Criteria Decision Analysis – State of the Art Surveys*. New York : Springer Science + Business Media Inc.

[6] Ghosal, S. (2013). Soft or hard: infrastructure matters in rural economic empowerment. *Journal of Infrastructure Development*, 5 (2), pp. 137-149.

[7] Goeldner, C. R. and Ritchie, J. R. (2009). *Tourism: principles, practices, philosophies*. (11th Edition), New Jersey, NJ: John Wiley & Sons.

[8] Hashemabadi, A. G. and Razmi, M. J. (2014). Studying Iran's Tourism Industry Position in Middle-East using Tourism Development Indicators and TOPSIS Method. *Advances in Management & Applied Economics*, 4 (5), pp. 85-97.

[9] Hwang, C. L. and Yoon, K. (1981). *Multiple Attribute Decision Making: Methods and Applications,* New York: Springer- Verlag.

[10] Jee, D. H. and Kang, J. K. (2000). A method for optimal material selection aided with decision making theory. *Materials and Design*, 21 (3), pp. 199-206

[11] Kuncová, M. and Doucek, P. (2011). Comparison of the Cluster Analysis and the Methods of the Multi-criteria Evaluation of Alternatives Used to Create a Groups of Countries Similar in the Take up of the Internet Services. *Proceedings of the 29th Conference Mathematical Methods in Economics* 2011, Janská Dolina, pp. 419–424.

[12] Kuncová, M., Hedija and V., Fiala, R. (2016). Firm Size as a Determinant of Firm Performance: The Case of Swine Raising, *AGRIS on-line Papers in Economics and Informatics*, 8(3), pp. 77 - 89.

[13] Kuncová, M. and Sekničková, J. (2013). Economic Analysis of Czech Regions by Econometric Models, MCDM and DEA. *Proceedings of the 31th Conference Mathematical Methods in Economics 2013*. Jihlava, College of Polytechnics, pp. 500–505.

[14] Lai,Y., J., and Liu, T., Y. (1994). TOPSIS for MODM. *European Journal of Operation Research*, 76(3), pp. 486–500.
[15] Latuszynska, A. (2014). Multiple-criteria Decision Analysis Using TOPSIS Method for Interval Data in Research Into the Level of Information Society Development. *Folia Oeconomica Stetinensia*, 13(2), pp. 63-76.
[16] Mangir, F. and Erdogan, S. (2011). Comparison of Economic Perforamnce among Six Countries in Global Financial Crisis: The Application of Fuzzy TOPSIS Method. *Economics, Management and Financial Markets*, 6 (2), pp. 122-136.
[17] Mardani, A., Jusoh, A., Zavadskas, E., K., Kazemilari, M., Ungku, N., U., A. and Khalifah, Z. (2016). Application of Multiple Criteria Decision Making Techniques in Tourism and Hospitality Industry: a Systematic Review. *Transformations in Business & Economics*, 15/1 (37), pp.192-213.
[18] Pearce, P. L. and Wu, M. Y. (2015). Soft infrastructure at tourism sites: identifying key issues for Asian tourism from case studies. *Tourism Recreation Research*, 40 (1), pp. 120-132.
[19] Sarkar, A. (2014). A TOPSIS Method to Evaluate the Technologies. *International Journal of Quality & Reliability Management*. 31(1), pp. 2-13.
[20] Soleymaninejad, M., Shadifar, M. and Karimi, A. (2016). Evaluation of Two Major Online Travel Agencies of US Using TOPSIS Method. Digital Technologies, 2(1), pp. 1-8
[21] Wang, T.-C. and Hsu, J., C. (2004). Evaluation of the Business Operation Performance of the Listing Companies by Applying TOPSIS Method. In: *Proceedings of the IEEE International Conference on Systems*, Man & Cybernetics. The Hague: IEEE, pp. 286–1291.
[22] WTTC, (2017). Travel & Tourism Economic Impact – Czech Republic. [online] Available at: https://www.wttc.org/-/media/files/reports/economic-impact-research/countries-2018/czechrepublic2018.pdf [Accessed 7 Feb. 2018].
[23] Yalcin, N., Bayrakdaroglu A. and Kahraman C. (2012). Application of Fuzzy Multi-Criteria Decision Making Methods for Financial Performance Evaluation of Turkish Manufacturing Industries, *Expert Systems with Applications*, 39, pp. 350-364.

Author's address

Ing. Martina Kuncová, Ph.D.
College of Polytechnics Jihlava, Department of Economic Studies
Tolstého 16, 586 01 Jihlava
Czech Republic
email: kuncova@vspj.cz

doc. Ing. Zuzana Tučková, Ph.D.
College of Polytechnics Jihlava, Department of Economic Studies
Tolstého 16, 586 01 Jihlava
Czech Republic
email: zuzana.tuckova@vspj.cz

EFFICIENCY OF FORMAL ADMINISTRATOR'S RULES IN EMERGENCY SERVICE SYSTEM REENGINEERING

Marek Kvet, Jaroslav Janáček,
University of Žilina, Faculty of Management Science and Informatics

Abstract
This paper deals with a case, when partial emergency system reengineering is suggested by one of the private service providers operating a considerable portion of the current service centers. The provider tries to maximize his profit subject to the system administrator's rules, which should protect public from worsening of their access to the service. We model the provider's behavior and study efficiency of the administrator's formal rules.

Keywords: *Location, Emergency medical service, System reengineering, Profit sharing, Formal rule efficiency*

JEL Classification: C61
AMS Classification: 90C27

1 INTRODUCTION

Emergency systems are designed to satisfy randomly emerging demands of public for mitigating consequences of accidents and other detrimental events. Deployment of a limited number service centers in the serviced geographical area plays substantial role in quality of the provided service, which is mostly represented by an average response time of the system. This way the emergency system design problem is transformed to the weighted p-median problem or other kind of location problem [1, 2, 8]. The weighted p-median problem was modelled either by the location-allocation or radial formulations [3, 4, 6, 10]. The initial service center deployment is usually suggested by so-called system administrator, who represents interests of public and thus he follows minimization of the average response time as the problem objective. The administrator usually supervises dispatching of emergency vehicles to individual users' demands. It is commonly supposed that each user demand is served from the nearest available service center. The service provision by emergency vehicles is performed by private providers, who own and run several service centers equipped with emergency vehicles.

As the distribution of service demands develops in time and space, the original service center deployment will cease to suit both serviced population and providers. These discrepancy can be mitigated by the system reengineering, when a given portion of service centers are reallocated. In some national or local emergency systems [5, 9], the system administrator is responsible for the reengineering. In other national systems, e.g. the emergency health care system of the Slovak Republic, the system administrator only defines some rules and a service provider is allowed to relocate his service centers [7]. In the recently mentioned emergency health care system, the profit of a provider is proportional to transportation performance necessary for the demand satisfaction.

In this paper, we focus on studying the efficiency of the administrator's rules, when the considered provider's objective of reengineering is to maximize his profit. The considered rules are that at most a given number of center locations can be changed and each center location can be moved only in a given radius from its original position.

In this paper, we use linear programming model of provider's reengineering his part of emergency service system to maximize his profit under rules imposed by the system administrator [7]. We perform a computational study, to find whether the administrator's rules are able to protect the public from lowering the system utility.

2 MODEL OF THE PROVIDER'S REENGENEERING

In this paper, we introduce J as a finite set of all system users. Let b_j denote a volume of expected demand of the user $j\in J$. Let I be a finite set of possible center locations. Symbol d_{ij} denotes the integer time-distance between locations i and j, where $i, j\in I\cup J$. The maximal relevant time is denoted by H. The current emergency service center deployment is described by two disjoint sets of located centers $I_L\subset I$ and $I_F\subset I$, where I_L contains p centers of the considered provider, who performs reengineering and I_F is the set of the centers belonging to the other providers.

The system administrator's rules are quantified by the following constants. The symbol D denotes the maximal time-distance between a current center location and the possible new location of the center. The integer w gives the maximal number of centers from I_L, which are allowed to change locations.

We will use several auxiliary structures. The symbol $N_t=\{i\in I\text{-}I_F\colon d_{ti}\le D\}$ denotes the set of all possible center locations, to which the center $t\in I_L$ can be moved. Similarly, set $S_i=\{t\in I_L\colon i\in N_t\}$ contains all centers of the considered provider, which can be moved to $i\in I_R$. The subset $I_R\subseteq I\text{-}I_F$ is defined as $I_R=\bigcup_{t\in I_L} N_t$. Realize that $t\in N_t$ and $i\in S_i$ for $t\in I_L$ and $i\in I_R$ and thus $I_L\subset I_R$ holds.

We introduce coefficients a^H_{ij} for each pair i, j $i\in I_R\cup I_F$ and $j\in J$. For the coefficient, $a^H_{ij}=1$ holds if and only if $d_{ij}\le H$. Otherwise $a^H_{ij}=0$.

The cost coefficients for $i\in I_R$ and $j\in J$ are defined as follows: $c_{ij}=0$ if $d_{ij}\ge min\{d_{tj}\colon t\in I_F\}$ and $c_{ij}=b_jd_{ij}$ otherwise.

Furthermore, two auxiliary structures $\{P_j\}$ and $\{R_j\}$ are introduced for every $j\in J$. The first structure is a system of ordered lists, where list P_j consists of $i\in I_R$ and it is ordered so that the following inequalities hold: $d_{P_j(1)j}\le d_{P_j(2)j}\le\cdots\le d_{P_j(|I_R|)j}$. An element R_j of the second structure is ordered list of subscripts from range 1, …, $|I_R|$, where $R_j(r)$ gives the minimal subscript, for which $d_{P_j(\mathrm{r})j}<d_{P_j(R_j(r))j}$.

We introduce the decision binary variable y_i defined for each $i\in I_R$, which takes the value of one, if a service center should be located at i and it takes the value of zero otherwise.

The reallocation variable $u_{ti}\in\{0, 1\}$ for $t\in I_L$ and $i\in N_t$ takes the value of one, if the service center at t should be moved to i and it takes the value of zero in the opposite case.

Finally, we introduce allocation variable $z_{ij}\in\{0, 1\}$ for $i\in I_R\cup I_F$ and $j\in J$, where $z_{ij}=1$ if a user demand located to j is serviced from the center location i.

The mathematical programming model of the provider's reengineering follows.

$$Maximize \sum_{j\in J}\sum_{i\in I_R} c_{ij}z_{ij} \tag{1}$$

$$Subject\ to \sum_{i\in I_R} y_i = p \tag{2}$$

$$\sum_{i\in I_L} y_i \ge p-w \tag{3}$$

$$\sum_{i\in N_t} u_{ti} = 1 \quad for\ t\in I_L \tag{4}$$

$$\sum_{t \in S_i} u_{ti} \le y_i \quad for\ i \in I_R \tag{5}$$

$$\sum_{i \in I_R} a_{ij}^H y_i + \sum_{i \in I_F} a_{ij}^H \ge 1 \quad for\ j \in J \tag{6}$$

$$\sum_{i \in I_R \cup I_F} z_{ij} = 1 \quad for\ j \in J \tag{7}$$

$$z_{ij} \le y_i \quad for\ j \in J, i \in I_R \tag{8}$$

$$1 - y_{P_j(k)} \ge \sum_{r=R_j(k)}^{|I_R|} z_{P_j(r)j} \quad for\ j \in J,\ k = 1, ..., |I_R| - 1 \tag{9}$$

$$y_i \in \{0, 1\} \quad for \quad i \in I_R \tag{10}$$

$$u_{ti} \in \{0, 1\} \quad for \quad t \in I_L, i \in N_t \tag{11}$$

$$z_{ij} \in \{0, 1\} \quad for\ i \in I_R \cup I_F,\ j \in J \tag{12}$$

The objective function (1) expresses the volume of transportation performance allocated to the considered provider, which is supposed to be proportional to the provider's profit. If a user is nearer to a center of other providers, the contribution to the considered provider's profit is zero. The misallocation of a user to a more distant center of the considered provider is prevented by constraints (9). Constraint (2) preserves constant number of centers belonging to the considered provider under reengineering and constraint (3) limits the number of changed center locations by the constant *w*. Constraints (4) allow moving the center from the current location *t* to at most one other possible location in the radius *D*. Constraints (5) enable to bring at most one center to a location *i* subject to condition that the original location of the brought center lies in the radius *D*. These constraints also assure consistency among the decisions on move and decisions on center location. Constraints (6) ensure that any user *j* lies in the radius *H* from a located center, i.e. maximal distance between a user and the nearest center is less than or equal to the value *H*.
Constraints (7) are commonly used allocation constraints, which assure that each user demand is allocated to exactly one center belonging either to the considered provider or to other providers. Link-up constraints (8) give relation between allocation variables z_{ij} and the location variables y_i, which model the decisions on locating service centers operated by the considered provider. Constraints (9) were developed to prevent the maximization process from allocating a demand to a more distant service center than the nearest one. The constraint formulated for location $P_j(k)$ and user *j* forbids allocation of user's *j* demand to every service center $P_j(r)$, which is more distant from the location *j* than the center location $P_j(k)$.

3 COMPUTATIONAL STUDY

To study presented approach to reengineering of the emergency service system, we performed series of numerical experiments, in which the optimization software FICO Xpress 8.3 (64-bit, release 2017) was used and the experiments were run on a PC equipped with the Intel® Core™ i7 5500U processor with the parameters: 2.4 GHz and 16 GB RAM.
Used benchmarks were derived from real emergency health care system, which was originally implemented in three selected regions of Slovak Republic. For each self-governing region, i.e. Trenčín (TN), Trnava (TT) and Žilina (ZA), all cities and villages with corresponding number b_j of inhabitants were taken into account. The coefficients b_j were rounded to hundreds. The set of communities represents both the set *J* of users' locations and the set *I* of possible center

locations. The cardinalities of these sets are 276, 249 and 315 respectively. The total number of located centers are 21, 18 and 29 respectively. The network distance from a user to the nearest located center was taken as the user´s disutility. It was assumed that the considered provider owns approximately half of the service centers.
An individual experiment was organized so that the current deployment of service centers for each self-governing region was studied first. Then, the reengineering was performed using the model (1)-(12) and the obtained results were compared to the current state to measure the possible improvement of studied system characteristics. The model was computed for different values of parameters w and D. The parameter w expresses the number of service centers, which can change their current location. Its value is given in percentage of all centers operated by the considered service provider. Parameter D limits the radius, in which the service center can be relocated. This way, 20 problems for all combinations of mentioned parameters were solved for each problem instance. The results obtained for the individual self-governing regions are presented in a triple of tables. The first table of each triple contains the average computational time in seconds. It must be noted that ten different instances were generated randomly for each self-governing region. These instances differ in the list of located service centers operated by the considered provider. The parameter w was set to 25, 50, 75 and 100 percent of the total number of centers operated by the considered provider. The parameter D took the value 5, 10, 15, 20 and 25. The second table of a triple contains the evaluation of average percentage worsening of the total transportation performance, i.e. elongation of the average time-distance from a user to the nearest located service center. The last table of the triple contains the average increase of the provider's profit given in percentage of the current profit. The values correspond to the improvement of objective function (1). The following Table 1, Table 2 and Table 3 correspond to the results obtained for the region of Trenčín (TN). The results for the region of Trnava (TT) are summarized in Table 4, Table 5 and Table 6. Finally, we present the results obtained for the self-governing region of Žilina (ZA) in Table 7, Table 8 and Table 9.

Table 1: Average computational times in seconds for different settings of parameters w and D in the self-governing region of Trenčín (TN).

		D				
		5	10	15	20	25
w	25%	0.3	2.0	8.0	9.0	20.3
	50%	0.3	2.9	12.7	44.0	95.2
	75%	0.3	3.5	15.4	40.0	302.8
	100%	0.3	3.5	16.5	48.2	410.0

Table 2: Average worsening of total transportation performance given in percentage of the current transportation performance – results for different settings of parameters w and D in the self-governing region of Trenčín (TN).

		D				
		5	10	15	20	25
w	25%	12.2	33.9	45.7	49.2	49.8
	50%	16.7	44.2	62.0	66.4	70.1
	75%	19.0	49.4	69.6	74.7	79.8
	100%	20.3	52.3	71.9	77.0	83.2

Table 3: Average increase of the provider's profit given in percentage of the current profit - results for different settings of *w* and *D* in the self-governing region of Trenčín (TN).

		D				
		5	10	15	20	25
w	25%	25.1	66.8	92.5	99.7	101.6
	50%	34.9	91.1	127.7	138.3	146.8
	75%	39.4	104.2	144.6	157.2	169.2
	100%	41.1	109.5	150.2	164.4	178.3

Table 4: Average computational times in seconds for different settings of parameters *w* and *D* in the self-governing region of Trnava (TT).

		D				
		5	10	15	20	25
w	25%	0.2	1.1	5.9	17.4	12.1
	50%	0.2	2.2	16.4	82.3	130.6
	75%	0.2	2.9	21.6	213.5	233.3
	100%	0.2	2.6	19.9	219.5	214.3

Table 5: Average worsening of total transportation performance given in percentage of the current transportation performance – results for different settings of parameters *w* and *D* in the self-governing region of Trnava (TT).

		D				
		5	10	15	20	25
w	25%	5.1	19.9	29.9	38.4	41.8
	50%	7.6	30.7	46.5	60.0	67.1
	75%	8.0	33.4	53.5	69.7	76.2
	100%	8.0	34.2	55.3	71.0	79.3

Table 6: Average increase of the provider's profit given in percentage of the current profit - results for different settings of *w* and *D* in the self-governing region of Trnava (TT).

		D				
		5	10	15	20	25
w	25%	15.4	41.8	65.3	82.5	96.2
	50%	23.4	66.9	104.9	135.1	154.2
	75%	24.5	77.3	124.5	158.3	178.9
	100%	24.5	80.5	132.0	165.3	188.7

Table 7: Average computational times in seconds for different settings of parameters *w* and *D* in the self-governing region of Žilina (ZA).

		D				
		5	10	15	20	25
w	25%	0.3	2.3	6.9	14.7	18.1
	50%	0.3	2.3	9.2	29.8	57.7
	75%	0.3	2.9	12.2	82.1	431.9
	100%	0.3	2.6	11.6	138.7	451.1

Table 8: Average worsening of total transportation performance given in percentage of the current transportation performance – results for different settings of parameters *w* and *D* in the self-governing region of Žilina (ZA).

		D				
		5	10	15	20	25
w	25%	12.4	33.8	38.6	44.3	42.8
	50%	18.2	45.9	54.7	61.2	60.6
	75%	19.3	49.5	62.2	70.2	71.0
	100%	19.3	49.7	64.0	73.1	74.6

Table 9: Average increase of the provider's profit given in percentage of the current profit - results for different settings of parameters *w* and *D* in the self-governing region of Žilina (ZA).

		D				
		5	10	15	20	25
w	25%	26.6	68.6	85.2	96.7	98.6
	50%	39.2	97.3	124.7	139.7	144.0
	75%	42.9	109.0	144.5	164.5	173.4
	100%	43.1	111.0	149.1	172.2	184.0

4 CONCLUSIONS

The paper deals with an approach to reengineering of an emergency service system, where one of the providers changes the service center deployment to maximize his profit depending on the transportation performance necessary for users' demand servicing. The computational study was focused on studying the efficiency of the administrator's rules imposed on provider's changes. The considered rules were formulated that at most a given number of center locations can be changed and each center location can be moved only in a given radius from its original position.

Having performed numerical experiments with benchmarks derived from current state of service centers deployment, we can conclude that the considered administrator's regulation rules are completely insufficient in preventing public from considerable worsening perceived system utility. It follows that if an emergency service provider is allowed to suggest some changes of his service centers deployment, then permission of these changes must be conditioned by evaluation of the change impact on system users. Future research may be aimed at usage of the suggested modelling technique in game modelling, in which different groups of providers compete for the profit under system administrator supervision.

Acknowledgements: This work was supported by the research grants VEGA 1/0342/18 "Optimal dimensioning of service systems", VEGA 1/0463/16 "Economically efficient charging infrastructure deployment for electric vehicles in smart cities and communities" and APVV-15-0179 "Reliability of emergency systems on infrastructure with uncertain functionality of critical elements".

References

[1] Brotcorne, L., Laporte, G. and Semet, F. (2003). Ambulance location and relocation models. *European Journal of Operational Research*, 147(3), pp. 451-463.

[2] Doerner, K. F., Gutjahr, W. J., Hartl, R. F., Karall, M. and Reimann, M. (2005). Heuristic Solution of an Extended Double-Coverage Ambulance Location Problem for Austria. *Central European Journal of operations research*, 13(4), pp. 325-340.

[3] Elloumi, S., Labbé, M. and Pochet, Y. (2004). A new formulation and resolution method for the p-center problem. *INFORMS Journal on Computing*, 16(1), pp. 84-94.
[4] García, S., Labbé, M. and Marín, A. (2011). Solving large p-median problems with a radius formulation. *INFORMS Journal on Computing*, 23(4), pp. 546-556.
[5] Guerriero, F., Miglionico, G. and Olivito, F. (2016). Location and reorganization problems: The Calabrian health care system case. *European Journal of Operational Research*, 250(3), pp. 939-954.
[6] Janáček, J. and Kvet, M. (2013). Public service system design with disutility relevance estimation. In: *Mathematical Methods in Economics 2013*, Jihlava, Czech Republic, pp. 332-337.
[7] Janáček, J. and Kvet, M. (2018). Reengineering of the Emergency Service System from the Point of Service Provider In: *ICORES 2018: 7th International Conference on Operations Research and Enterprise Systems*, Madeira, Portugal, pp. 224-231
[8] Jánošíková, Ľ. and Žarnay, M. (2014). Location of emergency stations as the capacitated p-median problem. In: *International scientific conference: Quantitative Methods in Economics-Multiple Criteria Decision Making XVII*, Virt, Slovakia, pp. 116-122.
[9] Reuter-Oppermann, M., van den Berg, P. L. and Vile, J. L. (2017). Logistics for Emergency Medical Service systems. *Health Systems*, 6(3), pp. 187-208.
[10] Sayah, D. and Irnich, S. (2016). A new compact formulation for the discrete p-dispersion problem. *European Journal of Operational Research*, 256(1), pp. 62-67.

Authors' addresses

Ing. Marek Kvet, PhD.
University of Žilina - Faculty of Management Science and Informatics
Univerzitná 8215/1, 010 26 Žilina, Slovakia
Email: marek.kvet@fri.uniza.sk

Prof. RNDr. Jaroslav Janáček, CSc.
University of Žilina - Faculty of Management Science and Informatics
Univerzitná 8215/1, 010 26 Žilina, Slovakia
Email: jaroslav.janacek@fri.uniza.sk

DOES COMPULSORY SCHOOLING HAVE AN IMPACT ON EARNINGS? EVIDENCE FROM POLAND

Jacek Liwiński, University of Warsaw

Abstract
In 1966 the minimum school-leaving age was increased from 14 to 15 years in Poland. This was a result of extending the primary school education from 7 to 8 years. At the same time, the reform did not affect the education system at post-primary levels, that is the system of secondary and higher education. In result, all education tracks were extended by one year. Using the regression discontinuity design and data from the Polish LFS (2001-2005), we find that the reform had no impact on men's and women's hourly earnings. A similar finding was reported earlier for a few Western European countries. However, our study is the first one to estimate the impact of the compulsory schooling extended in a centrally planned economy on the individuals' labour market outcomes in the period of economic transition.

Keywords: *education; schooling; earnings; regression discontinuity design*

JEL Classification: I21, I26, J24

1. INTRODUCTION

In recent years, several studies have been carried out on the impact of compulsory schooling on the individuals' labour market outcomes. These studies exploit the reforms of education increasing the minimum school-leaving age, as they can be treated as natural experiments and, thus, the causal effect of schooling can be identified. All of these studies concern market economies, and mostly the Western European countries. They found that extended compulsory schooling has on average no impact on earnings (Meghir and Palme, 2005, for Sweden; Pischke and von Wachter, 2008, for Germany; Oosterbeek and Webbink, 2007, for the Netherlands; Grenet, 2013, for France), or has a small positive impact (Chib and Jacobi, 2015, for the UK; Grenet, 2013, for England and Wales). However, the impact of such education reforms implemented in non-market economies on the individuals' labour market outcomes in the period of economic transformation has not been studied so far.

This study tries to identify the individuals' labour market outcomes of the reform of education increasing the minimum school-leaving age from 14 to 15 years in Poland in 1966. This was a result of extending the primary school education from 7 to 8 years. At the same time, the reform did not affect the education system at post-primary levels, that is the secondary and higher education. Thus, the reform extended all education tracks by one year. The authors of the reform argued that the general education should be extended because of a rapid technical progress. In their opinion young people should have a greater amount of general knowledge and skills, which will allow them to acquire professional competences in secondary and higher education, as well as to acquire and update qualifications during their career. Although the reform was implemented in 1966, when Poland was a centrally planned economy, we try to find out its labour market outcomes in the period 2001-2005, that is more than one decade after the economic transition started in Poland.

Using the regression discontinuity desing (RDD) method and data from the Polish LFS, we find that the reform had on average no impact on the individuals' hourly earnings. This finding is consistent with some of the empirical studies for other European countries.

This paper contributes to the existing literature as this is the first study which tries to identify the impact of the compulsory schooling extended in a centrally planned economy on the individuals' labour market outcomes in the period of economic transition. Moreover, this is the first study on the labour market outcomes of compulsory schooling in a Central and Eastern European (CEE) country.

The paper proceeds as follows. The second section describes the reform of education implemented in Poland in 1966. The third section contains a review of empirical literature. In sections 4 and 5 we present the method and data used in our analysis. In section 6 we present the results of our empirical analysis, and then we conclude in the last section.

2. THE NATURE OF THE 1966 REFORM

In 1948, shortly after the Second World War, the obligation to study in a seven-grade primary school was introduced in Poland. This was a fundamental change when compared to the pre-war period, when seven-grade schools existed almost exclusively in urban areas, while in rural areas children completed their education after the fourth grade of primary school. The educational authorities intended to introduce an eight-grade primary school already in 1945, but the implementation of this plan exceeded the capacity of the country, which was heavily damaged during the war, as the rebuilding was the first priority. Therefore, the reform was postponed. Ultimately, the law of 1961 determined the extension of primary school education from 7 to 8 years. The authors of the reform argued that the general education should be extended because of a rapid technical progress. In their opinion young people should have a greater amount of general knowledge, which will allow them to acquire professional knowledge in secondary and higher education, as well as to acquire and update qualifications during their career. The reform envisaged that the eighth grade will be created in the 1966/67 school year, and the first cohort that will start learning in it will be children born in 1952. At the same time, the minimum school leaving age was raised from 14 to 15 years, that is children had to complete the new primary school. Both before and after the introduction of the reform it was compulsory to complete a primary school, so the compulsory education was extended from 7 to 8 years as a result of the reform.

The structure of education above the primary level has not changed, so the reform basically led to the extension of all educational paths by one year. Both before and after the reform, primary school graduates could continue their education in three types of schools: in a four-year general secondary school, in a five-year vocational secondary school or in a three-year basic vocational school (see Figure 1). Then, the secondary school graduates could continue their education at higher education institutions. Thus, we can expect the reform to have a positive impact on the individuals' situation in the labor market regardless of the level of education achieved.

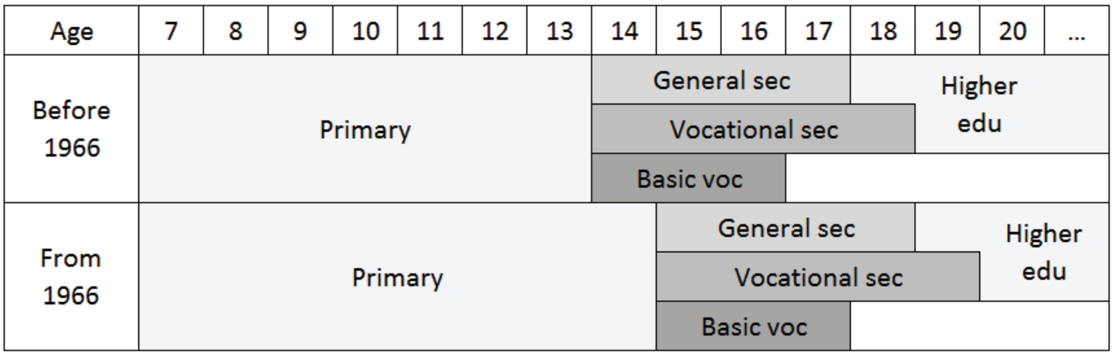


Figure 1. Changes to the Polish education system over time

Source: own elaboration.

The authors of the reform wanted to maintain the continuity of recruitment to basic vocational schools and at the same time they wanted to mitigate the increase in the number of students at primary schools in the 1966/67 school year. That is why the implementation of the reform was spread over two years. Pupils who finished the 7th grade of primary school in the 1965/66 school year were allowed to proceed to basic vocational schools, provided they were born before July 1, 1952. Those who wanted to study at a secondary school (general or vocational) or were born after 1 July 1952, regardless of the chosen educational path, had to continue their education in the 8th grade of a primary school. Thus, all the individuals born before January 1, 1952 were educated in the old system, those born in the period from January 1 to June 30, 1952 were educated in an old or new system, depending on what type of post-primary school they chose, while those born after June 30, 1952, they were covered by the new system. Obviously, the fact that a part of the 1952 school cohort was exempt from the reform makes identifying the causal effect of the reform more difficult. We will address this issue when discussing the method of analysis.

3. REVIEW OF LITERATURE

Most of the studies trying to identify the impact of the extension of compulsory schooling on wages were conducted for the Western European countries. In some of these countries, the reform was implemented gradually - first in a part of administrative units, and only then in the remaining ones - which allowed to identify the causal effect by using the difference-in-differences (DID) method. In these studies, pupils from the units in which the reform was introduced first are a treatment group, while those covered by the reform later are a control group (Fischer et al., 2016; Meghir and Palme, 2005; Pekkarinen et al., 2009; Pischke and von Wachter, 2008; Oosterbeek and Webbink, 2007). However, if the reform was introduced simultaneously throughout the country, the regression discontinuity design (RDD) method is used to identify its impact. This method involves comparing the years of schooling in formal education and the earnings of the first cohort affected by the reform and the last cohort preceding the reform.

The studies using the RDD method show that the reforms had a small positive impact on wages or that they had no impact at all. Three of these studies refer to the wage effects of increasing the minimum school leaving age from 14 to 15 years in the UK in 1947. Oreopoulos (2006) shows that this reform led to a 10% increase in earnings, while Devereux and Hart (2010) indicate that the increase in earnings was much smaller - only 3% - and it referred to men only. In turn, Chib and Jacobi (2015) using the fuzzy version of the RDD method (fuzzy regression discontinuity design) estimated the increase in earnings at 5-6%. Grenet (2013) conducted a comparative analysis of the effects of raising the school-leaving age from 15 to 16 years in France and in England and Wales. He found that in France the reform had no impact on wages, while in England and Wales the hourly rate increased by 6-7%. The author links the positive impact of the reform in England and Wales to the fact that it led to an increase in the percentage of young people completing formal education with a diploma confirming their competences, which was not the case in France. Relatively few studies refer to the educational reforms outside of Europe. Eble and Hu (2016) show that the extension of compulsory primary education by one year in China in 1980 led to an increase in wages on average by 2%, with a slightly higher wage premium obtained by individuals with low socio-economic background. Finally, Fuwa and Korwatanasakul (2015) report that the extension of compulsory primary schooling from 4 to 6 years, which was implemented in 1978 in Thailand, resulted in a ca. 8% increase in wages.

4. METHOD

Our goal is to examine the impact of the 1966 education reform on earnings. We expect that the additional year of primary school education will provide students with general skills, which, based on the model of skill formation (Cuhna, Heckman, 2007), will contribute to greater skills gains at subsequent stages of education. Based on the human capital theory, we expect that additional skills acquired during formal education will translate into higher productivity and, consequently, higher earnings.

As the 1966 reform of education was implemented simultaneously across the country, we used the regression discontinuity design (RDD) to identify its impact on earnings and employment. However, as the RDD requires a cutoff in the forcing variable, we need to address the issue of implementing the reform over a two-year period and adjust the evaluation method accordingly.

As for those born before January 1, 1952 or after June 30, 1952, it is clear that the former individuals remained in the old education system, while the latter were covered by the new system and, thus, they had to study for one year longer to achieve the same level as the former ones. The problem is the individuals born from January 1 till June 30, 1952, may have studied in an old or new system, depending on the type of post-primary school they chose. Therefore, there is no sharp cutoff. In order to eliminate this problem, we exclude individuals born between January 1 and June 30, 1952 from our analysis and we treat January 1, 1952 as a cutoff. In result, the estimation of the casual effect of the reform will be based on the difference between the labor market outcome of individuals born in the second half of 1952 and those born in 1951. We argue that the exclusion of individuals born in the first half of 1952 from our analysis will not bias the results, because there is no reason why the labor market outcomes of people born in the first or the second half of the same year should be different, conditional on being or not being covered by the reform. All individuals born in a given year start their school education on the same day and finish it on the same day, provided that they have chosen the same education track. If they enter the labor market at the same time, the age difference of a few months should not matter. But to be more convinced that our results are not affected by omitting the individuals born in the first half of 1952, we will repeat our analysis using a sample with all cohorts restricted to individuals born in the second half of a year only.

Following the standard approach to the regression discontinuity estimation we estimated three specifications: the first stage, the reduced form and, finally, the wage equation estimated by two-stage least squares (2SLS). We started our empirical analysis with estimating the first stage equation and the reduced-form effect of the raising of the minimum school-leaving age by means of a global or local polynomial approximation. This methodology involves using the whole sample and choosing a high-order polynomial to fit the relationship between the outcome variable Y_i (years of schooling, employment rate, earnings) and the forcing variable X_i (school cohort), allowing for an intercept shift at the cut-off, i.e. at the first cohort affected by the reform.

Our first stage equation takes the standard form:

$$S_i = \alpha_0 + \alpha_1 REF_i + f(X_i - C) + \alpha_2 SY_i + \varepsilon_i \qquad (1)$$

where: S_i stands for years of schooling, REF_i is a dummy variable equal to 1 for individuals covered by the education reform (born after June 30, 1952), and 0 otherwise (born up to 1951), $f(.)$ is a quartic polynomial function, X_i - individual's year of birth, C - the year of birth of the cut-off cohort (1952), SY_i - a set of dummy variables for each survey year, and ε_i is a random error. The key parameter in equation (1) is α_1, which measures the average treatment effect (ATT) of the increase in the minimum school-leaving age on the years of schooling of individuals born in 1952.

The reduced form is:

$$Y_i = \beta_0 + \beta_1 REF_i + f(X_i - C) + \beta_2 SY_i + v_i \qquad (2)$$

where: Y_i is our outcome variable, that is the employment rate or the logarithm of hourly earnings, and v_i is a random error. Here, the key parameter is β_1, which is the average treatment effect of the increase in a minimum school-leaving age on the outcome variable for individuals born in 1952.

Finally, in order to identify the causal effect of the reform we employed the instrumental variable (IV) method by estimating the following equation using two- stage least squares (2SLS):

$$Y_i = \gamma_0 + \gamma_1 S_i + f(X_i - C) + \gamma_2 SY_i + \mu_i \qquad (3)$$

where the years of schooling (S_i) were instrumented by the fact of being covered by the education reform (REF_i). The key parameter in equation (3), which is γ_1, can be interpreted as LATE, so it represents the impact of one additional year of compulsory education on compliers, that is on individuals who studied one year longer because of the reform. Such an interpretation requires the monotonicity assumption to be met (Imbens and Angrist 1994). This means that the increase in compulsory school-leaving age should prompt some of the individuals covered by the reform to study longer, but it should not induce anyone to shorten their schooling. It seems that in the case of the 1966 reform this condition should be met.

Since all individuals born in a given year start school education on the same day and, conditional on chosen education track, they complete education on the same day, they are, in fact, exposed to the same shocks. Thus, observations within cohorts are not independent of each other, which should be taken into account when estimating standard errors. Therefore, as suggested by Lee and Card (2008), robust standard errors were obtained by clustering at the cohort level.

5. DATA

Our empirical analysis is based on individual data from the Polish Labor Force Survey (LFS) for years 2001-2005.[1] The LFS is the best source of data for this type of analysis as this is the largest, representative sample survey of population, which provides information on years of schooling and wages. The analyzed period was determined by two data limitations. First, the information on the year when respondent achieved the present level of education was not available before 2001. Second, the information on the monthly date of birth is not available after 2005.

The sample was restricted to individuals born between 1945 and 1958 who completed formal education not later than at the age of 30. For reasons described in the methodological section we excluded those born in the first half of 1952. Besides, the sample is limited to the employed who reported the amount of earnings and working hours they had on the month prior to the survey.[2] Our analysis does not cover the self-employed, as they are not asked to report the amount of income in the LFS. To avoid using several observations for the same individual, we further restrict the sample to respondents who were interviewed in the first of the four survey waves.

We computed the years of schooling by subtracting seven years, that is the age at which children are obliged to start primary education in Poland, from the age at which the respondent achieved the present level of education. If the respondent completed more than one school at the same level of education, the age of completing the last one is taken into account. However, in case of dropping out of the last school attended, the period when he or she attended this school is not taken into account. Thus, the years of schooling may be underestimated for individuals who dropped out of the last school they attended.

6. RESULTS

As a starting point of our analysis, we estimated the wage equation using ordinary least squares (OLS) on the pooled sample of individuals born in the years 1945-1958. We regressed the hourly earnings on the years of schooling, the quartic polynomial of age and the dummy variables for each survey year. In result, we obtained a naive estimator of the wage premium from a year of schooling, which was further used as a benchmark for estimations of the causal effect of the reform. The results for men and women separately are presented in column 1 of Table 1. They show that one year of schooling is correlated with 5.5% higher hourly earnings for men and 7.7% higher earnings for women. But, obviously, these estimators are suspected to be heavily biased due to endogeneity.

---

[1] The Polish Labour Force Survey is carried out by the Central Statistical Office (GUS).

[2] The earnings were corrected for inflation using CPI.

Table 1. The estimation of the 1966 education reform impact on years of schooling, employment rate and hourly earnings

Gender		OLS	First stage	Reduced form	2SLS
		(1)	(2)	(3)	(4)
		Hourly earnings	Years of schooling	Hourly earnings	Hourly earnings
Men	Coefficient	0.058***	0,738***	0,004	0,005
	(std. error)	(0.003)	(0,138)	(0,023)	(0,029)
	N	7392	7392	7392	7392
Women	Coefficient	0.077***	0,181	-0,011	-0,062
	(std. error)	(0.002)	(0,165)	(0,021)	(0,133)
	N	7295	7295	7295	7295

Notes: The age and age square were additionally included in specification 1. Specifications 2-4 include the global quartic polynomial in school cohort, allowing for an intercept shift at the cut-off point (1952 school cohort). *** / ** / * denote 0.1%, 1% and 5% significance level, respectively.

Source: author's own analysis based on unit data from the Polish LFS for 2001-2005.

Then we estimated the first-stage equation. The results presented in column 2 of Table 1 show that the 1966 reform, which added one year to all the education tracks, had in fact a positive impact on the years of schooling of men only. Men covered by the reform studied on average 0.74 year more, while years of schooling of women remained unchanged as the estimator is positive (0.18) but insignificant. As the next step of our analysis we estimated the reduced-form equation. The results, presented in column 3 of Table 1, show that the reform had no impact on the hourly earnings of men and women. The estimator is positive for men and negative for women, but both are insignificant. Finally, column 4 in Table 1 presents the 2SLS estimates, showing no impact of the reform on earnings of men and women. This result was expected as the reduced form estimators were insignificant. So the only outcome of the reform that we have identified so far are increased years of schooling of men.

7. CONCLUSIONS

This study aimed to find out whether the education reform implemented in 1966 in Poland, which extended the minimum school leaving age from 14 to 15 years, had any impact on hourly earnings. Using the RDD method and the Polish LFS data, we find that the reform led to an increase in the years of schooling of men, but not women, and on average it had no impact on earnings of men and women.

Our basic result showing no impact of education reform on earnings is in line with the results of some previous studies (Meghir and Palme, 2005, for Sweden; Pischke & von Wachter, 2008, for Germany; Oosterbeek and Webbink, 2007, for the Netherlands; Grenet, 2013, for France). Some other studies, however, show that extending of compulsory education has a positive impact on women's earnings (Fischer et al., 2016, for Sweden), or men (Devereux and Hart, 2010, for the UK) or all individuals covered by the reform (Chib and Jacobi, 2015, for the UK; Grenet, 2013, for England and Wales). Grenet (2013) based on the results of his comparative study for France, and England and Wales puts forward the thesis that the positive wage effect of the extended compulsory education in England and Wales was a result of certifying the additional years of schooling, while in France, where the longer education was not certified, the reform had no impact on wages. Thus, it seems that the signalling of the years of schooling may be important. The reform implemented in 1966 in Poland extended the primary school education from 7 to 8 years, but people who completed a primary school in the old and in the new system obtained basically the same certificate. This certificate may also not play a signalling role in the labour market since employers usually pay attention to the highest level of education completed, while in the 1960s approximately 75% of people pursued further education after completing a primary school. Thus, the individuals who completed a 7-grade or an 8-grade primary schools are basically indistinguishable in the labour market.

References

[1] Chib, S. and Jacobi, L. (2015). Bayesian Fuzzy Regression Discontinuity Analysis and Returns to Compulsory Schooling. *Journal of Applied Econometrics*, 31(6), pp.1026-1047.

[2] Devereux, P. and Hart, R. (2010). Forced to be Rich? Returns to Compulsory Schooling in Britain. *The Economic Journal*, 120(549), pp.1345-1364.

[3] Eble, A. and Hu, F. (2016). The Importance of Educational Credentials: Schooling Decisions and Returns in Modern China, retrieved from: https://www.brown.edu/academics/economics/candidates/sites/brown.edu.academics.economics.candidates/files/Alex%20Eble%20-%20JMP.pdf

[4] Fischer, M., Karlsson, M., Nilsson, T., Schwarz, N. (2016). The Sooner the Better? Compulsory Schooling Reforms in Sweden, *IZA Discussion Paper* No. 10430.

[5] Fuwa, N. and Korwatanasakul, U. (2015). Revisiting the Returns to Education during the Rapid Structural and Rural Transformation in Thailand: a regression discontinuity approach, retrieved from: http://www3.grips.ac.jp/~esp/esp2/wp-content/uploads/2016/12/%E2%91%AAKorwatanasakul_HC20161.pdf

[6] Grenet, J. (2013). Is Extending Compulsory Schooling Alone Enough to Raise Earnings? Evidence from French and British Compulsory Schooling Laws. *Scandinavian Journal of Economics*, 115 (1), pp.176–210.

[7] Imbens, G. and Angrist, J. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2), pp.467−475.

[8] Lee, D.S. and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142 (2), pp.655–674.

[9] Meghir, C. and Palme, M. (2005). Educational Reform, Ability and Parental Background. *American Economic Review*, 95(1), pp.414−424.

[10] Oosterbeek, H. and Webbink, D. (2007). Wage Effects of an Extra Year of Basic Vocational Education, *Economics of Education Review*, 26(4), pp.408−419.

[11] Oreopoulos, P. (2006). Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter. *American Economic Review*, 96(1), pp.152−175.

[12] Pekkarinen, T., Uusitalo R. and Kerr S. (2009). School Tracking and Intergenerational Income Mobility: Evidence from the Finnish Comprehensive School Reform. *Journal of Public Economics*, 93(7-8), pp.965–73.

[13] Pischke, J.-S. and von Wachter, T. (2008). Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation. *Review of Economics and Statistics*, 90(3), pp.592−598.

Author's address: Jacek Liwiński, PhD., University of Warsaw, Faculty of Economic Sciences, Dluga 44/50, 00-241 Warszawa, Poland; email: jliwinski@wne.uw.edu.pl

COMBINATORIAL AUCTIONS AND DUALITY THEORY

Renata Majovska, University of Finance and Administration, Prague,
Petr Fiala, University of Economics, Prague

Abstract
Combinatorial auctions are those auctions in which bidders can place bids on combinations of items. Solving of combinatorial auctions can be modeled by linear programming problems with using of duality theory. The paper is devoted to analyzing an iterative approach to solving combinatorial auctions. In the iterative approach, there are multiple rounds of bidding and allocation and the problem is solved in an iterative and incremental way. There is a connection between efficient auctions and duality theory. The primal-dual algorithm can be taken as a decentralized and dynamic method of determine the pricing equilibrium. A primal-dual algorithm usually maintains a feasible dual solution and tries to compute a primal solution that is both feasible and satisfies the complementary slackness conditions. If such a solution is found, the algorithm terminates. Otherwise the dual solution is updated towards optimality and the algorithm continues with the next iteration. Several auction formats based on the primal-dual approach have been proposed.

Keywords: *combinatorial auction, duality theory, iterative approach, primal-dual algorithm*

JEL Classification: D44, C61
AMS Classification: 90C05

1 INTRODUCTION

Auctions are important market mechanisms for the allocation of goods and services. Design of auctions is a multidisciplinary effort made of contributions from economics, operations research, informatics, and other disciplines. The standard auction models can be generalized (Fiala, 2017). 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. However, alongside their advantages, combinatorial auctions raise a lot of questions and challenges (Cramton, Shoham and Steinberg, 2006).

The problem, called the winner determination problem, has received considerable attention in the literature. Iterative auctions are considered as an alternative for solving the winner determination problem. Iterative formats such as the English auction are very popular in electronic commerce applications. They allow bidders to learn about their competitors' bids, which is an important aspect if bidders' valuations are affiliated (Milgrom, 1987). In iterative auctions, bidders do not have to submit bids on all possible bundles at once but can bid only on a small number of bundles in each round.

Unfortunately, designing of iterative combinatorial auctions leads to a number of difficulties:

- threshold problem,
- exposure problem,
- ties,
- communicative complexity,
- determining feedback prices.

The well-known threshold problem refers to the difficulty that multiple bidders desiring small bundles that constitute a large bundle may have in outbidding a single bid for the large bundle. The exposure problem considers the risk of winning items at prices above the valuations, which usually happens if a bidder with a super-additive valuation of bundle wins only a part of this bundle. Though the exposure problem is usually typical only for pseudo-combinatorial auctions, it can also become relevant for combinatorial auctions in case of OR-bidding. Avoiding or resolving ties can become a problem, because allocations can be composed of multiple winners. Communicative complexity refers to problem that the amount of communication between the bidders and the auctioneer can become quite high. However, the most fundamental problem in the iterative combinatorial auctions design is determining feedback prices in each iteration. The primal-dual algorithm can be taken as a decentralized and dynamic method of determine the pricing equilibrium. Several auction formats based on the primal-dual approach have been proposed.

2 WINNER DETERMINATION PROBLEM

The typical combinatorial auction problem is so called winner determination problem. 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. Let us suppose that one seller offers a set M of m items, $j = 1, 2, \ldots, m$, to n potential buyers. Items are available in single units. A bid made by buyer i, $i = 1, 2, \ldots, n$, is defined as

$$B_i = \{S, v_i(S)\},$$

$S \subseteq M$, is a combination of items,
$v_i(S)$, is the valuation or offered price by buyer i for the combination of items S.

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 and that no buyer obtains more than one combination. Bivalent variables are introduced for model formulation:
$x_i(S)$ is a bivalent variable specifying if the combination S is assigned to buyer i ($x_i(S) = 1$).

The winner determination problem can be formulated as follows

$$\sum_{i=1}^{n} \sum_{S \subseteq M} v_i(S)\, x_i(S) \rightarrow \quad \max$$

subject to

$$\sum_{S \subseteq M} x_i(S) \leq 1, \quad \forall\, i,\; i = 1, 2, \ldots, n,$$

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

$$x_i(S) \in \{0, 1\}, \; \forall\; S \subseteq M, \; \forall\, i,\; i = 1, 2, \ldots, n.$$

The objective function expresses the revenue. The first constraint ensures that no bidder receives more than one combination of items. The second constraint ensures that overlapping sets of items are never assigned.

Complexity is a fundamental question in combinatorial auction design. The algorithms proposed for solving the winner determination problem are exact algorithms and approximate ones. Many researchers consider iterative auctions as an alternative.

3 PRICING SCHEMES

The key challenge in the iterative combinatorial auctions design is to provide information feedback to the bidders after each iteration (Pikovsky and Bichler, 2005). Pricing was adopted as the most intuitive mechanism of providing feedback. In contrast to the single-item single-unit auctions, pricing is not trivial for iterative combinatorial auctions. The main difference is the lack of the natural single-item prices. With bundle bids setting independent prices for individual items is not obvious and often even impossible. Different pricing schemes are introduced and discussed their impact on the auction outcome.

A set of prices $p_i(S)$, $i = 1, 2, \ldots, n$, $S \subseteq M$ is called:

- linear, if $\forall\, i,\ S$: $p_i(S) = \sum_{j \in S} p_i(j)$,
- anonymous, if $\forall\ k, l, S$: $p_k(S) = p_l(S)$.

Prices are linear if the price of a bundle is equal to the sum of the prices of its items, and anonymous if the prices of the same bundle are equal for every bidder. The non-anonymous prices are also called discriminatory prices. The following pricing schemes can be derived using the above definitions:

1. linear anonymous prices,
2. non-linear anonymous prices,
3. non-linear discriminatory prices.

The first pricing scheme is obviously the simplest one. Linear anonymous prices are easily understandable and usually considered fair by the bidders. The communication costs are also minimized, because the amount of information to be transferred is linear in the number of items. The second pricing scheme introduces the non-linearity property, which is often necessary to express strong super- or sub-additivity in the bidder valuations. Unfortunately, non-linear prices are often considered too complex and the communication costs also increase. If even non-linear anonymous prices are not sufficient to lead the auction to competitive equilibrium, the third pricing scheme can be used. However, discriminatory pricing introduces additional complexity and is often considered unfair by the bidders. A set of prices $p_i(S)$ is called compatible with the allocation $x_i(S)$ and valuations $v_i(S)$, if

$$\forall\, i, S : x_i(S) = 0 \Leftrightarrow p_i(S) > v_i(S) \text{ and } x_i(S) = 1 \Leftrightarrow p_i(S) \le v_i(S)$$

The set of prices is compatible with the given allocation at the given valuations if and only if all winning bids are higher than or equal to the prices and all loosing bids are lower than the prices (assuming the bidders bid at their valuations).

Compatible prices explain the winners why they won and the losers, why they lost. In fact, informing the bidders about the allocation $x_i(S)$ is superfluous, if compatible prices are communicated. However, not every set of compatible prices provides the bidder with meaningful information for improving bids in the next auction iteration. Another important observation is the fact that linear compatible prices are harder and often even impossible to construct, when the bidder valuations are super- or sub-additive.

A set of prices $p_i(S)$ is in competitive equilibrium with the allocation $x_i(S)$ and valuations $v_i(S)$, if

1. The prices $p_i(S)$ are compatible with the allocation $x_i(S)$ and valuations $v_i(S)$.
2. Given the prices $p_i(S)$, there exists no allocation with larger total revenue than the revenue of the allocation $x_i(S)$.

The idea behind this concept is to define prices characterizing the optimal allocation. The prices may not be too low to violate the compatibility condition 1, but they may not be too high to violate the condition 2. In general, one can show that the existence of competitive equilibrium prices implies optimality of the allocation and that the opposite is also true in case of non-linear discriminatory prices:

1. If an allocation $x_i(S)$ and prices $p_i(S)$ are in competitive equilibrium for the given valuations $v_i(S)$, this allocation is the optimal allocation.
2. For the optimal allocation $x_i(S)$ there always exist discriminatory non-linear competitive equilibrium prices $p_i(S)$. This is not always true for linear and anonymous non-linear prices.

4 PRIMAL-DUAL ALGORITHMS

One way of reducing some of the computational burden in solving the winner determination 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.

There is a connection between efficient auctions for many items, and duality theory. The Vickrey auction can be taken as an efficient pricing equilibrium, which corresponds to the optimal solution of a particular linear programming problem and its dual. The simplex algorithm can be taken as static approach to determining the Vickrey outcome. Alternatively, the primal-dual algorithm can be taken as a decentralized and dynamic method to determine the pricing equilibrium. A primal-dual algorithm usually maintains a feasible dual solution and tries to compute a primal solution that is both feasible and satisfies the complementary slackness conditions. If such a solution is found, the algorithm terminates. Otherwise the dual solution is updated towards optimality and the algorithm continues with the next iteration. The fundamental work (Bikhchandani and Ostroy, 2002) demonstrates a strong interrelationship between the iterative auctions and the primal-dual linear programming algorithms. A primal-dual linear programming algorithm can be interpreted as an auction where the dual variables represent item prices. The algorithm maintains a feasible allocation and a price set, and it terminates as the efficient allocation and competitive equilibrium prices are found.

For the winner determination problem we will formulate the LP relaxation and its dual.
Consider the LP relaxation of the winner determination problem (1):

$$\sum_{i=1}^{n} \sum_{S \subseteq M} v_i(S)\, x_i(S) \to \quad \max$$

subject to

$$\begin{aligned} &\sum_{S \subseteq M} x_i(S) \le 1, \quad \forall\, i,\ i = 1, 2, \dots, n, \\ &\sum_{i=1}^{n} \sum_{S \subseteq M} x_i(S) \le 1, \quad \forall\, j \in M, \\ &x_i(S) \ge 0, \ \forall\ S \subseteq M, \ \ \forall\, i,\ i = 1, 2, \dots, n. \end{aligned} \tag{2}$$

The corresponding dual to problem (2)

$$\sum_{i=1}^{n} p(i) + \sum_{j \in S} p(j) \rightarrow \quad \min$$

subject to

$$p(i) + \sum_{j \in S} p(j) \geq v_i(S) \quad \forall\, i,\ S, \tag{3}$$
$$p(i), p(j) \geq 0, \qquad \forall\, i,\ j,$$

The dual variables $p(j)$ can be interpreted as anonymous linear prices of items, the term $\sum_{j \in S} p(j)$ is then the price of the bundle S and $p(i) = \max_S\, [v_i(S) - \sum_{j \in S} p(j)]$ is the maximal utility for the bidder i at the prices $p(j)$.

Following two important properties can be proved for the problems (2) a (3):

1. The complementary-slackness conditions are satisfied if and only if the current allocation (primal solution) and the prices (dual solution) are in competitive equilibrium.
2. The formulation (2)-(3) is weak. For the optimal allocation there no always exist anonymous linear competitive equilibrium prices.

The formulation (2)-(3) can be strengthened. Additional variables $y(k)$ are introduced for each personalized item set partition k. In the integer version only one of the variables $y(k)$ can be set to 1, which means that the allocation $x_i(S)$ is compatible with the personalized partition k. The problem relaxation and its dual are then formulated as follows

$$\sum_{i=1}^{n} \sum_{S \subseteq M} v_i(S)\, x_i(S) \rightarrow \quad \max$$

subject to

$$\sum_{S \subseteq M} x_i(S) \leq 1, \quad \forall\, i,$$
$$x_i(S) \leq \sum_{k:[i,S] \in k} y(k), \quad \forall\, i,\ S, \tag{4}$$
$$\sum_{k} y(k) \leq 1,$$
$$x_i(S) \geq 0, y(k) \geq 0, \quad \forall\, i, S, k.$$

$$\sum_{i=1}^{n} p(i) + \pi \rightarrow \quad \min$$

subject to

$$p(i) + p_i(S) \geq v_i(S), \quad \forall\, i,\ S, \tag{5}$$
$$\pi - \sum_{[i,S] \in k} p_i(S) \geq 0, \quad \forall\, k,$$
$$p(i), p_i(S), \pi \geq 0, \qquad \forall\, i,\ S.$$

The dual variables $p_i(S)$ for each bundle S and each bidder i can be interpreted as discriminatory non-linear prices, $p(i) = \max_S\, [v_i(S) - p_i(S)]$ is the maximal utility for the bidder

i at the prices $p_i(S)$ and $\pi = \max_k \sum_{[i,S]\in k} p_i(S)$ is the maximal utility of the auctioneer at the prices $p_i(S)$.

Following two important properties can be proved for the problems (4) a (5):

1. The complementary-slackness conditions are satisfied if and only if the current allocation (primal solution) and the prices (dual solution) are in competitive equilibrium.
2. The formulation (4)-(5) is strong. It proves the existence of discriminatory non-linear competitive equilibrium prices for the optimal allocation.

5 AUCTION FORMATS

Several auction formats based on the primal-dual approach have been proposed in the literature. Though these auctions differ in several aspects, the general scheme can be outlined as follows:

1. Choose minimal initial prices.
2. Announce current prices and collect bids. Bids must be higher or equal than the prices.
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.

Concrete auction formats based on this scheme can be implemented in different ways (Parkes, 2001, Porter et al., 2003). The most important design choices are the following:

- bid structure,
- pricing scheme,
- price update rule,
- bid validity,
- feedback,
- way of computing a feasible primal solution in each iteration,
- stopping rule.

6 CONCLUSIONS

The paper is devoted to analyzing an iterative approach to solving combinatorial auctions. Iterative combinatorial auctions are a promising subject for research and for practical exploitations. The key challenge in the iterative combinatorial auctions design is to provide information feedback to the bidders. There is a connection between efficient auctions for many items, and duality theory. The primal-dual algorithm can be taken as a decentralized and dynamic method to determine the pricing equilibrium. Several auction formats based on the primal-dual approach have been proposed in the literature. Comparisons of iterative combinatorial auctions give us an opportunity to design modifications and new versions of these auctions. Auctions with complex bid structures are called multi-objective auctions, since they address multiple objectives in the negotiation space. Solving of multi-objective auctions can be based on applying duality theory for multi-objective linear programming problems (Fiala, 1981). Multi-objective optimization can be helpful for detailed analysis of combinatorial auctions. Auctions are used to solve a number of practical problems, for example, when ecological authority allocates subsidy to cover a part of costs (Sauer et al., 1998, Sauer et al., 2014).

Acknowledgements

This paper was written with the support of the Czech Science Foundation project No. P402/12/G097 DYME - *Dynamic Models in Economics* and by the grant No. IGA F4/57/2017, Faculty of Informatics and Statistics, University of Economics, Prague.

The paper is a result of institutional research project no. 7429/2018/08 "Analysis of ICT startups" supported by University of Finance and Administration, Prague.

References

[1] Bikhchandani, S. and Ostroy, J. M. (2002). The package assignment model. *Journal of Economic Theory*, 107(2), pp. 377-406.

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

[3] Fiala, P. (1981). Duality in linear vector optimization (in Czech). *Ekonomicko-matematický obzor*, 17(3), pp. 251-266.

[4] Fiala, P. (2017). Modelling and Solving of Multidimensional Auctions. *Ekonomický časopis*, 65(3), pp. 220-236.

[5] Milgrom, P. R. (1987). Auction theory. In: Bewley, T. (ed.) *Advances in Economic Theory: Fifth World Congress*, Cambridge: Cambridge University Press.

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

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

[8] Porter, D., Rassenti, S., Roopnarine, A. and Smith,V. (2003). Combinatorial auction design. In: *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*.

[9] Sauer, P., Dvorak, A., Mildeova, S. and Mokrisova, J. (1998). Contingent valuation analysis: Benefit from risk reduction of floods. *Politicka ekonomie*, 46(3), pp. 412-426.

[10] Sauer, P., Fiala, P., Dvorak, A., Kolinsky, O., Prasek, J., Ferbar, P. and Rederer, L. (2015). Improving Quality of Surface Waters with Coalition Projects and Environmental Subsidy Negotiation. *Polish Journal of Environmental Studies*, 24(3), pp. 1299-1307.

Authors' addresses

PaedDr. Renata Majovska, PhD.
University of Finance and Administration, Prague, Faculty of Economic Studies, Department of Computer Science and Mathematics
Estonska 500, 101 00 Praha 10
Czech Republic
email: renata.majovska@mail.vsfs.cz

Prof. RNDr. Ing. Petr Fiala, CSc., MBA
University of Economics, Prague, Faculty of Informatics and Statistics, Department of Econometrics
W. Churchill Sq. 4, 130 67 Prague 3
Czech Republic
email: pfiala@vse.cz

COMPARISON OF THE REGIONS OF THE CZECH AND SLOVAK REPUBLICS

Luboš Marek, University of Economics, Prague

Abstract
In this article, we compare the economic performance of the regions of the Czech Republic and the Slovak Republic. Two basic statistical indicators will be used as our tool for comparison. These are GDP per capita and the registered unemployment rate. Our data comes from the statistical offices of the two countries. Although the two offices each have the same data processing methodology, their data publishing timings differ. Thus some data is unfortunately published for differing time periods; this has shortened the time series used. We first analyzed each indicator separately. We evaluated its evolution over time and performed a region level comparison. Because this is time series data, we calculated and compared the growth tempos in the individual regions. The calculations were performed in MS Excel.

Keywords: *Wage, GDP, unemployment rate, Czech regions, Slovak regions*

JEL Classification: C500, O520
AMS Classification: 62-07

1 INTRODUCTION

In this article we compare the various Czech and Slovak regions based on selected macroeconomic indicators. Two basic economic indicators serve as our tool for comparison: GDP per capita and registered unemployment rate. We will first be comparing the regions individually based on individual indicators, and this both among each other and with a view to the development of selected indicators over time. Our data is taken from the Czech Statistical Office (2018) and Slovak Statistical Office (2018). Although the data is fully comparable, we encounter the problem of time series lengths. The time series the offices publish are of unequal lengths, and so in each case we have had to conform to the shorter of the two-time series. We have GDP per capita available for the 1995 - 2015 period, and registered unemployment data for 2001–2016. This is not a problem overall for individual comparisons, but it does make comparison of the two indicators at once more difficult.

2 METHODOLOGY

Most of the analyses were performed in MS Excel, and the tables and graphs presented are also from Excel. Because the Czech data is published in CZK (GDP per capita), we needed to convert this data into Euros before using it here. The exchange rate of the Czech National Bank (2018) was used. Because this data was collected over the course of time, the amounts within it are naturally affected e.g. by inflation. We do not take inflation into account in this article (we work with the typical prices for the given year) but including it in the calculations would not be problematic - however, it would not change the overall view of the comparison's results, as inflation evolved very similarly in the two countries.
Comparisons of the Czech or Slovak regions based on various criteria are occasionally published at conferences or in articles. In this context we can note e.g. (Marek, 2016) and (Marek, 2013). Both of these, however, compare Czech regions only. Nowhere in the literature could we find a comparison of the Czech and Slovak regions based on the mentioned criteria.

3 ANALYSIS

3.1 GDP per capita

Let us first assess the overall situation through the picture below.

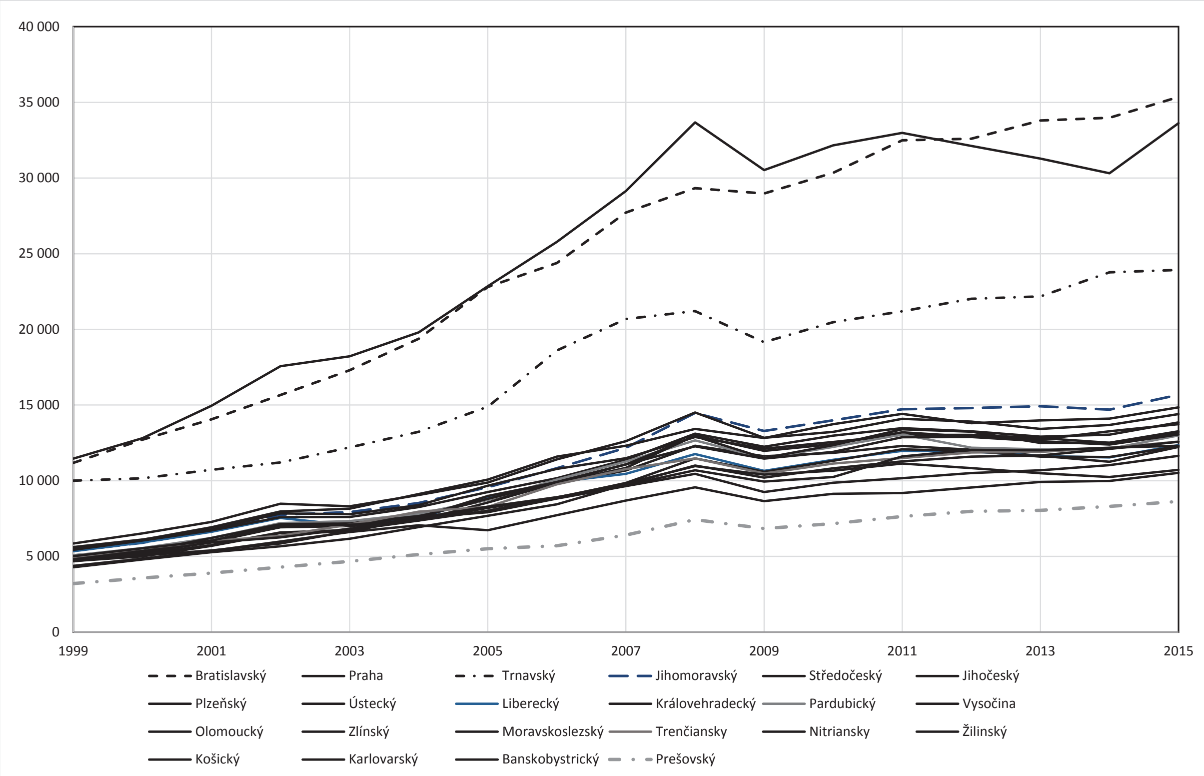


Figure 1: GDP per capita

The situation very much speaks for itself. The indicator is growing over the long term, essentially linearly. If we are to assess it by the last few years, the Bratislava region is faring best, with Prague very close behind. Somewhat isolated in third place is the Trnava region. The South Moravian and Central Bohemian regions can be found in fourth and fifth place. These are followed by a group of regions with similar values, with the Banská Bystrica and Prešov regions closing the list.

Now let us take a closer look at 2015.

Bratislava	35,352	Moravian-Silesian	13,392	Ústí nad Labem	12,353
Prague	33,928	South Bohemian	13,278	Nitra	12,308
Trnava	23,934	Vysočina	13,175	Košice	11,646
South Moravian	15,829	Pardubice	13,112	Karlovy Vary	10,810
Plzeň	14,987	Trenčín	12,590	Banská Bystrica	10,520
Central Bohemian	14,534	Žilina	12,575	Prešov	8,631
Hradec Králové	13,984	Liberec	12,468		
Zlín	13,864	Olomouc	12,382		

Table 1: GDP per capita - 2015

It can be said that the greater variations are in Slovakia. The top region - the Bratislava region - has a GDP per capita nearly 4x higher than the bottom region - the Prešov region. Moving on from this: as expected, the Bratislava region leads the comparison, followed by Prague, with both regions far above the other regions.

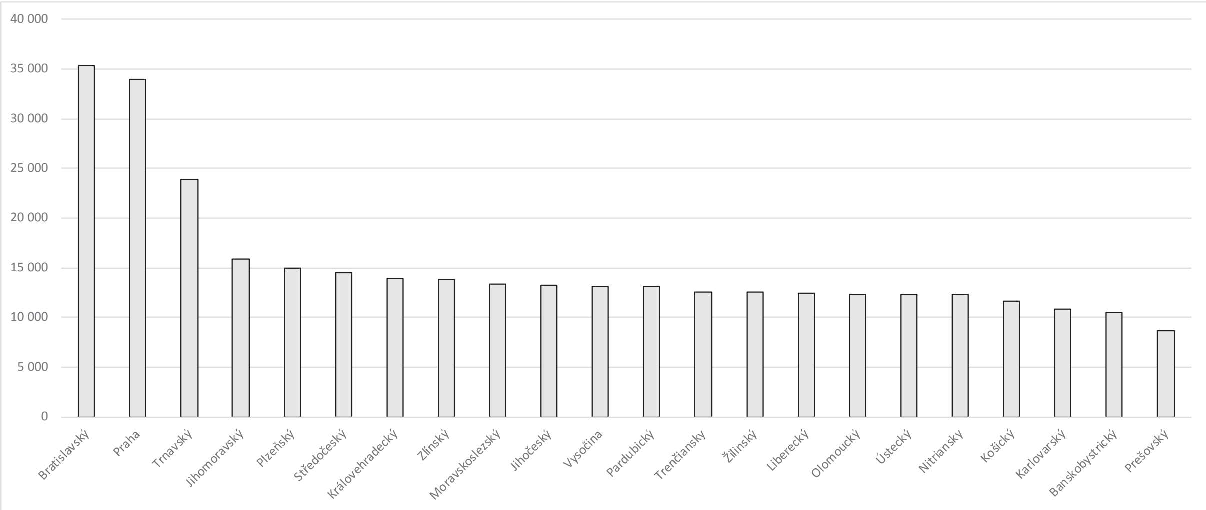


Figure 2: GDP per capita—2015

This graphical representation emphasizes the whole situation even further.
Because we are working with time series, we can calculate several of their descriptive characteristics (Hindls, et al., 2003). This is the average growth coefficient

$$\bar{k} = \sqrt[n-1]{\frac{y_n}{y_1}} \tag{1}$$

and the average absolute increase.

$$\bar{\Delta} = \frac{y_n - y_1}{n-1} \tag{2}$$

Bratislava	1.075	Prešov	1.064	Ústí nad Labem	1.057
Žilina	1.070	Olomouc	1.062	Banská Bystrica	1.056
Prague	1.070	Nitra	1.061	Trnava	1.056
South Moravian	1.069	Pardubice	1.061	South Bohemian	1.055
Zlín	1.067	Hradec Králové	1.060	Liberec	1.054
Moravian-Silesian	1.067	Trenčín	1.060	Karlovy Vary	1.049
Vysočina	1.065	Central Bohemian	1.058		
Plzeň	1.064	Košice	1.057		

Table 2: Average growth rate

As the table shows, the values in the different regions are very balanced. On average the Bratislava region is growing the fastest, followed by the Žilina and Prague regions (which turned out the same), with the South Moravian region just behind them. The Karlovy Vary region, meanwhile, fared the worst in terms of average growth rate. Nevertheless the relative differences among the regions are very small.
Table 3 shows how their average absolute increase turned out.

Bratislava	1,511	Hradec Králové	526	Ústí nad Labem	448
Prague	1,385	Žilina	518	Liberec	439
Trnava	871	Vysočina	518	Košice	428
South Moravian	644	Pardubice	495	Banská Bystrica	384
Plzeň	585	Trenčín	477	Karlovy Vary	356
Zlín	556	Olomouc	474	Prešov	339
Moravian-Silesian	536	Nitra	473		
Central Bohemian	535	South Bohemian	471		

Table 3: Average growth rate

The results in Table 3 correspond with those in Table 2. We can note that GDP per capita in the Bratislava region is growing by 7.5% per year on average, which represents an absolute growth of 1,511 EUR. The Žilina and Prague regions are growing on average by 7% per year, which in absolute terms means growth by 518 EUR for the Žilina region, while for the Prague region the same 7% growth means an absolute growth of 1,385 EUR. The Prešov region is in a special situation here: its average yearly growth is only 339 EUR, but despite this, it is growing relatively quickly, with an average growth tempo of 1.064% (ninth out of the 22 regions). We could continue with similar commentaries further, but a watchful reader can form their own conclusions.

3.2 General Unemployment Rate

Here again let us first look at the course taken by the general employment rate over the period from 2001 to 2016.

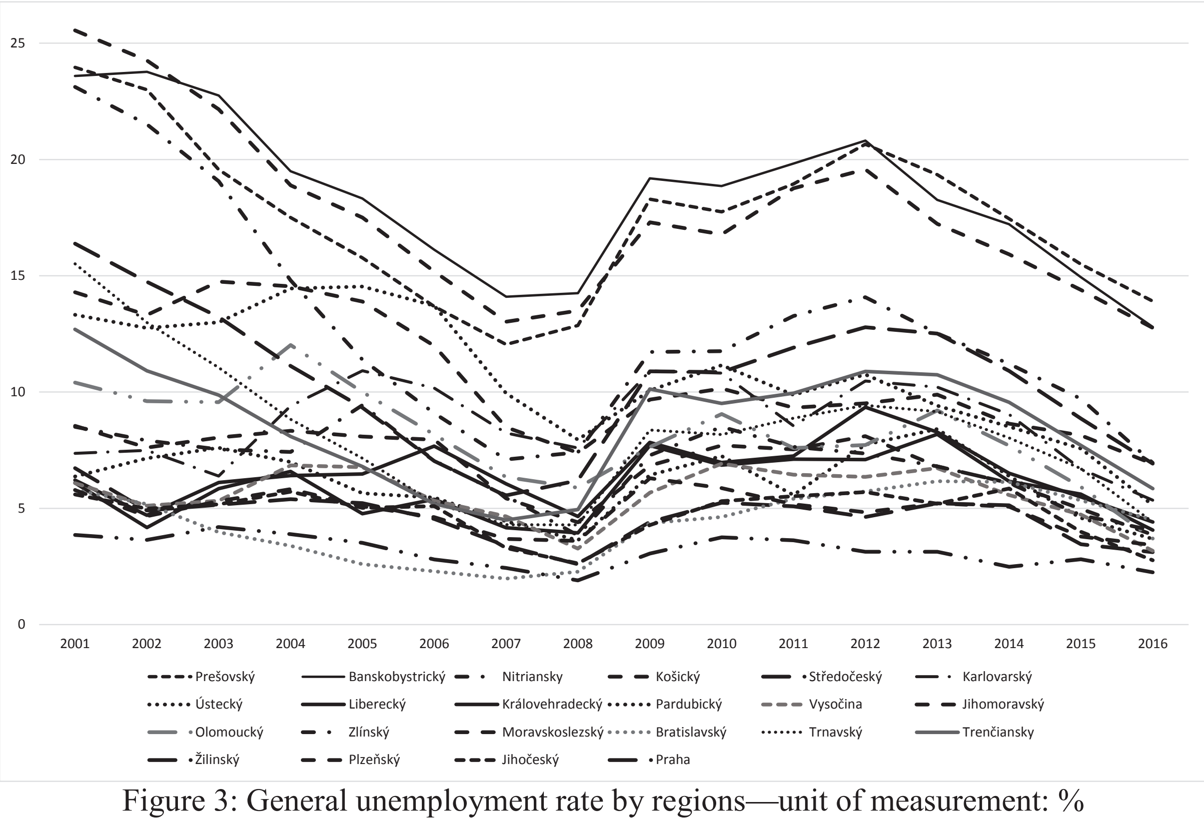


Figure 3: General unemployment rate by regions—unit of measurement: %

This graph is rather hard to navigate; the situation in the individual regions changes every year. The sad "top rankings" with the highest values go to the Prešov, Banská Bystrica, and Nitra regions. Prague, meanwhile, has since 2008 been faring best, with the lowest

unemployment rate. We can but regret that no data is available for 2017. Because this was a year of significant economic growth and strong labor demand, a significant drop in the unemployment rate in all regions can be expected here.
Now let us once again take a closer look at the last year analyzed, 2016.

Prague	2.25	Zlín	4.05	Moravian-Silesian	6.91
South Bohemian	2.77	Hradec Králové	4.06	Žilina	6.92
Central Bohemian	3.11	Liberec	4.41	Nitra	6.96
Vysočina	3.17	Trnava	4.41	Košice	12.76
Plzeň	3.42	Bratislava	4.51	Banská Bystrica	12.80
Pardubice	3.68	Ústí nad Labem	5.14	Prešov	13.91
Olomouc	3.70	Karlovy Vary	5.35		
South Moravian	3.86	Trenčín	5.85		

Table 4: General unemployment rate (%)—2016

Here again the whole situation stands out visually.

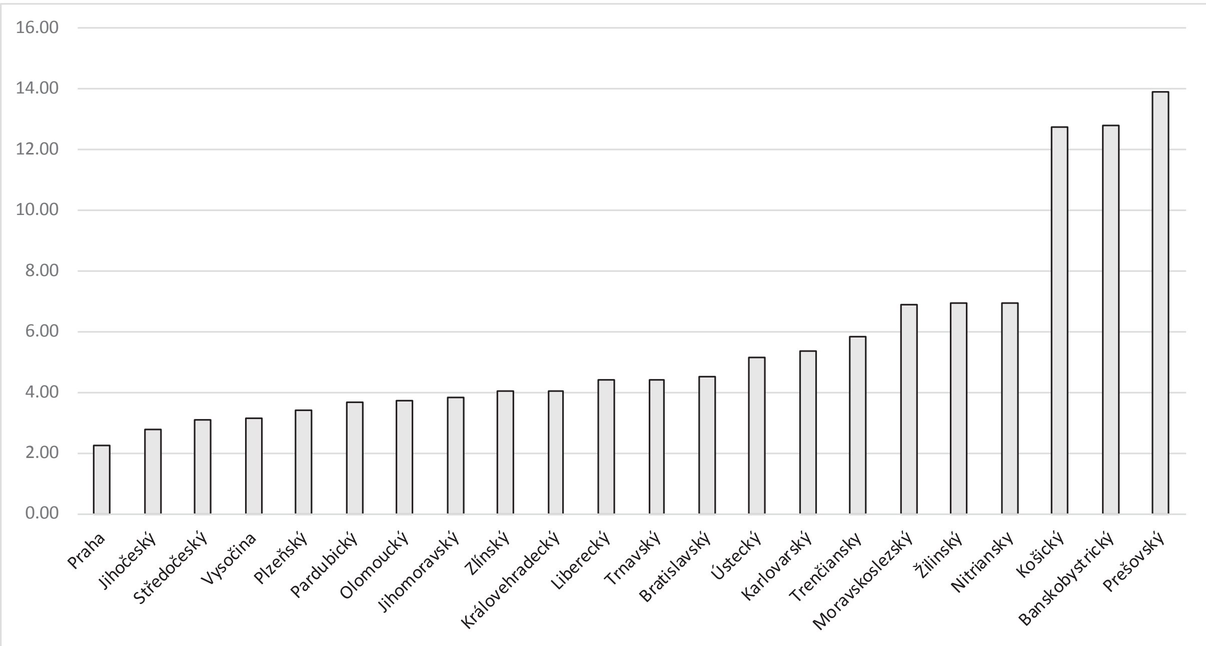


Figure 4: General unemployment rate (%)—2016

We will now attempt to assess the situation for the whole period observed. Let us take a look at the basic traits of the various time series under consideration. We first calculated the average unemployment rate (in %) for the whole period observed. We used the geometric mean for our calculation. As expected, the long-term situation is best in Prague and Bratislava. Meanwhile the average unemployment rates in the Banská Bystrica, Košice, and Prešov regions look menacing. They are in the 17 - 18% range and thus nearly 5x higher than in Prague and roughly 4.5x higher than in Bratislava. That means there are giant differences both among the regions being compared and among the regions of Slovakia. Out of the Slovak regions, if we rank from lowest unemployment to highest, then only Bratislava is within the top half of such a ranking, while the other 7 Slovak regions are in the bottom half of the regions compared.

From Figure 3 it is, however, clear that (with certain fluctuations) the situation is improving in the long term. The following table should serve as confirmation of this; it calculates the average growth rates (in our case these are more average decline rates).

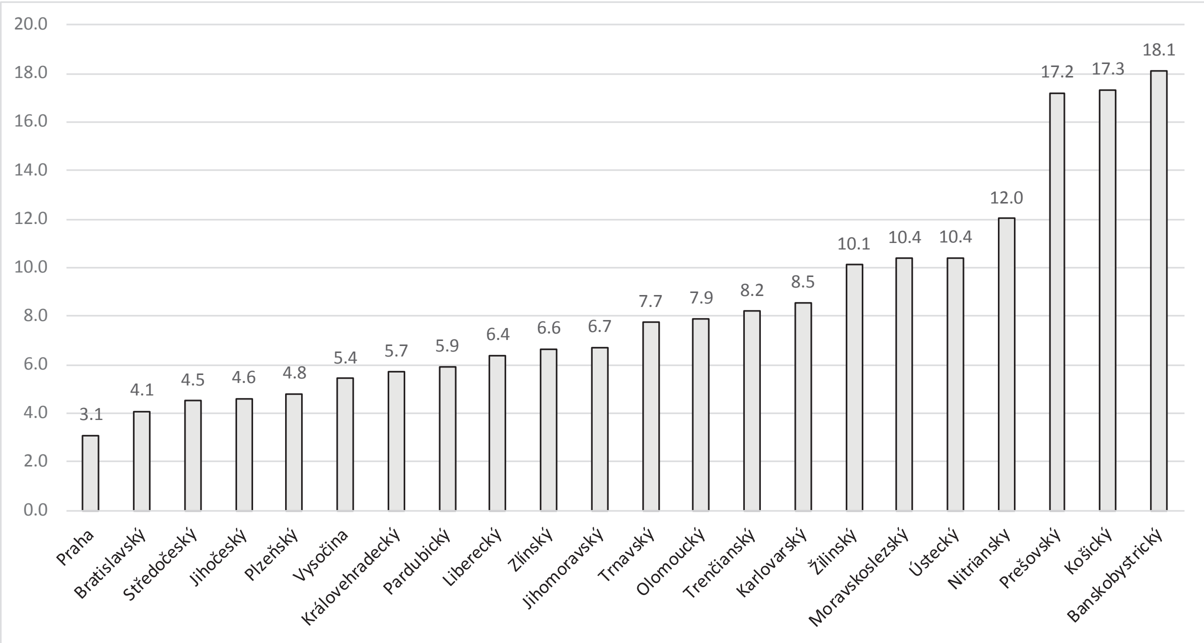


Figure 5: The average unemployment rate for 2001–2016

Trnava	0.920	Zlín	0.952	Prague	0.965
Nitra	0.923	Moravian-Silesian	0.953	Plzeň	0.966
Olomouc	0.933	South Bohemian	0.954	Hradec Králové	0.973
Ústí nad Labem	0.938	Košice	0.955	Liberec	0.977
Žilina	0.944	Vysočina	0.958	Karlovy Vary	0.979
South Moravian	0.948	Banská Bystrica	0.960	Bratislava	0.983
Trenčín	0.950	Pardubice	0.964		
Central Bohemian	0.950	Prešov	0.964		

Table 5: Average growth rate (%)

In all of the regions observed, the unemployment rate is falling. On average it is falling the fastest in the Trnava region: by 0.08% per year. It is falling the slowest in Bratislava: by 0.017% per year. These numbers are, however, relative, and only gain their meaning with a view to the absolute unemployment rate (Figure 3). The good news here is the overall downward trend in all of the regions.

4 CONCLUSIONS

In the present article, we performed a comparison of the regions of the Czech Republic and Slovak Republic based on two basic macroeconomic indicators - GDP per capita and the registered unemployment rate. We have reached the following conclusions:

- the situation varies widely among the individual regions,
- the two capitals have (as expected) a thoroughly privileged position; all of the observed and calculated indicators come out the most positively in these two regions,

- the Bratislava region has the highest GDP per capita, followed by Prague, with the Trnava region in third place. The Trnava region is unusual in that it lies alone between the values for the first two regions and those for the remaining 19 other regions. The Banská Bystrica and Prešov regions close off the list.
- for all regions, the average growth rates (GDP per capita) are comparable - but not, however, the average absolute growth rates, which differ significantly.
- the general unemployment rate is very different in the Czech regions vs. the Slovak regions. The Czech regions come out of the comparison well, and there are smaller differences among regions there as well. In Slovakia the differences are large, with Bratislava clearly dominating for lowest unemployment - roughly 4.5x lower than for the Banská Bystrica, Košice, and Prešov regions. Meanwhile, none of the Slovak regions (not even Bratislava) are in the "2016 Top Ten", that is, the ten regions with the lowest unemployment rate for 2016.

Speaking more broadly, the author has reached the conclusion that there is definitely room for a more detailed comparison of this kind. However, it would unfortunately be too expansive to fit within the bounds of this conference presentation. The comparison could also include other indicators (e.g. the average wage level) and more detailed comparisons. Likewise it would be good to use cluster analysis and evaluate the classification of regions into homogeneous groups from the standpoints of both individual indicators and their overall informational strength. This manifestly calls for the expanding of this contribution and the composing of a detailed article.

Acknowledgments: This work was supported by the Czech Science Foundation project No. P402/12/G097 "DYME – Dynamic Models in Economics".

References

[1] Czech National Bank, (2018) *Czech National Bank*. [Online] Available at: http://www.cnb.cz/cs/financni_trhy/devizovy_trh/kurzy_devizoveho_trhu/index.html [Accessed 15 Jan. 2018].

[2] Czech Statistical Office, (2018) *Czech Statistical Office*. [Online] Available at: https://www.czso.cz/csu/czso/databaze-registry [Accessed 15 Jan. 2018].

[3] Hindls, R., Hronová, S. and Seger, J. (2003) *Statistika pro ekonomy*. Praha: Professional Publishing.

[4] Marek, L. (2016) *Economic disparities of Czech Regions*. Karviná, School of Business Administration in Karvina.

[5] Marek, L. (2013) *Regional Disparities in Wages*. Brno, Mendelova univerzita.

[6] Slovak Statistical Office, (2018) *Slovak Statistical Office*. [Online] Available at: http://datacube.statistics.sk [Accessed 15 Jan. 2018].

Author's address

Doc. RNDr. Luboš Marek, CSc.
University of Economics, Prague, Faculty of Informatics and Statistics, Department of Statistics and Probability
nám. W. Churchilla 4, 130 67 Praha 3
Czech Republic
email: marek@vse.cz

SOCIAL ACCOUNTING MATRIX MULTIPLIERS FOR SLOVAKIA

Veronika Mitkova, Faculty of Social and Economic Sciences, Comenius University in Bratislava

Abstract
The paper presents the social accounting matrix for Slovakia, year 2011 and concerns to the sectoral analysis of agriculture, mining and extraction, textiles and clothing, light manufacturing, heavy manufacturing, utilities and construction and services sectors; their inter and intra-industry linkages and their effects. The second part of the paper deals with the social accounting matrix unconstrained multiplier formula and the effects of exogenous shock in external demand in the sectors to these sectors production, the households, the factors, the government, the savings and the rest of the world. Based on the work of Breinsinger, Thomas and Thurlow (2009) the general formula for multiplier was derived and applied to the eight sectors economy. As the most influential sector in the Slovak economy was identified the services sector (including trade) followed by the utilities and construction, the agriculture and light manufactures sectors. The total multiplier effects on income and GDP do not reach the same value as the initial shock due to the high tax burden.

Keywords: *Social Accounting Matrix, Unconstrained Multiplier, Slovakia, Demand Shocks*

JEL Classification: E160
AMS Classification: 15A99

1 ANALYSIS OF THE SOCIAL ACCOUNTING MATRIX

The paper starts with short and simple analysis of production, trade, consumption shares to identify the structure of flows in the economic system of The Slovak Republic. The unconstrained social accounting matrix analysis is based on the input-output model principles enlarging the analysis of the linkage effects to the complex economic system including production sectors as well as factors, households, government and rest of the world. The goal of this paper is to analyze the external demand shocks throughout the sectors.

1.1 The social accounting matrix

The social accounting matrix (SAM) was built up on the Global Trade Analysis Project (GTAP) dataset (Aguiar, Narayanan and McDougall, 2016), based on year 2011, reformulated and aggregated to the International Food Policy Research Institute (IFPRI) standards. The SAM structure used in this paper is as follows in the Table 3.

Activities and commodities sectors are assumed identical, 57 original sectors were aggregated to 7 new as follows in Table 1. Sectoral aggregation consists of two sectors: Slovakia and the rest of the world (ROW). Factors aggregation includes clerks, service/shop workers, agricultural and unskilled, technicians, assoc. professional, officials and managers to labor factor and land, capital and natural resources to capital factor. Since the GTAP database output gives detailed tax structure which is not desired for this research, the taxes yields were completely assigned to the government. Only one aggregated household was assumed.

Table 1: Sectoral aggregation

Agriculture	paddy rice, wheat, cereal grains nec, vegetables, fruit, nuts, oil seeds, sugar cane, sugar beet, plant based fibers, crops nec, cattle, sheep, goats, horses, animal products, raw milk, wool, silk-worm cocoons, cattle meat, sheep

	meat, goat meat, horse meat, meat products nec, vegetable oils and fats, dairy products, processed rice, sugar, food products nec, beverage and tobacco products
Mining and extraction	forestry, fishing, coal, oil, gas, minerals nec
Textiles and clothing	textiles, wearing apparel
Light manufacturing	leather products, wood products, paper products, publishing, metal products, motor vehicles and parts, transport equipment nec, manufactures nec
Heavy manufacturing	petroleum, coal products, chemical, rubber, plastic products, mineral products nec, electronic equipment, machinery and equipment nec
Utilities and construction	electricity, gas manufacture, distribution, water, construction
Services	trade, transport nec, sea transport, air transport, communication, financial services nec, insurance, business services nec, recreation and other services, public administration, defense, health, education, dwellings

All monetary values of the data in the SAM are in $US millions and the base year is 2011. The aggregated SAM (in Table 3) is balanced throughout the accounts.

1.2 The production shares

Table 2 shows that the services sector constitutes the main part of the total GDP (50.1%) since there is included a trade sub-sector; it is followed by manufactures sectors (total of 27.2%). The extraction and the textile sectors contribution to the GDP are almost omissible. Capital is used more intensively in average; expect for the textiles sector which is the most labor-intensive sector (24% capital vs. 76% labor); a significant difference is in the utilities sector (63.7% vs. 36.3% for labor). The production technologies of each sector are a significant part of SAM; enables to estimate the linkages of particular sectors. The intermediate output shares show that individual sectors are self-sufficient: mostly heavy manufactures, light manufactures, services, agriculture and textiles with share of output on own inputs more than 30%. It may be identified that light manufactures and utilities sectors pull the heavy manufactures sector as its strong downstream sectors. The textiles sector intensively uses labor inputs, (nearly one third production shares) while services sector labor (20.3%) and capital (23.9%) leading to the highest value-added ratio sector.

Table 2: Production shares, Slovakia, 2011

Production shares		AgrA	ExtrA	TextA	LMfgA	HMfgA	UtilA	SerA	total
GDP sectoral shares		6.5%	1.4%	1.2%	10.9%	16.3%	13.8%	50.1%	100.0%
VA shares	lab	48.7%	49.9%	76.0%	46.0%	51.7%	36.3%	45.9%	46.1%
	cap	51.3%	50.1%	24.0%	54.0%	48.3%	63.7%	54.1%	53.9%
Activity production shares	AgrC	33.1%	1.1%	0.5%	0.2%	0.2%	0.2%	1.8%	3.0%
	ExtrC	0.6%	14.6%	0.6%	1.9%	11.0%	3.9%	0.8%	4.2%
	TextC	0.1%	0.2%	30.9%	0.5%	0.2%	0.1%	0.3%	0.6%
	LMfgC	5.0%	3.2%	4.2%	36.6%	5.2%	6.0%	4.0%	9.8%
	HMfgC	9.9%	9.7%	9.2%	25.8%	49.6%	18.1%	10.0%	24.0%
	UtilC	2.1%	5.3%	2.3%	2.4%	4.1%	17.8%	4.4%	5.4%
	SerC	16.1%	13.9%	12.5%	10.1%	9.5%	15.1%	34.5%	20.1%
	lab	16.0%	25.9%	30.3%	10.3%	10.4%	14.1%	20.3%	15.2%
	cap	16.9%	26.0%	9.6%	12.1%	9.7%	24.7%	23.9%	17.8%

Table 3: Social accounting matrix for Slovakia, 2011, IFPRI standards

	Activties							Commodities							Factors		Hhlds	Gov	S-I	ROW	Total
	AgrA	ExtrA	TextA	LMfgA	HMfgA	UtilA	SerA	AgrC	ExtrC	TextC	LMfgC	HMfgC	UtilC	SerC	lab	cap					
AgrA								16 455													16 455
ExtrA									2 073												2 073
TextA										2 665											2 665
LMfgA											40 382										40 382
HMfgA												70 010									70 010
UtilA													28 963								28 963
SerA														91 373							91 373
AgrC	5 245	23	13	88	157	68	1 663										12 753	7	35	2 781	22 832
ExtrC	97	306	14	755	7 147	1 108	739										595	0	1	399	11 161
TextC	19	4	758	206	130	39	295										2 326	5	3	1 841	5 626
LMfgC	799	67	102	14 303	3 395	1 708	3 617										5 707	1	1 740	27 126	58 565
HMfgC	1 574	204	225	10 092	32 346	5 192	9 172										9 001	9	7 237	36 462	111 515
UtilC	328	112	57	943	2 690	5 117	3 989										3 425	3	12 987	1 651	31 300
SerC	2 556	292	308	3 942	6 191	4 339	31 557										23 461	18 065	427	5 150	96 287
lab	2 535	544	743	4 033	6 794	4 033	18 552														37 234
cap	2 673	546	234	4 725	6 350	7 077	21 853														43 459
Hhlds															33 879	29 809		13 373		243	77 304
Gov	631	-26	211	1 294	4 810	280	-64	962	190	218	1 129	3 307	583	1 415	3 355	2 883	10 285				31 463
S-I														694		10 767	9 752			1 218	22 431
ROW								5 414	8 899	2 744	17 054	38 200	1 755	2 806							76 872
Total	16 455	2 074	2 665	40 381	70 010	28 961	91 373	22 831	11 162	5 627	58 565	111 517	31 301	96 288	37 234	43 459	77 305	31 462	22 430	76 871	

1.3 The commodity shares

Most of Slovakia's foreign exchange is generated by a heavy manufactures sector with almost 50% of total imports and exports (as seen in the Table 4); followed by light manufactures with 22.2% shares on total imports and 36% shares on total exports. The extraction sector faces the most competition from imports (near 80%) and the textile sector has also high (48.8%) import penetration. Textiles sector exports three quarters of its total production, light manufactures (69.4%) and heavy manufactures (55.9%) exports more than half of production. Utilities and services sectors operate mainly domestically.

Table 4: Trade shares, Slovakia, 2011

Commodity shares		AgrC	ExtrC	TextC	LMfgC	HMfgC	UtilC	SerC	total
trade shares	Imports	7.0%	11.6%	3.6%	22.2%	49.7%	2.3%	3.7%	100.0%
	Exports	3.7%	0.5%	2.4%	36.0%	48.4%	2.2%	6.8%	100.0%
trade intensities	Import penetration	23.7%	79.7%	48.8%	29.1%	34.3%	5.6%	2.9%	22.8%
	Export intensity	17.6%	19.0%	75.0%	69.4%	55.9%	5.8%	5.6%	30.8%

Table 5 shows the demand shares of intermediate demand and parts of final demand; in the upper part of the table by commodity; in the lower part by demand source. The main part of the intermediate consumption demands heavy manufactures (35.8%), services (30%) and light manufactures (14.6%). Households spend their income mostly on services (41%) and agricultures (22.3%). The government consumption is composed solely of services. More than half of investments flow to utilities and one third to heavy manufactures; while the investments to automotive sector appear here. Exported are mostly heavy (48.8%) and light (36%) manufactures. The second part of the Table 5 shows that more than half of the agriculture output is consumed by households, 31.8% flows to intermediates and 12.2% is exported. More than 90% of extraction output is used in another sectors'. Light and heavy manufactures, utilities and services; for each sector more than 40% of total output is used as intermediates.

Table 5: Demand shares, Slovakia, 2011

Demand shares		Interm. demand	Private cons.	Govern. cons.	Invest-ment	Exports	Total
Demand shares (by commodity)	AgrC	4.4%	22.3%	0.0%	0.2%	3.7%	6.8%
	ExtrC	6.2%	1.0%	0.0%	0.0%	0.5%	3.3%
	TextC	0.9%	4.1%	0.0%	0.0%	2.4%	1.7%
	LMfgC	14.6%	10.0%	0.0%	7.8%	36.0%	17.4%
	HMfgC	35.8%	15.7%	0.0%	32.3%	48.4%	33.1%
	UtilC	8.1%	6.0%	0.0%	57.9%	2.2%	9.3%
	SerC	30.0%	41.0%	99.9%	1.9%	6.8%	28.5%
Demand shares (by demand source)	AgrC	31.8%	55.9%	0.0%	0.2%	12.2%	100.0%
	ExtrC	91.1%	5.3%	0.0%	0.0%	3.6%	100.0%
	TextC	25.8%	41.3%	0.1%	0.1%	32.7%	100.0%
	LMfgC	41.0%	9.7%	0.0%	3.0%	46.3%	100.0%
	HMfgC	52.7%	8.1%	0.0%	6.5%	32.7%	100.0%
	UtilC	42.3%	10.9%	0.0%	41.5%	5.3%	100.0%
	SerC	51.1%	24.4%	18.8%	0.4%	5.3%	100.0%
	total	48.6%	17.0%	5.4%	6.7%	22.4%	100.0%

2 UNCONSTRAINED MULTIPLIER EFFECTS

Social accounting matrix multipliers analysis is an extensive analysis of classical Leontief input-output model in which the household and government are considered as endogenous sectors. SAM multipliers models were used to deal with the electricity industry analysis in Spain (Duarte, Langarita and Sanchez-Choliz, 2017), to analyze the interconnection between the forest sectors and the rest of the economy for the Amazon area (Ferreira and Fachinello, 2015). This methodology was used to quantify the social impacts and develop a social valuation metrics by McBain and Alsamawi (2014). Seung (2014) estimated the spillover effects between Alaska and the rest of the United States by an inter-regional social accounting matrix and multiplier decomposition. Civardi, Pansini, and Lenti (2010) propose the methodology based on the decomposition of the elements of the global multiplier matrix by dividing the impact of exogenous injections into different "microscopic" effects. The research was conducted on 2000 Vietnamese SAMs.

2.1 The assumptions and multiplied formula

The unconstrained multiplier models according to Breinsinger, Thomas and Thurlow (2009) assume fixed set of prices; so the demand changes lead solely to the changes in output. Let's assume unlimited factors supply in the economy to allow this mechanism. It means any shift in the demand side leads to the shift in the supply side. The input coefficients are assumed to be fixed with no influence of exogenous demand shocks to the structural relationships among sectors. The linkage effects are linear and there are no behavioral changes.

The original work assumes 2 sectors economy; we extend it to n sectors.

Let:

E_i	exogenous component of demand, $i = 1,2,\ldots,n$
X_{ij}	gross output of each activity, $i, j = 1,2,\ldots,n$
Y_j	total household income (equal to total factor income), $j = 1,2,\ldots,n$
Z_{ij}	total demand for each commodity, $i, j = 1,2,\ldots,n$
a_{ij}	technical coefficients, $i, j = 1,2,\ldots,n$
b_i	share of domestic output in total demand, $i = 1,2,\ldots,n$
c_i	household consumption expenditure shares, $i = 1,2,\ldots,n$
v_j	share of value-added or factor income in gross output, $j = 1,2,\ldots,n$

Total demand of sector Z is composed of intermediate demand, final demand and exogenous demand:

$$\sum_{j-1}^{n} Z_{ij} = \sum_{j=1}^{n} a_{ij}X_{ij} + c_i \sum_{j=1}^{n} Y_j + E_i \qquad i = 1,2,\cdots,n \tag{1}$$

The gross output X is a part of the total demand:

$$\sum_{j=1}^{n} X_{ij} = b_i \sum_{j=1}^{n} Z_{ij} \qquad i = 1,2,\cdots,n \tag{2}$$

The household's total income Y depends on shares on factors in each sector:

$$\sum_{j=1}^{n} Y_j = \sum_{j=1}^{n} v_j X_{ij} \quad i = 1,2,\cdots,n \tag{3}$$

Thus the total income identity is:

$$\sum_{j=1}^{n} Y_j = b_i \sum_{j=1}^{n} v_j Z_{ij} \quad i = 1,2,\cdots,n \tag{4}$$

Than (1) may be written as

$$\sum_{j=1}^{n} Z_{ij} = b_i \sum_{j=1}^{n} a_{ij} Z_{ij} + c_i b_i \sum_{j=1}^{n} v_j Z_{ij} + E_i \quad i = 1,2,\cdots,n \tag{5}$$

$$\sum_{j=1}^{n} Z_{ij} \left(1 - b_i \sum_{j=1}^{n} a_{ij} - c_i b_i \sum_{j=1}^{n} v_j \right) = E_i \quad i = 1,2,\cdots,n \tag{6}$$

re-written in matrix form:

$$(\mathbf{I} - \mathbf{M})\mathbf{Z} = \mathbf{E} \tag{7}$$

where **M** is the coefficient matrix derived from the SAM by dividing each column through by its column totals.

Than the multiplier is

$$\mathbf{Z} = (\mathbf{I} - \mathbf{M})^{-1}\mathbf{E} \tag{8}$$

2.2 The exogenous shock effects and linkages

Let’s suppose the exogenous shock in volume 1 mil $US in each sector. The multiplier effects are summarized in the Table 6 and total multiplier effects in the Table 7. An exogenous shock of increase in demand for services by 1 mil $US leads to a total increase of output by 1.81 mil $US. The multiplier takes account the total effect of all rounds and linkages in the economy. The second highest effect on output is in the utilities (1.75 mil $US) and agriculture (1.4 $US) sectors. The multiplier effect higher than one is also in the light manufacture (1.39 $US) and heavy manufacture (1.17$US) sectors. The exogenous demand increase has only a little total effect in the sectors of textiles (0.77 $US) and extraction (0.29 $US). The significant differences across sectors have to be taken into account when considered the total impact of exogenous demand increase or decrease.

The effects on GDP and income have therefore the same schedule. The effect on labor and capital is consistently the greatest in services, utilities and agriculture leading to the highest total effect on GDP and income in these sectors. Neither total multiplier effects reaches 1 $US indicating great tax burden across sectors. Concerning to the government; the highest multiplier effect is in the agriculture and heavy manufactures sectors.

Table 6: The multiplier effects

	AgrC	ExtrC	TextC	LMfgC	HMfgC	UtilC	SerC
AgrA	0.94	0.00	0.00	0.01	0.00	0.01	0.03
ExtrA	0.00	0.19	0.00	0.01	0.02	0.01	0.01
TextA	0.00	0.00	0.55	0.00	0.00	0.00	0.00
LMfgA	0.06	0.01	0.03	0.93	0.05	0.08	0.07
HMfgA	0.12	0.03	0.06	0.23	0.92	0.22	0.16
UtilA	0.04	0.02	0.02	0.04	0.05	1.13	0.08
SerA	0.24	0.05	0.11	0.17	0.14	0.28	1.46
AgrC	1.30	0.00	0.01	0.01	0.01	0.01	0.04
ExtrC	0.02	1.03	0.01	0.05	0.10	0.07	0.03
TextC	0.00	0.00	1.16	0.01	0.00	0.00	0.01
LMfgC	0.08	0.01	0.04	1.35	0.07	0.12	0.10
HMfgC	0.19	0.04	0.10	0.36	1.46	0.36	0.26
UtilC	0.04	0.02	0.02	0.05	0.05	1.22	0.09
SerC	0.25	0.05	0.11	0.18	0.14	0.30	1.54
lab	0.22	0.07	0.19	0.16	0.13	0.25	0.34
cap	0.24	0.07	0.09	0.18	0.14	0.38	0.40
hous	0.36	0.11	0.23	0.27	0.22	0.49	0.58
gov	0.15	0.03	0.12	0.11	0.14	0.12	0.11
s-i	0.06	0.02	0.02	0.05	0.04	0.10	0.11
row	0.43	0.84	0.62	0.57	0.61	0.30	0.20

Table 7: The total multiplier effects

	AgrC	ExtrC	TextC	LMfgC	HMfgC	UtilC	SerC
Output	1.40	0.29	0.77	1.39	1.17	1.75	1.81
GDP	0.46	0.14	0.28	0.34	0.27	0.63	0.74
Income	0.36	0.11	0.23	0.27	0.22	0.49	0.58

3 CONCLUSIONS

The social accounting matrix for Slovakia, year 2011 was constructed in this paper and the sectoral analysis was performed with special emphasis to the sectors inter and intra-industry linkages and their effects. The services sector (including trade sub-sector) constitutes the main part of the total GDP with more than fifty percents. The significant part of the intermediate consumption demands heavy manufactures, services and light manufactures (total of more than 80%). Households spend their income mostly on services and agricultures. The government consumption is composed solely of services. More than half of investments flow to utilities and one third to heavy manufactures. Exported are mostly heavy and light manufactures with more than 75% share on the total export. More than half of the agriculture output is consumed by households and more than 90% of extraction output is used in another sectors'.

The general formula for the unconstrained multiplier was derived and applied to the Slovak social accounting matrix. The increase of exogenous demand in each sector was executed and analyzed. Identified was the most important services sector, utilities and construction sector and manufactures sector. The high tax burden in each sector lowers the multiplier effect on GDP and income leading to the values less than the initial shock.

Acknowledgements: This paper research was supported by the Slovak Research and Development Agency, project No. APVV-14-0020: Pareto Optimal Interactions Modeling in the Financial and Business Adaptive Processes and by the the Slovak Ministry of Education, Science and Sport by grant VEGA 1/0340/16: Modeling of financing and cooperation efficiency of the Slovak economy sectors and regions.

References

[1] Aguiar, A., Narayanan, B. and McDougall, R. (2016). An Overview of the GTAP 9 Data Base. *Journal of Global Economic Analysis*, 1(1), 181-208.

[2] Breinsinger, C., Thomas, M. and Thurlow J. (2009). *Social Accounting Matrices and Multiplier Analysis.* Food Security in Practice Technical Guide 5. Washington, D.C.: International Food Policy Research Institute.

[3] Civardi, M., Pansini R.V. and Lenti, R.T. (2010). Extensions to the Multiplier Decomposition Approach in the SAM Framework: An Application to Vietnam. *Economic System Research,* 22(2), pp. 111-128.

[4] Duarte, R., Langarita, R. and Sanchez-Choliz, J. (2017). The Electricity Industry in Spain: A structural Analysis Using a Disaggregated Input-output Model. *Energy*, 141, pp. 2640-2651.

[5] Ferreira Filho, J. B. S.; Fachinello, A. L. (2015). Employment and Income Generation in the Brazilian Amazon Forest: A Social Account Matrix Based Multiplier Approach. *International Forestry Review*, 17(1), pp. 85-96.

[6] McBain, D. and Alsamawi, A. (2014). Quantitative Accounting for Social Economic Indicators. *Natural Resources Forum,* 38(3), pp. 193-202.

[7] Seung, C.K. (2014). Measuring Spillover Effects of Shocks to the Alaska Economy: An Inter-regional Social Accounting Matrix (IRSAM) Model Approach. *Economic Systems Research* 26(2), pp. 224-238.

Author's address

Ing. Veronika Miťková, PhD.
Comenius University in Bratislava, Faculty of Social and Economic Sciences,
Mlynské luhy 4, 821 05 Bratislava, Slovak Republic
email: veronika.mitkova@fse.uniba.sk

MACROECONOMIC DEVELOPMENT AND INVESTMENT OPPORTUNITIES IDENTIFICATION – HISTORICAL CONFRONTATION

Vladimír Mlynarovič, Faculty of Social and Economic Sciences, Comenius University in Bratislava

Abstract:
The paper presents multiple criteria approach for identification of investment opportunities among government bonds. It is based on a confrontation of their spreads and level and trends characteristic of the macroeconomic development of corresponding countries. The spread is defined as the difference between the yield to maturity of the specific country government bond and Germany government bond with the same maturity. The OECD macroeconomic data are used for macroeconomic ranking of selected countries. The paper also suggests an optimization approach to find weights of importance for selected macroeconomic criteria.

Keywords: *macroeconomic criteria, spreads, multiple criteria approach, criteria importance weights*

JEL Classification: C61
AMS Classification: 90C29

1 INTRODUCTION

There are many applications of known MCDM techniques that as the main and only result offer an outranking of examined variants or alternatives. In many cases a sensitivity analysis on changes in inputs is missing as well. The question about usefulness of such results can arise. In the paper we present extended version of the model presented in (Mlynarovič and Tuš, 2010) were we also use such techniques to rank countries on the base of their macroeconomic characteristics. But this result is confronted with the spreads of the countries with the goal to identify promising investment opportunities among government bonds. The paper also suggests some optimization approach to criteria weights modeling. The paper confronts its investments recommendation with known actual government bond yields as well.

2 MODEL

Suppose we have responses a_{ij}^t of different variants (countries) i, $i = 1, \ldots, m$, on different objectives (macroeconomic characteristics) j, $j = 1, \ldots, n$, in specific time periods (years) t, $t = 1, \ldots, T$. For these responses we define two types of macroeconomic criteria, the criteria of level and the criteria of trend. Let $[d_l, h_l]$, $1 \le d_l \le h_l \le T$, is the time period for level criteria and $[d_r, h_r]$, $1 \le d_r < d_l \le h_l \le h_r \le T$, is the time period for trend criteria, then

$$c_{ij}^l = \frac{\sum_{t=d_l}^{h_l} a_{ij}^t}{h_l - d_l + 1} \qquad (1)$$

is the level value of criterion j, $j = 1, \ldots, n$, for country i, $i = 1, \ldots, m$, and

$$c_{ij}^{r} = c_{ij}^{l} - \frac{\sum_{t=d_r}^{h_r} a_{ij}^{t}}{h_r - d_r + 1} \tag{2}$$

is the trend value of criterion $j, j = 1, \ldots, n$, for country i, $i = 1, \ldots, m$.

So we have a multiple criteria decision making problem with m alternatives and $2n$ criteria and we can use some of known methods, (e.g. Brans, Mareschal and Vincke, 1984, Mlynarovič, 1998 or Mlynarovič and Hozlár, 1989) to rank the alternatives. But in this application we suggest to use so called $z - scores$ instead of level and trend values of criteria. The advantageous of such approach is the fact that in this case we take into account not only the criteria values, but also their variability.

Let μ_j^l and σ_j^l are the mean and the standard deviation of $c_{ij}^l, i = 1, \ldots, m$, and μ_j^r and σ_j^r are the mean and the standard deviation of $c_{ij}^r, i = 1, \ldots, m$, then

$$s_{ij}^{l} = \frac{c_{ij}^{l} - \mu_j^{l}}{\sigma_j^{l}}, \quad i = 1, \ldots m, \ j = 1, \ldots n \tag{3}$$

are $z - scores$ for level criteria and

$$s_{ij}^{r} = \frac{c_{ij}^{r} - \mu_j^{r}}{\sigma_j^{r}}, \quad i = 1, \ldots m, \ j = 1, \ldots n \tag{4}$$

are $z - scores$ for trend criteria.

Countries macroeconomic outranking can be written as a multiple criteria decision making problem

$$\text{"max"} \quad \left\{ \mathbf{s}_i = \left(\mathbf{s}_i^l, \mathbf{s}_i^r\right) = \left(s_{i1}^l, \ldots s_{in}^l, s_{i1}^r, \ldots s_{in}^r\right) \quad i = 1, \ldots m \right\} \tag{5}$$

where, without loss of universality, it is assumed that “the more the better” is applied for all criteria. Let

$$\mathbf{w}^l = \left(w_1^l, w_2^l, \ldots, w_n^l\right), \quad \mathbf{w}^r = \left(w_1^r, w_2^r, \ldots, w_n^r\right) \tag{6}$$

are criteria weights where

$$\sum_{j=1}^{n} w_j^l + \sum_{j=1}^{n} w_j^r = 1 \tag{7}$$

$$w_j^l, w_j^r \geq 0, \ j = 1, 2, \ldots, n \tag{8}$$

then PROMETHEE II net flow for country i can be expressed as

$$\Phi_i(\mathbf{w}^l, \mathbf{w}^r) = \sum_{j=1}^{n} w_j^l \Phi_{ij}^l + \sum_{j=1}^{n} w_j^r \Phi_{ij}^r, i = 1, 2, \ldots, m \quad (9)$$

where Φ_{ij}^r and Φ_{ij}^l are contributions of the criterion *j* to the net flow of the country *i*.

Suppose that application of the PROMETHEE II method provides the values of net flows Φ_i, $i = 1, \ldots, m$, for which

$$\Phi_1 > \Phi_2 > \ldots \Phi_i > \ldots \Phi_m \quad (10)$$

It is known that in the PROMETHEE terminology it means that the country 1 is the best one and the country *m* is the worst one.

Let p_i, $i = 1, \ldots, m$, is the spread for country *i* defined as the difference between the yield to maturity of the specific country government bond and Germany government bond with the same maturity. It holds the higher value of p_i then higher probability of default for country *i*. Suppose that r_i, $i = 1, .., m$, is the yield to maturity of government bonds offered by the country *i*. In the world of ideal information for given PROMETHEE results should by hold

$$p_1 < p_2 < \ldots p_i < \ldots < p_m \quad (11)$$

$$r_1 < r_2 < \ldots r_i < \ldots < r_m \quad (12)$$

because the worse country the higher return must be offered. In this situation the potential investor has only to decide how much risk he is willing to accept. In practical situations one can hardly expect such unambiguous result and possible contradictions between macroeconomic results and spreads provide a space for identification good investment opportunities.

A very often question in such type of applications is the question about weights of the criteria. In a spite of the fact that PROMETHHE II provides an analytical tool for analysis of weights sensitivity (Mareschal,1988) specific values of weights are still questionable. A naïve approach with all weights equal or expert values of the weights miss analytical support. In the next we suggest an optimization approach to the weights specification. Suppose firstly that for period h_r (< *T*) the investment opportunities are identify and we know the that r_i, $i = 1, .., m$, is the yield of government bonds of the country *i* in period from h_r to *T*. Then for the period from 1 to h_r the criterion weights can be found as solution of the following optimization problem:

$$\min \quad \Phi_m(\mathbf{w}^l, \mathbf{w}^r) \quad (13)$$

subject to

$$\Phi_k(\mathbf{w}^l, \mathbf{w}^r) \geq \Phi_{k+1}(\mathbf{w}^l, \mathbf{w}^r), k = 1, 2, \ldots, m-1 \quad (14)$$

$$\sum_{j=1}^{n} w_j^l + \sum_{j=1}^{n} w_j^r = 1, \quad w_j^l, w_j^r \geq 0, j = 1, 2, \ldots, n \quad (15)$$

or, in the case the problem has no feasible solution, one can use a goal programming approach in the form

$$\min \quad \sum_{k=1}^{m=1} d_k \tag{16}$$

subject to

$$\Phi_k\left(\mathbf{w}^l, \mathbf{w}^r\right) - \Phi_{k+1}\left(\mathbf{w}^l, \mathbf{w}^r\right) + d_k \geq 0, \; k = 1, 2, \mathrm{K}\;, m-1 \tag{17}$$

$$\sum_{j=1}^{n} w_j^l + \sum_{j=1}^{n} w_j^r = 1, \quad w_j^l, w_j^r \geq 0, \; j = 1, 2, \mathrm{K}\;, n, \tag{18}$$

$$d_k \geq 0, \; k = 1, 2, \mathrm{K}\; m-1$$

The resulting optimal criteria weights ate the used for macroeconomic analysis for whole the whole period from 1 to T and for the identification attractive investment opportunities at the period (year) T.

3 DECISION SUPPORT SYSTEM AND ITS APPLICATIONS

The decision support system developed in excel environment realizes PROMETHEE II outranking method including sensitivity analysis for the importance weights for the selected criteria. The user can select: the period for level criteria, the period for trend criteria, countries from the list, criteria from the list and assign the importance weights, control parameters for PROMETHEE II method.

Table 1: Countries outranking and government bonds yields

Country	Promethee II Ranking	Spreads	Macroeconomics	Government Bonds Attractiveness	10years Bonds Performance
Sweden	0.1620	32.4	Good	high	29.16%
Poland	0.1595	159.7	Good	very high	23.63%
Slovakia	0.1193	134.9	Good	very high	32.51%
Denmark	0.1131	61.9	Good	high	-
Finland	0.0740	11.3	Good	low	-
Germany	0.0599	0	Good	none	31.10%
Netherlands	0.0586	14.3	Good	low	31.33%
Italy	0.0583	119.4	Good	very high	18.36%
United Kingdom	0.0563	0	Good	none	-
Luxembourg	0.0505	0	Good	none	-
Austria	0.0479	39.4	Good	high	33.36%
Belgium	0.0426	45.7	Good	high	31.28%
France	0.0201	19.5	Good	low	27.37%
Portugal	-0.0265	148.9	Bad	high	1.21%
Czech Republic	-0.0336	132.9	Bad	high	36.35%
Estonia	-0.0395	-	Bad	-	-
Slovenia	-0.0433	107.3	Bad	high	33.01%
Hungary	-0.0756	-	Bad	-	44.67%
Bulgaria	-0.0964	-	Bad	-	-
Ireland	-0.1481	153.1	Bad	high	39.09%
Spain	-0.1800	117.4	Bad	high	23.65%
Greece	-0.3795	-	Bad	high	-

The OECD (Eurostat database, 2017) regularly provides the historical and prognostic yearly macroeconomic data about: demand and output, wages, costs, unemployment and inflation, key supply side data, saving, fiscal balances and public indebtedness, interest rates and exchange rates, external trade and payments and other background data. These data for the period from 1992 to 2015 were used to construct the following macroeconomic criteria:

1. Gross value added at 2000 basic prices excluding FISIM: total economy
2. Private final consumption expenditure at 2000 prices
3. Gross fixed capital formation at 2000 prices: total economy
4. Unemployment rate: total:- Member States: definition EUROSTAT
5. Harmonized consumer price index (All-items)
6. Saving rate, gross: households and NPISH (Gross saving as percentage of gross Disposable income)
7. National income at current market prices
8. Final consumption expenditure of general government at 2000 prices
9. Budget deficit as % of GDP
10. Implicit interest rate: general government:- Interest as percent of gross public debt of preceding year Excessive deficit procedure (based on ESA 1995)
11. General Government expenditures as % of GDP
12. Nominal long-term interest rates
13. Total exports of goods: - Foreign trade statistics
14. Total imports of goods: - Foreign trade statistics
15. Current account balances as a percentage of GDP
16. Export market growth in goods and services

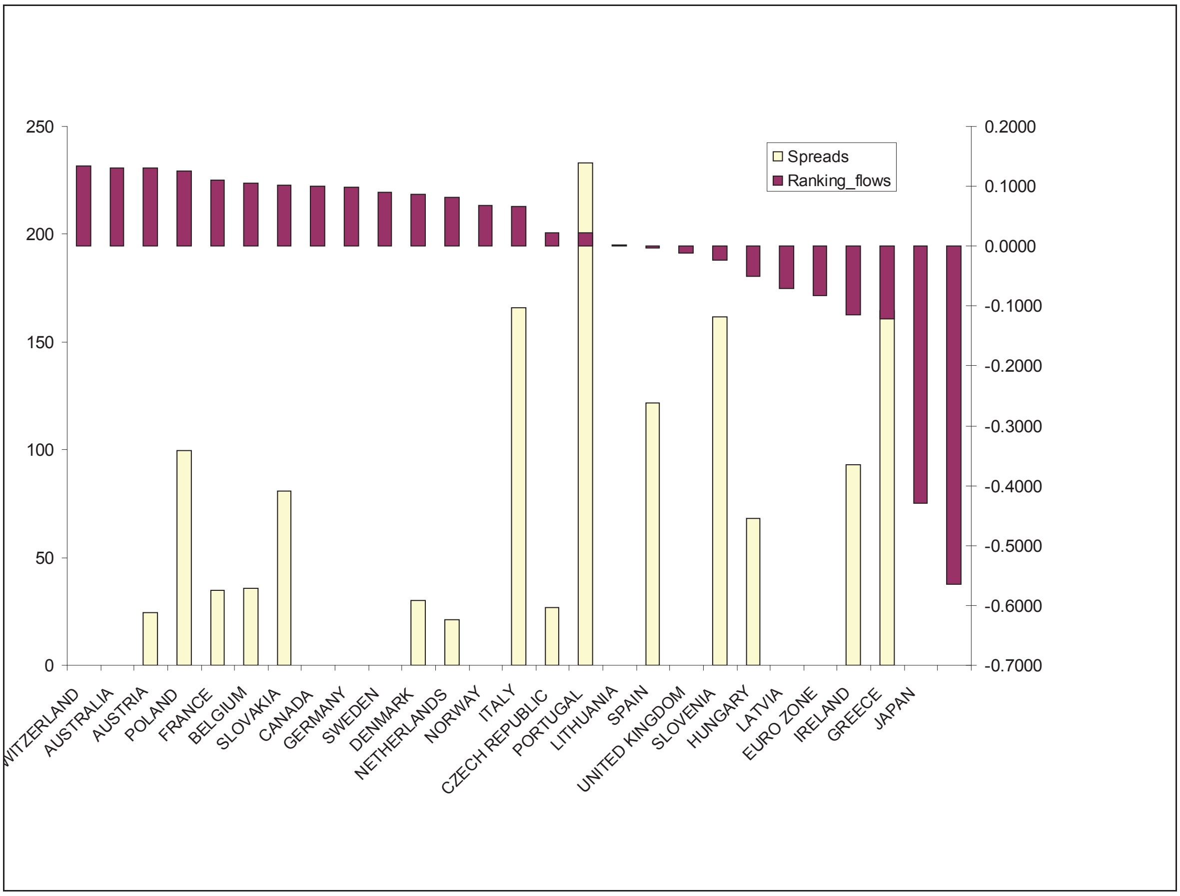


Figure 1: Results for spreads owing to Germany bonds

Macroeconomic results and investment recommendations derived from these data are finally confronted with spreads in the year 2010 and government bond yields in the period from 2010 to 2015 (Bloomberg, 2016) as it is presented at the Table 1. These government bonds yields were then used for weights optimization for macroeconomic data for period from 1992 to 2010 and used for investment opportunities identification in the year 2015 as it is shown in Figure 1.

4 CONCLUSIONS

In the paper it is shown that a combination of multiple criteria macroeconomic analysis with criteria weights optimization and historical government bonds yields in the form of the spreads via benchmark can by looking forward approach for investment recommendations.

Acknowledgements

This paper research was supported by the Slovak Research and Development Agency, project No. APVV-14-0020: "Pareto Optimal Interactions Modeling in the Financial and Business Adaptive Processes", and by the Slovak Ministry of Education, Science and Sport, project No. VEGA 1/0340/16: "Modelovanie efektívnosti financovania a kooperácie sektorov a regiónov slovenskej ekonomiky."

References

[1] Brans, J.P., Mareschal, B. and Vincke, Ph. (1984). PROMETHEE – A New Family of Outranking Methods in Multicriteria Analysis. In: Brans, J.P. (ed.) *Operational Research 84*. Amsterdam: North Holland, pp. 477-490.

[2] Mareschal, B. (1988). Weight Stability Intervals in Multicriteria Decision Aid. *EJ of OR*, 33, pp. 54 – 64.

[3] Mlynarovič, V. (1998). *Modely a metódy viackriteriálneho rozhodovania.* Bratislava: Ekonóm.

[4] Mlynarovič, V. and Hozlár, E. (1989). Metódy triedy PROMETHEE pre komplexné vyhodnocovanie variantov a ich aplikácia. *Ekonomicko – matematický obzor*, 25(4), pp. 435-452.

[5] Mlynarovič, V. and Tuš, B. (2010). Investment opportunities identification - a multiple criteria approach. In: Quantitative Methods in Economics (Multiple Criteria Decision Making XV). Bratislava: Ekonóm, pp. 139-147.

[6] Bloomberg (2017). *Bloomberg Professional*. [Online]. Available at: Subscription service (Accessed: June, 2017).

[7] Eurostat, (2017). *Database – Eurostat*. [Online] Available at: http://ec.europa.eu/eurostat/data/database (Accessed: 7 June, 2017).

Author's address

doc. Ing. Vladimír Mlynarovič, CSc.
Fakulta sociálnych a ekonomických vied Univerzity Komenského v Bratislave
Mlynské luhy 4, 821 05 Bratislava
email: vladimir.mlynarovic@fses.uniba.sk

THE APPLICATION OF MULTIPLE CRITERIA DECISION MAKING METHODS IN TRANSPORT

Ivana Olivková, VŠB – Technical University of Ostrava

Abstract

This paper deals with the application of the multicriteria decision making methods in specific decision situation in the public transport. In the transport practice there are often cases when it is necessary to decide on the variants of fundamental importance the earliest opportunity but the availability, quantity and quality of information for decision usually does not correspond to their real needs. In this case there are needed detail knowledges about individual variants and that is why the evaluation is carried out by experts and can be objectified their evaluation. For solving such type of the task can be employed the methods based on the thresholds of sensitivity. These methods are used to rank a finite number of decision variants especially in a case when neither a direct evaluation nor the utility theory gives satisfactory results. From this group of methods have been chosen AGREPREF method and approximation of fuzzy relations method. These methods are due to its nature suitable for evaluating variants with a set of qualitative criteria.

Keywords: *Multiple criteria decision making, Methods based on the thresholds of sensitivity, Transport.*

JEL Classification: C44
AMS Classification: 90C15

1 INTRODUCTION

The multicriteria decision making, as the name itself indicates, deals with the evaluation of particular variants according to several criteria. The term “variant” designates each of the solutions of the decision problem. The “criterion” is a property that is being evaluated with the given variant. To each criterion is assigned weight that expresses the importance of particular criteria with regard to the others.

Multiple criteria decision making (MCDM) is considered as a complex decision making tool involving both quantitative and qualitative factors. In recent years, several MCDM techniques and approaches have been suggested to choosing the optimal options. Their application depends on the type, completeness and on level of detail of available information.

In recent years several previous studies have employed MCDM tools and applications to solve area problems in transport. For example the AHP method was used for selection the best support design for the main transport road (Yavuz, Iphar and Once, 2008). In the field of transport sustainability was used ELECTRE method (Bojković, Anić and Pejčić-Tarle, 2010). For selection if the best transportation policy was applied TOPSIS method (Ülengin *et al.*, 2016).

In some practical situations, there are often cases when it is necessary to decide on the variants of fundamental importance but the availability, quantity, and quality of information for decision usually does not correspond to their real needs. Therefore, for evaluation difficult quantifiable criteria is important expert assessment. For solving such type of the task can be employed the methods based on the thresholds of sensitivity.

2 METHODOLOGY

The initial step of each MCDM method is to form an evaluating matrix, the elements of which reflect the evaluation of particular criteria $K=\{K_1, K_2, \ldots, K_m\}$ for each variant $X = \{x_1, x_2, , x_n\}$. The matrix Y consists then of elements y_{ij} where $i = 1,\ldots,n$ variants and $j = 1,\ldots, m$ criteria.

The evaluating matrix:

$$Y = \begin{array}{c c} & \begin{array}{ccc} K_1 & \ldots & K_m \end{array} \\ \begin{array}{c} x_1 \\ \vdots \\ x_n \end{array} & \begin{bmatrix} y_{11} & \cdots & y_{1m} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nm} \end{bmatrix} \end{array} \tag{1}$$

To determine the optimal variant $x^* \epsilon X$, it is sufficient to arrange the variants from X on the basis of their overall evaluation with respect to the criteria from K. The variant that occupies the first place in this arrangement is then the optimal variant (Fotr and Píšek, 1986).

Let's assume that is set the relative importance of the criteria in the form of weight (2):

$$v = (v_1, v_2, \ldots, v_m), \sum_{i=1}^{m} v_i = 1, v_i \geq 0 \tag{2}$$

where $v_1, v_2, \ldots, v_m$ are weights of criteria.

The aim of methods based on the thresholds of sensitivity is creation of preference variants matrix V. For every couple of variants x_j and x_k is aggregate the criteria which:

- prefer variant x_j before variant x_k, set their indexes mark as I_{jk}
- prefer variant x_k before variant x_j, set their indexes mark as I_{kj}
- both of variants have equivalent values and from viewpoints of these criteria are indifferent, set their indexes mark as $I_{j\sim k}$

Degree of preference variant x_j before x_k is:

$$V_{jk} = \sum_{i \in I_{jk}} v_i \tag{3}$$

Degree of preference variant x_k before x_j is:

$$V_{kj} = \sum_{i \in I_{kj}} v_i \tag{4}$$

Degree of indifference variant x_j and x_k is:

$$V_{j\sim k} = \sum_{i \in I_{j\sim k}} v_i \tag{5}$$

From sum of degree of preference is formed preference of variant matrix $V=[V_{jk}]$. This matrix is default for following methods (Fotr and Píšek, 1986).

2.1 AGREPREF method

The basic idea of this method is to compare variants according to the preference from best to worst. The first step is the creation preference relation matrix $P=[P_{jk}]$ from the matrix V using (6).

$$P_{jk} = \begin{cases} 1, if\ V_{jk} > V_{kj} \\ 0, if\ V_{jk} \leq V_{kj} \end{cases} \quad (6)$$

P is a square matrix with zeros on the main diagonal. This matrix is then processed. The procedure is as follows.
First is calculated value *Dj*, which indicated difference between count of variants, before which is variant preferred, and count of variants, which are preferred before given variant (7).

$$D_j = \sum_k P_{jk} - \sum_k P_{kj} \quad (7)$$

According of decreasing values of characteristic D_j are gradually rearranged both rows and columns of the matrix *P*. Thus calculated matrix is symbolized by $R=[R_{jk}]$ (indexes *j,k* no longer correspond identification of variant by numbers, but these are indexes row and column of elements of the matrix *R*). The matrix *R* is modified to ensure its transitivity. If $r_{jk} = 1$ and $r_{kl} = 1$, have to be $r_{jl} = 1$. If $r_{jl} = 0$, this element is equal to *1*.

The obtained matrix is again rearranged according to the recalculated characteristic *D* and modified to meet the asymmetry conditions. If $r_{jk} = 1$, it have to be ensured that element $r_{kj} = 0$. According of decreasing values D_j are again gradually rearranged rows. Final matrix R has the following characteristics:

- in lower triangular submatrix are zeros;
- in the upper triangular submatrix are numbers one;
- on the main diagonal are zeros.

In the matrix R is resulting arrangement of variants from best variant to worst variant.

2.2 Approximation of fuzzy relations method

This method is based on the idea of sorting the variants from the best to the worst according to decreasing row sums of the *V** matrix. Practical procedure begins by modifying preferences of variants matrix *V* using relation (8).

$$V^*_{jk} = \frac{1}{2} + \frac{1}{2}\left(V_{jk} - V_{kj}\right) \quad (8)$$

Thus calculated strict preference matrix is symbolized $V^*=[V^*_{jk}]$. This step ensures strict asymmetry. The row sums of the *V** matrix are calculated. According of decreasing values of characteristic *Sj* are gradually rearranged both rows and columns of the *V** matrix.

Thus calculated matrix is symbolized by $T=[T_{jk}]$. Then it is necessary to determine whether the matrix *T* fulfill the condition of transitivity, i.e. if for all *(j, k, l)*, where $j > k > l$ is valid. If this is not the case, the matrix *T* must be approximated by *W* matrix.

$$W = \frac{1}{2}(W^1 + W^2) \quad (9)$$

where

$$W^1 = \left[W^1_{jk} \right], \quad W^1_{jk} = \max_{\substack{p \geq j \\ q \leq k}} T_{pq} \quad (j \leq k) \quad (10)$$

$$W^1_{kj} = 1 - W^1_{jk} \quad (11)$$

$$W_{jj}^{1} = 0,5 \tag{12}$$

$$W^2 = \left[W_{jk}^{2} \right], \quad W_{jk}^{2} = \min_{\substack{p \le j \\ q \ge k}} T_{pq} \quad (j \le k) \tag{13}$$

$$W_{kj}^{2} = 1 - W_{jk}^{2} \tag{14}$$

$$W_{jj}^{2} = 0,5 \tag{15}$$

In the matrix *W* is resulting arrangement of variants from best variant to worst variant.

3 THE APPLICATION OF METHODS

MCDM methods are applied to select the most suitable fare collection systems in public transport. The possibilities of modern technology in the fare collection systems in public transport are discussed for example in publication (Olivková, 2016). Based on these sources are identified following variants:

- Check in - is based on the principle of login of the passenger through smart cards
- Check in/Check out - is based on the principle of login and logout of the passenger through smart cards
- Be-In/Be-out - is system eliminating any further manipulation of the media when getting in and out; the presence of a passenger in the vehicle is detected with Be-In/Be-Out system through a specific personal smart card
- SMS ticket - is a specific type of electronic fare collection system that allows electronic ordering of tickets through SMS sent from a mobile phone
- NFC (Near Field Communication) - is a new technology that combines smart card (RFID) and mobile phone

Aim is to find a variant with the best evaluation of all criteria, eventually these variants appropriately arrange.

The data for the evaluation was obtained by survey of selected group of evaluators (experts on traffic issues) from Brno and Ostrava from January to April 2017. The evaluators were informed about the variants of fare collection systems and based on their expertise they evaluated variants through criteria.

Results of analysis can be exported from Excel table. Table 1 represents criteria and variants of solution. Each criterion is represented by one column, and each row represented one variant. Results in table 1 shows average value of scores of individual variants according to selected criteria. Rating scale is in the range from 0 to 1 b. The higher is number, the better is variant evaluated. Table 1 is the default matrix for solving.

Criteria / Variants	Simplicity and comfort	Operation costs	Multipurpose	Safety	Demands of applications	Speed
Check-In	0,68	0,67	0,59	0,48	0,63	0,76
Check-In/Check-Out	0,62	0,65	0,55	0,43	0,61	0,51
Be-In/Be-Out	0,89	0,51	0,41	0,57	0,42	0,98
SMS	0,72	0,46	0,43	0,67	0,81	0,43
NFC	0,78	0,61	0,92	0,75	0,84	0,81

Table 1 Matrix Y

The second step is determined the weight of each criteria. Weights of criteria reflect their importance from the point of view of the evaluators. For the determination of weight is applied (2). Weights is calculated on the basis of scoring method. Scoring range is from 1 to 10. The weights are normalized (the sum of set of weights is equal to one). The normalization of criteria weights is done according to (16):

$$v_i = \frac{b_i}{\sum_{i=1}^{m} b_i} \tag{16}$$

v_i is weight of criteria and b_i is score of criteria

Average value of scores and average values of the weights of individual criteria are displayed in the Table 2.

Criteria / Calculation	Simplicity and comfort	Operation costs	Multipurpose	Safety	Demands of applications	Speed
Score	4,8	4,6	7,8	8,6	5,6	3,6
Weight	0,14	0,13	0,22	0,25	0,16	0,10

Table 2 Weights of criteria

The next step is creation of preference variants matrix *V*. The procedure is described in chapter 2. From sum of the degree of preference V_{jk} and V_{kj} is formed preference of variant matrix $V=[V_{jk}]$. Elements of matrix *V* are displayed in table 3.

Variants	Check-In	Check-In/ Check-Out	Be-In/ Be-Out	SMS	NFC
Check-In	0	1	0,51	0,45	0,13
Check-In/ Check-Out	0	0	0,51	0,45	0,13
Be-In/ Be-Out	0,49	0,49	0	0,37	0,24
SMS	0,55	0,55	0,63	0	0
NFC	0,87	0,87	0,76	1	0

Table 3 Preference of variant matrix V

Preference of variant matrix *V* is the default for following method.

3.1 The application AGREPREF method

The first step of AGREPREF method is creation of preferences relation matrix *P* from matrix *V* and calculation values D_j. The procedure is described in chapter 2.1. Elements of matrix P and values D_j are displayed in the table 4.

Variants	Check-In	Check-In/ Check-Out	Be-In/ Be-Out	SMS	NFC	D_j
Check-In	0	1	1	0	0	0
Check-In/ Check-Out	0	0	1	0	0	-2
Be-In/ Be-Out	0	0	0	0	0	-4
SMS	1	1	1	0	0	2
NFC	1	1	1	1	0	4

Table 4 Preferences relation matrix *P*

P is square matrix with zeros on the main diagonal. This matrix is then processed.
According of decreasing values of characteristic D_j are gradually rearranged both rows and columns of the matrix *P*. This calculated matrix *R* is displayed in table 5.

Variants	NFC	SMS	Check-In	Check-In/ Check-Out	Be-In/ Be-Out	D_j
NFC	0	1	1	1	1	4
SMS	0	0	1	1	1	2
Check-In	0	0	0	1	1	0
Check-In/ Check-Out	0	0	0	0	1	-2
Be-In/ Be-Out	0	0	0	0	0	-4

Table 5 Matrix R

The obtained matrix meets the asymmetry conditions. In the matrix *R* is resulting arrangement of variants from best variant to worst variant.

3.2 The application approximation of fuzzy relations method

The first step of approximation of fuzzy relations method is the creation strict preference matrix *V** from matrix *V* and calculation values of row sum *Sj*. The procedure is described in chapter 2.2. Elements of matrix *V** and values S_j are displayed in the table 6.

Variants	Check-In	Check-In/ Check-Out	Be-In/ Be-Out	SMS	NFC	S_j
Check-In	0,5	1	0,51	0,45	0,13	2,59
Check-In/ Check-Out	0	0,5	0,51	0,45	0,13	1,59
Be-In/ Be-Out	0,49	0,49	0,5	0,37	0,24	2,09
SMS	0,55	0,55	0,63	0,5	0	2,23
NFC	0,87	0,87	0,76	1	0,5	4

Table 6 Matrix *V**

According of decreasing values of characteristic *Sj* are gradually rearranged both rows and columns of the *V** matrix. This calculated matrix is displayed in table 7.

Variants	NFC	Check-In	SMS	Be-In/ Be-Out	Check-In/ Check-Out
NFC	0,5	0,87	1	0,76	0,87
Check-In	0,13	0,5	0,45	0,51	1
SMS	0	0,55	0,5	0,63	0,55
Be-In/ Be-Out	0,24	0,49	0,37	0,5	0,49
Check-In/ Check-Out	0,13	0	0,45	0,51	0,5

Table 7 Matrix *T*

The matrix *T* does not fulfil the condition of transitivity. It must be approximated by the matrix *W*. Matrix *W* is created based on the arithmetic mean of matrixes W^1 and W^2. This calculated matrix is displayed in table 8.

Variants	NFC	Check-In	SMS	Be-In/ Be-Out	Check-In/ Check-Out	U_j
NFC	0,5	0,82	0,88	0,88	0,94	4,01
Check-In	0,19	0,5	0,50	0,57	0,94	2,69
SMS	0,12	0,50	0,5	0,57	0,59	2,28
Be-In/ Be-Out	0,12	0,43	0,43	0,5	0,50	1,98
Check-In/ Check-Out	0,07	0,07	0,41	0,50	0,5	1,54

Table 8 Matrix *W*

In the matrix *W* is resulting arrangement of variants from best variant to worst variant.

4 CONCLUSIONS

This article deals with the application of MCDM methods in specific decision situation in the public transport. The main aim of this paper is to use this method on large datasets and verify the accuracy of the method based on the thresholds of sensitivity. The basis for this group of methods is to determine the preferential relationships of all pairs of variants with respect to the individual criteria.

When comparing the preferential arrangement of variants of individual methods, it is clear that the order is not identical. However, according to both methods, the best variant of fare collection systems in public transport is NFC system. Check-In and SMS variants of fare collection systems in public transport are suitable too. The decision-maker could choose one of these variants as an alternative to the NFC variant. But variants Be-In / Be-Out and Check-In / Check-Out appear to be less suitable and in making further decisions would no longer be taken into account.

The main difference between methods based on the thresholds of sensitivity to other MCDM methods is that it is not possible to obtain a numerical overall evaluation of the individual variants. Advantage of this approach is the possibility of solving of the decision problems:

- with multiple criteria
- with more decision variants and
- with different type of all kind of input data.

Acknowledgements: This paper was supported by the internal research project SP2018/110 Research and Development in Transport.

References

[1] Bojković, N., Anić, I. and Pejčić-Tarle, S. (2010). One solution for cross-country transport sustainability evaluation using a modified ELECTRE method. *Ecological Economics*, 69(5), pp. 1176–1186. DOI: 10.1016/j.ecolecon.2010.01.006, ISSN: 0921-8009.

[2] Fotr J., and Píšek M. (1986). *Exaktní metody ekonomického rozhodování*. Academia: Praha.

[3] Olivková, I. (2016). Methodology of Multi-criteria Evaluation of Variants of Check-in Systems in Public Transport. In: *Proceedings of International Scientific Conference Quantitative Methods in Economics – Multiple Criteria Decision Making XVIII*. Bratislava: Letra Interactive, pp. 264 – 270. ISBN 978-80-972328-0-1.

[4] Ülengin, F., Kabak, Ö., Önsel, Ş., Ülengin, B. and Aktaş, E. (2010). A problem-structuring model for analyzing transportation–environment relationships. *European Journal of Operational Research*, 200(3), pp. 844–859. DOI: 10.1016/j.ejor.2009.01.023, ISSN: 0377-2217.

[5] Yavuz, M., Iphar, M. and Once, G. (2008). The optimum support design selection by using AHP method for the main haulage road in WLC Tuncbilek colliery. *Tunnelling and Underground Space Technology*, 23(2), pp. 111–119. DOI: 10.1016/j.tust.2007.02.001, ISSN: 0886-7798.

Author's address

doc. Ing. Ivana Olivková, Ph.D.
VŠB–Technical University of Ostrava, Faculty of Mechanical Engineering, Institute of Transport, 17. listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic
email: ivana.olivkova@vsb.cz

ON THE STRUCTURE OF INDEPENDENT SETS IN TRIP DIGRAPH AND THEIR APPLICATIONS IN OPTIMUM VEHICLE SCHEDULING

Stanislav Palúch, Štefan Peško, Tomáš Majer,
University of Žilina, Faculty of Managements Science and informatics

Abstract
Vehicle scheduling problem is to arrange a given set of trips into minimum number of running boards fulfilling several additional constraints. There are several mathematical models for just mentioned problem - e.g. mathematical programming model, bipartite digraph model and trip digraph model. Trip digraph is an acyclic transitive digraph with vertex set equal to the set of all trips and arcs representing linkability of trips. Every directed path in trip digraph represents a running board. Maximum independent sets in trip digraph can be used for exact decomposition of vehicle scheduling problem. This paper studies the structure of maximum independent sets in an acyclic transitive digraph and shows how they can be used for simplification of corresponding models.

Keywords: *Vehicle scheduling, mathematical programming, digraph*

JEL Classification: C610
AMS Classification: 0C850, 90C05, 65K05

1 INTRODUCTION

1.1 The state of art and motivation

The first person to use independent sets in scheduling was K. Vašek in (Vašek, 1985) and then D. Engelthaller in (Engelthaller, 1987) where he proposed a man machine approach to minimizing number of vehicles by small changes of the set of trips. Independent sets were used for decomposition of real scheduling instances, since size of RAM was small at those times to contain data of whole instance.

The sizes of RAM are sufficiently large nowadays so decomposition of scheduling instances is not necessary. On the other hand, the performance and efficiency of solvers skyrocketed such that many problems can be solved today exactly in reasonable time. However, the computation time can depend on the number of decision variables. This paper shows, how independent set can reduce the number of variables of several scheduling problems.

1.2 Fundamental notions

Fundamental elements of a regional or municipal regular passenger transport are trips. A trip is defined by its departure time, departure place, arrival time and arrival place. A precedence relation $\prec$ on the given set of trips T can be introduced as follows: We will say that the trip $t_i \in T$ precedes the trip $t_j \in T$ and write $t_i \prec t_j$ if the same bus can move from arrival place of the trip t_i sufficiently early before departure time of the trip t_j. Sometimes we will say that trip t_j is linkable behind trip t_i if $t_i \prec t_j$. Some expenses $c(i,j)$ can be associated with a linkage $t_i \prec t_j$ as the cost of dead mileage from arrival place of t_i to departure place of t_j and penalties for line change, waiting time etc. Precedence relation on the set T of trips is ireflexive and transitive.

Trips of a given set T are arranged into daily schedules of work of vehicles called running boards. A running board is a sequence of trips $(t_1, t_2, \dots, t_r)$ such that $t_1 \prec t_2 \prec \dots \prec t_r$. The cost $c(t_1, t_2, \dots, t_r)$ of running board $(t_1, t_2, \dots, t_r)$ is

$$c(t_1, t_2, \dots, t_r) = \sum_{i=1}^{r-1} c(t_i, t_{i+1})$$

Fundamental vehicle scheduling problem is to arrange all trips of a given set of trips T into minimum number of running boards with minimum total cost of all running boards. The set of all running boards is called a bus schedule.

Let us have a bus schedule. We will say that the trip t_j is linked behind trip t_i and write $t_i \preccurlyeq t_j$, if the trip t_j immediately follows after trip t_i in a running board of the given bus schedule.

Remark. A bus schedule is fully defined by the set of ordered pairs

$$\{(t_i, t_j) \mid t_i \in T, t_j \in T, t_i \preccurlyeq t_j \}$$

Several approaches and many mathematical models are used for modeling of vehicle scheduling – a graph theory approach (Czimmermann and Peško, 2004), (Palúch, 2001) and (Peško, 2001) using flow model, model of graph coloring (Palúch, 2003) and also integer programming models (Palúch and Majer, 2016), (Palúch, Peško and Majer, 2016), (Peško, Palúch and Majer, 2016) used for more complex problems.

2 MATHEMATICAL MODELS FOR FUNDAMENTAL VEHICLE SCHEDULING PROBLEM

2.1 A graph theory model

Let us have a set of trips $T = \{t_1, t_2, \dots, t_n\}$ with precedence relation $\prec$, and linkage cost $c(t_i, t_j)$. Let us define the set A as follows:

$$A = \{(i,j) \mid i = 1,2, \dots, n, \qquad j = 1,2, \dots, n, \qquad t_i \prec t_j\}$$

A trip digraph is a digraph $G = (V, A, c)$ with vertex set $V = (1,2, \dots, n)$, arc set A and arc cost $c(i,j) = c(t_i, t_j)$. A trip digraph is a transitive acyclic digraph.

Agreement. We will often write $i \prec j$ instead of $(i,j) \in A$ since (by our opinion) first notation describes better real essence of relation “to be linkable”. We will say that $i, j \in V$ are comparable if $i \prec j$ or $j \prec i$. Otherwise we will say that i, j are incomparable.

Every directed path in trip digraph G represents a feasible running board with cost equal to the length of corresponding path. To find an optimum bus schedule means to find a minimum number of directed paths with minimum total length containing all vertices of G. This problem is polynomial solvable by methods similar to Ford Fulkerson algorithms for max-flow- min-cost problem.

It follows from Dilworth theorem that the minimum number of bus schedules containing all trips is equal to the cardinality of maximum independent set in corresponding trip digraph G. Let $I = \{s_1, s_2, \dots, s_r\}$ be a maximum cardinality independent set in trip digraph G, $s \in V$. Then there are no two vertices $s_p \in I$, $s_q \in I$ such that it holds $s_q \prec s$ together with $s \prec s_p$, otherwise it would emerge from transitivity that $s_q \prec s_p$ what is in contradiction of independence of I.

The set V is decomposed into three disjoint subsets U, I, W such that for arbitrary $s \in U$ there exists at least one $s_p \in I$ such that $s \prec s_p$ and for every $s \in W$ there exists at least one $s_q \in I$ such that $s_q \prec s$.

No directed path of an optimum solution covering trip digraph can contain an arc $(i,j) \in A$ such, $i \in U$ and $j \in W$.

This is because every path of optimum solution has to contain exactly one vertex of maximum cardinality independent set I. Let us have a path μ of optimum solution containing an arc $(i,j) \in A$ such that $i \in U$ and $j \in W$. Path μ contains also one vertex $s_p \in I$. Digraph G is transitive that is why every two vertices of this path μ are comparable. Vertex s_p lies on μ in front of arc (i,j) or behind it. If s_p lied on μ in front of arc (i,j) then it would hold $s_p \prec i$ what would imply that $i \in W$, what is in contradiction with assumption that $i \in U$. If s_p lied on μ behind arc (i,j) then it would hold $j \prec s_p$ and then $j \in U$ what contradicts to assumption that $j \in W$.

Theorem. Let T be a set of trips, let $G = (V, A, c)$ be corresponding trip digraph. Let I be a maximum cardinality independent set in G, let sets U, W are defined as above. Assign $\bar{G} = (V, A \setminus U \times W, c')$. Let P be an optimum solution for an objective function having as the main criterion minimization of the number of vehicles in digraph $\bar{G}$. Then P is an optimum solution for the same objective function in digraph G.

2.2 Integer programming model

Let us have a set of trips $T = \{t_1, t_2, \dots, t_n\}$ with precedence relation $\prec$, let

$$A = \{(i,j) \mid i \in S, j \in S, t_i \prec t_j\}. \tag{1}$$

Denote $c_{ij} = c(t_i, t_j)$. Let x_{ij} be a binary decision variable defined for every pair $(i,j) \in A$ with the following meaning:

$$x_{ij} = \begin{cases} 1, & \text{if } t_i \preccurlyeq t_j \text{ i.e. if the trip } t_j \text{ is linked behind trip } t_i \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

Essential integer linear programming model for fundamental vehicle scheduling problem is:

$$\text{Maximize} \quad \sum_{(i,j)\in A} (K - c_{ij}) x_{ij} \quad \text{where } K \text{ is a large number} \tag{3}$$

$$\text{subject to:} \quad \sum_{i \in T} x_{ij} \le 1 \qquad \text{for } j \in T \tag{4}$$

$$\sum_{j \in T}^{n} x_{ij} \le 1 \qquad \text{for } i \in T \tag{5}$$

$$x_{ij} \in \{0,1\} \qquad \text{for } i \in T, \quad j \in T \tag{6}$$

Problem (3) – (6) is a special instance of assignment problem, where condition (6) can be replaced by constraint

$$0 \le x_{ij} \qquad \text{for } \ i \in T, \quad j \in T \tag{7}$$

and resulting problem (3) – (6), (7) is a linear programming problem which is solvable in very short time even for enormous number of trips.

Mathematical model (3) – (6), (7) amended with additional constraints or changed objective function (3) can be used for flexible bus scheduling (Palúch, Peško and Majer, 2016), (Peško, 2001), scheduling with roundabout routes to depot (Palúch and Majer, 2016). An extension of model (3) – (6), (7) is contained in (Peško, Palúch and Majer, 2016).

Vehicle and driver are closely related in Slovak and Czech conditions therefore running boards of vehicles should comply to several constraints laid to crews (Černá, 2004). However, integrated vehicle and crew scheduling is studied also in other countries (Prata, 2016).

Additional constraints have to be added to the model (3) – (6) in order to obtain running boards feasible also to crews – mainly to ensure safety break, meal break and roundabout routes of vehicles into depot. Procedure for computing a feasible bus schedule starts with solution of problem (3) – (6). Then such parts of running boards are identified that do not comply with desired requirements. Such infeasible parts can by prohibited by constraints formulated as linear constraints. Extended model is no longer polynomial; however, it is computable in reasonable time even for real world instances. Resulting bus schedule is again analyzed and if any infeasible parts are found additional corresponding constrains are added to recent model, etc. Procedure continues until a feasible solution is obtained.

Minimization of the number of vehicles remains still the main objective. Let I be a maximum cardinality independent set in corresponding trip digraph $G = (V, A, c)$, let

$$U = \{i \mid i \in V \text{ such that there exists } s \in I \text{ such that } i \prec s\} \tag{8}$$
$$W = \{j \mid j \in V \text{ such that there exists } s \in I \text{ such that } s \prec j\} \tag{9}$$

Then no solution having minimum number of running boards – no path of path covering of G – can contain arc (i, j) such that $i \in U$ and $j \in W$, i.e. $(i, j) \in U \times W$.
Therefore it suffices to define decision variables x_{ij} only for $(i, j) \in A \setminus U \times W$ and maximize objective function

$$\sum_{(i,j) \in A \setminus U \times W} (K - c_{ij}) x_{ij}$$

Reduction of the number of variables can lead to decrease computing time.

Similarly, mathematical models for problems treated in (Palúch and Majer, 2016), (Palúch, Peško and Majer, 2016), (Peško, Palúch and Majer, 2016), (Peško, 2001) and similar can arise from model (3) – (6), (7) amended by another set of constraints. If minimization of used vehicles remains still the main objective, new extended models can use reduced set of variables.

Trip digraph can contain even more maximum cardinality independent sets. Therefore, it is convenient to study their structure in acyclic transitive digraphs.

3 STRUCTURE OF MAXIMUM CARDINALITY INDEPENDENT SETS

Definition. Let $G = (V, A)$ be a digraph. We will say that $M \subseteq V$ is an independent set, if for arbitrary $u \in M$, $v \in M$ $(u, v) \notin A$. We will say that $M \subseteq V$ is a maximum cardinality independent set, if M is independent set with maximum cardinality.

Theorem. Dilworth. Let $G = (V, A)$ be an acyclic transitive digraph. The minimum number of path in G covering all vertices of V is equal to the cardinality of maximum cardinality independent set in G.

Agreement about notation. We will use symbol $\mathcal{M}$ for the set of all maximum cardinality independent sets in an acyclic transitive digraph G.

Theorem. Let $G = (V, A)$ be an acyclic transitive digraph, let $M \in \mathcal{M}$ be a maximum cardinality independent set in G. Then it holds for every vertex $v \in V \setminus M$ exactly one of following assertions:

a) or there exists a $x \in M$ such that $(x, v) \in A$ (i.e. $x \prec v$)

b) or there exists a $y \in M$ such that $(v, y) \in A$ (i.e. $v \prec y$)

but a) and b) cannot hold simultaneously.

Proof.

Suppose that no one from a) and b) holds. Then the set $M \cup \{v\}$ is an independent set what contradicts with maximality of M.

Suppose that there exist $x \in M, y \in M$ such that $x \prec v \prec y$. Digraph G is transitive, therefore $x \prec y$ what contradicts to independence of M.

Definition. Let $G = (V, A)$ be an acyclic transitive digraph, let $M \in \mathcal{M}$ be a maximum cardinality independent set in G.

We will say that the vertex $v \in V \setminus M$ lies in front of M and write $v \prec M$ if there exists a $x \in M$ such that$(v, x) \in A$ i.e. $v \prec x$.

We will say that the vertex $v \in V \setminus M$ lies behind M and write $M \prec v$, if there exists a $y \in M$ such that $(y, v) \in A$ i.e. $y \prec v$.

We will say that the vertex $v \in V \setminus M$ is comparable with M if $v \prec M$ or $M \prec v$.

Remark. Let $M \in \mathcal{M}$ be a maximum cardinality independent set in an acyclic transitive digraph G. Then every vertex is comparable with the set M.

Next definition specifies an ordering on the set of all maximum independent sets in G.

Definition. Let $G = (V, A)$ be an acyclic transitive digraph, let $U, W \in \mathcal{M}$ are two maximum independent sets in G. We will say that the set U precedes the set W and write $U \preccurlyeq W$ if for arbitrary $u \in U \setminus W$ there exists a $w \in W$ such that $u \prec w$.

Remark. Observe that it holds $U \preccurlyeq U$ for arbitrary maximum cardinality independent set $U \in \mathcal{M}$ since $U \setminus U = \emptyset$ and therefore it fulfils the above definition.

Remark. Relation $U \preccurlyeq W$ on $\mathcal{M}$ can be equivalently defined as follows:

$$U \preccurlyeq W \text{ if for every } u \in U \setminus W \text{ it holds } u \prec W.$$

Lemma. Let $G = (V, A)$ be an acyclic transitive digraph, let $U, W \in \mathcal{M}$ are two maximum cardinality independent sets in G. Then $U \preccurlyeq W$ if and only if for every $w \in W \setminus U$ it holds $U \prec w$.

Proof.
Let $U \preccurlyeq W$. Let $w \in W \setminus U$ such that it does not hold $U \prec w$. Then it must hold $w \prec U$ – i.e. there exists a $x \in U$ such that $w \prec x$. It cannot be $x \in W$ otherwise W would contain two comparable elements x and w. We have $x \in U$, $x \notin W$ i.e. $x \in U \setminus W$. Since $U \preccurlyeq W$ there exists a $z \in W$ such that $x \prec z$.
For $w \in W \setminus U$ such that it does not hold $U \prec w$ we have found a $x \in U$ and $z \in W$ such that $w \prec x$ and $x \prec z$ what implies $w \prec z$ what is contradiction with assumption that W is an independent set.

Theorem. Let $G = (V, A)$ be an acyclic transitive digraph, let $U, W \in \mathcal{M}$ are two maximum cardinality independent sets in G. If $U \preccurlyeq W$ and $W \preccurlyeq U$ then $U = W$.

Proof.
Let U, W are two maximum cardinality independent sets in G, let $U \preccurlyeq W$ and $W \preccurlyeq U$ and let $U \neq W$. Then both sets $U \setminus W$ and $W \setminus U$ are nonempty sets since U and W have the same cardinality.
Let $x \in U \setminus W$. Since $U \preccurlyeq W$ there exists a $w \in W$ such that $x \prec w$. It holds $w \notin U$ otherwise U would contain two comparable elements x and w. $w \in W \setminus U$ and since $W \preccurlyeq U$ there exists $z \in U$ such that $w \prec z$. It holds $x \prec w$ and $w \prec z$ and from transitivity of relation $\prec$ it follows that $x \prec z$.
If there exists a $x \in U \setminus W$ (provided $U \preccurlyeq W$ and $W \preccurlyeq U$) then we have found a $z \in U$ with $w \prec z$ what is in contradiction with independence of U. Therefore $U \setminus W = \emptyset$ and $U = W$.

Theorem. Relation $\preccurlyeq$ is a transitive relation on the set $\mathcal{M}$ of all maximum cardinality independent sets in an acyclic transitive digraph G.

Proof.
Let U, W, Z are maximum independent sets in G, let $U \preccurlyeq W \preccurlyeq Z$. Let $u \in U \setminus Z$. We are looking for such $z \in Z$ for which it holds $u \prec z$.
Case $u \in W$.
Then $u \in W \setminus Z$ and since $W \preccurlyeq Z$ there exists a $z \in Z$ with $u \prec z$.
Case $u \notin W$.
Then $u \in U \setminus W$ and since $U \preccurlyeq W$ there exists a $w \in W$ with $u \prec w$. If also $w \in Z$ it suffices to set $z = w$. If $w \notin Z$ then $w \in W \setminus Z$ and since $W \preccurlyeq Z$ exists a $z \in Z$ with $w \prec z$. We have $u \prec w$ and $w \prec z$ and from transitivity $u \prec z$.

Definition. Let $G = (V, A)$ be an acyclic transitive digraph, let $U, W \in \mathcal{M}$ are two maximum cardinality independent sets in G.
The meet $U \wedge W$ of U, W is defined as

$$U \wedge W = (U \cap V) \cup \{\, x \mid x \in U \cup W \text{ such that there exists } y \in U \cup W \text{ with } x \prec y \}$$

The join $U \vee W$ of U, W is defined as

$$U \vee W = (U \cap V) \cup \{\, x \mid x \in U \cup W \text{ such that there exists } y \in U \cup W \text{ with } y \prec x \}$$

Remark. It holds: If $U, W \in \mathcal{M}$, $U \preccurlyeq M$ then $U \wedge W = U$, $U \vee W = W$.

Theorem. Let $\mathcal{M}$ be the set of all maximum cardinality independent sets in an acyclic transitive digraph G, let $U, W \in \mathcal{M}$. Then $U \wedge W \in \mathcal{M}$ $\quad U \vee W \in \mathcal{M}$, i.e. $U \wedge W$ and $U \vee W$ are also maximum cardinality independent sets in digraph G.

Proof.
We will prove that $U \wedge W$ is an independent set.
Let $i, j \in U \wedge W$. If both $i, j \in U$ or $i, j \in W$ then i, j are not comparable.
Suppose $i \in U \setminus W$ and $j \in W \setminus U$, $i, j \in U \wedge W$ and $i \prec j$. Then there exists $w \in W$, $u \in U$ such that $i \prec w$, $j \prec u$. If $i \prec j$ then it follows from $i \prec j, j \prec u$ that $i \prec u$ what is I n contradiction with independence of U. It can be proved similarly that $U \vee W$ is an independent set.
Let $x \in (U \cup W) \setminus (U \cap V)$ then there exists a $y \in U \cup W$ such that $x \prec y$ or $y \prec x$. Therefore sets $\{ x \mid x \in U \cup W \ \exists\, y \in U \cup W \ \ x \prec y \}$ and $\{ x \mid x \in U \cup W \ \exists\, y \in U \cup W \ y \prec x \}$ have together $|(U \cup W) \setminus (U \cap V)|$ elements therefore

$$|U \wedge W| + |U \vee W| = 2|(U \cap W)| + |(U \cup W) \setminus (U \cap W)| = |U| + |W|$$

Both independent sets have together double elements as one maximum independent set and since no one can have more both have the same number of vertices.

Corollary. The set $\mathcal{M}$ of all maximum cardinality independent sets in an acyclic transitive digraph G with join operation $U \wedge W$ and meet operation $U \vee W$ is a lattice.

4 CONCLUSIONS

This article showed how a maximum cardinality independent set can be used to reduce the number of variables in integer linear programming models for solving advanced vehicle scheduling problems. Structure of maximum cardinality sets was studied. It was shown that operations joint and meet can be defined on the set of maximum cardinality independent sets – with mentioned operation it becomes a lattice. Every maximum independent set divides the set of trips into two subsets, such that no directed couple containing trips from both subsets can be used in optimum solution with minimum number of trips, what can reduce the set of decision variables in corresponding integer linear program.

Acknowledgements: This work was supported by Slovak Research and Development Agency – grant APVV-14-0658 ”Optimization of Municipal and Regional Public Transport” and also by Scientific Grant Agency of Ministry of Schools and Sport SR and Slovak Academy of Science, grant VEGA 1/0582/16 „Economic optimization of processes on networks“.

References

[1] Prata, B. A. (2016). A multiobjective metaheuristic approach for the integrated vehicle and crew scheduling. *Journal of Transport Literature*, [online] 10(2). Available at: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2238-10312016000200010#c01 [Accessed 30 Apr. 2018].

[2] Czimmermann, P. and Peško, Š. (2004). Autobusové rozvrhy s dvoma typmi vozidiel (Bus Schedules with Two Types of Vehicles). *3-rd International Conference APLIMAT 2004*, pp. 321-326.

[3] Černá, A. (2001). Optimization of Regional Bus Transport. In: *Proc. of International Conference Transportation Science*. Praha: Fakulta Dopravní ČVUT, pp. 70-75.
[4] Engeltahller, D. (1987). *Uplatnenie systémov človek – stroj pri tvorbe obehového rozvrhu v osobnej doprave.* Žilina: VŠDS, PhD. Thesis.
[5] Palúch, S. (2001). Graph Theory Approach to Bus Scheduling Problem. *Studies of the faculty of management science and informatics*. 9(3), pp. 53-57.
[6] Palúch, S. (2003). Bus Scheduling as a Graph Coloring Problem. *Communications*, 5(4), pp. 16-20.
[7] Palúch, S. and Majer, T. (2016). Vehicle scheduling with roundabout routes to depot. In: *Mathematical methods in economics - MME 2016*, Liberec: Technical university of Liberec, pp. 623-628.
[8] Palúch, S., Peško, Š and Majer, T. (2016). An exact solution of the minimum fleet size problem with flexible bus trips. In: *Quantitative methods in economics (multiple criteria decision making XVIII)*. [online] Bratislava: Letra Interactive, pp. 278-282. Available at: http://www.fhi.sk/files/katedry/kove/ssov/proceedings/Zbornik2016.pdf [Accessed 30 Apr. 2018].
[9] Peško, Š., Palúch, S. and Majer, T. (2016). A group matching model for a vehicle scheduling problem. In: *Quantitative methods in economics (multiple criteria decision making XVIII)*. [online] Bratislava: Letra Interactive, pp. 298-302. Available at: http://www.fhi.sk/files/katedry/kove/ssov/proceedings/Zbornik2016.pdf [Accessed 30 Apr. 2018].
[10] Peško, Š. (2001). The Minimum Fleet Size Problem for Flexible Bus Scheduling. *Studies of the Faculty of Management Science and Informatics*, 9(2), pp. 59-65.
[11] Vašek, K. (1985). Problém dekompozície pri optimalizácii obehových rozvrhov. *Zborník prác VÚD 1985*, 45(1), pp. 148 – 154.

Author's address

|RNDr. Stanislav Palúch, CSc.
University of Žilina, Faculty of Managements Science and Informatics,
Department of Mathematical Methods and Operations Research
Univerzitná 8215/1, 010 26 Žilina
Slovak Republic
email: stanislav.paluch@fri.uniza.sk

Ing. Tomáš Majer, PhD.
University of Žilina, Faculty of Managements Science and Informatics,
Univerzitná 8215/1, 010 26 Žilina
Tel: +421-41-513 4251
email: tomas.majer@fri.uniza.sk

|RNDr. Štefan Peško, CSc.
University of Žilina, Faculty of Managements Science and Informatics,
Department of Mathematical Methods and Operations Research
Univerzitná 8215/1, 010 26 Žilina
Slovak Republic
email: stefan.pesko@fri.uniza.sk

PORTFOLIO SELECTION MODEL BASED ON CVaR PERFORMANCE MEASURE

Juraj Pekár, University of Economics in Bratislava
Ivan Brezina, University of Economics in Bratislava
Ivan Brezina Jr., Pan-European University

Abstract
All investors are looking for investment instruments to guarantee maximal return and maximal liquidity by minimal risk; each investor must indeed understand the specific approach of risk, return and liquidity. The quantification the risk of investment is one of the most important task to be done while investing. The measure of investment represents the results quantification reached by chosen investors strategy; means the statistical summarization of reached returns. It is very important to consider the estimated level of risk. Decision-making plays the important role and is determined by the manager and his risk-based approach. In this paper, the performance measures based on Value at Risk are presented, together with the theory of the portfolio selection model based on CVaR. At first, the conditions of performance measure coherence are provided. Next, the performance measures based on Value at Risk and last, but not least, portfolio selection model based on CVaR performance.

Keywords: *portfolio performance measure, coherent performance measure, portfolio selection model*

JEL Classification: *C61, C02*
AMS Classification: *13P25, 74P05, 74P10*

1 INTRODUCTION

The investment risk can be described as the volatility of return measure. It represents the danger which may influence the measure of expected investor returns. By the analysis of risk and volatility we are able to estimate the interval representing the probability of portfolio assets increase.

Performance measure of investment (Pekár et al. 2016) is the quantification of achieved results by chosen strategy of investor. It is the statistical summarization of achieved returns measure, of potential risk estimate and of the investors approach to act with risk. Hereby we would like to stress the theoretical optimization model of portfolio selection based on CVaR performance measure. The one-criteria decision making is considered here, as the performance measure is based on the portfolio return and risk values.

2 COHERENT PERFORMANCE MEASURES

The topic of risk and expected returns measure is closely related to portfolio performance measure. There are some conditions to be considered in each acceptable portfolio performance measure. Rachev et al. (Rachev et al. 2008) define one category of performance measures. It is known as the coherent performance measure.

Let the function $G: X \to R, X \in A$ (where A is the set of all acceptable portfolios) and the function $G(X)$ we can define: $G(X) = \frac{V(X)}{\rho(X)}$, where $V(X)$ represents expected return measure and $\rho(X)$ risk measure, which are define for the random variable X. This function we call coherent performance measure, if:

1. The Function of expected return $V(X)$ satisfied the superadditivity condition $V(X_1 + X_2) \geq V(X_1) + V(X_2)$ for $\forall X_1, X_2 \in A$ and the function of risk measure $\rho(X)$ satisfied the superadditivity condition $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$ for $\forall X_1, X_2 \in A$.
2. In the case that risk measure and expected return measure are positive $X_1 \leq X_2$ and $G(X_1) \leq G(X_2)$ for $\forall X_1, X_2 \in A$. In the case that risk measure and expected return measure are negative $G(X_1) \geq G(X_2)$ for $\forall X_1, X_2 \in A$.

The alternative definition of return measure coherence function $G(X) = \frac{V(X)}{\rho(X)}$ is the coherence of expected return measure $V(X)$ and the coherence of risk measure $\rho(X)$.

We talk about the coherent expected return measure, if the following axioms are fulfilled. We will build on axiom according to Rachev et al. (Rachev et al. 2008), that the expected return measure must be:

1. superadditive: $V(X_1 + X_2) \geq V(X_1) + V(X_2), \forall X_1, X_2 \in A$
2. positive homogeneous: $V(cX) = cV(X), c \geq 0, \forall X \in A$
3. monotone: if $X_1 \geq X_2$, then $V(X_1) \geq V(X_2), \forall X_1, X_2 \in A$
4. translational invariant: $V(X + c) = V(X) + c, \forall c \in R, \forall X \in A$

We talk about the coherent risk measure, if the following axioms are fulfilled. We will build on axiom according to Artzner et al. (Artzner et al. 1999), that the risk measure must be:

1. superadditive: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2), \forall X_1, X_2 \in A$
2. positive homogeneous: $\rho(cX) = c\rho(X), c \geq 0, \forall X \in A$
3. monotone: if $X_1 \leq X_2$, then $\rho(X_1) \geq \rho(X_2), \forall X_1, X_2 \in A$
4. translational invariant: $\rho(X + c) = \rho(X) - c, \forall c, \forall X \in A$

3 PERFORMANCE MEASURE BASED ON VALUE AT RISK

In the literature we can find performance measure based on Sharpe ratio (see Sharpe 1994). According to Sharpe we use VaR and CVaR as the risk measure instead of standard deviation.

We assume the existence of the continuous random variable X and its probability density function $f(x)$, $x \in X$. Then the expected value of return $E(X)$ will be:

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx \tag{1}$$

Dowd (Dowd 2000) used VaR as the risk measure. According to it we can define Value at Risk (VaR):

$$P\left(-X < VaR\right) = 1 - \alpha \tag{2}$$

where X – random variable represented the value of profit and loss,

α – significance level.

Then we can define the performance of portfolio as the division of the expected value of return $E(X)$ minus riskless measure r_f and VaR value.

$$S_{VaR}\left(X\right) = \frac{E\left(X\right) - r_f}{VaR_\alpha(X)} = \frac{\int_{-\infty}^{+\infty} x \cdot f(x)dx - r_f}{P\left(-X < VaR\right) = 1 - \alpha} \tag{3}$$

where $E(X)$ is the expected value of return,
r_f is the riskless asset return,
$VaR_\alpha(X)$ is the value of VaR at α significance level.

We can use the CVaR risk measure too. CvaR is conditional expected value, which we get as the weighted average of values X. These values are less than VaR. If $P\left(-X < VaR\right) = 1 - \alpha$ and if X have the probability density function $f(x)$, then:

$$CVaR_\alpha\left(X\right) = -\frac{1}{\alpha} \int_{-\infty}^{VaR_\alpha} x \cdot f\left(x\right) dx \tag{4}$$

Then the portfolio performance we can define as the division of the expected value of return $E(X)$ minus riskless measure r_f and CVaR value.

$$S_{CVaR}\left(X\right) = \frac{E\left(X\right) - r_f}{CVaR_\alpha(X)} = \frac{\int_{-\infty}^{+\infty} x \cdot f(x)dx - r_f}{-1/\alpha \int_{-\infty}^{VaR_\alpha} x \cdot f\left(x\right) dx} \tag{5}$$

where $E(X)$ is the expected value of return,
r_f is the riskless asset return,
$CVaR_\alpha(X)$ is the value of CVaR at α significance level.

If we assume the existence of discrete random variable, then we can calculate the performance measure. Let the vector **r** represents discrete random variable X, which express portfolio return measure in the concrete period. The expected return of this investment will be:

$$E(\mathbf{r}) = \sum_{t=1}^{T} p_t r_t \, . \tag{6}$$

CVaR risk measure we define:

$$\begin{aligned} CVaR_\alpha\left(X\right) &= VaR_\alpha - \frac{1}{\alpha} E\left[\left|\mathbf{r} + VaR_\alpha\right|_-\right] \\ CVaR_\alpha\left(X\right) &= VaR_\alpha + \frac{1}{\alpha} \sum_{t=1}^{T} p_t \max(-\left(VaR_\alpha + r_t\right), 0) \end{aligned} \tag{7}$$

Then the portfolio performance we can define as the division of the expected value of return $E(X)$ minus riskless measure r_f and CVaR value for the discrete random variable.

$$S_{CVaR}(\mathbf{r}) = \frac{E(\mathbf{r}) - r_f}{CVaR_\alpha(\mathbf{r})} = \frac{\sum_{t=1}^{T} p_t r_t - r_f}{VaR_\alpha - 1/\alpha \sum_{t=1}^{T} p_t \max(-(VaR_\alpha + r_t), 0)} \tag{8}$$

4 PORTFOLIO SELECTION MODEL BASED ON CVaR PERFORMANCE MEASURE

In previous sections we analyse the performance measure, which is define as the division of additional return and CVaR. In this section we create the portfolio selection model based on this performance measure.

Let r_{jt} be the t-th component ($t = 1, 2, ..., T$) of discrete random variable, which represents vector of returns $\mathbf{r}_j$ for $j=1,2, ..., n$. We assume, that we can explain the expected value of random variable as the weight average from this data. We assume the same probability for states, which may occur. Let w_j be the proportion of j-th asset in overall investment. The CVaR function for discrete random variable is:

$$CVaR_\alpha(\mathbf{w}) = VaR_\alpha + \frac{1}{\alpha} \sum_{t=1}^{T} p_t \max\left(-\left(VaR_\alpha + \sum_{j=1}^{n} r_{jt} w_j \right), 0 \right) \tag{9}$$

Than the portfolio selection objective function based on CVaR risk measure is:

$$\min \left\{ VaR_\alpha + \frac{1}{\alpha} \sum_{t=1}^{T} p_t \max\left(-\left(VaR_\alpha + \sum_{j=1}^{n} r_{jt} w_j \right), 0 \right) \right\} \tag{10}$$

For linear transformation of written function we need to replace $\max(-VaR_\alpha - \sum_{j=1}^{n} r_{jt} w_j, 0)$.

For this transformation we use the variables z_t, where $z_t \geq 0$ ($t = 1, 2,...T$), which are the difference between VaR a portfolio return in state t, if the return is less than VaR, in other case they are equal to zero.

The goal of the optimisation is to minimize the investment risk to such performance measure, which is defined as the division of additional return and CVaR. The problem is then:

$$\max \left\{ \frac{\sum_{j=1}^{n} E_j w_j - r_f}{\left(VaR_\alpha + \frac{1}{\alpha} \sum_{t=1}^{T} p_t z_t \right)} \right\}$$

$$z_t + \sum_{j=1}^{n} r_{jt} w_j + VaR_\alpha \geq 0, t = 1, 2, ...T \tag{11}$$

$$\sum_{j=1}^{n} w_j = 1$$

$$w_1, w_2, ...w_n \geq 0, z_1, z_2, ...z_T \geq 0$$

The denominator can't be equal to zero. It makes the problem unsolvable. In the denominator are the return values and the coherent risk measure CVaR. It means, that we can apply the property of the translational invariance $\rho(X+c)=\rho(X)-c, \forall c, \forall X \in A$, which can be used by input data transformation.
Up to now we assume the existence of the j-th asset (t = 1,2,...,T) return r_{jt}. We realize the transformation of returns. Let $h = \max\{r_{jt}, j=1,2,...n, t=1,2,...T\}$, then $r^*_{jt} = r_{jt} - h$, j = 1,2,...,n, t =1,2,...,T. Because the property of the translational invariance is valid, the decrease of all returns leads to the increase of the risk. The risk increases by the value h for all assets.
This transformation ensures, that the VaR value will be non-negative, because all return values are non-positive. The formula $\frac{1}{\alpha}\sum_{t=1}^{T} p_t z_t$ represents the sum of non-negative components. The value of denominator $VaR_\alpha + \frac{1}{\alpha}\sum_{t=1}^{T} p_t z_t$ is then positive. It means, that we can find the solution of the problem:

$$
\begin{aligned}
&\max\left\{\frac{\sum_{j=1}^{n} E_j w_j - r_f}{\left(VaR_\alpha + \frac{1}{\alpha}\sum_{t=1}^{T} p_t z_t\right)}\right\} \\
&z_t + \sum_{j=1}^{n} r^*_{jt} w_j + VaR_\alpha \geq 0, t=1,2,...T \\
&\sum_{j=1}^{n} w_j = 1 \\
&w_1, w_2, ...w_n \geq 0, z_1, z_2, ...z_T, VaR_\alpha \geq 0
\end{aligned}
\tag{12}
$$

The result is the solution of this problem, which consist of asset proportions of invested capital value in optimal portfolio.

5 CONCLUSION

Topic in this paper is widely discussed by many authors. They composed several types of performance measure able to meet the requirements of practice to simplify the calculation and the economic implementation at the same time. On the other hand, the role of presented measures should meet the criteria important to facilitate the discussed. The main goal is nowadays to find a suitable, generally accepted, performance measure.

In this paper, presented approach can be used in the portfolio selection process. This portfolio selection could be done by presented portfolio selection model. As first, it is necessary to analyse the input data of the implementation model. This analysis should be performed by statistical methods and the portfolio selection depends on this input data analyse. The whole process of portfolio selection is very complicated and this paper represents one option (out of many) which may affect the final investor decision.

Acknowledgements

This paper is supported by the Grant Agency of Slovak Republic – VEGA, grant no. No. 1/0351/17 Application of selected models of game theory to solve some economic problems of Slovakia.

References

[1] Artzner, P., Delbaen, F., Eber, J. M. and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance,* 9(3), pp. 203-228.
[2] Dowd, K. (2000). Adjusting for Risk: An improved Sharpe Ratio. *International Review of Economics and Finance*, 9(3), pp. 209-222.
[3] Pekár, J., Čičková, Z. and Brezina, I. (2016). Portfolio performance measurement using differential evolution. *Central European Journal of Operations Research*, 24(2), pp. 421-433.
[4] Rachev, S., Ortobelli S., Stoyanov, S., Fabozzi , F. J. and Biglova, A. (2008). Desirable Properties of an Ideal Risk Measure in Portfolio Theory, *International Journal of Theoretical and Applied Finance*, 11(1), pp. 19-54.
[5] Sharpe, W.F. (1994). The Sharpe Ratio. *Journal of Portfolio Management*, 21(1), pp. 49-58.

Author's address

prof. Mgr. Juraj Pekár, PhD.
University of Economics in Bratislava, Faculty of business informatics, Department of Operations research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
email: juraj.pekar@euba.sk

prof. Ing. Ivan Brezina, CSc.
University of Economics in Bratislava, Faculty of business informatics, Department of Operations research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
email: ivan.brezina@euba.sk

Ing. Ivan Brezina Jr., PhD.
Pan-European university, Faculty of Economics, Department of Economics
Tematínska 10, 851 05 Bratislava
Slovak Republic
email: brezina.ivan@yahoo.com

OPTIMIZATION OF INTERNATIONAL TRUCK TRANSPORT

Jan Pelikán, University of Economics Prague

Abstract
Due to globalization, the international trade requires transporting of a lot of goods for a long distance. Therefore the truck transport is largely used. Due to higher operating costs, it is necessary to utilize the maximum capacity and to optimize routes of these vehicles. Logistic companies offer two types of organization of truck transport. First of them (marked as FTL transport) provides the full truck capacity to a single transporter (trading factory, manufacturer). Cost is derivative of route length. The second option is the utilization of cross-dock. Transporter transports goods to the warehouse (cross-dock), then the logistic company sorts goods according to the destination and places it in a truck, which is going to appropriate destination. In this case, the cost is derivative of the quantities of goods and destination. This article is a case study, where the transport of predetermined goods at a certain time interval is optimized. Both types of organization of truck transport are usable. A original mathematical model based on real data is designed and a heuristic method proposed.

Keywords: Truck loading problem, Integer programming, Heuristic methods

JEL Classification: C44
AMS Classification: 90C15

1 INTRODUCTION

In international trade, truck transport realizes considerable part of transporting goods. This type of transport can transport a lot of goods and is flexible to changes in needs and places of transport. When we utilize the maximum capacity of the truck, cost per unit is also favorable. If the volume of transport almost matches truck capacity, it is advantageous to hire a truck for this transport (FTL full truck load). Otherwise, when the volume of transport required does not allow full usage of the truck´s capacity, it is more advantageous to share a truck with requirements of other transporters (LTL less than truck load), which is enabled by cross-dock. Transporter transports goods to the warehouse (cross-dock), followed by sorting goods according to the destination and transporting goods along with consignments from other transporters to the destination. When using LTL transport, the cost is derivative on the volume of transport requirement. When transport is FTL type, then logistic company calculates costs per whole truck regardless of whether it is full or not.

This article is a case study of a company producing the computers components, which supplies goods to foreign purchasers. The destination is either customs warehouse or a warehouse of a retail chain abroad (east contries: Poland, baltic states, Russia...). For the transport of goods is used an external logistic company, which enables both types of transport (direct or cross-dock). The problem relates to the transport of deliveries, consignments, parcels at a certain time interval. For each consignment is given volume and time, when it is ready for transport and the destination (country, place). It is needed to ship the consignment until determined time from the time when it is ready for transport.

Some of them will lead up directly to the purchaser (FTL transport), others will carry consignments to the cross-dock. Concurrently needs to be determined, where to load which consignments. The goal is to minimize total shipping costs, consisting of FTL transport cost,

transport cost to the cross-dock and transport cost from the cross-dock to the purchaser (LTL transport cost).
In literature, this problem is called truck loading problem. In the scientific literature, there are several papers about the truck loading problem. Their contents are different from the problems solved in this article. Many of these sources solve the problem of cyclic drives of the trucks loaded with different types of products, e. g. the oil products, so that the time interval between two deliveries is maximal [1] [3]. Another author [2] proposed a way of periodical loading of carriers with more compartments on the load area where the compartments have the various capacity and different types of goods are considered. Yuceer [4] proved this problem to be NP-hard, proposed a mathematical model, analyzed its properties and proposed a solving algorithm.

2 TRUCK TRANSPORT OPTIMIZATION MODEL

Given the time interval <0, *T*> along with the number of parcels *n,* that should be transported from manufacturer to customers. We assume a discrete time at which we optimize truck loading, i.e., at time 0, 1, 2, ..., *T*, so time unit is 1 (one shift, day, working day).
For each parcel, the time, when it is ready to be transported, is known. For parcels, it is possible to postpone their transport over a specific number of days (in our model we assume the delay of two days is maximal).
We know the volume and the weight of a parcel, but weight of parcels are important for the model (goods are computers components). The company providing the transport offers two options: direct delivery to a customer or transport via the cross-dock. The transport is realized by trucks with given capacity.
The aim is to determine the way of transport with the minimal transportation costs. It means to set the time of departing for each truck, also to decide what to load on each truck and whether it is a direct or cross-dock type of truck. The number of available trucks is not limited, but for purposes of the formulation of the model, we will assume a certain maximal number of trucks per day. This amount matches the need of the case study. The capacity of each truck is 26 pallets (europalets).
The mathematical model contains following parameters and variables.

Parameters of the model:

$\overline{t_i}$ the time when the parcel *i* is ready for transport,
q_i the volume of the parcel *i*, measured by the number of pallets,
$d_i^{\,s}$ the binary parameter equals to 1 if the *i*-th parcel should be delivered to the destination *s*,
c_s is the transportation costs of one truck going direct from manufacture to destination *s,*
c_0 is the transportation costs of one truck going from manufacture to cross-dock*,*
$\overline{c_s}$ is the transportation costs of one pallet going from cross-dock to destination *s.*

Variables of the model:

x_{itk} a binary variable, equals to 1 if the parcel *i* is sent off at the time *t* in a direct truck *k* otherwise, it equals to 0
x^0_{itk} a binary variable, equals to 1 if the parcel *i* is sent off at the time *t* in a truck *k* to the cross-dock otherwise, it equals to 0
y^s_{tk} a binary variable, equals to 1 if the *k*-th truck is sent in the time *t* directly to the destination *s*,
y^0_{tk} a binary variable equals to 1 if the *k*-th truck is sent in the time *t* to the cross-dock.

The mathematical model:

$$\sum_{t,k,s} c_s y_{tk}^s + c_0 \sum_{t,k} y_{tk}^0 + \sum_{i,t,k,s} \overline{c_s}\, x_{itk}^0\, q_i \;\rightarrow\; \min \tag{1}$$

$$\sum_i x_{itk} q_{i\leq} \leq 26 y_{tk}^s, \quad \forall t,k \tag{2}$$

$$\sum_i x_{itk}^0 q_{i\leq} \leq 26 y_{tk}^0, \quad \forall t,k \tag{3}$$

$$\sum_{tk} x_{itk} + \sum_{tk} x_{itk}^0 = 1, \;\; \forall i \tag{4}$$

$$x_{itk} = x_{itk}^0 = 0, \;\; \forall i,t,k, \quad \forall t > \bar{t}_i + 2, \;\; t < \bar{t}_i \tag{5}$$

$$x_{itk} \leq d_i^s y_k^s \;\; \forall i,t,k,s \tag{6}$$

$$y_{tk}^0 + \sum_s y_{tk}^s \leq 1 \; \forall t,k \tag{7}$$

$$x_{itk}, x_{itk}^0, \; y_{tk}^0, \; y_{tk}^s \text{ are binary} \quad \forall i,t,k,s. \tag{8}$$

Inequality (2) assures that if the truck k does not go to directly to customers in the time t, then it is not possible to use it for transport of the parcels. At the same, this inequality limits the amount of the load by the capacity of the truck. Similarly, (3) is valid for the trucks going the cross-dock. Thanks to (5), the parcel is not allowed to be loaded earlier than it is available, or later than it is feasible, which is more than two days later. Constraint (6) does not allow loading a parcel into a truck that will not arrive at the destination of the parcel. The truck k can go directly to the destination s (y^s_{tk}=1) or go to cross-dock (y^0_{tk}=1) in the time t, so the inequality (7) has to hold.
The object function (1) consists of three terms. The first term $\sum_{t,k,s} c_s y_{tk}^s$ means costs of all truck which go direct from producer to destination. Costs of truck going from producer to cross-dock is expressed by the second term $c_0 \sum_{t,k} y_{tk}^0$. Last term $\sum_{i,t,k,s} \overline{c_s}\, x_{itk}^0\, q_i$ in (1) contents costs of transport of all parcels transported from cross-dock to their destination and depends on quantity of parcels q_i.
The model is linear with binary variables, but it can contain a large number of binary variables in the real world instances in case of a big number of parcels and a long time horizon, even though the condition (5) fixes many of them to zero.

3 NUMERICAL EXPERIMENTS

Described model of operating the international truck transport is based on a case study and proposed models were verified on the data from the involved company. The time interval for this computation was 14 days, during which approximately 1000 parcels should be transported to 10 destinations. A mathematical model (1)-(7) contains 60 thousand binary variables and app. 100 thousand constraints. The model was proved on PC (InteCore2Quad, 2,83GHz, CPLEX 12.0). The computation was interrupted after 30 mins., a good feasible solution was obtained with gap 3,5%.

4 OPTIMIZATION ON AN INFINITE TIME INTERVAL

A mathematical model of truck transport optimization on a finite interval provides a solution, where at the end of the interval the task decides to dispatch all parcels regardless of efficiency thus regardless of the insufficient use of the truck, because the goal of this optimization is to transport all parcels from manufacturer to purchasers during this time interval, i.e. to the time

T. In reality, parcels which are ready to transport in the time *T* and *T-1* can delay, so it is possible to transport them in the time *T+1* or *T+2* when the capacity of the truck is used more efficiently. Because of that, it is possible to purpose following heuristic.

Let´s solve the task of optimization of truck transport gradually on overlapping time intervals <0, *T*>, <*T*/2, 3/2*T*>, <T, *2T*>,… using mathematical model (1)-(8). As a result of that, we obtain an information about dispatched trucks and their loading with consignments. The way to use these results is following.

Let´s start with solving the task on the interval <0, *T* >. The result will be used only for dispatching trucks on the interval <0, *T/2*>. We add parcels, which were not shipped during this interval to parcels which are ready to transport in the time interval <*T*/2, *3/2T*> and then solve the task on this interval. From the optimal solution obtained from the model (1)-(8), we use the information about shipped trucks and parcels from the interval <*T/2, T*>. We move parcels, which are ready to transport in the interval <*T*, *3/2T*> along with delayed parcels (ready to transport in the time *T* and *T-1* and unshipped until the time *T*) to the optimization of truck transport on the interval <*T*, *2T*>. We use the model for parcels from this interval, which are ready for transport to which we add delayed parcels, which are ready to transport in the time *T* and *T*-1.

This procedure can be repeated when there is new information about parcels, which are ready to transport. It is achieved a reduction of inefficiency in solving the optimization of truck transport occurring at the end of the interval.

5 THE TASK OF TRUCK TRANSPORT WITHOUT INFORMATION ABOUT FUTURE SHIPMENTS

In practical applications, it is possible that we have do not have information about parcels that we will need to transport in future, we only have information about parcels that we currently need to transport, i.e. in the time t, information about parcels in the time t+1,t+2,.. is missing. We assume that parcels which are ready to transport in the time t can delay maximally for a certain time. For example with a maximum delay of 2 days, parcels which are ready to transport in the time t can be sent off in the time t or t+1 or eventually t+2. The problem occurs when there are in the time t parcels ready to transport and parcels which were ready to transport in the time t-1 and t-2, but they were a delayed due to the inefficiently loaded truck.

We need to decide wheater we dispatch the truck in the time t and with which parcels loaded onto it and eventually which type of transport to use (FTL type or transport via cross-dock). The capacity of the truck is still a restrictive factor. The essential matter is then what percentage of the full capacity of the truck is economical sufficient for sending the truck off, remembering that it is not possible to wait for too long for the parcels until their sending off. The presented mathematical model and historical data can be used to establish the load limit. Question is wheater to ship the truck even if it is not full and how full (percentage) it needs to be to send it off. This percentage is necessary to dispatch the truck, and it is a subject of simulation on historical data, or determined as a result of the optimization model with the use of historical data.

6 CONCLUSION

This article proposes original model for optimization of the truck transport of real case study. The model is based on a set of parcels given in advance, which should be transported. The optimization solves a decision whether a parcel ready for the delivery should be loaded onto a truck and sent off or, if the truck is not fully loaded, postpone the delivery to the time when the truck is loaded more efficiently with new incoming parcels. The model contains a significant number of variables depending mostly on the number of parcels in the given time interval. Therefore, the applicability of the model will be higher with a smaller number of parcels.

Acknowledgments: Supported by the grant No. 16-00408S of the Czech Grant Agency.

References

[1] Bukchin Y., and Sarin S.C.(2004): A cyclic policy for the loading of multiple products on a vehicle with different compartment sizes. In: *IIE Transactions* 30/7, pp.641-653.
[2] Brown G.G. and Graves G.W. (1981): Real-time dispatching of petroleum tank trucks. In: *Management Science* 27, pp. 19-32.
[3] Ronen D. (1994): Dispatching petroleum products. In: *Operations Research* 43, pp. 379-387.
[4] Yuceer, U. (1997): A multi-product loading problem: A model a solution method. In: *European Journal of Operational Research*, 101, pp. 519-631.

Author's address
Prof. RNDr. Jan Pelikán, CSc.
University of Economics Prague, Faculty of informatics and statistics, Department of Econometrics
W. Churchill sq. 4, Prague 3
Czech Republic
email: pelikan@vse.cz

NEW APPROACH TO UNIFORM WORKLOAD DISTRIBUTION PROBLEMS

Štefan Peško, Tomáš Majer,
University of Žilina, Faculty of Managements Science and Informatics

Abstract
In this paper, we study problems occurring in uniform scheduling of workload distribution. Given a nonnegative real matrix with daily loads of vehicles as columns, we need to minimize an irregularity measure of vector of row sums (workloads) by permuting matrix columns. For two columns matrix it is known that the minimization problem for minimum difference of row sums is polynomial solvable. For general matrices, this minimization problem is NP-hard. We study new MILP and MIQP models for small numbers of different elements of matrix. We show that this approach effectively solves the instances of medium size. This approach can be used as a heuristic method also for instances with big numbers of different elements via assignment this to small number of classes, which are represented by average value of elements in class. Exact and heuristic solutions are discussed. Computational experiments with Gurobi solver are presented.

Keywords: *Uniform scheduling, workload distribution, MILP and MIQP, Gurobi*

JEL Classification: C02, C61, C65
AMS Classification: 90B35, 11C99, 05D40

1 INTRODUCTION

We deal with solving an NP-hard problem that occurs in fair scheduling of workload distribution. It means that we need to split total workload to given number of approximately equal parts. The level of fairness is expressed by irregularity measure. The lower measure implies the higher regularity. The goal is to minimize the total cost of weighted parts of this fair schedule.

First formulations under the name Matrix Permutation Problem (MPP) can be found in (Černý, 1985, Tegze and Vlach, 1986). They are motivated by Peško's mathematical interpretation of a real problem, when it was needed to assign weekly work schedule to a given number of drivers as evenly as possible. The goal is to find a minimally varied set of row-sums of columns permuted matrix. The approaches based on graph theory solving graph version of the MPP are studied in (Czimmermann, 2013, Czimmermann and Peško, 2007). In (Peško and Černý, 2006) is present different real situations where making managerial decisions can be influenced by the methodology of fair assignment. For reference, 32 MPPs are defined and examined in work (Dell'Olmoa, 2005). They are characterized by particular measure of set uniformity to be optimized. 21 of the studied problems can be solved by linear time algorithms, 7 require more complex algorithms but can still be solved in polynomial time, and 3 are proved to be NP-hard. The possibility of based on aggregate two-column sub-problems for a fair scheduling and workload distribution can be found in paper (Peško and Hajtmánek, 2016). The stochastic algorithm for uniform workload distribution problems is studied in paper (Peško and Kaukič, 2015) and in interval arithmetic in paper (Peško, 2006).

New interesting results can be found in paper (Jakobsons and Wang, 2016) where arrangement increasing function with matrix input over intra-column permutations was studied. Authors show that many classical optimization problems, including stochastic crew scheduling and assembly of reliable system, have objective functions with this structure. In the recent time the investigation of the problem continues, under the name of Matrix Arrangement Problem (MAP) (Jakobsons and Wang, 2016) or Block Rearranging Algorithm (BRA) in the papers (Boud, Jakobsons and Vanduffel, 2017, Puccetti and Ruschendorf, 2012).

2 MATHEMATICAL FORMULATION

Throughout this paper we will denote by $A = (a_{ij})$ a nonnegative real matrix with m rows and n columns. Also, we will denote by $I = \{1,2,\dots,m\}, J = \{1,2,\dots,n\}$ the sets of rows and columns indices. For each column $j \in J$ of the matrix $\boldsymbol{A}$ we will use notation π_j for the permutation of rows in that column. Now, let as denote by $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ the vector of permutations of all columns of A and by A^π the permuted matrix itself. The element in row i and column j of matrix A^π is thus $a_{\pi_{j(i)},j}$ for $i \in I, j \in J$. We will use notation S^π for the vector of row sums of permuted matrix, so

$$S^\pi = (s_1^\pi, s_2^\pi, \dots, s_m^\pi), \text{ where } s_i^\pi = \sum_{j \in J} a_{\pi_{j(i)},j} \quad \text{for } i \in I.$$

Let us have some irregularity measure f (see (Boudt, Jakobsons and Vanduffel 2017), (Černý, 1985), (Peško and Černý, 2006), (Tegze and Vlach, 1986)), i.e. the nonnegative real function defined on set of nonnegative vectors in R^m. The following optimization problem we will call **uniform workload distribution problem** (**UWDP**):

$$f(S^{\pi *}) = \min \{ f(S^\pi) : \pi \in \prod_{mn} \},$$

where $\prod_{mn}$ is the set of all $n-$tuples of permutations of the row indices I.

In this paper we will consider two basic irregularity measures defined on the m dimensional vector $(x_1, x_2, \dots, x_m)$

$$f_{dif}(x_1, x_2, \dots, x_m) = max_{i \in I} x_i - min_{i \in I} x_i, \tag{1}$$

and

$$f_{var}(x_1, x_2, \dots, x_m) = \sum_{i \in I} (x_i - \bar{x})^2, \ \bar{x} = \frac{1}{m} \sum_{i \in I} x_i. \tag{2}$$

Now we begin with the MILP formulation of the **UWDP** for object function f_{dif} from (Peško and Kaukič, 2015).

3 MILP AND MIQP MODELS

We define the binary variables $x_{ijk}, i \in I, j \in J, k \in K$. Value of x_{ijk} is equal to 1 if $\pi_{j(i)} = k$ and is zero otherwise. The element a_{ij}^π of permutated matrix A^π can be expressed by the sum $a_{ij}^\pi = \sum_{k \in K} a_{kj} x_{ijk}$. We also introduce the real variables $z_i, i \in I$ for i-the row sum of permuted matrix and the variables z_L, z_U for (estimated) lower and upper bounds of sums S_L, S_U.

We minimized, in MILP model (Peško and Kaukič, 2015), the maximal difference (1) between profits of drivers in workload distribution. This can be formulated and solved as a very special version of three-dimensional MILP problem (**ML1**):

$$
\begin{array}{llr}
z_U - z_L \to minimum & & (3)\\
\text{s.t.} \quad \sum_{k\in I} x_{ijk} \;=\; 1 & \forall (i,j) \in I \times J, & (4)\\
\sum_{i\in I} x_{ijk} \;=\; 1 & \forall (k,j) \in I \times J, & (5)\\
\sum_{(j,k)\in J\times I} a_{kj}x_{ijk} \;=\; z_i & \forall i \in I, & (6)\\
x_{ijk} \;\in \{0,1\} & \forall (i,j,k) \in I \times J \times K, & (7)\\
z_L \le z_i \le z_U & \forall i \in I, & (8)\\
S_L \le z_L \le z_U \le S_U & & (9)
\end{array}
$$

The objective function (3) gives the value of f_{dif}. Constraints (4), (5) and (7) define assignment polytope for every column of matrix. Constraints (6), (8) and (9) give the lower and the upper bounds of row sums of optimal solution.

This model has m^2n bivalent variables (7), $m + 2$ nonnegative variables and $2(mn + m + 1)$ constraints. Our computation experiments with the **ML1** showed that model can be used only for matrices of small size. We also found the matrix of size 10×6 with uniformly distributed elements for which we have not found optimal solution roughly for 20 hour of computing via Gurobi MILP solver.

So, we begin to study new models for **small numbers of different elements** of matrix. Denote I, J the indices of rows and columns of given matrix $A = (a_{ij}), i \in I, j \in J$ but with small numbers of different elements from set $D = \{d_k : k \in K\}$, where $K = \{1,2,..,\kappa\}, \kappa \ll mn$ is an index set of different elements. Denote matrix $F = (f_{kj}), k \in K, j \in J$ the number of elements with value d_k in column $j \in J$ of matrix A. The optimal rearrangement matrix will be called $A^* = (a^*_{ij})$.

We will use binary variables $y_{ijk}, i \in I, j \in J, k \in K$ with value $y_{ijk} = 1$ iff $a^*_{ij} = d_k$. The variables z_L, z_U are minimum and maximum row sums of solution as in model M1. Now we can formulate following MILP problem (**ML2**):

$$
\begin{array}{llr}
z_U - z_L \to minimum & & (10)\\
\text{s.t.} \quad \sum_{k\in I} y_{ijk} \;=\; 1 & \forall (i,j) \in I \times J, & (11)\\
\sum_{i\in I} y_{ijk} \;=\; f_{kj} & \forall (j,k) \in J \times K, & (12)\\
\sum_{(j,k)\in J\times I} d_{kj}y_{ijk} \;\le\; z_U & \forall i \in I, & (13)\\
\sum_{(j,k)\in J\times I} d_{kj}y_{ijk} \;\ge\; z_L & \forall i \in I, & (14)\\
y_{ijk} \;\in \{0,1\} & \forall (i,j,k) \in I \times J \times K, & (15)\\
S_L \le z_L \le z_U \le S_U & & (16)
\end{array}
$$

The objective function (10) again gives the value of f_{dif}. Constraint (11) chose exactly one value from feasible values. Constraint (12) ensure required numbers of values in columns.

Constraints (13), (14) and (16) give the lower and the upper bounds of row sums of optimal solution.

The model **ML2** has $mn\kappa$ bivalent variables, $2(m+1)$ nonnegative variables and $n(m+\kappa)+5$ constraints.
Ideal value of row-sums is equal $\bar{a}=\frac{1}{m}\sum_{(i,j)\in I\times J}a_{ij}$ and so after using irregularity measures (2) of row sums f_{var} we get following MIQP model (**MQ**):

$$\sum_{i\in I}(z_i-\bar{a})^2\rightarrow minimum \tag{17}$$

$$\text{s.t.}\quad \sum_{k\in I}y_{ijk}=1 \qquad \forall(i,j)\in I\times J, \tag{18}$$

$$\sum_{i\in I}y_{ijk}=f_{kj} \qquad \forall(j,k)\in J\times K, \tag{19}$$

$$\sum_{(j,k)\in J\times I}d_{kj}y_{ijk}=z_i \qquad \forall i\in I, \tag{20}$$

$$y_{ijk}\in\{0,1\} \qquad \forall(i,j,k)\in I\times J\times K, \tag{21}$$

$$S_L\le z_i\le S_U \tag{22}$$

The objective function (17) gives the value of f_{var}. Constraints (18), (19) and condition (21) coincide with constraints (11), (12) and (15). Constraint (20) defines corresponding row sums.

The model **MQ** has $mn\kappa$ bivalent variables, m nonnegative variables and $n(m+\kappa)+m$ constraints. Next computation experiments compare and summarize the experience with the new models **ML2** and **MQ**.

4 COMPUTATIONAL EXPERIMENTS

Our experiments and simulations were conducted on HP XW600 Workstation (8-core Xenon 3GHz, Ram 16GB) with OS Linux (Debian jessie). We used Python interface to commercial mathematical programming solver Gurobi.

First experiment was tried for random (uniformly) generated integer–valued matrices with three sets of different values $K_t=\{1,\dots,t\}, t\in\{10,20,30\}$ and $D_t=\{0,\dots,t-1\}$ We achieve solution times given (in seconds) below in table 1. We can see that it can be quickly resolved where model **ML1** fails via model **ML2**. We found with surprise that the bottom and upper boundaries of row-sums S_L and S_U have almost no impact on the time of calculation.

Next experiment was done with matrices $B=\left(a_{ij}^2-\min\{a_{ij}^2:\mathrm{i}\in\mathrm{I}\}\right)$, where A are uniformly generated integer–valued matrices as in first case. Then the number of different elements in the matrix B is $|D_t|\ge t$ and the set possible values D_t positive standard deviation. The solution times are in table 2.

High quality solutions offer also model **MQ** with quadratic object function in some cases comparable to model **ML2**.

TABLE 1
Computation times of models for random matrices A

$m \times n$	t	ML1	ML2	MQ	$m \times n$	t	ML1	ML2	MQ
10×14	10	3	< 1	1	10×28	10	12	1	1
20×14	10	72	< 1	< 1	20×28	10	393	2	51
30×14	10	402	6	55	30×28	10	312	9	298
10×14	20	2.5	2	2	10×28	20	11	4	5
20×14	20	44	4	33	20×28	20	146	31	76
30×14	20	72	18	317	30×28	20	>1200	37	>1200
10×14	30	3	2	7	10×28	30	7	2	8
20×14	30	31	7	36	20×28	30	178	29	419
30×14	30	469	31	>1200	30×28	30	227	73	806

TABLE 2
Computation times of models for random matrices *B*

$m \times n$	t	$\|D_t\|$	$std(D_t)$	ML1	ML2	MQ
10×14	10	18	26.5	17	3	3
20×14	10	25	25.0	185	28	29
30×14	10	18	26.5	>1200	44	43
10×14	20	49	115.3	256	9	8
20×14	20	38	112.8	>1200	129	>1200
30×14	20	20	113.5	>1200	186	>1200
10×14	30	81	256.2	381	23	>1200
20×14	30	70	247	>1200	>1200	>1200

5 CONCLUSIONS

The **UWDP** problems are of great practical significance. We have seen that the traditional mathematical programming solver has difficulty with known model **ML1** but solve well with new models **ML2** and **MQ** for instances with small numbers of different elements of matrix. In the real-word applications integer valued matrices can be achieved by quantifying the profit into several integer levels. This approach can be used as a heuristic method also for instances with big numbers of different elements via assignment this to small number of classes, which are represented by average value of elements in class.

Acknowledgements: The research was supported by the research grants APVV-14-0658 Optimization of urban and regional public personal transport and VEGA 1/0582/16 Economic optimization of networks processes.

References

[1] Boudt, K., Jakobsons, E. and Vanduffel, S. (2017). Block Rearranging Elements within Matrix Columns to Minimize the Variability of the Row Sums. *SSRN,* [online]. Available at: http://dx.doi.org/10.2139/ssrn.2634669 [Accessed 8 Apr. 2018].

[2] Czimmermann, P. (2013). The Most Regular Perfect Matchings in Graph. *Studies of the University of Žilina, Mathematical Series*, 16(1), pp. 15-18.

[3] Czimmermann, P., Peško, Š. and Černý, J. (2006). Uniform workload distribution problem. *Communications,* 18(1A), pp. 55-58.

[4] Černý, J. (1985). How to minimise non-uniformity. *PMFA XXX* , pp. 145-150.

[5] Dell'Olmoa, P., Hansenb, P., Pallottinoc, S. and Storchid G. (2005). On Uniform k-Partition Problems, *Discrete Applied Mathematics*, 150(1–3), pp. 121-139.

[6] Jakobsons, E. and Wang, R. (2016). Negative Dependence in Matrix Arrangement Problems. *SSRN* [online], Available at: http://dx.doi.org/10.2139/ssrn.2756934 [Accessed 8 Apr. 2018].
[7] Peško, Š. and Černý, J. (2006). Uniform Splitting in Managerial Decision Making, *E+M, Economics and Management,* IX(4), pp. 67-71.
[8] Peško, Š. (2006). The Matrix Permutation Problem in Interval Arithmetic, *Journal of Information, Control and Management Systems*, 4, pp. 35-40.
[9] Peško, Š. and Kaukič, M. (2015). Stochastic algorithm for uniform workload distribution problems. In: *Informatics, 13th international scientific conference on informatics*, Poprad, pp. 206-210.
[10] Peško, Š. and Hajtmanek, R. (2016). The treatment of uncertainty in uniform workload distribution problems. In: *Mathematical methods in economics 2016, 34th international conference*, Liberec, pp. 669-674.
[11] Puccetti, G. and Ruschendorg, L. (2012) Computation and sharp bounds on expected value of a supermodular function of risk with given marginal. *Computations in Statistics – Simulation and Computation*, 44(3), pp.705-718.
[12] Tegze, M. and Vlach, M. (1986). On the matrix permutation problem, *Zeitschrift für Operations Research*, 30(3), pp. A155-A159.

Author's address
doc. RNDr. Štefan Peško, CSc.
University of Žilina, Faculty of Management and Informatics
Department of Mathematical Methods and Operations Research
Univerzitná 8215/1, 010 26 Žilina, Slovak Republic
email: stefan.pesko@fri.uniza.sk

Ing. Tomáš Majer, PhD.
University of Žilina, Faculty of Management and Informatics
Department of Mathematical Methods and Operations Research
Univerzitná 8215/1, 010 26 Žilina, Slovak Republic
email: tomas.majer@fri.uniza.sk

MULTIPLE CRITERIA ANALYSIS OF DIGITAL ECONOMY IN THE EUROPEAN UNION COUNTRIES

Michał Bernard Pietrzak, Nicolaus Copernicus University in Toruń,
Bartosz Ziemkiewicz, Nicolaus Copernicus University in Toruń

Abstract

The objective of the article is compare the European Union member states in regard to their achievements in creating conditions for digital economy development. The research took a multiple-criteria perspective as the phenomenon of digital economy was measured with application of ten diagnostic variables. The data for the research was obtained from Eurostat. The analysis was conducted for the years 2011 and 2017. In the research taxonomic measure of development obtained with the Zdzisław Hellwig method was used. Based on the method the ranking and grouping of the countries was possible. The grouping of the economies was done with natural breaks method. The research confirmed significant disparities not only between the old and new EU member states, but also significant differences in the level of development of the digital economy among EU-15. Additionally, the research also proved high stability of disparities between the European economies.

Keywords: *Hellwig's method, taxonomic measure of development, digital economy, European Union*

JEL Classification: C38, O33
AMS Classification: 91C15

1 INTRODUCTION

Development of digital economy has been considered as a source of important potential for all European countries, which was stressed in the Lisbon Strategy and Europe 2020 plan (Balcerzak, 2015). From macroeconomic point of view it was believed that the digital economy can be used by less developed EU members for obtaining quicker convergence process and must be utilised by the whole European economy for keeping and improving its international competitiveness. Form microeconomic perspective building effective digital economy creates important opportunities for quicker development of small and middle sized enterprises that diminish the negative influence of many traditional growth obstacles and enable to take advantage of economies of scale based on lower transaction costs (Ivanová, 2017). Therefore, the main objective of the article is to compare the EU member states in regard to their achievements in creating conditions for the digital economy development in the years 2011 and 2017. In the research taxonomic measure of development (TMD) and natural breaks method were applied. Eurostat was the source of the data. The article is a continuation of previous research of the authors regarding the determinants influencing development of the digital economy in Visegrad countries at regional level (Balcerzak and Pietrzak, 2017a; 2017b).

2 PROCEDURE FOR ASSESING TAXONOMIC MEASURE OF DEVELOPMENT

Most of the economic phenomena have complex and multiple-criteria character. This means that a proper description of variability of these phenomena requires consideration of many variables, where each individual variable describes specific aspect of a given phenomenon (Balcerzak, 2016a; Keshavarz Ghorabaee et al., 2016; Trąpczyńsk et al., 2016; Balcerzak et al., 2017; Daragahi, 2017; Cano et al., 2017; Mandić et al., 2017; Vavrek et al., 2017; Chrupała-Pniak et al. 2017; Meluzin et al., 2017; 2018; Pilelienė and Grigaliūnaitė, 2017; Shpolianska et al. 2017; Hilkevics and Hilkevica 2017; Monni et al., 2017). Therefore, these scientific problems can be described with application of multidimensional analysis tools. Among most common examples of economic problems related to the analysis of multidimensional phenomena one can find level of socio-economic development (Zavadskas et al., 2016; Bartkowiak-Bakun, 2017; Kuc, 2017a; 2017b; Cyrek, 2017), technological potential, competitiveness and innovativeness (Balcerzak, 2016b; 2016c; Kaynak et al., 2017; Zygmunt, 2017; Kruk and Waśniewska, 2017), level of entrepreneurship (Pietrzak et al., 2017, Rogalska, 2017), or quality of institutions (Balcerzak, 2009; Pilc, 2017; Żelazny and Pietrucha, 2017). Taxonomic measure of development (TMD) is a common tool applied for research of that kind. In order to use this measure, the analysed phenomenon is broken down into a set of economic aspects, where each aspect describes a different part of the economic phenomenon. For each aspect, a set of diagnostic variables can be defined, characterizing the selected aspect and allowing its description. Then, based on the adopted diagnostic variables, a synthetic variable TMD is determined, which takes into account the influence of all determinants of the economic phenomenon under research. This means that the estimated values of TMD allow for the synthetic evaluation of socio-economic phenomena.

The procedure for assessing the TMD was provided by Hellwig (1972). It can be described with the following steps[1]:

1. Determining the research problem and socio-economic phenomenon related to the problem.
2. Establishing a set of economic objects that will be analysed O_i.
3. Selection of a set of diagnostic variables X_i that describe various aspects of the phenomenon and determining their character (stimulant, dis-stimulant)[2].
4. Normalization of diagnostic variables X_i with the application of a chosen method (eg standardization, unitarization method), which enables to obtain a set of normalized variables Z_i.
5. Determination of the pattern of development P_i for all the diagnostic variables, where for stimulants a maximum value of the variable is taken.
6. Selecting the distance measure and determining the distance D_i of each object from the pattern of development (eg Euclidean distance, urban distance).
8. Determination of the TMD for the object $\mathrm{O_i}$ based on the formula:

$$\mathrm{TMD_i} = 1 - \frac{\mathrm{D_i}}{\mathrm{D_N}} \tag{1}$$

where: $\mathrm{D_N}$ is a variable that normalise TMD values in the range from zero to one.

---

[1] In the article the most important step of the procedure are described. The detailed presentation of the method is available in the seminal paper by Hellwig (1972).

[2] For stimulants X_{Sj} for every two values x_{Sij}, x_{Skj} that refer to objects O_i, O_k , the relation $x_{Sij} > x_{Skj} \rightarrow O_i\ \theta\ O_k$ is fulfilled, where θ means that object O_i is preferred to O_k. For dis-stimulants X_{Sj} for every two values x_{Sij}, x_{Skj} that refer to objects O_i, O_k, the relation $x_{Sij} < x_{Skj} \rightarrow O_i\ \pi\ O_k$ is fulfilled, where π means that object O_k is preferred to object O_i. (Balcerzak and Pietrzak, 2017a).

3 EMPIRICAL RESEARCH

The procedure for assessing the TMD, which was described in previous section, was used in the analysis of the development level of the digital economy in the EU. The study covered all 28 countries. With the help of the TMD measure, the level of digital economy in all the countries was estimated in the years 2011 and 2017. The obtained values of the digital economy level in two separate periods allowed to examine dynamics of the phenomenon in the period 2011-2017. In the first step of the study, a set of diagnostic variables was established. The adopted diagnostic variables are presented in Table 1. The diagnostic variables were selected based on the previous literature review (see Balcerzak and Pietrzak, 2017a; Wierzbicka, 2018). Table 1 presents also the nature of the variables – all of them were classified as stimulants. Data concerning the variables was obtained from the Eurostat database.

Table 1. Diagnostic variables applied for assessing of the level of digital economy in the EU countries

Variable	Name	Unit	Character
x1	Households - level of internet access: all type	% of households	stimulant
x2	Households - level of internet access: broadband	% of households	stimulant
x3	Households - availability of computers	% of households	stimulant
x4	Individuals - mobile internet access	% of individuals	stimulant
x5	Individuals - frequency of computer use	% of individuals	stimulant
x6	Individuals - computer use	% of individuals	stimulant
x7	Individuals - frequency of internet use	% of individuals	stimulant
x8	Internet purchases by individuals	% of individuals	stimulant
x9	E-government activities of individuals via websites	% of individuals	stimulant
x10	E-commerce purchases	% of individuals	stimulant

Source: own work based on Balcerzak and Pietrzak, 2017a.

Then, the standardization[3] of diagnostic variables was made and on the basis of the obtained values, the constant pattern of development for 2011 as well as for 2017 was taken. Adopting a fixed value of the pattern for subsequent years allowed to compare the obtained results in time. In the last stage the distances of the countries from the established pattern of development were calculated, and finally the values of the TMD were determined. The TMD values closer to unity indicate higher level of the digital economy development in a selected country. The TMD values also allowed to group the countries to four classes based on the natural breaks method[4]. The percentage change of the TMD in the period 2011-2017 was also calculated. The results are presented in Table 2 and Figure 1.

Table 2. the level of development of digital economy in the EU countries

Country	2011			2017			% change
	TMD	**Rank**	**Class**	**TMD**	**Rank**	**Class**	
Denmark	0,748	3	4	0,801	1	4	7,03%
Sweden	0,784	2	4	0,798	2	4	1,79%
Finland	0,793	1	4	0,761	3	4	-4,01%
Netherlands	0,669	6	4	0,737	4	4	10,12%
United Kingdom	0,706	5	4	0,712	5	4	0,87%

[3] The standardisation of the variables was based on the mean value and standard deviation.

[4] The main idea of the natural breaks method is based on the minimization of variance for objects from the chosen subsets and maximization of variance between the subsets (Jenks, 1967).

Luxembourg	0,658	7	4	0,670	6	4	1,85%
Germany	0,734	4	4	0,643	7	3	-12,42%
Estonia	0,513	13	3	0,609	8	3	18,72%
Austria	0,645	8	4	0,605	9	3	-6,21%
Czech Republic	0,504	15	3	0,571	10	3	13,38%
Belgium	0,546	10	3	0,561	11	3	2,89%
Malta	0,459	17	3	0,532	12	3	16,02%
Ireland	0,537	11	3	0,527	13	3	-1,86%
France	0,547	9	3	0,512	14	3	-6,41%
Hungary	0,477	16	3	0,511	15	3	7,12%
Slovakia	0,511	14	3	0,504	16	3	-1,36%
Slovenia	0,526	12	3	0,474	17	2	-9,86%
Spain	0,419	18	2	0,455	18	2	8,55%
Latvia	0,381	20	2	0,454	19	2	19,15%
Lithuania	0,396	19	2	0,420	20	2	6,07%
Cyprus	0,337	23	2	0,411	21	2	21,85%
Poland	0,367	21	2	0,359	22	2	-2,19%
Portugal	0,312	24	2	0,289	23	1	-7,37%
Italy	0,280	25	2	0,265	24	1	-5,45%
Greece	0,232	26	2	0,247	25	1	6,41%
Croatia	0,338	22	2	0,237	26	1	-29,97%
Romania	0,031	28	1	0,141	27	1	351,96%
Bulgaria	0,146	27	1	0,123	28	1	-15,63%

Source: own estimation based on Eurostat data.

In class 1, which was grouping the countries with the lowest level of the digital economy development, in 2011 there were only two countries: Romania and Bulgaria. In 2017, Croatia, Greece, Italy and Portugal have been degraded from class 2, as a result they have also been assigned to class 1. One should notice here very high growth of the TMD measure for Romania in 2017. For Bulgaria relatively slow changes in the situation has resulted in the decline in the obtained TMD measure by 15.63% in the analysed period. As a result, in the last year of the research Romania overtook Bulgaria in terms of the level of development of the analysed phenomenon. However, it should be emphasized that the two countries are characterized by very low level of the digital economy development and they distinguish disadvantageously in this respect from other EU countries.

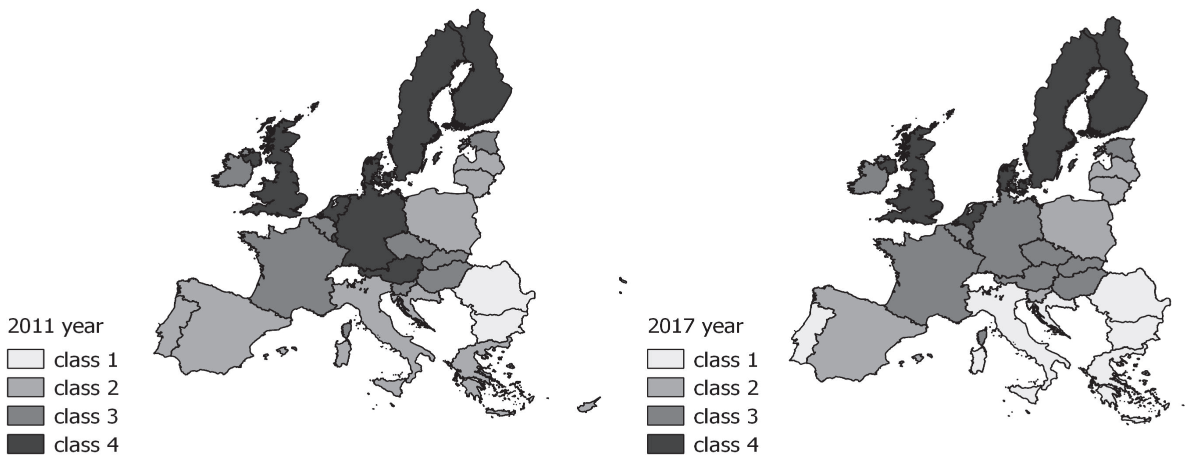
In class 2, also with a low level of digital economy development, in 2011 the countries of southern Europe and selected countries of Central and Eastern Europe dominate. In the class 2 one could see the following economies: Cyprus, Portugal, Italy, Greece, Spain, Latvia, Lithuania, Poland and Croatia. In 2017 class 2 was enlarged to include Slovenia and reduced by Croatia, Greece, Italy and Portugal, which was mentioned above.

In class 3 with an average digital economy level in 2011, there were selected countries of the so-called „old“ Union, Central and Eastern Europe and Malta. Malta, as the only southern European country, is characterized by a significantly higher level of the phenomenon. Among the countries of the „old“ union one could find here: Belgium, Ireland and France. In this class the following Central and Eastern European countries were grouped: Estonia, Czech Republic, Hungary, Slovakia and Slovenia. In 2017, Germany and Austria were also assigned to class 3. However, in 2017 Slovenia was degraded from class 3 to class 2. It is necessary to emphasize the high position of Estonia among the countries of Central and Eastern Europe. It can be stated that Estonia, due to the cultural closeness with the Scandinavian countries, may in the future probably again increase the level of digital economy and get closer to such countries as Sweden or Finland.

Class 4 with the highest level of the digital economy development in 2011 grouped only countries from the so-called „old“ Union. The following countries have been assigned to this class: Denmark, Sweden, Finland, the Netherlands, United Kingdom, Luxembourg, Germany and Austria. One should mention here the Scandinavian countries, which are characterized by the highest level of the digital economy development within the entire European Union. As it was mentioned, in 2017 Germany and Austria were removed from class 4.

In the case of TMD dynamics in 2011-2017, only for 10 in 28 countries the percentage change exceeded 10%. The negative percentage change occurred only in three countries: Germany (-12.42%), Bulgaria (-15.63%) and Croatia (-29.97%). Positive percentage changes occurred in the case of seven economies Holland (10.12%), Estonia (18.72%), Czech Republic (13.38%), Malta (16.02%), Latvia (19.15%), Cyprus (21.58%), and Romania (351.96%). Although many countries experienced significant percentage changes in the digital economy level, it should be emphasized that the established classes in 2017 have a similar structure to the classes from the year 2011. This factor is especially important form the policy point of view. In the begging of the XXI century it was commonly believed that the digital economy phenomenon can be used by relatively less developed European countries to increase the speed of closing their development gap. However, the current research confirms that using the potential of the digital economy is not automatic and can be only obtained under condition of good long term policy, as it was in the case of Estonia.

Figure 1. The level of TMD for digital economy in the year 2011 and 2017



Source: own calculations based on Eurostat data.

4 CONCLUSIONS

The main aim of the article was to analyse the level of digital economy development in the European Union countries. Based on the suggestion of the European Commission the phenomenon was analysed as a multiple-criteria problem. Therefore the method of taxonomic measure of development proposed by Zdzisław Hellwig was used. In the research the period 2011-2017 was analysed, which in the case of the digital economy development, can be considered as a long enough time to catch eventual structural changes. In the research the ranking and grouping of countries were proposed.

The main contribution of the paper relates to grouping of the European countries into 4 classes, which confirms structural diversity of the European Economy. Additionally, the research showed that the structural disparities between the countries are relatively stable, which means that the problem of digital trap at macroeconomic level can be also important issue in the European Union.

References

[1] Balcerzak, A. P. (2009). Effectiveness of the Institutional System Related to the Potential of the Knowledge Based Economy. *Ekonomista*, 6, pp. 711-739.

[2] Balcerzak, A. P. (2015). Europe 2020 Strategy and Structural Diversity Between Old and New Member States. Application of Zero-unitarization Method for Dynamic Analysis in the Years 2004-2013. *Economics & Sociology*, 8(2), pp. 190-210.

[3] Balcerzak, A. P. (2016a). Fiscal Burden in the European Union Countries. *Economic Annals XXI*, 161(9-10), pp. 4-6.

[4] Balcerzak, A. P. (2016b). Technological Potential of European Economy. Proposition of Measurement with Application of Multiple Criteria Decision Analysis. *Montenegrin Journal of Economics*, 12(3), pp. 7-17.

[5] Balcerzak, A. P. (2016c). Multiple-criteria Evaluation of Quality of Human Capital in the European Union Countries. *Economics & Sociology*, 9(2), pp. 11-27.

[6] Balcerzak, A. P. and Pietrzak, M. B. (2017a). Digital Economy in Visegrad Coutnries. Multiple-criteria decision analysis at Regional Level in the Years 2012 and 2015, *Journal of Competitiveness*, 9(2), pp. 5-18.

[7] Balcerzak, A. P. and Pietrzak, M. B. (2017b). Digital economy in Polish regions. Proposal of measurement via TOPSIS with generalized distance measure GDM. In M. Papież and S. Śmiech (Eds.). *The 11th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings*. Cracow: Foundation of the Cracow University of Economics, pp. 21-28.

[8] Balcerzak, A. P., Kliestik, T., Streimikiene, D. and Smrčka L. (2017). Non-parametric approach to measuring the efficiency of banking sectors in European Union Countries. *Acta Polytechnica Hungarica*, 14(7), pp. 51-70.

[9] Bartkowiak-Bakun, N. (2017). The diversity of socioeconomic development of rural areas in Poland in The Western Borderland and the problem of post-state farm localities. *Oeconomia Copernicana*, 8(3), pp. 417-432.

[10] Cano, J., Campo, E. and Gómez-Montoya, R. (2017). International Market Selection Using Fuzzy Weighing and Monte Carlo Simulation. *Polish Journal of Management Studies*, 16(2), pp. 40-50.

[11] Chrupała-Pniak, M., Grabowski, D. and Sulimowska-Formowicz, M. (2017). Trust in Effective International Business Cooperation: Mediating Effect of Work Engagement. *Entrepreneurial Business and* Economics Review, 5(2), pp. 27-50.

[12] Cyrek, M. (2017). Social efficiency of employment in three sectors — a comparison of Polish regions. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(3), pp. 417-432.

[13] Daragahi, G. A. (2017). The Impact of Innovation on Customer Satisfaction: A Study of the Cosmetics Producer in Tehran. *International Review*, 1-2, pp. 121-132.

[14] Hilkevics, S. and Hilkevica, G. (2017). New information technologies use for Latvian stock companies financial health evaluation. *Entrepreneurship and Sustainability Issues,* 5(2), pp. 178-189.

[15] Hellwig, Z. (1972). Procedure of Evaluating High-Level Manpower Data and Typology of Countries by Means of the Taxonomic Method. In: Z. Gostowski (Ed.). *Towards a System of Human Resources Indicators for Less Developed Countries, Papers Prepared for a UNESCO Research Project.* Wrocław: Ossolineum-The Polish Academy of Sciences, pp. 115-134.

[16] Ivanová, E. (2017). Barriers to the development of SMEs in the Slovak Republic. *Oeconomia Copernicana*, 8(2), pp. 255-272.

[17] Jenks, G. F. (1967). The Data Model Concept in Statistical Mapping. *International Yearbook of Cartography*, 7, pp. 186-190.

[18] Kaynak, S., Altuntas, S. and Dereli, T. (2017). Comparing the innovation performance of EU candidate countries: an entropy-based TOPSIS approach. *Economic Research-Ekonomska Istraživanja*, 30(1), pp. 31-54.

[19] Keshavarz Ghorabaee, M., Zavadskas, E. K., Amiri, M. and Antucheviciene, J. (2016). Evaluation by an Area-based Method of Ranking Interval Type-2 Fuzzy Sets (EAMRIT-2F) for Multi-criteria Group Decision-making. *Transformations in Business & Economics*, 15(3)(39), pp. 76-95.

[20] Kruk, H. and Waśniewska, A. (2017). Application of the Perkal method for assessing competitiveness of the countries of Central and Eastern Europe. *Oeconomia Copernicana*, 8(3), pp. 337-352.

[21] Kuc M. (2017b). Is the Regional Divergence a Price for the International Convergence. The Case of the Visegrad Group. *Journal of Competitiveness*, 9 (4), pp. 50-65.

[22] Kuc, M. (2017a). Social convergence in Nordic countries at regional level. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(1), pp. 25-41.

[23] Mandić, K., Delibašić, B., Knežević, S. and Benković, S. (2017). Analysis of the efficiency of insurance companies in Serbia using the fuzzy AHP and TOPSIS methods. *Economic Research-Ekonomska Istraživanja*, 30(1), pp. 550-565.

[24] Meluzin, T. , Pietrzak, M. B., Balcerzak, A. P., Zinecker, M., Doubravský, K. and Dohnal, M. (2017). Rumours Related to Political Instability and their Impact on IPOs: The Use of Qualitative Modeling with Incomplete Knowledge. *Polish Journal of Management Studies*, 16 (2), pp. 171-187.

[25] Meluzin, T., Zinecker, M., Balcerzak, A. P., Doubravský, K., Pietrzak, M. B. and Dohnal, M. (2018), The timing of initial public offerings – non-numerical model based on qualitative trends. *Journal of Business Economics and Management,* 19(1), pp. 109-125.

[26] Monni, S.; Palumbo, F. and Tvaronavičienė, M. (2017). Cluster performance: an attempt to evaluate the Lithuanian case. *Entrepreneurship and Sustainability Issues* 5(1), pp. 43-57.

[27] Pietrzak, M. B., Balcerzak, A. P., Gajdos, A. and Arendt, Ł (2017). Entrepreneurial environment at regional level: the case of Polish path towards sustainable socio-economic development. *Entrepreneurship and Sustainability Issues*, 5(2). pp. 190-203.

[28] Pilc, M. (2017). Cultural, political and economic roots of the labor market institutional framework in the OECD and post-socialist countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(4), pp. 713-731.

[29] Pilelienė, L. and Grigaliūnaitė, V. (2017). Colour temperature in advertising and its impact on consumer purchase intentions. *Oeconomia Copernicana*, 8(4), 657-669.

[30] Rogalska, E. (2017). Multivariate analysis of entrepreneurship in Poland at regional level. In T. Kliestik (Ed.). *17th International Scientific Conference Globalization and Its Socio-Economic Consequences. University of Zilina, The Faculty of Operation and Economics of Transport and Communication, Department of Economics, 4th – 5th October 2017. (Part V.).* Zilina: University of Zilina, pp. 2198-2202.

[31] Shpolianska, I., Dolzhenko, A. and Glushenko, S. (2017). Fuzzy models of web services selection in the development of service-oriented information systems for small and medium enterprises. *Ekonomicko-manazerske spectrum*, 11(2), pp. 64-75.

[32] Trąpczyński, P., Jankowska, B., Dzikowska, M. and Gorynia, M. (2016). Identification of Linkages between the Competitive Potential and Competitive Position of SMEs Related to their Internationalization Patterns Shortly after the Economic Crisis. *Entrepreneurial Business and Economics Review,* 4(4), pp. 29-50.

[33] Wierzbicka, W. (2018). Information infrastructure as a pillar of the knowledge-based economy — an analysis of regional differentiation in Poland. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *13*(1), pp. 123-139.

[34] Vavrek R., Adamisin P. and Kotulic R. (2017). Multi-criteria evaluation of municipalities in Slovakia – case study in selected districts. *Polish Journal of Management Studies*, 16(2), pp. 290-301.

[35] Zavadskas, K. E., Govindan, K., Antucheviciene, J. and Turskis, Z. (2016). Hybrid multiple criteria decision-making methods: a review of applications for sustainability issues. *Economic Research-Ekonomska Istraživanja*, 29(1), pp. 857-887.

[36] Zygmunt, A. (2017). Innovation activities of Polish firms. Multivariate analysis of the moderate innovator countries. *Oeconomia Copernicana*, 8(4), pp. 505-521.

[37] Żelazny, R., & Pietrucha, J. (2017). Measuring innovation and institution: the creative economy index. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(1), pp. 43-62.

Author's address

Michał Bernard Pietrzak, PhD
Nicolaus Copernicus University in Toruń, Department of Econometrics and Statistics, ul. Gagarina 13a, 87-100 Toruń, Poland
Mail: michal.pietrzak@umk.pl

Bartosz Ziemkiewicz, PhD
Nicolaus Copernicus University in Toruń, Department of Probability Theory and Stochastic Analysis, ul. Chopina 12/18, 87-100 Toruń, Poland
Mail: bartek@mat.umk.pl

THE MIQP MODEL FOR INVENTORY MANAGEMENT PROBLEM IN PRODUCTION SCHEDULING OF FRUIT JUICE BEVERAGES

Dražen Popović, Nenad Bjelić, Milorad Vidović, Gordana Radivojević, Branislava Ratković,
University of Belgrade, Faculty of Transport and Traffic Engineering, Department of Logistics

Abstract
A production lot-sizing and scheduling problem are among the most significant areas of management in most manufacturing companies. Usually there are several production lines with capacity constraints and setup times (cleaning, changeover etc). The production plan must define the quantity of products that will be produced on each production line in given planning horizon. The products consumption is stochastic value and therefore safety stocks are required to reduce the possibility of stock-outs. On the other hand, products shelf life can be very important for retailers since they must sell these products to final customers. Considering available stock of products at the start of planning horizon, as well as expected consumption, production plan must solve lot-sizing and scheduling problem so that the stock level of each product does not fall under safety stock level or go over maximum desirable stock regarding shelf life. Additionally, production facility has limited warehouse capacity for product storage. In this paper, we propose the mixed integer quadratic programming (MIQP) model to solve lot-sizing and scheduling problems with the main objectives: minimizing the violation of minimum and maximum stock level per each product, minimizing the violation of maximum stock level of all products, and finally minimizing total stock in the system.

Keywords: *inventory management, MIQP, production, lot-sizing and scheduling problem, fruit juice beverages.*

JEL Classification: C, L
AMS Classification: 90-08

1 INTRODUCTION

Manufacturing companies with production lines setup must constantly make production plans where number of products must be allocated on different production lines in observed planning horizon. This problem is known as lot-sizing and scheduling problem. As defined by Clark et al. 2011 production planning and scheduling seeks to efficiently allocate resources while fulfilling customer requirements and market demand, often by trading-off conflicting objectives. The main objective is to determine the moment of production as well as the production duration (which determines the produced quantity) on multiple production lines.

The beverage manufacturing industry is highly dependent on the efficient production planning and recently a number of research papers were published on this topic. Ferreira et al. 2010 developed MILP model for solving integrated lot sizing and scheduling decisions in production planning of small-scale soft drink plant, that was primarily focused on bottling phase (which is often a bottleneck in production). Similar problem was observed by Ferreira et al. 2012 but with somewhat different approach. They proposed four single-stage formulations to solve the synchronised two-stage lot sizing and scheduling problem, where the first stage's syrup lots in tanks is synchronized with the second stage's soft drink lots on bottling lines. Another research paper on the topic of lot sizing and scheduling was written by

Baldo et al. 2014 in the brewery industry, where authors presented MIP-based heuristic which produces relatively good-quality solutions for real-world problem instances. Most recent research paper on the topic of production scheduling was published by Pagliarussi et al. 2017 where authors developed MILP model for solving lot sizing and scheduling of fruit juice beverages production, in multi-period planning horizon and with multiple production lines.

In this paper we observe fruit juice beverage production with inventory management of finished goods as primarily optimization segment in MIQP model, which is described in detail in Section 2. In Section 3 we present the mathematical formulation, while test instances and computational results are given in Section 4. Concluding remarks are presented in Section 5.

2 PROBLEM DESCTIPTION

Usual problem setup consists of several production lines with capacity constraints and setup times (cleaning, changeover etc). The production plan must define the quantity of products that will be produced on each production line in given planning horizon, in other words for each product start time of production and the production duration must be determined. Production must satisfy expected products demand/consumption and this is the main generator of production. The products consumption is stochastic value and therefore safety stocks are required to reduce the possibility of stock-outs. On the other hand, products shelf life can be very important for retailers since they must sell these products to final customers. In that regard, high stock levels can reduce shelf life of products which can lead to obsolete products (these products cannot be delivered to retailers and therefore must be destroyed which incurs additional costs). Considering available stock of products at the start of planning horizon, as well as expected consumption, production plan must solve lot-sizing and scheduling problem so that the stock level of each product does not fall under safety stock level or go over maximum desirable stock regarding shelf life. Additionally, production facility has limited warehouse capacity for product storage. In the case of warehouse capacity overflow, excessive quantities can be stored at outsourced warehouse which incurs additional costs (more expensive than the storage in the production facility warehouse).

In this paper, we describe the MIQP model to solve lot-sizing and scheduling problem with stock level and shelf life constraints with the main objective of minimizing the violation of minimum and maximum stock level per each product, as it is presented in Figure 1 where C_i denotes expected daily consumption of product i. S_i represents safety stock as a function of the number of daily consumptions and M_i represents maximum desirable stock level regarding shelf life again as a function of the number of daily consumptions. For example, if a product's shelf life is 200 days with expected daily consumption of 2 pallets and retailer wants a minimum of 80% product's shelf life on delivery, then M_i is equal to 40 days of 2 pallets consumption (80 pallets of maximum desirable stock level of product i). We use quadratic formulation of these first two sub-functions because we want to evenly distribute any possible violations of safety and shelf life stock levels among all products. For example, it is better that all 10 products have 5% safety stock level violation then the case where 1 product has 50% violation of stock level (by use of the quadratic function former case is taken as better solution in MIQP).

Secondly, objective function of proposed MIQP should minimize the violation of maximum stock level of all products (regarding the available warehouse capacity at production facility) for the purpose of indirectly minimizing cost of outsourced warehousing. Finally, objective function should minimize total stock in the system, since larger stock level incurs larger inventory related cost mainly due to equipment and opportunity costs.

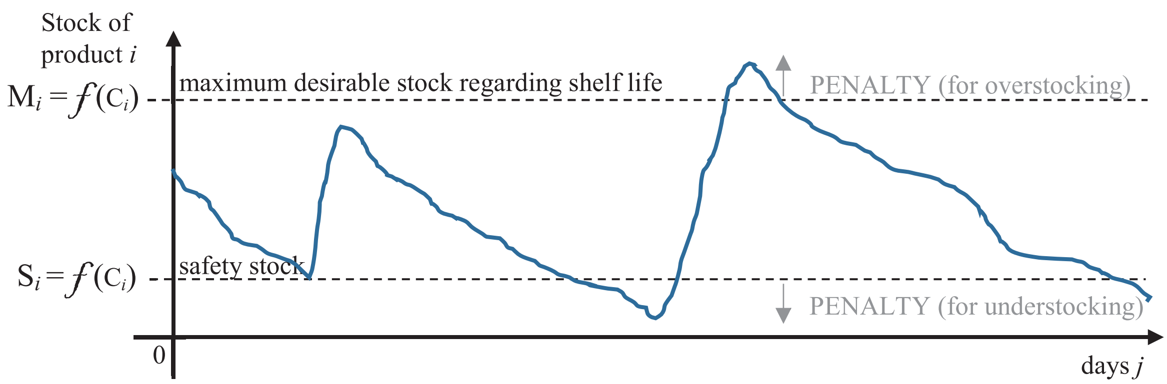


Figure 1. - Maximum and minimum desirable stock level of a product, based on safety stock and required shelf life

3 MATHEMATICAL FORMULATION

We developed multi objective MIQP model for solving lot sizing and scheduling in production planning of fruit juice bottling phase with multiple production lines in multi period planning horizon. Product i can have different format of package v (e.g. 0.2, 0.33, 0.5, 1, 1.5, 2 litres) and different package type (e.g. PET, Terta pak, glass). In our model, each package type requires separate production line k. Changing the production from one to another product requires changeover time on the production line denoted as U. This time can be additionally prolonged by value F if the two successive products have different formats of package. Finally, in each day disinfection of production line is required, denoted as D. Production is organized in two shifts, 16 h a day, 7 days a week. In the following we describe indices, parameters and variables of proposed MILP model.

Indices:

- i - article, $i \in I = (1, 2, \ldots, |I|)$
- j - day in observed planning horizon, $j \in J = (1, 2, \ldots, |J|)$
- v - format of packaging, $v \in V = (0.2, 0.33, 0.5, 1, 1.5, 2)$ in litres
- k - production line (product type group), $k \in K = (1, 2, \ldots, |K|)$

Parameters:

- C_i - expected consumption of product i of format v from group k
- z_{i0}^{vk} - stock level of product i of format v from group k at the start of the planning horizon (j=0)
- P_i^{vk} - hourly production capacity of product i of format v from group k
- D - duration of daily disinfection of production line
- F - duration of product format changeover on production line
- U - duration of product changeover on production line
- Q_{max} - available warehouse capacity for finished goods (palletized products)
- R_{max} - maximal storage time of product in outbound warehouse (from production to delivery)
- R_{min} - minimal number of consumption days that must be satisfied by safety stock level
- M_i - maximum desirable stock level of product i regarding shelf life (stock as a function of the number of daily consumptions)
- S_i - safety stock level of product i (stock as a function of the number of daily consumptions)

Decision variables:

x_{ij}^{vk} - production hours of product i in day j of format v from group k
z_{ij}^{vk} - stock level of product i in day j of format v from group k
O_j - violation of total stock level in day j (overflow regarding to the Q_{max})
A_{ij} - violation of M_i for product i in day j (as percentage of M_i)
B_{ij} - violation of minimal safety stock level for product i in day j (as percentage of S_i)
y_j^{vk} - binary variable that is equal to 1 if at least one product of format v is being produced in day j at production line k
y_j^k - number of product format changeover in day j on production line
t_{ij}^{vk} - binary variable that is equal to 1 if at product i of format v is being produced in day j at production line k
t_j^k - number of product changeover in day j on production line k

Objective function (1) has four weighted sub-functions that are focused on inventory management of finished goods (products): a) quadratic value minimization of overstocking the maximum shelf life related stock level; b) quadratic value minimization of understocking of safety stock level; c) minimization of total stock overflow regarding warehouse capacity; and d) minimization of overall stock level in the system.

$$\min \rightarrow \mathrm{w}_1 \cdot \sum_{i \in I} \sum_{j \in J} A_{ij}^2 + \mathrm{w}_2 \cdot \sum_{i \in I} \sum_{j \in J} B_{ij}^2 + \mathrm{w}_3 \cdot \sum_{j \in J} O_j + \mathrm{w}_4 \cdot \sum_{i \in I} z_{ij} \tag{1}$$

Constraints:

$$O_j \geq \sum_{i \in I} z_{ij} - Q_{\max} \quad \forall j \in J \tag{2}$$

$$z_{ij} = z_{ij-1} + x_{ij}^{vk} \cdot P_i^{vk} - C_i \quad \forall i \in I, j \in J \tag{3}$$

$$A_{ij} \geq \frac{z_{ij} - M_i}{M_i} \quad \forall i \in I, j \in J \tag{4}$$

$$M_i = R_{\max} \cdot C_i \quad \forall i \in I \tag{4.1}$$

$$B_{ij} \geq \frac{S_i - z_{ij}}{S_i} \quad \forall i \in I, j \in J \tag{5}$$

$$S_i = R_{\min} \cdot C_i \quad \forall i \in I \tag{5.1}$$

$$y_j^{vk} \geq \frac{x_{ij}^{vk}}{24} \quad \forall i \in I, j \in J, v \in V, k \in K \tag{6}$$

$$y_j^k \geq \sum_{v \in V} y_j^{vk} - 1 \quad \forall j \in J, k \in K \tag{6.1}$$

$$t_{ij}^{vk} \geq \frac{x_{ij}^{vk}}{24} \quad \forall i \in I, j \in J, v \in V, k \in K \tag{7}$$

$$t_j^k \geq \sum_{i\in I}\sum_{v\in V} t_{ij}^{vk} - 1 \quad \forall j \in J, k \in K \tag{7.1}$$

$$\sum_{i\in I}\sum_{v\in V} x_{ij}^{vk} \leq 16 - D - y_j^k \cdot F - t_j^k \cdot U \quad \forall j \in J, k \in K \tag{8}$$

$$\begin{aligned} & t_{ij}^{vk}, y_j^{vk} = \{0,1\} \quad \text{binary} \\ & O_j, A_{ij}, B_{ij} \geq 0 \quad \text{real positive} \\ & x_{ij}^{vk}, t_j^k, y_j^k \geq 0 \quad \text{integers} \\ & z_{ij} \text{ - real positive } (-\infty, +\infty) \end{aligned} \tag{9}$$

Constraints (2) defines total stock overflow of warehouse capacity per each day j of planning horizon. Stock level of product i per each day j is defined by constraints (3). Equations (4.1) defines M_i maximal desirable stock of product i regarding shelf life, while constraints (4) defines violation of M_i for product i in day j. Equations (5.1) defines S_i safety stock level of product i, while constraints (5) defines violation of S_i for product i in day j. Number of product changeovers in day j for production line k are defined by constraints (6) and (6.1). Number of product format changeovers in day j for production line k are defined by constraints (7) and (7.1). Constraints (8) defines that a sum of all production runs in day j for production line k must be lower or equal than available working hours in a single day (two shifts of total 16 h), considering all changeovers and cleansing time. Constraints (9) defines the nature of the variables. It should be noted that stock variable z_{ij} can have negative value, in which case some order quantities are unfulfilled and can be satisfied in the following days when sufficient stocks become available (backorder policy). These negative stocks are penalized by second sub-function of understocking (B_{ij}).

4 TEST INSTANCES AND COMPUTATIONAL RESULTS

This research was inspired by real life case in one of the leading fruit juicy companies with production located in Serbia. Therefore, we have generated test instances according to real case data, but somewhat simplified and reduced in size for the purpose of initial MIQP model formulation. Large real-life scale problems could not be solved to optimality, primarily due to products and planning horizon (about 150 products and 1 month of planning horizon can be the case in real life production). Our intention was to firstly develop MIQP model, and then in the later research to develop suitable heuristic approach for solving real life scale instances.

We generated instances with |I|=20 products that can be packed in 3 different packaging types (and therefore |K|=3 separate production lines). Each production line runs in two shifts of 16h in total, where disinfection is D=4 h, format changeover is F=3 h, and product changeover is U=3 h. The first line produces PET packaging with 3 formats (0.3, 0.5, 1.5 litres), second line produces TETRA PAK packaging with 2 formats (1, 2 litres), and third line produces GLASS packaging with 1 format (0.2 litres). Accordingly, production capacities for three production lines and its formats are [[2.4, 4.0, 7.2], [2.8, 3.2], [0.4]], given in pallets per hour. Minimal number of consumption days that must be satisfied by safety stock level is R_{min}=5 days, while maximal storage time of product in outbound warehouse R_{max} = 20 days. The available warehouse capacity for finished goods is Q_{max}=400 pallet locations. One input data example is presented in Table 1 (data are randomly generated by exponential and gauss distribution according to reduced real-life data). Mathematical model was implemented by the CPLEX 12.6. on the Intel(R) Core(TM) i5-3470 CPU @ 3.20GHz with 8 GB RAM. In Table 2. we present optimal solution of test instance which is given in Table 1.

Table 1. Input data example of one test instance

Product i	Production line k	Format v	Daily consumption C_i	Starting stock z_i	Packaging type
1	1	0.3	1.58	10.50	PET
2	1	0.3	0.22	1.52	PET
3	1	0.5	1.57	8.07	PET
4	1	0.5	2.45	16.47	PET
5	1	0.5	0.56	2.77	PET
6	1	1.5	16.06	80.91	PET
7	1	1.5	0.60	3.02	PET
8	1	1.5	34.53	189.74	PET
9	2	1	0.10	0.74	TETRA_PAK
10	2	1	1.36	7.78	TETRA_PAK
11	2	1	0.34	1.58	TETRA_PAK
12	2	1	0.47	2.18	TETRA_PAK
13	2	2	0.21	1.16	TETRA_PAK
14	2	2	4.68	31.17	TETRA_PAK
15	2	2	0.86	4.56	TETRA_PAK
16	2	2	3.72	22.79	TETRA_PAK
17	3	0.2	0.84	4.60	GLASS
18	3	0.2	0.50	3.70	GLASS
19	3	0.2	3.45	18.22	GLASS
20	3	0.2	0.10	0.61	GLASS

Table 2. Optimal solution of test instance from Table 1 (w_1=1000, w_2=10000, w_3=1, w_4=0.01)

	x_{ij} - production hours							z_{ij} - stock level							Safety stock S_i	Shelf life stock M_i
i \ j	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*1*	*2*	*3*	*4*	*5*	*6*	*7*		
1	-	-	-	4	-	-	-	8.9	**7.3**	**5.8**	13.8	12.2	10.6	9.1	7.9	31.5
2	-	-	-	1	-	-	-	1.3	**1.1**	**0.8**	3.0	2.8	2.6	2.4	1.1	4.5
3	2	-	-	-	-	-	1	14.5	12.9	11.4	9.8	8.2	**6.7**	9.1	7.9	31.4
4	3	-	-	-	-	-	-	26.0	23.6	21.1	18.6	16.2	13.7	**11.3**	12.3	49.1
5	1	-	-	-	-	-	-	6.2	5.6	5.1	4.5	4.0	3.4	**2.8**	2.8	11.2
6	-	8	-	-	2	4	-	**64.9**	106.4	90.3	**74.3**	**72.6**	85.4	**69.3**	80.3	321.2
7	-	1	-	-	-	-	-	**2.4**	9.0	8.4	7.8	7.2	6.6	6	3.0	12.0
8	-	-	12	-	7	5	5	**155.2**	**120.7**	**172.5**	**138.0**	**153.9**	**155.3**	**156.8**	172.7	690.7
9	-	-	-	1	-	-	-	0.6	0.5	**0.4**	**3.1**	**3.0**	**2.9**	**2.8**	0.5	2.0
10	-	1	-	2	-	-	1	**6.4**	7.9	**6.5**	10.7	9.4	8.0	9.5	6.8	27.2
11	1	-	-	-	-	-	-	4.0	3.7	3.4	3.0	2.7	2.3	2	1.7	6.8
12	1	-	-	-	-	1	-	4.5	4.0	3.6	3.1	2.6	4.9	4.5	2.4	9.5
13	-	1	-	-	-	-	-	**1.0**	3.9	3.7	3.5	3.3	3.1	2.9	1.0	4.1
14	-	-	4	-	1	2	1	26.5	**21.8**	29.9	25.3	23.8	25.5	24	23.4	93.5
15	1	-	-	-	1	-	-	6.9	6.0	5.2	4.3	6.6	5.8	4.9	4.3	17.3
16	-	1	3	-	2	-	1	19.1	**18.5**	24.4	20.7	23.4	19.6	19.1	18.6	74.5
17	-	7	-	4	-	-	3	**3.8**	5.7	4.9	5.6	4.8	4.0	4.3	4.2	16.8
18	-	-	-	5	-	-	-	3.2	2.7	**2.2**	3.7	3.2	2.7	**2.2**	2.5	9.9
19	12	-	12	-	12	12	6	19.6	**16.1**	17.5	**14.0**	**15.4**	**16.7**	**15.7**	17.2	69.0
20	-	2	-	-	-	-	-	0.5	1.2	1.1	1.0	0.9	0.8	0.7	0.5	2.0
							$\sum$	375.5	378.6	**418.2**	367.8	376.2	380.6	359.4		

5 CONCLUSIONS

In this research paper we have presented MIQP model for solving a production lot-sizing and scheduling problem in fruit juice beverage production with inventory management. We have presented one random instance and its optimal solution obtained from MIQP model. Model was able to optimally solve this small-scale instance in 117.4 seconds and it is not capable of solving larger real-life scale problems to optimality in reasonable computational time (even some of the random generated small-scale instances of the same size cannot be solved to optimality in 1000 seconds), where computational time raises exponentially with the increase of problem dimensions (mainly dependent on the size of product assortment and length of planning horizon).

We can see from the results in Table 2 that production plan has relatively small deviance from stock level restrictions (bolded values in Table 2): daily average of shelf life stock level violation per product is 1.4%, daily average of safety stock level violation per product is 2.6%, and daily average violation of total stock level is 2.6 pallets (only happened once with 18.3 pallets overflow). Off course, these deviations are also dependent on the input data especially in the case when this model is introduced in the production for the first time, as well as on the quality of the proposed MIQP.

Further research direction should be focused on developing suitable heuristic approach that can obtain good quality solutions in reasonable computational time.

Acknowledgements: This work was supported by the Ministry of Education, Science and Technical development of the Government of the Republic of Serbia through the project TR36006, for the period 2011-2018.

References

[1] Alistair Clark, Bernardo Almada-Lobo, and Christian Almeder (2011). Lot sizing and scheduling: industrial extensions and research opportunities. *International Journal of Production Research*, 49(9), pp. 2457-2461.

[2] Deisemara Ferreira, Alistair R. Clark, Bernardo Almada-Lobo, and Reinaldo Morabito (2012). Single-stage formulations for synchronised two-stage lot sizing and scheduling in soft drink production. *International Journal of Production Economics*, 136(2), pp. 255-265.

[3] Deisemara Ferreira, Reinaldo Morabito, and Socorro Rangel (2010). Relax and fix heuristics to solve one-stage one-machine lot-scheduling models for small-scale soft drink plants. *Computers & Operations Research*, 37(4), pp. 684-691.

[4] Marina Sanches Pagliarussi, Reinaldo Morabito, and Maristela Oliveira Santos (2017). Optimizing the production scheduling of fruit juice beverages using mixed integer programming models. *Gestão & Produção*, 24(1), pp. 64-77.

[5] Tamara A. Baldo, Maristela O. Santos, Bernardo Almada-Lobo, and Reinaldo Morabito (2014). An optimization approach for the lot sizing and scheduling problem in the brewery industry. *Computers & Industrial Engineering,* 72, pp. 58–71.

Author's address

Dražen Popović, Assistant Professor, Nenad Bjelić, Assistant Professor, Milorad Vidović, Professor, Gordana Radivojević, Professor, Branislava Ratković, Assistant Professor
University of Belgrade, Faculty of Transport and Traffic Engineering, Department of Logistics
Vojvode Stepe 305, Belgrade
Republic of Serbia
emails: d.popovic@sf.bg.ac.rs, n.bjelic@sf.bg.ac.rs, m.vidovic@sf.bg.ac.rs, g.radivojevic@sf.bg.ac.rs, b.ratkovic@sf.bg.ac.rs

HOME EQUITY LOAN AS A CAUSE OF ASYMMETRIC RELATIONSHIP BETWEEN REAL ESTATE PRICE AND CONSUMPTION

Vlastimil Reichel, Masaryk University

Abstract
The aim of the paper is to point out the cause of the asymmetric relationship between real estate price and consumption. Generally, when prices raise consumption decreases. However, real estate is not only source of housing. It is also a source of wealth and as such can be converted into liquidity at a suitable time by using a home equity loan. Unfortunately, the best time for this convertibility is the moment least suitable for the authorities who are responsible for monetary control. Firstly, I refer to the channel of a transmission mechanism by which central bank can disrupt the economy at the time of boom. Secondly, I introduce and verify a mechanism that may cause the rise of consumption depending on the real estate price increase in boom times. Last but not least, I mention why there could be the relationship between the price of real estate and the consumption problematic when setting monetary policy. The cross-correlation analysis is implemented on data of the Czech Republic which cover the period between 2006Q1 and 2017Q4. The outcome of the work results in the fact that a mechanism which can pump up a large amount of financial resources to the economy at an inappropriate moment can be a reason why the effectiveness of restrictive measures of the monetary authorities in the Czech Republic is slowing down.
Keywords: *Transmission mechanism, Home equity loan, Cross-correlation analysis*

JEL Classification: E320, E580
AMS Classification: 62M10

1 INTRODUCTION

Central banks of advanced economies tend to respond to economic fluctuations through different channels of transmission mechanisms to properly implement their monetary policy. If the central bank targets inflation, then it can be expected to do everything they can keep inflation at a predetermined level. Its main objective - the price stability - the CNB achieves by altering the setting of monetary conditions using its instruments, especially the basic interest rates. Open market operations increase / decrease the market short-term interest rate and thereby influence the behavior of economic agents who are deciding on their investments or their future consumption. In other words, if there is an unexpected deviation from the inflation target, for example by rising property prices, the central bank will adjust its policy to prevent this growth. The intervention ultimately affects consumption or investment and is reflected in the output of the economy (CNB, 2018).

One channel that describes this mechanism is the Balance sheet channel (Mischkin, 1996). This channel is using Client Nominal Interest Rate. If the central bank, in response to the unexpected increase in price, increases the nominal interest rates, commercial banks respond by increasing their client's nominal interest rates. This ultimately means a drop in demand for new loans. The decline in new loans leads to a reduction in the net assets of economic agents, and consumer or investment spending decreases. The reduction may also be accelerated by the existence of the External Financial Premium, which commercial banks charge as a form of bonus associated with uncertainty in the market (Mischkin, 1996)(Iacoviello and Minetti, 2008).

2 TRANSMISSION MECHANISM AND HOME EQUITY LOAN

The home equity loan is the name of a mortgage loan that is not assigned. This is a non-purpose loan secured by a lien of the property. The debtor can use the received funds to satisfy all the needs; which mean that the main purpose is not for real estate purchase, land, reconstruction, etc.

The question remains, what role the home equity loan has in the transmission mechanism and how it could link the price of real estate and consumption. It is known that the growth of property prices discourages economic agents from buying them. The logical consequence of this change in behavior is that economic agents start demanding fewer loans, covered by real estate, and pressure on demand for real estate to ease. In addition to the monetary restriction, this effect is reinforced by an increase in interest rates for mortgage banks. The central bank can intervene if perceives that the economy (or property prices) is growing fast. Monetary restrictions indirectly cause a decrease in consumption through the transmission mechanism (Mischkin, 1996).

This creates a place for introducing the home equity loan into the transmission mechanism as a type of financial friction (Guerrieri and Iacoviello, 2017). As has been said, the home equity loan is a non-purpose loan and can be drawn to any financial activity of economic agents. It means that some economic agents while drawing on this loan perceive the rise in property prices rather positively because the higher price of real estate means the possibility to borrow a higher amount of non-purpose funds. In other words, the wealth due to the ownership of the property has increased, and they use it as a secured loan to transform into consumption (Iacoviello, 2012). By considering a larger number of similarly minded agents, the economy is getting a lot of extra money, even if it can be tightened by the central bank. This mechanism was modeled using the DSGE model with occasional binding constraints (Reichel and Hloušek, 2015). Unfortunately, there was a lack of evidence that home equity loans are being used in a boom more than in a recession in the Czech Republic. To find out how much money have accumulated in the economy, the information from penize.com were used. In 2006, this web site reported (Peníze, 2006) that one of the mortgage banks provided nearly five thousand of home equity loans per year with an average loan amount of half a million crowns. By taking the average amount and multiply it by the total number of the home equity loans that were provided on the market in 2006, we receive a loan worth more than CZK 8 billion. This number is quite high if we imagine that the first home equity loans were only released in the Czech Republic in 2004 and how the real estate market has been under development so far.

Since most mortgage banks make sure that home equity loan is not used to finance investment spending, these loans are consumed. Put another way, households have borrowed money used, for example, to buy a car, finance studies; settle the joint property of spouses in divorce proceedings, etc. Therefore it is rather clear that the household mortgage consumption is increasing by home equity loan. Followingly will be determined why this increase can be initiated at the wrong time and thus disrupt monetary policy measures.

Guerrieri and Iacoviello (2017) have shown an asymmetrical relationship between real estate and consumption prices on US data. The results of the work show that a higher price of real estate has motivated US households to use home equity loans excessively, thus distorting the financial market, and the Fed's subsequent reactions have not prevented the onset of the financial crisis. It is worth mentioning that the amount of purpose or non-purpose borrowed mortgages for the most of the time with the price of real estate is counter-cyclical and slightly overtaking it. However, in times of excessive real estate prices, the number of drawn non-purpose mortgages increases sharply and becomes pro-cyclical and synchronous with the property price. Even though the central bank tends to tighten the economy by increasing the interest rates, consumption is growing. This creates a space for financial friction, which can directly face the classic direction of consumption when the central bank restrains.

3 METHODOLOGY

The methodology can be written in two steps: (i) data collection and their preparation and (ii) a cross-correlation calculating.

3.1 Data

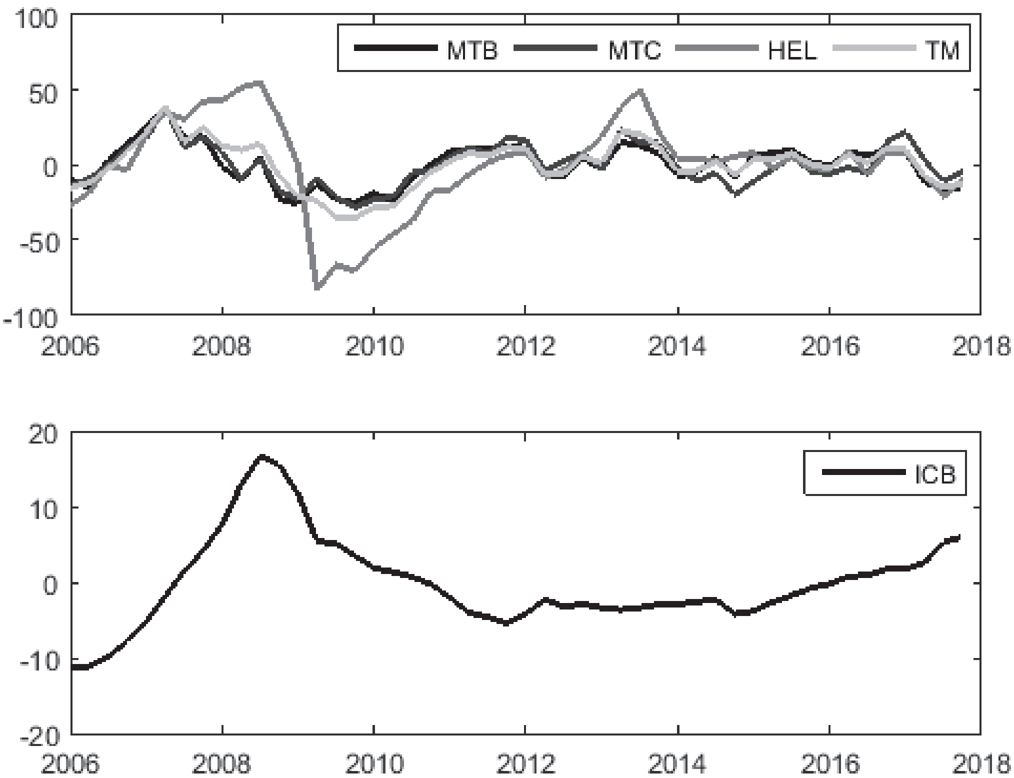
The time series of the flat price index and the development of the number of mortgage loans according to the purpose of their drawing were used in this paper. The length of time series covers the period from 2006:Q1 to 2017:Q4. A more detailed description of the time series is given in the Table 1.

Table 1 Time series description

Time series		Data source	Frequency
The offering price of flats	ICB	CZSO	quarterly
Volume of mortgage to buying	MTB	FH	monthly
Volume of mortgage to construct	MTC	FH	monthly
Volume of home equity loan	HEL	FH	monthly
Volume of Total of mortgage	TM	FH	monthly

Source: Czech Statistical Office (2018), Fincentrum hypoindex (2018)

Figure 1 Time series gaps



Source: Own construction

The MTB, MTC, HEL and TM variables have been transformed from monthly to quarterly frequency and seasonally adjusted. All variables were stationarised using the Hodrick-Prescott filter (with lambda=1600, see Hodrick and Prescott 1981). The gaps gained can be seen in Figure 1.

3.2 The cross-correlation approach

The cross-correlation approach is a method of estimating the degree where two series are correlated. The theory and algorithm were adopted from Bourke (1996). Consider two series x_i and y_i where $i = 0,1,2... N - 1$. Here x_i is the vector of all variables mentioned above and y_i is house price. The cross-correlation $r(d)$ at delay d is defined as

$$r(d) = \frac{\sum_i[(x_i - \bar{x}) * (y_{i-d} - \bar{y})]}{\sqrt{\sum_i(x_i - \bar{x})^2} * \sqrt{\sum_i(y_{i-d} - \bar{y})^2}},$$

where $\bar{x}$ and $\bar{y}$ are the samples mean of the corresponding series. If the above is computed for all delays $d = 0,1,2, ... N - 1$ then it results in a cross-correlation series of twice the length of the original series.

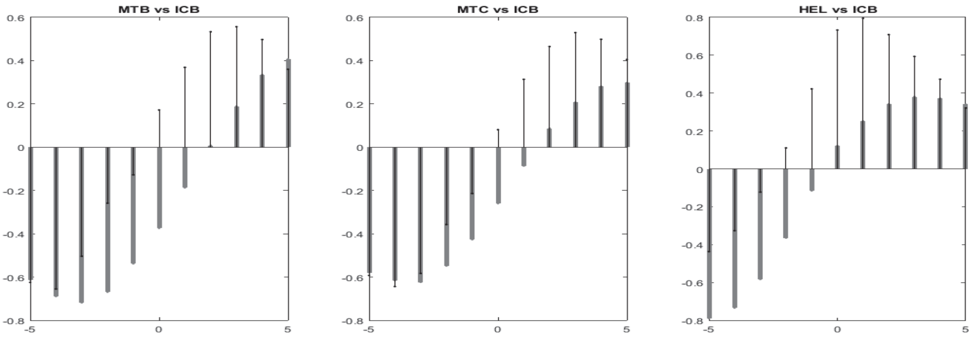
The range of delays d and so the length of the cross-correlation series can be less than N, for example the aim may be to test correlation at the short delays only. The denominator in the expression above serves to normalize the correlation coefficients such as $-1 \leq r(d) \leq 1$, where the bounds indicate maximum correlation and "0" indicates no correlation. A high negative correlation indicates a high correlation of the inverse of one of the series.

4 RESULTS

The second chapter described the positive link between the price of real estate and the amount of non-purpose loans. Now it will be shown, with the data of the Czech Republic, if the claim could be considered valid even within this economy. Figure 1 illustrates the development of gaps in selected time series. The picture shows clear pattern in the development of time series HEL and ICB during the boom in 2008 were pro-cyclical. Both of them reached their peak in the same quarter. In comparison with the MTB and MTC time series to ICB, is obvious the decrease in a few quarters earlier than HEL is obvious. As far as the overall development of time series gaps is concerned, it is not possible to identify the pro-cyclicality of the individual series over the whole period.

To determine whether the series are procyclical or counter-cyclical and delayed or overtaking, the cross-correlation calculation method was used. The results of this method are captured in Figure 2. The y-axis stands for the amount of calculated correlation, the x-axis changed one of the time series, where x = 0 indicates that the time series are not shifted in the calculation. Cross-correlations have been calculated over the two periods, particularly the boom period and the consequent fall from 2006:Q1 to 2010:Q4 marked by black color and the period from 2006:Q1 to 2017:Q4 marked by grey. It is seen from the Figure 2 that the relations between all types of mortgages and the real estate prices during the boom period remain the same - the gray and black columns overlap, only the strength of the relationship changes in the boom period.

Figure 2 Crosscorrelation coefficients



Source: Own construction

The values of all cross-correlations are captured in Table 2. The highest correlation values are underlined. The highest correlation values between the ICB and the MTB were negative on the whole sample 2006:Q1 - 2017:Q4 and were shifted by three periods. This signifies the counter-cyclical relationship between MTB (or MTC) mortgage loans and the real estate price, with real estate prices lagging behind the number of mortgages taken over three periods. Unfortunately, the differences between the values of the correlation coefficient between ICB and MTB (or MTC) are statistically insignificant between shifts t-5 to t-1. For this reason, it cannot be clearly determined what period is variable MTB (or MTC) versus variable ICB shifted.

In the boom period 2006:Q1 - 2010:Q4 very similar results of the correlation coefficients between the variables ICB and MTB (or MTC) are determined. Behavior is again counter-cyclical, and MTB (or MTC) leads the ICB by three to five quarters with the highest correlation values in the time t-4 (again, it cannot be determined clearly because of the statistical significance of correlation differences). The difference between the whole sample and the boom period is only observed between the ICB and HEL variables. In contrast, for the HEL mortgage loans, counter-cyclical and lead behavior throughout the whole period becomes pro-cyclical and synchronous behavior just as higher real estate prices rises (in the boom period). This means, in the case of excessive real estate prices, number of the home equity loans is rising.

Table 2: Cross-correlation between house price and the other variable

Variable	t-5	t-4	t-3	t-2	t-1	t	t+1	t+2	t+3	t+4	t+5
Period: 2006:Q1 - 2017:Q4											
MTB	-0,61	-0,69	-0,72	-0,67	-0,54	-0,37	-0,19	0,01	0,19	0,34	0,41
MTC	-0,58	-0,62	-0,62	-0,55	-0,43	-0,26	-0,09	0,09	0,21	0,28	0,30
HEL	-0,79	-0,73	-0,58	-0,36	-0,12	0,12	0,25	0,34	0,38	0,37	0,34
Period: 2006:Q1 - 2010:Q4											
MTB	-0,62	-0,65	-0,50	-0,26	-0,13	0,17	0,37	0,53	0,56	0,50	0,36
MTC	-0,59	-0,64	-0,58	-0,36	-0,21	0,08	0,31	0,46	0,53	0,50	0,41
HEL	-0,44	-0,33	-0,12	0,11	0,42	0,73	0,79	0,71	0,59	0,47	0,32

Source: Own construction

5 DISCUSSION

Prices for apartments in the Czech Republic have risen steadily since their minimum in 2012. The dynamics of growth slowed down in comparison to 2015, but the price of small apartments scored increase by an additional percent. According to the Remax real estate agency (Remax, 2018), the average prices of apartments in Brno (disposition 1 + 1) has increased by 69% from its minimum 1.28 million in 2012 to 2.16 million Czech crowns. The average price for the apartments with 2 + 1 disposition it has risen by 45% (average bid price rose from 2.07 to 3 million). The price for the apartments with 4 + 1 disposition prices increased by 55% since the minimum in 2012 (the average price rose from 3.65 million to 5.64 million).

The CNB responded to this increase by tightening the supervision of mortgage banks. Firstly has recommended to mortgage banks to reduce the loan to value (LTV) ratio. As the impact of these measures was insufficient, the CNB tried to enforce the amendment of the CNB Act. This amendment would allow the CNB to gain control over LTV, Debt Service To Income (DSTI) and Debt to Income (DTI) indicators. In the case of non-compliance with these conditions, the CNB could impose a sanction to mortgage banks. When the CNB's amendment law was not approved by the government, the CNB responded by raising the interest rates. The effects of monetary restrictions moderated the pressure on real estate demand and the rise in real estate prices slowed down according to the indicators. The number of newly drawn mortgage loans has decreased as was expected. However, the composition of mortgage loans by purpose has changed. The change in the number of mortgages by purpose is shown in Table 3. It is clear that the number of MTB mortgage rates has fallen, MTC remains relatively stable and HEL has risen. The question is how this scenario is consistent with the scenario of economic agents' behavior in 2008 which was described in chapter 2.

Table 3 Development of the amount of mortgage loans

	2017:M9	2017:M10	2017:M11	2017:M12	2018:M1
Total	8 235	8 667	9 466	9351	7783
MTB	5023 (61%)	5287 (61%)	5679 (60%)	5237 (56%)	4202 (54%)
MTC	1647 (20%)	1647 (19%)	1798 (19%)	1777 (19%)	1712 (22%)
HEL	1565 (19%)	1733 (20%)	1989 (21%)	2337 (25%)	1869 (24%)

Source: Fincentrum Hypoindex (2018)

6 CONCLUSIONS

The paper discussed the relationship between the price of real estate and the drawing of mortgage loans. The paper was supported the claim that the real estate is not only a source of housing but also the wealth that can be converted into liquidity at a suitable time using a home equity loan. Unfortunately, the best time to convert is when the authorities who are responsible for monetary control are the least suitable. An example supporting this claim was demonstrated by the recent economic crisis in the US.

By using the appropriate data describing the Czech economy, the relevance of this topic to the economy has been determined. It was referred to the existence of a mechanism in the Czech Republic, which is able to pump up a large amount of funds into the economy at an inappropriate moment and thereby slow down the effectiveness of restrictive measures by the monetary control authorities. Ultimately was discussed if the current real estate prices in the Czech Republic may be comparable to the real estate price growth in the pre-crisis period of 2008.

The paper supported the presumption of the existence of credit constraints. Further work will focus on modeling the transmission mechanism in a DSGE model with credit constraints. Immediately there is an excessive growth in property prices in the Czech Republic. This was the main reason for writing this paper. The paper helps to understand the relationships of selected variables in the boom of the economy.

Acknowledgements: This work was supported by funding of specific research at Faculty of Economics and Administration, Masaryk University, project MUNI/A/0966/2017. This support is gratefully acknowledged.

References

[1] Guerrieri, L. and Iacoviello M. (2017). Collateral Constraints and Macroeconomic Asymmetries, *Journal of Monetary Economics*, 90(9), pp.28-49.
[2] Hodrick, R. and Prescott, E. (1981). Post-war U.S. business cycles: an empirical investigation. *Discussion Papers 451*, Northwestern University, Center for Mathematical Studies in Economics and Management Science.
[3] Iacoviello, M. and Minetti, R. (2008). The Credit Channel of Monetary Policy: Evidence from the Housing Market, *Journal of Macroeconomics*, 30(1), pp. 69-96.
[4] Iacoviello, M. (2012). Housing Wealth and Consumption, *International Encyclopedia of Housing and Home*, pp. 673-678.
[5] Mishkin, F. S. (1996). The Channels of Monetary Tra0nsmission: Lessons for Monetary Policy, *NBER Working Papers 5464*, National Bureau of Economic Research, Inc.
[6] Reichel, V. and Hloušek, M. (2015). Asymmetric behaviour of collateral constraint: applications on the Czech economy. In Martinčík, D., Ircingová, J., Janeček P.,. *Proceedings of 33nd International Conference Mathematical Methods in Economics. Plzeň*: University of West Bohemia. pp. 707-712.
[7] Bourke, P. (1996). CrossCorrelation [online]. Available at: http://paulbourke.net/miscellaneous/correlate/ [Accessed 7 Mar. 2018].
[8] Czech Statistical Office, (2018). Database - CZSO [online]. Available at: https://www.czso.cz/csu/czso/ceny_bytu [Accessed 8 Mar. 2018].
[9] Czech National Bank, (2018). CNB [online]. Available at: https://www.cnb.cz/cs/menova_politika/ [Accessed 9 Mar. 2018].
[10] Fincentrum Hypoindex, (2018). Database - Fincentrum Hypoindex [online]. Available at: https://www.hypoindex.cz/hypoindex-vyvoj [Accessed 10 Mar. 2018].
[11] Peníze, (2006). Peníze [online]. Available at: https://www.penize.cz/18102-americka-hypoteka-jak-se-vam-libi-penize-na-cokoli [Accessed 11 Mar. 2018].
[12] Remax, (2018). Remax [online]. Available at: https://www.remaxpro.cz/kompletni-informace-o-vyvoji-cen-nemovitosti-v-brne-a-okoli-i-ctvrtleti-2017/ [Accessed 12 Mar. 2018].

Author's address
Ing. Mgr. Vlastimil Reichel
Masaryk University
Lipová 507/41a
602 00, Brno
Czech Republic
Email: reichel.v@mail.muni.cz

CLUSTER ANALYSIS OF ENTREPRENEURIAL ENVIRONMENT IN POLISH REGIONS

Elżbieta Rogalska,University of Warmia and Mazury in Olsztyn

Abstract
Level of entrepreneurship and conditions for growth of small and medium enterprises are currently considered as important determinants of economic development. The level of entrepreneurship and entrepreneurial conditions are especially important for growth and economic stability at regional level. The main objective of the article is to analyse similarities between Polish regions in regard to entrepreneurial conditions at regional level. The research was done at NUTS 2 level with application of data provided from Statics Poland (Local data bank). The study was conducted for the years 2011 and 2017. The analysed phenomenon was considered as a multiple-criteria problem. Thus, the Ward's clustering method was applied in order to group the regions into homogeneous classes. The conducted research confirmed the disparities between regions in Poland and showed their relatively high stability in time. From government policy perspective this factor can be considered as a confirmation of relatively low effectiveness of regional policy in Poland.

Keywords: *Cluster analysis, Ward's method, entrepreneurship, Poland*

JEL Classification: C38, R11
AMS Classification: 91C20

1 INTRODUCTION

Quality of entrepreneurial environment and good conditions for growth of small and medium enterprises (SMEs) have been commonly considered as important determinants of sustainable development both at regional and national level (Czarniewski, 2016; Begonja et al., 2016; Kozubíková et al., 2017; Pietrzak and Balcerzak, 2016a; 2016b; Dobeš et al., 2016; Erkut, 2016; Zygmunt, 2017; Fabuš, 2017; Dvorsky et al., 2017; Petrenko et al., 2017; Vinczeova and Kascakova, 2017; Pietrzak et al., 2017). In all European economies SMEs are responsible for creation of majority of jobs and high percentage of GDP. The economic role of SMEs is especially important in Central and Eastern European Economies (Kljucnikov and Belas, 2016; Tomovska Misoska et al., 2016; Simionescu et al., 2017). In the case of regional development SMEs are usually the main economic stabilizer of socio-economic welfare and quality of life of local communities. Therefore, the research concerning entrepreneurial conditions at regional level should be considered as an important scientific task.

The main aim of the research is to analyse similarities between Polish regions in regard to entrepreneurial conditions at regional level. In the current paper the analysis has been done at NUTS 2 level. Based on previous research the analysed phenomenon was considered as a multiple-criteria problem (Rogalska, 2017), which justified application of taxonomic tools. In the article, the Ward's clustering method was applied, which enabled to group the regions into homogeneous subsets. The research was conducted for the years 2011 and 2017. In the analysis the data provided by Statics Poland (Local data bank) was used.

2 CLUSTER ANALISIS: IDEA AND METHODS

For most of socio-economic phenomena one can distinguish a set of economic aspects, where each of which describes different properties of a given scientific problem (Balcerzak and Pietrzak, 2016; 2017; Meluzin et al., 2017; 2018; Stankeviciene et al., 2017; Ahmed et al., 2017). This means that economic phenomena usually have a multidimensional character (Mardani et al., 2016; Keshavarz Ghorabaee et al. 2016; 2017), which should be analysed with application of taxonomic tools. The usefulness of taxonomic approach can be seen on the example of research concerning a level of socio-economic development (Bartkowiak-Bakun, 2017) or a level of international competitiveness (Pietrzak and Balcerzak, 2016c; Cheba and Szopik-Depczyńska, 2017; Kruk and Waśniewska, 2017). Tools that enable to conduct effective research in regard to such phenomena are provided by a multiple-criteria decision analysis methodologies (Balcerzak et al., 2017; Cano et al., 2017; Stanickova, 2017; Zemlickienė et al., 2018).

The tasks of ordering and classifying of objects are important objectives in the case of multidimensional comparative analysis. The ordering relates to objects evaluation on a basis of a value of obtained synthetic variable (for example taxonomic measure of development), and then creation of rankings of objects (Balcerzak, 2016a; 2016b; Kuc, 2017a, 2017b). On the other hand, the classification relates to creating homogenous subsets of objects, where a criterion of similarity between objects, which is based on a distance measure between them, is usually applied (Rollnik-Sadowska and Dąbrowska, 2018).

Numerous classification methods have been developed that allow to select groups of objects that are similar to each other in terms of the analysed phenomenon. The most popular once are hierarchical clustering methods (agglomeration methods), where a hierarchy of objects is created due to their similarity. In the hierarchy there is the lowest level, where all objects are treated as separate. Then, from analysed objects groups are created, the number of which decreases as a part of transition to a higher level of the hierarchy. This is due to the fact that the number of objects is increased within groups. At the last level of the hierarchy, all objects form one common group. Among these methods, one can distinguish: single linkage method, the furthest neighbour method, the unweighted centroid method, the weighted average method, the average method and the Ward's method. The Ward's method is currently considered as one of the most commonly applied tools in the case of socio-economic research. The Ward's method differs significantly from the other mentioned methods in the approach to the analysis of variance (Ward, 1963). In this method, groups (clusters) are determined on the basis of criterion of minimizing the sum of the squares of distance between objects, which are included in the cluster. The obtained results are presented in the form of a dendrogram, where groups of objects are presented, depending on the distance between the groups (Reiff et al., 2016; Mačerinskienė and Aleknavičiūtė, 2017).

3 EMPIRICAL RESEARCH

The Ward's method was used to establish the hierarchical structure for the phenomenon of entrepreneurial conditions at regional level as part of cluster analysis. The study covered 16 voivodships in Poland (NUTS 2), and the analysis was carried out for the years 2011 and 2017. The selected set of diagnostic variables was used to describe the entrepreneurial conditions at regional level (Table 1). The diagnostic variables were selected based on literature review provided by Rogalska (2017) and mostly limited by availability of data at regional level. Data regarding the diagnostic variables were obtained from the Statics Poland

(Local data bank) (https://bdl.stat.gov.pl/BDL/start). The data was standardised with application of classic standardisation.

Table 1. Diagnostic variables describing entrepreneurial environment at regional level

Variable	Description
X_1	Number of entities included in the REGON registration per 10 thousand inhabitants
X_2	Natural persons conducting economic activity per 1 thousand inhabitants
X_3	Share of new registered companies form creative industry in the total number of new registered commercial law companies
X_4	Capital expenditures in enterprises per capita
X_5	Gross value of fixed assets in enterprises per capita

Source: own work based on Statics Poland (Local data bank) and Rogalska (2017).

On the basis of the obtained results after application the Ward's method the dendograms for the year 2011 and 2017 were created, which represent the hierarchical structure of regions in regard to entrepreneurial conditions in Poland at NUTS 2 level (see Figure 1 and 2). The conducted analysis enabled to identify clusters of regions similar to each other based on the diagnostic variables. For the years under evaluation, the division of regions into 5 groups was proposed. The grouping of NUTS 2 to specific groups is also presented in Table 2.

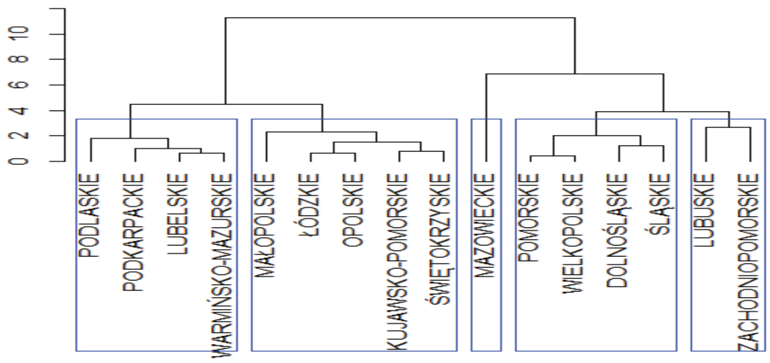


Figure 1. Dendogram based on the diagnostic variables for entrepreneurial conditions for the year 2011
Source: own work based on Statics Poland (Local data bank) (https://bdl.stat.gov.pl/BDL/start.

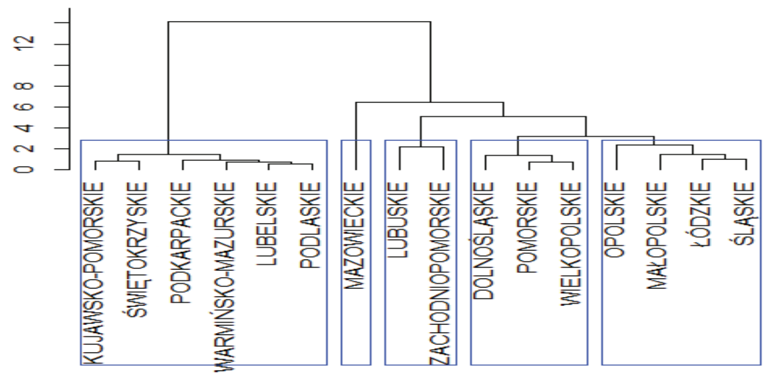


Figure 2. Dendogram based on the diagnostic variables for entrepreneurial conditions for the year 2017
Source: own work based on Statics Poland (Local data bank) (https://bdl.stat.gov.pl/BDL/start).

Table 2. Grouping of NUTS 2 regions in the years 2011 and 2017

NUTS 2 Region	2011	NUTS 2 Region	2017
MAZOWIECKIE	1	MAZOWIECKIE	1
WIELKOPOLSKIE	2	WIELKOPOLSKIE	2
DOLNOŚLĄSKIE	2	DOLNOŚLĄSKIE	2
POMORSKIE	2	POMORSKIE	2
ŚLĄSKIE	2	ŚLĄSKIE	4
LUBUSKIE	3	LUBUSKIE	3
ZACHODNIOPOMORSKIE	3	ZACHODNIOPOMORSKIE	3
MAŁOPOLSKIE	4	MAŁOPOLSKIE	4
ŁÓDZKIE	4	ŁÓDZKIE	4
OPOLSKIE	4	OPOLSKIE	4
KUJAWSKO-POMORSKIE	4	KUJAWSKO-POMORSKIE	5
ŚWIĘTOKRZYSKIE	4	ŚWIĘTOKRZYSKIE	5
LUBELSKIE	5	LUBELSKIE	5
PODKARPACKIE	5	PODKARPACKIE	5
PODLASKIE	5	PODLASKIE	5
WARMIŃSKO-MAZURSKIE	5	WARMIŃSKO-MAZURSKIE	5

Source: own work based on Statics Poland (Local data bank) (https://bdl.stat.gov.pl/BDL/start.

Cluster analysis allowed to identify 5 groups of NUTS 2 regions, where for both years the first group was formed only by Mazowieckie voivodship. The voivodship is characterized by the best entrepreneurial conditions, which can be explained by the fact that in the voivodship the capital city of Poland (the most developed metropolis) is located.

The second cluster with a good entrepreneurial conditions is composed of Wielkopolskie, Dolnośląskie, Pomorskie and Śląskie NUTS 2. In these regions there are located the largest urban centres in Poland such as cities of Poznań, Wrocław, Katowice and Gdańsk. It is worth to emphasize that in the year 2017 the Śląskie region is not classified in that group. It could be found in the fourth cluster with regions characterised by relatively worse entrepreneurial conditions. The third and fourth clusters are formed by regions with average entrepreneurial conditions. One should notice relatively big similarities between Lubuskie and Zachodniopomorskie, which form third cluster and Małopolskie, Łódzkie and Opolskie grouping fourth cluster.

The last fifth cluster with the worst entrepreneurial conditions is formed by NUTS 2 from Eastern Poland: Lubelskie, Podkarpackie, Podlasie and Warmińsko-Mazurksie. This is due to the weakest socio-economic situation of the Eastern part of country. In 2017, Świętokrzyskie and Kujawsko-Pomorskie regions were also assigned to this group. It should be emphasized that in this cluster there are also relatively large urban centres, such as Białystok, Kielce, Lublin, Rzeszów, Bydgoszcz and Toruń. However, these municipal centres do not have sufficient potential to create strong growth areas and their impact is only local. In addition, these centres tend to develop more slowly than urban centres belonging to other clusters. Therefore, one can expect that these factors may contribute to the consolidation of relatively weak entrepreneurial conditions for the NUTS 2 regions in the fifth cluster.

The research on the tendencies relating to entrepreneurial conditions at regional level in Poland requires a deeper analysis. However, the conducted study has clearly indicated the spatial dichotomy of this phenomenon in Poland. The obtained results are consistent with previous research relating to ordering to the NUTS 2 regions (Rogalska, 2017). Relatively weak entrepreneurial conditions for the voivodships form the fifth cluster seem to be permanent phenomenon, which is directly related to relatively worse socio-economic situation of Eastern Poland (Pietrzak and Balcerzak, 2016d, 2017). The established disparities confirm that without effective government regional policy differences in the level of socio-economic development between Polish regions can increase in the future.

4 CONCLUSIONS

The current research concentrated on the problem of classification of NUTS 2 regions in Poland in the years 2011 and 2017 in regard to entrepreneurial conditions. The timespan of the analysis was limited by the availability of data at regional level, but it can be still considered as long enough to verify the potential influence of regional policy. In the research the Ward's clustering method was applied.

The conducted research confirmed that in spite of significant funds devoted to regional policy, which should help in reducing the disparities between Polish NUTS 2 regions, the regional structure of the economy has been relatively stable for the analysed period. The research confirmed that structural regional disparities in regard to entrepreneurial conditions should be an object of special interests for Polish government, as in the future it can become a significant growth obstacle for the economy.

References

[1] Ahmed, R. R., Vveinhardt, J. and Streimikiene, D. (2017). Interactive digital media and impact of customer attitude and technology on brand awareness: evidence from the South Asian countries. *Journal of Business Economics and Management*, 18(6), pp. 1115-1134.

[2] Balcerzak, A. P. (2016a). Technological Potential of European Economy. Proposition of Measurement with Application of Multiple Criteria Decision Analysis. *Montenegrin Journal of Economics*, 12(3), pp. 7-17.

[3] Balcerzak, A. P. (2016b). Multiple-criteria Evaluation of Quality of Human Capital in the European Union Countries. *Economics & Sociology*, 9(2), pp. 11-27.

[4] Balcerzak, A. P., Kliestik, T., Streimikiene, D. and Smrčka L. (2017). Non-parametric approach to measuring the efficiency of banking sectors in European Union Countries. *Acta Polytechnica Hungarica*, 14(7), pp. 51-70.

[5] Balcerzak, A. P. and Pietrzak, M.B. (2016). Quality of Institutions for Knowledge-based Economy within New Institutional Economics Framework. Multiple Criteria Decision Analysis for European Countries in the Years 2000–2013. *Economics & Sociology*, 9(4), pp. 66-81.

[6] Balcerzak, A. P. and Pietrzak, M. B. (2017). Digital Economy in Visegrad Coutnries. Multiple-criteria decision analysis at Regional Level in the Years 2012 and 2015, *Journal of Competitiveness*, 9(2), pp. 5-18.

[7] Bartkowiak-Bakun, N. (2017). The diversity of socioeconomic development of rural areas in Poland in The Western Borderland and the problem of post-state farm localities. *Oeconomia Copernicana*, 8(3), pp. 417-432.

[8] Begonja, M., Čićek, F., Balboni, B. and Gerbin, A. (2016). Innovation and business performance determinants of SMEs in the Adriatic region that introduced social innovation. *Economic Research-Ekonomska Istraživanja*, 29(1), pp. 1136-1149.

[9] Cano, J., Campo, E. and Gómez-Montoya, R. (2017). International Market Selection Using Fuzzy Weighing and Monte Carlo Simulation. *Polish Journal of Management Studies*, 16(2), pp. 40-50.

[10] Cheba, K. and Szopik-Depczyńska, K. (2017). Multidimensional comparative analysis of the competitive capacity of the European Union countries and geographical regions. *Oeconomia Copernicana*, 8(4), pp. 487-504.

[11] Czarniewski, S. (2016). Small and medium-sized enterprises in the context of innovation and entrepreneurship in the economy. *Polish Journal of Management Studies*, 13(1), pp. 30-39.

[12] Dobeš K., Kot S., Kramoliš J. and Sopková G. (2017). The Perception of Governmental Support in The Context of Competitiveness of SMEs in the Czech Republic. *Journal of Competitiveness*, 9(3), pp. 34-50.

[13] Dvorsky, J., Sopkova, G. and Janoskova, M. (2017). Evaluation of the Social Environment and Access to Financial Resources for Business: Case study of the Czech and Slovak Republic. *Ekonomicko-manazerske Spectrum*, 11(1), pp. 62-73.

[14] Erkut, B. (2016). Entrepreneurship and Economic Freedom: Do Objective and Subjective Data Reflect the Same Tendencies? *Entrepreneurial Business and Economics Review*, 4(3), pp. 11-26.

[15] Fabuš, M. (2017). Current development of business environment in Slovakia and Czech Republic. *Entrepreneurship and Sustainability Issues,* 5(1), pp. 127-137.

[16] Keshavarz Ghorabaee, M., Zavadskas, E.K., Amiri, M. and Antucheviciene, J. (2016). Evaluation by an Area-based Method of Ranking Interval Type-2 Fuzzy Sets (EAMRIT-2F) for Multi-criteria Group Decision-making. *Transformations in Business & Economics*, 15(3)(39), pp. 76-95.

[17] Keshavarz Ghorabaee, M., Amiri, M., Kazimieras Zavadskas, E. and Antucheviciene, J. (2017). Supplier evaluation and selection in fuzzy environments: a review of MADM approaches. *Economic Research-Ekonomska Istraživanja*, 30(1), pp. 1073-1118.

[18] Kljucnikov, A. and Belas, J. (2016). Approaches of Czech entrepreneurs to debt financing and management of credit risk. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 11(2), pp. 343-365.

[19] Kruk, H. and Waśniewska, A. (2017). Application of the Perkal method for assessing competitiveness of the countries of Central and Eastern Europe. *Oeconomia Copernicana*, 8(3), pp. 337-352.

[20] Kuc, M. (2017a). Social convergence in Nordic countries at regional level. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(1), pp. 25-41.

[21] Kuc M. (2017b). Is the Regional Divergence a Price for the International Convergence. The Case of the Visegrad Group. *Journal of Competitiveness*, 9(4), pp. 50-65.

[22] Kozubíková L., Homolka L. and Kristalas. (2017). The Effect of Business Environment and Entrepreneurs' Gender on Perception of Financial Risk in The SMEs Sector. *Journal of Competitiveness*, 9(1), pp. 36-50.

[23] Mačerinskienė, I. and Aleknavičiūtė, R. (2017). National intellectual capital influence on economic growth in the European Union countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(4), pp. 573-592.

[24] Mardani, A., Jusoh, A., Zavadskas, E.K., Kazemilari, M., Ungku, N.U.A. and Khalifah, Z. (2016). Application of Multiple Criteria Decision Making Techniques in Tourism and Hospitality Industry: a Systematic Review. *Transformations in Business & Economics*, 15(1)(37), pp. 192-213.

[25] Meluzin, T. , Pietrzak, M. B., Balcerzak, A. P., Zinecker, M., Doubravský, K. and Dohnal, M. (2017). Rumours Related to Political Instability and their Impact on IPOs: The Use of Qualitative Modeling with Incomplete Knowledge. *Polish Journal of Management Studies*, 16(2), pp. 171-187.

[26] Meluzin, T., Zinecker, M., Balcerzak, A. P., Doubravský, K., Pietrzak, M. B. and Dohnal, M. (2018), The timing of initial public offerings – non-numerical model based on qualitative trends. *Journal of Business Economics and Management,* 19(1), pp. 109-125.

[27] Petrenko, E., Iskakov, N., Metsyk, O. and Khassanova, T. (2017). Ecosystem of entrepreneurship: risks related to loss of trust in stability of economic environment in Kazakhstan. *Entrepreneurship and Sustainability Issues,* 5(1), pp. 105-115.

[28] Pietrzak, M. B. and Balcerzak, A. P. (2016a). A Spatial SAR Model in Evaluating Influence of Entrepreneurship and Investments on Unemployment in Poland. In M. Reiff and P. Gezik (Eds.). *Proceedings of the International Scientific Conference Quantitative Methods in Economics Multiple Criteria Decision Making XVIII*. Vratna: Letra Interactive, pp. 303-308.

[29] Pietrzak, M. B. and Balcerzak, A. P. (2016b). Assessment of Socio-Economic Sustainability in New European Union Members States in the years 2004-2012. In: M. Papież and S. Śmiech (Eds.). *The 10th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena*. *Conference Proceedings*. Cracow: Foundation of the Cracow University of Economics, pp. 120-129.

[30] Pietrzak, M. B. and Balcerzak, A. P (2016c). Quality of Human Capital and Total Factor Productivity in New EU Member States. In: T. Loster & T. Pavelka (Eds.). *The 10th International Days of Statistics and Economics. Conference Proceedings. September 8-10*, *2016*. Prague: Libuse Macakova, Melandrium, pp. 1492-1501.

[31] Pietrzak. M. B. and Balcerzak, A. P. (2016d). Socio-economic Sustainability in Poland. SEM Analysis at Regional Level. In: T. Kliestik (Ed.). *16th International Scientific Conference Globalization and Its Socio-Economic Consequences. University of Zilina, The Faculty of Operation and Economics of Transport and Communication, Department of Economics, 5th – 6th October 2016*. *(Part IV.).* Zilina: University of Zilina, pp. 1704-1712.

[32] Pietrzak, M. B and Balcerzak, A. B. (2017). Economic development of Polish voivodeships in the years 2010-2014. Application of taxonomic measure of development with entropy weights. In M. Papież and S. Śmiech (Eds.), *The 11th Professor Aleksander Zelias International Conference on Modelling and Forecasting of*

Socio-Economic Phenomena. Conference Proceedings. Cracow: Foundation of the Cracow University of Economics, pp. 310-318.

[33] Pietrzak, M. B., Balcerzak, A. P., Gajdos, A. and Arendt, Ł (2017). Entrepreneurial environment at regional level: the case of Polish path towards sustainable socio-economic development. *Entrepreneurship and Sustainability Issues,* 5(2), pp. 190-203.

[34] Reiff, M., Surmanová, K., Balcerzak, A.P. and Pietrzak, M.B. (2016). Multiple Criteria Analysis of European Union Agriculture. *Journal of International Studies*, 9(3), pp. 62-74.

[35] Rogalska, E. (2017). Multivariate analysis of entrepreneurship in Poland at regional level. In: T. Kliestik (Ed.). 17th International Scientific Conference Globalization and Its Socio-Economic Consequences. University of Zilina, The Faculty of Operation and Economics of Transport and Communication, Department of Economics, 4th – 5th October 2017. (Part V.). Zilina: University of Zilina, pp. 2198-2202.

[36] Rollnik-Sadowska, E., Dąbrowska, D. (2018). Cluster analysis of effectiveness of labour market policy in the European Union. *Oeconomia Copernicana*, 9(1), pp. 143-158.

[37] Simionescu, M., Lazányi, K., Sopková, G., Dobeš, K. and Balcerzak, A. P. (2017). Determinants of Economic Growth in V4 Countries and Romania. *Journal of Competitiveness*, 9(1), pp. 103-113.

[38] Stanickova, M. (2017). Can the implementation of the Europe 2020 Strategy goals be efficient? The challenge for achieving social equality in the European Union. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(3), pp. 383-398.

[39] Stankeviciene, J., Kraujaliene, L. and Vaiciukeviciute, A. (2017). Assessment of technology transfer office performance for value creation in higher education institutions. *Journal of Business Economics and Management*, *18*(6), pp. 1063-1081.

[40] Tomovska Misoska, A., Dimitrova, M. and Mrsik, J. (2016). Drivers of entrepreneurial intentions among business students in Macedonia. *Economic Research-Ekonomska Istraživanja*, 29(1), pp. 1062-1074.

[41] Vinczeova, M. and Kascakova, A. (2017). How do Slovak small and medium-sized enterprises make decision on sources of finance? *Ekonomicko-manazerske spektrum*, 11(2), pp. 111-121.

[42] Ward, J. H., Jr. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58, pp. 236-244.

[43] Zemlickienė, V., Bublienė, R. and Jakubavičius, A., (2018). A model for assessing the commercial potential of high technologies. *Oeconomia Copernicana*, 9(1), pp. 29-54.

[44] Zygmunt, J. (2018). Entrepreneurial activity drivers in the transition economies. Evidence from the Visegrad countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 13(1), 89-103.

Author's address

Elżbieta Rogalska, PhD.
University of Wamia and Mazury, Department of Microeconomics, ul. Plac Cieszyński 1, 10-719 Olsztyn, Poland
Email: ela.n.rogalska@gmail.com

PRICE POLICY IN GAMES IN SPATIAL COMPETITION

Allan Jose Sequeira Lopez, Zuzana Čičková, University of Economics in Bratislava

Abstract
Spatial competition in game theory studies situations where two or more individuals or firms compete to attract customers, making decisions regarding the location in a region with a certain spatial structure. This paper is focused on price competition in spatial game and analysis of situation when each company has its own pricing strategy and the customer´s preferences are limited by the maximum amount they are willing to pay for the goods at given price. The price comprises unit price and transport costs and the spatial price discrimination will be applied. The paper presents the formulation of the mentioned situation as a two-player game with a constant and non-constant sum. This approach is illustrated in the form of illustrative examples.

***Keywords*:** spatial competition, pricing policy, game theory, localization

JEL classification: C72, C61, D19, D21, D43, D61, D58
AMS Classification: 91A24, 91A70, 91A10.

INTRODUCTION

The game theory is a discipline of applied mathematics focused on decision making situations involving more than one entity (individuals, companies, states, political parties, etc.) with the assumptions and rules of the conflict (game) being precisely defined.

Our paper is focused on formulating and solving a specific game in the field of spatial games. Spatial games are part of the analysis of spatial competition issues. The concept of spatial competition can be identified with a competitive position. Such games generally include localization decisions, pricing decisions and production levels for a number of competing businesses that provide goods or services to a group of consumers or users. Network competition is basically concerned with interdependence between economic entities in an environment of imperfect competition.

Location issues have been investigated for centuries, but only after the post-WWII development of operational research these issues began to be considered relevant. Already in 1924, Fetter formulated the Law of Market Areas (Fetter, 1924) where he analyzed how consumers evaluate purchase prices and transport costs for each of the manufacturing companies before deciding to buy the goods from one of them. As a result, the market share threshold between these companies is defined by the location of those consumers for whom the total purchase price is indifferent to one of the companies or another. This work had an important impact on work on the theory of network competition, but was not as important as the subsequent work of Hotelling (1929).

The beginnings of the spatial competition analysis date back to Cournot's work (Cournot, 1838), who developed a model with companies competing by the quantities offered, and Bertrand (1883), who based its approach on a model with companies competing by prices. The works of Cournot and Bertrand were further followed by the American economist and statistician Harold Hotelling (1895–1973) who laid the foundation for the mainstream of spatial competition research in his well known article Stability in Competition in 1929 (Hotelling, 1929). Hotelling modelled a linear market in which competing companies sell a homogeneous product where consumers are evenly distributed in the space and demand is inelastic which means that each consumer gets just one unit of goods. The product is homogeneous in both companies and therefore, in the first stage, they differ only in placement and in the second stage in the price. Further expansions of this approach can be found in (Gabszwicz and Thisse, 1986), (Arguedas and Hamid, 2008), and (Hamid et al., 2011).

Our paper is focused on a specific spatial game that takes into account the two sellers' price differentials. The aim is to point out the difference in the formulation depending on whether or not we consider lost or missing demand. Thus the game can be formulated as a matrix or bimatrix game. We also illustrate the formulations by examples.

1 TWO-PLAYER GAMES WITH ZERO AND NON-ZERO TOTALS

The aim of the paper is to analyze the possibilities of network game formulation as a two-player game with a constant and non-constant sum. A two-player game is a game of two participants (players), each player chooses independently (without information about the choice of an opponent) one of the finite number of behavioural scenarios (strategies). It is assumed that players are intelligent and want to achieve the best result. In games with a constant sum (matrix games) we assume that the interests of the players are diametrically opposed, on the other hand, in games with non-constant sum (bimatrix games) it is not the case. The question is which strategy a player has to choose so that choosing another strategy can not improve his result (Goga, 2013) (Chobot et al., 1991).

Two-player games can be formalized as follows: Let $P = \{1,2\}$ be a set of players where each player has a finite set of strategies (X – player 1, Y – player 2}, i.e. player 1 chooses $\mathbf{x} \in X$, player 2 chooses $\mathbf{y} \in Y$. A set of all results of the game can be defined as $(x, y) \in X\text{x}Y$. The individual elements of the sets X and Y can be arranged by a finite number of non-negative numbers (elements of the set $X\colon i = 1,2, \dots m$ and elements of the set $Y\colon j = 1,2, \dots n$). The values of the game for player 1 can be indicated in the matrix $\mathbf{A}_{mxn} = \{a_{ij}\}$, where a_{ij} indicates the player's payment at the result (i, j), i.e. if $a_{ij} > 0$, the player collects and if $a_{ij} < 0$, the player pays. The values of the game for player 2 can be indicated in the matrix $\mathbf{B}_{nxm} = \{b_{ji}\}$, where b_{ji} indicates the player's payment at the result (i, j). The constant sum can be characterized as follows: $\mathbf{B}_{nxm} = \mathbf{C}_{nxm} - \mathbf{A}^{\mathrm{T}}_{mxn}$, where $\mathbf{C}_{nxm} = \{c\}$ with c being the constant independent of the strategy choice.

The solution to the game is to identify equilibrium strategies for both players. The status is defined as equilibrium, if the system has a tendency to remain in such state under certain conditions (only such set of strategies can be considered as a satisfactory result, if any effort to unilaterally violation automatically leads to damage to a player attempting to do so).

Game solutions are based on the following assumptions: Both players have complete information about the model of the conflict situation, i.e. they know the payment matrix $\mathbf{A}_{mxn} = \{a_{ij}\}$, players are intelligent, i.e. the players want to maximize the payout and know that so does the opponent, the players are careful, i.e. they try to minimize the risk. The solution to the game lies in identifying the equilibrium point in pure strategies (Chobot, et al., 1991). (Goga, 2013), (Dlouhý and Fiala, 2007), if such a solution does not exist, we are looking for a solution in mixed strategies. The solution then consists of assuming a repetition of the game and determining the probability distribution of strategies (mixed strategies).

The mixed strategy of player 1 is the the m-dimensional vector $x, \sum_{i=1}^{m} x_i = 1, x_i \geq 0, i = 1,2, \dots m$ and the mixed strategy of player 2 is the the n-dimensional vector $y, \sum_{j=1}^{n} y_j = 1, y_j \geq 0, j = 1,2, \dots n$. Mixed strategies can then be identified using mathematical models. In case of games with a constant sum these are simple tasks of linear programming (Chobot et al., 1991). In case of non-constant sum games, the solution is more complex because it is a non-linear programming task. The solution lies in the identification of the equilibrium points satisfying Karush-Kuhn-Tucker conditions (Čičková and Zagiba, 2018).

2 SPATIAL COMPETITION GAME AS A TWO-PLAYER GAME

In this section we introduce a specific spatial competition game formulation as well as a solution of illustrative examples. We consider a duopoly market type. The proposed game is based on the following assumptions: Let $N = \{1,2, \dots n\}$ is a set of customers and let be given a finite continuous oriented edgewise-rated graph $G = (V, H)$, where V represents a non-empty n-element set of graph nodes, and $H \subset VxV$ represents a set of edges $h_{ij} = (v_i, v_j)$ from the node v_i to the node v_j, with each oriented edge h_{ij} being assigned a real number $o(h_{ij})$ called a valuation or also the edge value of h_{ij}. The network game is formulated in a so-called complete or a complete weighted graph $\bar{G} = (V, \bar{H})$ with the same set of nodes as graph G, where $\bar{H}$ is the set of edges between each pair of nodes v_i and v_j, their valuation being equal to the minimum distance between the nodes v_i and v_j in the original graph $i, j \in N$. If d_{ij} represents the minimum distance (the shortest path length) between nodes v_i and v_j, then the matrix $\mathbf{D}_{nxn} = \{d_{ij}\}$ is the shortest distance matrix.

Let´s assume two companies (players), $P = \{1,2\}$ offering a homogeneous product (goods or service) that have the ability to build their operations in one of the nodes that are simultaneously the seat of the customers, i.e. in any of the elements of the set $N = \{1,2, \dots n\}$. The demand of each node is constant. Although both players offer an identical product in an unlimited amount, the product price is different. Let us denote $p^{(1)}$ the product price for player 1 and $p^{(2)}$ the product price for player 2. We do not consider any capacity limitations; every customer can buy a product at any company. However, the customer only considers buying one unit of goods (services).

In the first case, we do not consider the "lost" demand, i.e. the customer always makes the purchase. Customers, however, take into consideration the total price of the product consisting of both the purchase price of the product and the price of the transport to a chosen company. Transport costs are rated as t/unit of distance. Then for player 1 the cost matrix can be defined as $\mathbf{N}^{(1)} = \{n_{ij}^{(1)}\}, i, j = 1,2, \dots n$ with elements are defined as follows:

$$n_{ij}^{(1)} = t * d_{ij} + p^{(1)}, i, j = 1, 2 \dots n \tag{1}$$

Analogical for the player 2 we specify the matrix $\mathbf{N}^{(2)} = \left\{n_{ij}^{(2)}\right\}, i, j = 1,2, \dots n$:

$$n_{ij}^{(2)} = t * d_{ij} + p^{(2)}, i, j = 1, 2 \dots n \tag{2}$$

Factors (1) and (2) take into account both the total transport costs of the transport from the i −th customer to the respective company as well as the purchase price of the product.

Furthermore, we assume the following: If a player 1 builds a service place in the j −th node, he will get the customer from the i −th node just if $n_{ij}^{(1)} < n_{ij}^{(2)}$, $i, j = 1, 2 \dots n$, if $n_{ij}^{(1)} = n_{ij}^{(2)}$ the players share the demand equally, otherwise the customer from the i −th node is served by the player

2. Then it is possible to define the elements of the payment matrix of player 1 (**A**) in the form of the following pseudo code:

LOOP (k,i,j = 1,2,...n) **DO**
IF $n_{ij}^{(1)} < n_{ij}^{(2)}$ **DO** $a_{ij} = a_{ij} + 1$; (3)
ELSEIF $n_{ij}^{(1)} = n_{ij}^{(2)}$ **DO** $a_{ij} = a_{ij} + 0,5$;
ENDIF

Obviously, under the given assumptions, it is possible to formulate a game with a constant sum, where the constant of the game $c = n$ (since players share a constant demand of n nodes). The solution to the game in pure strategies is to identify the saddle point of matrix **A**. If such a point does not exist, equilibrium mixed strategies can be determined by solving the problem of linear programming (Chobot et al., 1991).

In the second case, we also consider the "lost" demand, i.e. customers have a fixed price that they are willing to pay for the product. This maximum price includes both the transport costs and the purchase price. Let´s denote this price as T_{max}. Then, the matrix of player 1 payouts ($\mathbf{A} = \{a_{ij}\}, i, j = 1,2, \ldots n$) can be expressed in the form of the following pseudocode:

LOOP (k,i,j = 1,2,...n) **DO**
IF $n_{ij}^{(1)} < n_{ij}^{(2)}$ **and** $n_{ij}^{(1)} \leq T_{MAX}$ **DO** $a_{ij} = a_{ij} + 1$; (4)
ELSEIF $n_{ij}^{(1)} = n_{ij}^{(2)}$ **and** $n_{ij}^{(1)} \leq T_{MAX}$ **DO** $a_{ij} = a_{ij} + 0,5$;
ENDIF

Then it is also possible to quantify the payout matrix of player 2 ($\mathbf{B} = \{b_{ij}\}, i, j = 1,2, \ldots n$) as follows:

LOOP (k,i,j = 1,2,...n) **DO**
IF $n_{ij}^{(2)} < n_{ij}^{(1)}$ **and** $n_{ij}^{(2)} \leq T_{MAX}$ **DO** $b_{ij} = b_{ij} + 1$; (5)
ELSEIF $n_{ij}^{(2)} = n_{ij}^{(1)}$ **and** $n_{ij}^{(2)} \leq T_{MAX}$ **DO** $b_{ij} = b_{ij} + 0,5$;
ENDIF

Obviously, this is now a game with an non-constant sum. If there is no solution of the game in pure strategies, we will determine equilibrium mixed strategies of the players based on a nonlinear model (Čičková and Zagiba, 2018). In the next section, we will illustrate this approach by solving illustrative examples.

Examples

Let's assume the existence of five potential customers $N = \{1, 2, \ldots 5\}$, each of them is located in the unique node of a graph G. At the same time, we assume the form of a duopolistic market where each of the companies can build a branch office in any node of this graph $i = 1,2, \ldots 5$. Each company represents a player, where each player maximizes the number of nodes that are served. Although both players offer a homogeneous product, its price is $p^{(1)}$= 1 for player 1 and $p^{(2)}$= 2,1 for player 2. Let the shortest distance matrix $\mathbf{D} = \{d_{ij}\}, i, j = 1,2, \ldots 5$ between all the nodes of the network be as follows:

$$\mathbf{D} = \begin{bmatrix} 0 & 3 & 4 & 3 & 2 \\ 3 & 0 & 7 & 6 & 5 \\ 4 & 7 & 0 & 3 & 2 \\ 3 & 6 & 3 & 0 & 1 \\ 2 & 5 & 2 & 1 & 0 \end{bmatrix}$$

We assume unit transport costs $t = 1$.

Based on (1) and (2) it is possible to quantify matrix elements $\mathbf{N}^{(1)}$ and $\mathbf{N}^{(2)}$ as follows:

$$\mathbf{N}^{(1)} = \begin{bmatrix} 1 & 4 & 5 & 4 & 3 \\ 4 & 1 & 8 & 7 & 6 \\ 5 & 8 & 1 & 4 & 3 \\ 4 & 7 & 4 & 1 & 2 \\ 3 & 6 & 3 & 2 & 1 \end{bmatrix} \text{ and } \mathbf{N}^{(2)} = \begin{bmatrix} 2.1 & 5.1 & 6.1 & 5.1 & 4.1 \\ 5.1 & 2.1 & 9.1 & 8.1 & 7.1 \\ 6.1 & 9.1 & 2.1 & 5.1 & 4.1 \\ 5.1 & 8.1 & 5.1 & 2.1 & 3.1 \\ 4.1 & 7.1 & 4.1 & 3.1 & 2.1 \end{bmatrix}$$

If we do not consider lost demand, it is possible to quantify the elements of matrix **A** for player 1 as follows:

$$\mathbf{A} = \begin{bmatrix} 5 & 4 & 4 & 4 & 2 \\ 1 & 5 & 2 & 2 & 2 \\ 3 & 4 & 5 & 4 & 1 \\ 3 & 4 & 4 & 5 & 5 \\ 3 & 4 & 4 & 5 & 5 \end{bmatrix}$$

Since matrix **A** does not have a saddle point, we are looking for a solution to the game in mixed strategies. The game is formulated as a linear programming task and solved with GAMS (solver Gurobi 3.0.1). Obviously, the constant of this game is $c = 5$. The mixed equilibrium strategy of player 1 is represented by the vector $\mathbf{x}^{(0)} = (0,4;0;0;0.6;0)^T$ i.e. player 1 should invest 40 % of the funds in node 1 and 60 % in branch office in node 4. Mixed strategy of the player 2 represent dual $\mathbf{y}^{(0)} = (0,6;0;0;0;0,4)^T$ i.e. player 2 should invest 60 % of his capital in node 1 and a 40 % in branch office in node 5. The average value of the game would be 3,8 for player 1 (3,8 serviced node) and average value of the game for player 2 will be $c - 3{,}8 = 5 - 3{,}8 = 1{,}2$.

Let´s set the maximum purchase price of the product. Let´s $T_{max} = 4$. In such case, if the total price (transport cost and product price) exceeds 4 MU, the customer does not buy the product. Under these conditions, the elements of matrix **A** (for player 1, based on 4) as well as elements of matrix **B** (for player 2, based on 5) must be determined:

$$\mathbf{A} = \begin{bmatrix} 4 & 3 & 4 & 3 & 2 \\ 1 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 2 & 1 \\ 3 & 4 & 3 & 4 & 4 \\ 3 & 4 & 3 & 4 & 4 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 2 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \end{bmatrix}$$

Again, the game does not have a solution in pure strategies, therefore we are looking for a solution in mixed strategies (Čičková and Zagiba, 2018). We used GAMS (solver Minos 5.51). The solution of the game is as follows: The mixed equilibrium strategy of player 1 is represented by the vector $\mathbf{x}^{(0)} = (0;0;0;1;0)^T$ i.e. player 1 should invest all of the funds in node 3. Mixed strategy of the player 2 is represented by vector $\mathbf{y}^{(0)} = (0,5;0,5;0;0;0)^T$ i.e. player 2 should invest 50 % of his capital in node 1 and 50 % in branch office in node 2. The value of the game would be 3,5 for player 1, and for player 2 the value of the game is 1.

3 CONCLUSION

This contribution is focused on a specific problem of spatial competition. The problems of spatial competition are characterized by conflicting character; therefore, the Game theory can be used to solve them. The problem is formulated for duopoly (on the supply side). The issues are analyzed in the transport network with individual buyers located in the individual nodes of such network. The sellers decide on their position while trying to respect the behaviour of buyers who minimize both the costs associated with the transport price and the transport costs. In our paper we formulate a game considering the "lost" demand as well as a game where we do not consider the "lost" demand. Both types of games were also illustrated by the formulation of the example. The GAMS professional software, which ranks among the powerful optimization computing environments, was used to solve the games mentioned above.

References

[1] Arguedas, C. and Hamid, H. (2008). A note on product differentiation under concave cost. Madrid: Universidad Europea de Madrid.

[2] Bertrand, J. (1883). Book review of theorie Mathematique de la richesse sociale and of recherches sur les principles mathematiques de la theorie des richesses. Andrew F. Daughety ed. Cambridge : Cambridge University Press.

[3] Cournot, A. A. (1838). Recherches sur les principes Mathématiques de la théorie des richesses (Researches on the Mathematical Principles of the Theory of Wealth). Paris: L´Universite Royale de France.

[4] Čičková, Z. and Zagiba, M. (2018). Podmienky Optimálnosti Karusha-Kuhna Tuckera a Bimaticové Hry. Bratislava: Ekonóm.

[5] Dlouhý, M. and Fiala, P. (2007). Úvod do teórie hier. Praha: Oeconomica.

[6] Fetter, F. A. (1924). The Economic Law of Market Areas. The Quarterly Journal of Economics, 38(3), pp. 520-529.

[7] Gabszwicz, J. and Thisse, J.F. (1986). On the Nature of Competition with Differentiated Products. The Economic Journal, 96(381), pp. 160-172.

[8] Goga, M. (2013). Teória Hier. Bratislava: Iura Edition, spol. s r. o..

[9] Hamid, H., Bustamante, M. and Sanz, M. (2011). Revisiting price Equilibrium Existence in the Linear-city Model of Spatial Competition. Regional Science, 90(1), pp. 179-196.

[10] Hotelling, H. (1929). Stability in Competition. The Economic Journal, pp. 41-57.

[11] Chobot, M., Turnovec, F. and Ulašín, V. (1991). Teória hier a rozhodovania. Bratislava: Vydavateľstvo Alfa.

Acknowledgements: This work was supported by the project VEGA 1/0351/17 Application of Selected Models of Game Theory in Solving Some Economic Problems of Slovakia.

Author’s address
Ing. Allan Jose Sequeira Lopez.
University of Economics in Bratislava, Faculty of business informatics, Department of Operations research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
email: allan.lopez@euba.sk

doc. Ing. Zuzana Čičková, PhD.
University of Economics in Bratislava, Faculty of business informatics, Department of Operations research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
email: cickova@euba.sk

HIDE AND SEEK – AN APPLICATION OF VEHICLE ROUTING PROBLEM IN GEOCACHING

Veronika Skocdopolova, University of Economics, Prague,
Tomas Frolik, University of Economics, Prague

Abstract
Geocaching is a widespread outdoor leisure activity, in which a Global Positioning System plays its role. This paper presents a model, that can help with planning routes of geocachers, because most of the players try to create a route, where they search for caches, as short as possible according to some structure of found geocaches. The model is based on vehicle routing problem and its result is a plan for few days with respect to specific requirements of types of geocaches. First chapter brings short introduction about the geocaching game, and briefly describes the theory that is used in the second part, where the model is presented.

Keywords: *travelling salesman problem, vehicle routing problem, geocaching*

JEL Classification: C61
AMS Classification: 90B06, 90C09, 90C90

1 INTRODUCTION

Operation research models can be used not only in business activities where the cost should be minimized, or the profit maximized. We can use them also in leisure time activities. This paper presents a utilization of modified vehicle routing problem in geocaching.

1.1 Geocaching

Geocaching is a widespread outdoor leisure activity based on Global Positioning System (GPS). The game was created in the USA right after the 1st May 2000 when the US president Bill Clinton announced that the "Selective Availability" of GPS would be discontinued as a result of the 1996 executive order. This fact allowed civilian users of GPS to receive a non-degraded signal globally. First caches appeared in the US, followed by New Zealand and Chile. First cache in Europe was hidden in Ireland in June 2000.

Geocache (or shortly cache) is traditionally a waterproof container with a log book and pencil. This container is hidden. Coordinates of the container and other details of the cache and its location are posted on a website (e.g. geocaching.com). The players of the game (geocachers or just cachers) obtain the coordinates form the website and seek out the cache using their GPS receiver. (Geocaching, 2018) (GeoWiki, 2018)

1.2 Vehicle Routing Problem

The Vehicle Routing Problem (VRP) is a NP-hard problem (Cerny, 2012). It is a modification of Travelling Salesman Problem (TSP). The objective of the TSP is to find the shortest cycle in the given graph that includes each node of the graph once only (the salesman has to visit all his customers and then go back home). Such a cycle is called Hamiltonian cycle. The VRP designs several routes from a single depot to a number of customers. The basic constraints of VRP are that each node (customer) is visited exactly once by exactly one vehicle and that each route starts and ends at the same depot. The objective is to design optimal delivery routes. The most common additional constraints of the VRP are linked to vehicle capacity restrictions, total time restrictions, time windows or precedence relations (Watson et al, 2014). If there are different vehicles available with different capacity and cost, the heterogeneous fleet vehicle routing problem is considered. The basic VRP model is defined in (Eiselt and Sandblom, 2000).

2 PROBLEM DESCRIPTION AND MODEL FORMULATION

The problem this paper deals with is to plan a trip for a geocacher. The geocacher wants to spend few days with caches seeking. At first the cacher has to select some region and possible caches that he or she is interested in. Each cache has its specific code, terrain and difficulty ranking and also favorite points. The cacher has to specify e.g. how many days he or she will spend with geocaching (how many routes he needs to plan), how many kilometers is able to walk (or ride a bike) each day, how many caches he or she would like to find each day and in total. According to terrain and difficulty ranking the cacher can estimate how much time he or she will need for seeking the cache out.

The cacher may want to find a special type of cache called multicache. Multicache (or puzzle cache) means that the cacher has to find several usual caches which help him or her to find the final cache. In this case, the cacher has to specify the multicache and list all its predecessors. Once we have all the necessary input data we can formulate the following mathematical model.

The model operates with following variables and parameters:

- x_{ijk} – binary variable that is equal 1 when the route goes from the cache i to the cache j on the day k, and 0 otherwise; $i, j = 1, 2, \ldots, n,$ where n is number of caches plus the starting point;
- y_{ik} – binary variable that is equal 1 when the cache i is visited on the day k, and 0 otherwise;
- u_{ik} – secondary variable that represents the order of visited caches i on the day k;
- D_{ij} – distance between caches i and j;
- V – maximal amount of caches in one route (one day);
- $DAYM$ – daily minimum – minimal amount of caches found in one day;
- TM – minimal amount of caches found during the whole trip (all days);
- FP_i – favorite points of the cache i, it is a kind of evaluation of cache interest;
- FPM – minimal amount of favorite points of caches found during the whole trip (all days);
- DIF_i – difficulty rating of the cache i;
- $DIFM$ – minimal average difficulty rating of caches found during the whole trip (all days);
- TER_i – terrain rating of the cache i;
- $TERM$ – minimal average terrain rating of caches found during the whole trip (all days);
- DM – minimal distance of each route;
- T_{ij} – time needed to travel between caches i and j;
- S_i – time needed to seek the cache i;
- TT – total time – maximal time for each route (day);
- M_i – binary value that is equal 1 if the cache i is a mystery cache, and 0 otherwise;
- MM – minimal amount of mystery caches found during the whole trip (all days);
- WF_i – binary value that is equal 1 if the cache i is worth finding, and 0 otherwise;
- WFM – minimal amount of worth finding caches found during the whole trip (all days);

The objective of the model is to minimize the total distance for the trip:

$$z = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{K} D_{ij}\ x_{ijk} \tag{1}$$

subject to several constraints.

If the cache i is visited on the day k, (y_{ik} = 1) there has to be one route to the place i on the day k:

$$\sum_{j=1}^{n} x_{ijk} = y_{ik} \qquad\qquad i = 2, 3, \ldots, n,\ \forall k. \tag{2}$$

If there is a route to the place i, there has to be route outside this place:

$$\sum_{i=1}^{n} x_{ijk} = \sum_{i=1}^{n} x_{ijk} \qquad\qquad \forall j, \forall k. \tag{3}$$

Each cache can be visited at most once during the trip:

$$\sum_{k=1}^{K} y_{ik} \leq 1 \qquad\qquad i = 2, 3, \ldots, n, \tag{4}$$

All the routes have to start and end in the starting point. The following constraints (5)–(7) are so-called sub-tour breaking constraints, where the variable u_{ik} represents the order of visited caches:

$$u_{ik} + 1 - 2V(1 - x_{ijk}) \leq u_{jk} \qquad\qquad \forall i, \forall k, \text{j}= 2, 3, \ldots, n, \tag{5}$$

$$u_{ik} \leq V \qquad\qquad i = 2, 3, \ldots, n,\ \forall k, \tag{6}$$

$$u_{1k} = 0 \qquad\qquad \forall k. \tag{7}$$

Each day at least *DAYM* caches have to found:

$$\sum_{i=2}^{n} y_{ik} \geq DAYM \qquad\qquad \forall k. \tag{8}$$

During the whole trip at least *TM* caches have to be found:

$$\sum_{i=2}^{n} \sum_{k=1}^{K} y_{ik} \geq TM \tag{9}$$

During the whole trip caches with at least *FPM* favorite points in total have to be found:

$$\sum_{i=2}^{n} \sum_{k=1}^{K} y_{ik} FP_i \geq FPM \tag{10}$$

The average difficulty ranking of all the found caches has to be at least *DIFM*:

$$\sum_{i=2}^{n} \sum_{k=1}^{K} y_{ik} DIF_i \geq DIFM \sum_{i=2}^{n} \sum_{k=1}^{K} y_{ik} \tag{11}$$

The average terrain ranking of all the found caches has to be at least *TERM*:

$$\sum_{i=2}^{n} \sum_{k=1}^{K} y_{ik} TER_i \geq TERM \sum_{i=2}^{n} \sum_{k=1}^{K} y_{ik} \tag{12}$$

The daily distance that the cacher undertakes each day has to be at least *DM*:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ijk}D_{ij} \geq DM \qquad \forall k. \tag{13}$$

The time the cacher spend travelling among caches and seeking them each day must not exceed *TT*:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_{ijk}T_{ij} + \sum_{i=2}^{n} y_{ik}S_i \leq TT \qquad \forall k. \tag{14}$$

During the whole trip at least *MM* mystery caches have to be found:

$$\sum_{i=2}^{n}\sum_{k=1}^{K} y_{ik}M_i \geq MM \tag{15}$$

There can be some caches that are worth finding though they are further from the route or with a lower difficulty or terrain ranking than is the average. During the whole trip at least *WFM* worth finding caches have to be found:

$$\sum_{i=2}^{n}\sum_{k=1}^{K} y_{ik}WF_i \geq WFM \tag{16}$$

Before a multicache can be found ($i \in multi2$), several usual caches ($i \in multi1$) have to be found. The constraint (17) assures, that if the multicache is among the found caches, all the usual caches have to be found too:

$$\sum_{i\in multi1}\sum_{k=1}^{K} y_{ik} = |multi1| \sum_{i\in multi2}\sum_{k=1}^{K} y_{ik} \tag{17}$$

$$u_{ik} \leq u_{lm} \qquad i \in multi1, k \leq m, l \in multi2. \tag{18}$$

The variables x_{ijk} and y_{ik} are binary variables:

$$x_{ijk} = 0(1) \qquad \forall i, \forall j, \forall k, \tag{19}$$

$$y_{ik} = 0(1) \qquad \forall i, \forall k. \tag{20}$$

3 CONCLUSIONS

The presented model was successfully tested on data of selected caches from the region of Bohemian Paradise. Among the selected caches there was one multicache called *Humprecht from all the world side*. To find such a cache the cacher has to find 4 usual caches, each in the direction of one world side. After that the cacher is able to find the final cache. The data set consisted of 82 different caches (+ the starting point; *i*, *j* = 1, 2, …, 83) and the routes were planned for 5 days (*k* = 1, 2, …, 5). The model was solved on a computer with 8 GB RAM and 2,6 GHz processor using the Gurobi solver in MPL environment. The solving process collapsed after approx. 12 hours without achieving optimal solution. Only feasible solution was found. The gap between the best feasible solution and bound of the objective function was approx. 16 %. As the best feasible solution was found after approx. 6 hours and wasn't improved during next 6 hours, we assumed that the found solution can be close to optimality.

The presented model is a leisure time utilization of vehicle routing problem modified with several quality requirements. The standard VRP model was also extended for planning routes for given number of days.

Acknowledgements: This work was supported by the Faculty of Informatics and Statistics, University of Economics, Prague, grant no. IP400040.

References

[1] Cerny, M. (2012). *Computations, Volume III* (In Czech). Prague: Professional Publishing.
[2] Eiselt, H. A, and Sandblom, C.-L. (2000). *Integer Programming and Network Models.* Berlin: Springer-Verlag.
[3] Watson, M., Lewis, S., Cacioppi, P., and Jayaraman, J. (2014). *Supply Chain Network Design: Applying Optimization and Analytics to the Global Supply Chain*. New Jersey: Pearson.
[4] Geocaching, (2018). *Database of cache information*. [online] Available at: https://www.geocaching.com/play [Accessed 8 Mar. 2018].
[5] GeoWiki, (2018). *Additional information about geocaching*. [online] Available at: http://wiki.geocaching.cz/ [Accessed 8 Mar. 2018].

Author's address

Ing. Veronika Skočdopolová, PhD.
University of Economics, Prague, Faculty of Informatics and Statistics, department of econometrics
W. Churchill sq. 4, 130 67 Prague 3
Czech Republic
email: veronika.skocdopolova@vse.cz

Bc. Tomáš Frolík
University of Economics, Prague, Faculty of Informatics and Statistics, department of econometrics
W. Churchill sq. 4, 130 67 Prague 3
Czech Republic
email: xfrot03@vse.cz

CENTRAL MOMENTS AND RISK-SENSITIVE OPTIMALITY IN MARKOV REWARD CHAINS

Karel Sladký,
Institute of Information Theory and Automation of the AS CR

Abstract There is no doubt that usual optimization criteria examined in the literature on optimization of Markov reward processes, e.g. total discounted or mean reward, may be quite insufficient to characterize the problem from the point of the decision maker. To this end it is necessary to select more sophisticated criteria that reflect also the variability-risk features of the problem (cf. Cavazos-Cadena and Fernandez-Gaucherand (1999), Cavazos-Cadena and Hernández-Hernández (2005), Howard and Matheson (1972), Jaquette (1976), Kawai (1987), Mandl (1971), Sladký (2005),(2008),(2013), van Dijk and Sladký (2006), White (1988)). In the present paper we consider unichain Markov reward processes with finite state spaces and assume that the generated reward is evaluated by an exponential utility function. Using the Taylor expansion we present explicit formulae for calculating variance and higher central moments of the total reward generated by the Markov reward chain along with its asymptotic behavior and the growth rates if the considered time horizon tends to infinity.

Keywords Discrete-time Markov reward chains, exponential utility, moment generating functions, formulae for central moments.

JEL Classification C44, C61
AMS Classification 90C40

1 Formulation and Notation

In this note, we consider at discrete time points Markov reward process $X = \{X_n, n = 0, 1, \ldots\}$ with finite state space $\mathcal{I} = \{1, 2, \ldots, N\}$, matrix of transition probabilities $P = [p_{ij}]$ and transition reward matrix $R = [r_{ij}]$, i.e. reward r_{ij} is accrued to a transition from state i to state j. The symbol E_i denotes the expectation if $X_0 = i$; $\mathrm{P}(X_m = j)$ is the probability that X is in state j at time m. Moreover, I denotes an identity matrix and e is reserved for a unit column vector.

Recall that $P^* := \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} P^k$ (with elements p^*_{ij}) exists, and if P is aperiodic then even $P^* = \lim_{k\to\infty} P^k$ and the convergence is geometrical. Moreover, if P is unichain, i.e. P contains a single class of recurrent states, then $p^*_{ij} = p^*_j$, i.e. limiting distribution is independent of the starting state (see e.g. Puterman (1994), Ross (1983)).

In what follows, the reward generated by the Markov chain X is evaluated by an exponential utility function, say $u^\gamma(\cdot)$, i.e. utility function with constant risk sensitivity $\gamma \in \mathbb{R}$, where

$$u^\gamma(\xi) := \begin{cases} \operatorname{sign}(\gamma)\, \exp(\gamma\xi) & \text{if } \gamma \neq 0 \\ \xi & \text{for } \gamma = 0. \end{cases} \tag{1}$$

Obviously, $u^\gamma(\cdot)$ is strictly increasing and convex, if $\gamma > 0$, for $\gamma = 0$ (the risk neutral case) $u^\gamma(\xi) = \xi$ is linear, if $\gamma < 0$ then $u^\gamma(\cdot)$ is concave (see e.g. Howard and Matheson (1972), Jaquette (1976), Sladký (2008),(2013)).

Observe that $U^{(\gamma)}(\xi) := \mathsf{E}\exp(\gamma\xi)$ is also the moment generating function of ξ provided the expectation is finite for $|\gamma| < h$, and some $h > 0$. As it is well-known (see e.g. Gut(2004), Th. 8.3) then

$$\mathsf{E}\exp[\gamma\xi]^k < \infty \text{ for all } k = 1, 2, \ldots \qquad \mathsf{E}\,\xi^n = \frac{\mathrm{d}^n}{\mathrm{d}\gamma^n} U^{(\gamma)}(\xi)|_{\gamma=0}, \quad (n = 1, 2, \ldots) \tag{2}$$

and the Taylor expansion around $\gamma = 0$ reads

$$U^{(\gamma)}(\xi) = 1 + \mathsf{E}\sum_{k=1}^{\infty} \frac{(\gamma\xi)^k}{k!} = 1 + \sum_{k=1}^{\infty} \frac{\gamma^k}{k!} \cdot \mathsf{E}\,\xi^k \quad \text{for} \quad |\gamma| < h. \tag{3}$$

Considering Markov reward chains, let

$$\xi^{(n)} = \sum_{k=0}^{n-1} r_{X_k, X_{k+1}} \quad \text{be the (random) total reward received in the } n \text{ next transitions of the considered Markov chain } X.$$

Supposing that $X_0 = i$, on taking expectation we have for the first and second moment of $\xi^{(n)}$

$$V_i(n) := \mathsf{E}_i\,(\xi^{(n)}) = \mathsf{E}_i \sum_{k=0}^{n-1} r_{X_k, X_{k+1}}, \qquad S_i(n) := \mathsf{E}_i\,(\xi^{(n)})^2 = \mathsf{E}_i\,(\sum_{k=0}^{n-1} r_{X_k, X_{k+1}})^2 \tag{4}$$

hence the corresponding variance (i.e. the second central moment)

$$\sigma_i(n) := \mathsf{E}_i\,[\xi^{(n)} - V_i(n)]^2 = S_i(n) - [V_i(n)]^2. \tag{5}$$

Similarly, if the chain starts in state i for the expected utility we have

$$U_i(\gamma, n) := \mathsf{E}_i\,[\exp(\gamma\xi^{(n)})] = \mathsf{E}_i \exp[\gamma\,(r_{i,X_1} + \xi^{(1,n)}_{X_1})], \tag{6}$$

where $\xi^{(m,n)}_{X_m}$ (for $m < n$) is the reward obtained in the interval $[m, n]$ starting with state X_m.

In what follows let, $U(\gamma, n)$ be the (column) vector of expected utilities with elements $U_i(\gamma, n)$. Conditioning in (6) on X_1, from (6) we immediately get the recurrence formula

$$U_i(\gamma, n+1) = \sum_{j\in\mathcal{I}} p_{ij} \cdot \mathrm{e}^{\gamma r_{ij}} \cdot U_j(\gamma, n) = \sum_{j\in\mathcal{I}} q_{ij} \cdot U_j(\gamma, n) \quad \text{where} \quad U_j(\gamma, 0) = 1 \tag{7}$$

$$U(\gamma, n+1) = Q \cdot U(\gamma, n) \quad \text{with } U(\gamma, 0) = e, \text{ where } Q = [q_{ij}] = P \otimes R \tag{8}$$

the symbol $\otimes$ is used for the Hadamard product of matrices, i.e. $q_{ij} := p_{ij} \cdot \mathrm{e}^{\gamma r_{ij}}$

The paper is organized as follows. Section 2 contains basic formulae for calculating higher and higher central moments. Explicit formulae for higher central moments can be found in section 3, growth rates of central moments are discussed in section 4. Conclusions are made in the last section.

2 Exponential Utility and Higher Moments

Recall that by (7) $U_i(\gamma, n) = \mathsf{E}_i\,[\exp(\gamma\xi^{(n)})]$ is also the moment generating function of $\xi^{(n)}$. Hence (cf. (2)) for some $h > 0$ and any $|\gamma| < h$

$\frac{\mathrm{d}}{\mathrm{d}\gamma}\mathsf{E}_i\,[\exp(\gamma\xi^{(n)})] = \mathsf{E}_i\,\xi^{(n)}[\exp(\gamma\xi^{(n)})]$, hence for $k = 0,1,2\ldots,\ n = 0,1,2\ldots$

$$M_i^{(k)}(n) := \mathsf{E}_i\left(\xi^{(n)}\right)^k = \frac{\mathrm{d}^k}{\mathrm{d}\gamma^k}\mathsf{E}_i\,[\exp(\gamma\xi^{(n)})]|_{\gamma=0} \text{ is the } k\text{th moment of } \xi^{(n)} \tag{9}$$

and (cf. (3)) the Taylor expansion around $\gamma = 0$ reads

$$U_i(\gamma,n) = 1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}M_i^{(k)}(n) \quad \text{for} \quad |\gamma| < h. \tag{10}$$

This along with the identities (6), (7) will be extremely useful for finding explicit formulas of moments of $\xi^{(n)}$. In particular, since for $|\gamma| < h$

$$\mathrm{e}^{\gamma r_{ij}} = 1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}[r_{ij}]^k$$

from (7),(10) we immediately get

$$1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}M_i^{(k)}(n+1) = \sum_{j\in\mathcal{I}}p_{ij}\left(1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}[r_{ij}]^k\right)\left(1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}M_j^{(k)}(n)\right). \tag{11}$$

Similarly on introducing the moment generating function for the central moments of $\xi^{(n)}$ by

$$\tilde{U}_i(\gamma,n) := \mathsf{E}_i\,[\exp(\gamma(\xi^{(n)} - \mathsf{E}_i\,\xi^{(n)})] \quad \text{for all} \quad i\in\mathcal{I} \tag{12}$$

for the kth central moment of $\xi^{(n)}$ we have

$$\widetilde{M}_i^{(k)}(n) := \mathsf{E}_i\left[\xi^{(n)} - \mathsf{E}_i\,\xi^{(n)}\right]^k = \frac{\mathrm{d}^k}{\mathrm{d}\gamma^k}\mathsf{E}_i\,[\exp(\gamma(\xi^{(n)} - \mathsf{E}_i\,\xi^{(n)})]|_{\gamma=0} \tag{13}$$

and the Taylor expansion around $\gamma = 0$ for $|\gamma| < h$ reads

$$\tilde{U}_i(\gamma,n) = 1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}\cdot\widetilde{M}_i^{(k)}(n). \tag{14}$$

For what follows, let for real g, w_i's ($i\in\mathcal{I}$)

$$\tilde{\varphi}_{ij}(w,g) := r_{ij} - g + w_j - w_i \quad \text{where} \quad c = \max_{i\in\mathcal{I}}|w_i|.$$

Then $r_{X_k,X_{k+1}} = \tilde{\varphi}_{X_k,X_{k+1}}(w,g) + g + w_{k+1} - w_k$ and from the first part of (7) we arrive at we arrive at

$$V_i(n) = ng + w_i + \mathsf{E}_i\,[\sum_{k=0}^{n-1}\varphi_{X_k,X_{k+1}}(w,g) - w^{\gamma}_{X_n}]. \tag{15}$$

It is well-known that for unichain transition matrix P w_i's and g can be selected such that if

$$\sum_{i\in\mathcal{I}}p_{ij}\varphi_{ij}(w,g) = 0 \quad \text{for all} \quad i\in\mathcal{I} \quad \text{then} \quad V_i(n) = ng + w_i + \mathsf{E}_i\,w_{X_n} = 0$$

Similarly, from (7) we conclude that

$$U_i(\gamma,n) = \mathsf{E}_i\,\mathrm{e}^{\gamma\sum\limits_{k=0}^{n-1}r_{X_k,X_{k+1}}} = \mathrm{e}^{\gamma[ng+w_i^{\gamma}]}\times\mathsf{E}_i\,\mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1}\tilde{\varphi}_{X_k,X_{k+1}}(w^{\gamma},g) - w^{\gamma}_{X_n}]}. \tag{16}$$

Observe that the first term on the RHS of (15) is non-random and hence

$$\mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1} \tilde\varphi_{X_k,X_{k+1}}(w^\gamma,g)-c]} \le \frac{U_i(\gamma,n)}{\mathrm{e}^{\gamma[ng+w_i^\gamma]}} \le \mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1} \tilde\varphi_{X_k,X_{k+1}}(w^\gamma,g)+c]} \tag{17}$$

If w_i^γ's and g are selected such that for any $i \in \mathcal{I}$

$$\sum_{j\in\mathcal{I}} p_{ij}\, \mathrm{e}^{\gamma[r_{ij}-w_i^\gamma+w_j^\gamma-g]} = 1 \Longleftrightarrow \sum_{j\in\mathcal{I}} p_{ij}\, \mathrm{e}^{\gamma[r_{ij}+w_j^\gamma]} = \mathrm{e}^{\gamma[g+w_i]} \tag{18}$$

then for $X_0 = \ell$, $k = 0,1,\dots$ $\mathsf{E}_\ell\, \{\exp[\gamma\tilde\varphi_{X_k,X_{k+1}}(w^\gamma,g)]|X_k = m\} = 1$.
Since for $v_i := \mathrm{e}^{\gamma w_i}$, $\rho := \mathrm{e}^g$ the RHS of the second equality of (18) can be also written as $\sum_{j\in\mathcal{I}} q_{ij}\, v_j = \rho v_i$, the (column) vector v (with elements v_i's) is the Perron eigenvector of a nonnegative matrix Q with elements $q_{ij} = p_{ij}\, \mathrm{e}^{\gamma r_{ij}}$ and ρ is the spectral radius of Q. It is well known that if Q is irreducible then the Perron eigenvector v can be selected strictly positive. Observe that Q is irreducible if the matrix P is irreducible. Moreover, v can be selected strictly positive if the matrix P is only unichain and the risk sensitive coefficient γ is sufficiently close to null (cf. Gantmakher (1959)).

Since for suitably selected g, w_i's $\mathsf{E}\, \mathrm{e}^{\gamma\varphi_{X_{n-1},X_n}(w,g)+c}|X_{n-1} = i] = \mathrm{e}^{\gamma c}$ holds, then for the RHS of (18) we can conclude that

$$\mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1} \tilde\varphi_{X_k,X_{k+1}}(w,g)-c]} = \mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-2} \tilde\varphi_{X_k,X_{k+1}}(w^\gamma,g)-c]}$$

and on iterating the above displayed formula we can conclude that

$$\mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1} \tilde\varphi_{X_k,X_{k+1}}(w^\gamma,g)-w']} = \mathsf{E}_i\, \mathrm{e}^{\gamma[-w']}.$$

Inserting into (18) we arrive at the following bounds on $U_i(\gamma,n)$

$$\mathrm{e}^{\gamma[ng+w_i^\gamma]-c} \le U_i(\gamma,n) \le \mathrm{e}^{\gamma[ng+w_i^\gamma]+c} \tag{19}$$

For the central moments similarly to (12), (13) we can conclude from (18), (19) that

$$\tilde U_i(\gamma,n) := \mathsf{E}_i\, \mathrm{e}^{\gamma[\xi^{(n)}-(ng-w_i+w_{X_n})]} = \sum_{j\in\mathcal{I}} p_{ij} \mathrm{e}^{\gamma(r_{ij}-g+w_i-w_j)} \tilde U_j(\gamma,n-1) \tag{20}$$

where

$$\tilde U_j(\gamma,n-1) = \mathsf{E}_j \mathrm{e}^{\gamma[\xi^{(1,n)}-(n-1)g+w_j-w_{X_n}]}.$$

In analogy to (11) we get

$$1+\sum_{k=1}^{\infty}\frac{\gamma^k}{k!}\widetilde{M}_i^{(k)}(n+1) = \sum_{j\in\mathcal{I}} p_{ij}\big(1+\sum_{k=1}^{\infty}\frac{\gamma^k}{k!}[r_{ij}-(g+w_i-w_j)]^k\big)\times\big(1+\sum_{k=1}^{\infty}\frac{\gamma^k}{k!}\widetilde{M}_j^{(k)}(n)\big). \tag{21}$$

3 Higher central moments: Explicit Formulas

Similarly as in the previous section our analysis based on (12),(13),(14) and (19) enables to generate recursively all central moments of $\xi^{(n)}$. Recall that

$$\widetilde U_i(\gamma,n+1) = \sum_{j\in\mathcal{I}} p_{ij}\, \mathrm{e}^{\gamma\,[r_{ij}-(g+w_i-w_j)]}\widetilde U_j(\gamma,n), \qquad \widetilde U_j(\gamma,n) = 1+\sum_{k=1}^{\infty}\frac{\gamma^k}{k!}\widetilde{M}_j^{(k)}(n)$$

$$\mathrm{e}^{\gamma\,[r_{ij}-(g+w_i-w_j)]} = 1+\sum_{k=1}^{\infty}\frac{\gamma^k}{k!}[r_{ij}-(g+w_i-w_j)]^k$$

By comparing in (20) the terms γ^k $(k = 1, 2, \ldots)$ we obtain the following recursive formulas for the central moments (obviously, the first central moment $\widetilde{M}_i^{(1)}(n) \equiv 0$ for all n).

In particular,

$$\text{For } k = 1: \widetilde{M}_i^{(1)}(n+1) = \sum_{j\in\mathcal{I}} p_{ij}\, \widetilde{M}_j^{(1)}(n) \Rightarrow \widetilde{M}_j^{(1)}(n) = 0. \tag{22}$$

$$\text{For } k = 2: \widetilde{M}_i^{(2)}(n+1) = \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]^2 + \sum_{j\in\mathcal{I}} p_{ij}\, \widetilde{M}_j^{(2)}(n). \tag{23}$$

$$\begin{aligned}\text{For } k = 3: \widetilde{M}_i^{(3)}(n+1) &= \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]^3 \\ &+3 \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]\, \widetilde{M}_j^{(2)}(n) + \sum_{j\in\mathcal{I}} p_{ij}\, \widetilde{M}_j^{(3)}(n).\end{aligned} \tag{24}$$

$$\begin{aligned}\text{For } k = 4: \widetilde{M}_i^{(4)}(n+1) &= \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]^4 \\ &+ 6 \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]^2\, \widetilde{M}_j^{(2)}(n) \\ &+ 4 \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]\, \widetilde{M}_j^{(3)}(n) \; + \; \sum_{j\in\mathcal{I}} p_{ij}, \widetilde{M}_j^{(4)}(n)\end{aligned} \tag{25}$$

In general:

$$\begin{aligned}\widetilde{M}_i^{(s)}(n+1) &= \sum_{j\in\mathcal{I}} p_{ij}\; \{[(r_{ij} + w_j) - (g + w_i)]^s\} \\ &+ \sum_{j\in\mathcal{I}} p_{ij} \left\{ \sum_{k=1}^{s-1} \begin{pmatrix} s \\ k \end{pmatrix} [(r_{ij} + w_j) - (g + w_i)]^k\, \widetilde{M}_j^{(s-k)}(n) \right\} + \sum_{j\in\mathcal{I}} p_{ij}\, \widetilde{M}_j^{(s)}(n)\end{aligned} \tag{26}$$

that can be also written as

$$\widetilde{M}_i^{(s)}(n+1) = \sum_{k=0}^{s} \begin{pmatrix} s \\ k \end{pmatrix} \sum_{j\in\mathcal{I}} p_{ij} \left\{ [(r_{ij} + w_j) - (g + w_i)]^k\, \widetilde{M}_j^{(s-k)}(n) \right\} \tag{27}$$

4 Growth rates of the central moments

To begin with, let us recall the result for the growth rate of the variance of the total reward. In virtue of (23) we immediately conclude that the variance (i.e. the central second moment) of the total reward grows linearly over time and the growth rate $g^{(2)}$ of $\widetilde{M}_i^{(2)}(n)$ in (24) can be found as a solution of

$$\widetilde{M}_i^{(2)}(n) = n g^{(2)} + w_i^{(2)} \qquad \text{where} \tag{28}$$

$$g^{(2)} + w_i^{(2)} = s_i^{(2)} + \sum_{j\in\mathcal{I}} p_{ij} w_j^{(2)}, \quad s_i^{(2)} = \sum_{j\in\mathcal{I}} p_{ij} [(r_{ij} + w_j) - (g + w_i)]^2. \tag{29}$$

To establish the growth rate of $\widetilde{M}_i^{(3)}(n)$, it suffices to insert into (24) from (23). Since $\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,g^{(2)}=0$ we can conclude that

$$\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,(ng^{(2)}+w_j^{(2)}) = \sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,w_j^{(2)}.$$

On inserting into (23) we have

$$\begin{aligned}\widetilde{M}_i^{(3)}(n+1) &= \sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]^3 + 3\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,w_j^{(2)} \\ &\quad + \sum_{j\in\mathcal{I}} p_{ij}\,\widetilde{M}_j^{(3)}(n). \end{aligned} \tag{30}$$

Hence using the same arguments as for the second central moment we can conclude that

$$\widetilde{M}_i^{(3)}(n) = ng^{(3)} + w_i^{(3)} \quad \text{where} \quad g^{(3)} + w_i^{(3)} = s_i^{(3)} + \sum_{j\in\mathcal{I}} p_{ij} w_j^{(3)} \tag{31}$$

$$s_i^{(3)} = \sum_{j\in\mathcal{I}} p_{ij}\left\{[(r_{ij}+w_j)-(g+w_i)]^3 + 3\,[(r_{ij}+w_j)-(g+w_i)]\,w_j^{(2)}\right\}. \tag{32}$$

Conclusion: The growth of the second and the third central moments of the total reward is linear over time.

Unfortunately, this approach cannot be directly applied for finding formulas for the higher central moments of exponential utility functions as the following analysis can show. In particular, to establish the growth rate of $\widetilde{M}_i^{(4)}(n)$, we insert into (25) from (24) and (23). After some algebra, since $\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,g^{(3)}=0$, we arrive at

$$\begin{aligned}\widetilde{M}_i^{(4)}(n+1) &= \sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]^4 \;+\; 6\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]^2\,(ng^{(2)}+w_j^{(2)}) \\ &+ 4\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,w_j^{(3)} \;+\; \sum_{j\in\mathcal{I}} p_{ij}\,\widetilde{M}_j^{(4)}(n). \end{aligned} \tag{33}$$

Observe that the above equations (33) for the fourth central moments $\widetilde{M}_i^{(4)}(n)$ have the same structure as recursive equations (23),(24) for second a third central moments $\widetilde{M}_i^{(2)}(n)$, $\widetilde{M}_i^{(3)}(n)$. Unfortunately, the second term on the RHS of (33) contains also n, hence analogy with formulas for calculating average reward of a (unichain) Markov decision process cannot be used.

5 Conclusions

In the paper explicit formulae for calculating variance and higher central moments of the total reward generated by the Markov reward chain were obtained along with their growth rates if the considered time horizon tends to the infinity. In particular, it is shown that not only the variance (i.e. the second central moment) but also the skewness (i.e. the third central moment) grow linearly over time. Unfortunately, the linear growth rate does not hold for higher central moments.

Acknowledgements: This work was supported by the Czech Science Foundation under Grant 18-02739S.

References

[1] Cavazos-Cadena, R. and Fernandez-Gaucherand, F. (1999). Controlled Markov Chains with Risk-sensitive Criteria: Average Cost, Optimality Equations and Optimal Solutions. *Mathematical Methods of Operations Research*, 49(2), pp. 299-324.

[2] Cavazos-Cadena, R. and Hernández-Hernández, D. (2005). A Characterization of the Optimal Risk-sensitive Average Cost Infinite Controlled Markov Chains. *Annals of Applied Probability*, 15, pp. 175–212.

[3] Gantmakher, F. R. (1959). *The theory of matrices.* London: Chelsea.

[4] Gut, A. (2004). *Probability: a graduate course.* New York: Springer.

[5] Howard, R. A. and Matheson, J. (1972). Risk-sensitive Markov Decision Processes. *Management Science*, 18, pp. 356–369.

[6] Jaquette, S. C. (1976). A Utility Criterion for Markov Decision Processes. *Management Science*, 23, pp. 43–49.

[7] Kawai, H.(1987). A Variance Minimization Problem for a Markov Decision Process. *European Journal on Operational Research*, 31, pp. 140–145.

[8] Mandl, P. (1971). On the Variance in Controlled Markov Chains. *Kybernetika*, 7, pp. 1–12.

[9] Puterman, M. L. (1994). *Markov decision processes – discrete stochastic dynamic programming.* New York: Wiley.

[10] Ross, S. M. (1983). *Introduction to stochastic dynamic programming.* New York: Academic Press.

[11] Sladký, K. (2005). On Mean Reward Variance in Semi-Markov Processes. *Mathematical Methods of Operations Research*, 62, pp. 387–397.

[12] Sladký, K. (2008). Growth Rates and Average Optimality in Risk-sensitive Markov Decision Chains. *Kybernetika*, 44(2), pp. 205-226.

[13] Sladký, K.(2013). Risk-sensitive and Mean Variance Optimality in Markov Decision Processes. *Acta Oeconomica Pragensia*, 7, pp. 146–161.

[14] van Dijk, N.M. and Sladký, K. (2006). On the Total Reward Variance for Continuous-time Markov Reward Chains. *Journal of Applied Probability*, 43, pp. 1044-1052.

[15] White, D. J. (1988). Mean, Variance and Probability Criteria in Finite Markov Decision Processes: A Review. *Journal of Optimization Theory and Applications*, 56(1), pp. 1-29.

Author's address

Ing. Karel Sladký, CSc.
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Department of Econometrics
Pod Vodárenskou věží 4, 182 08 Praha 8
Czech Republic
e-mail: sladky@utia.cas.cz

TWO HEURISTICS FOR VEHICLE ROUTING WITH THREE KINDS OF CARRIERS: A COMPUTATIONAL STUDY

Petr Štourač, Department of Econometrics, University of Economics, Prague, Czech Republic,
Michal Černý, Department of Econometrics, University of Economics, Prague, Czech Republic,
Miroslav Rada, Department of Econometrics, University of Economics, Prague, Czech Republic

Abstract
We consider the following version of Vehicle Routing. We are given a family of indivisible batches with known weights located in City 1 (depot), where three carriers are located. The batches are to be delivered into known destinations on a given graph with weighted edges, where the weights correspond to distances. There are no restrictions on delivery times or assumptions on distances. Carrier I has a fleet of vehicles with known capacities which can be used in the TSP-style (Travelling Salesman Problem, i.e., a vehicle can serve a sequence of nodes and returns to City 1) and the price of the delivery is proportional to the total distance traveled by all vehicles. Carrier II works in a similar way, with the exception that he has a single vehicle with an unlimited capacity and that the price is proportional to the distance traveled and weight transported. Carrier III is a kind of `post-office' - the costs are proportional to the weight transported only. The question is how to assign the batches to the carriers in order to minimize overall shipping costs. We design two heuristic algorithms for the problem. We compare their performance to the exact approach: we formulate a natural Integer Linear Programming (ILP) model and use CPLEX to solve it.

Keywords: *Vehicle routing, binary integer programming, heuristic.*

JEL Classification: C61
AMS Classification: 90C59

1 INTRODUCTION

The „Vehicle Routing Problem“ (VRP) is one of the most widely studied topics in Operations Research. The VRP studies how to serve, from a central depot, a set of customers using a fleet of vehicles with varying capacities. This problem was first formulated in article (Clarke and Wright, 1964). A methodology for classifying the literature of the Vehicle Routing Problem is presented in (Eksioglu, Vural and Reisman, 2009; Braekers, Ramaekers and Van Nieuwenhuyse, 2016). An effective heuristic algorithm for the Traveling Salesman Problem is described in (Lin and Kernighan, 1973), an efficient insertion heuristics for Vehicle Routing Problem are described in (Campbell and Savelsbergh, 2004).

We consider Vehicle Routing with more kinds of carriers distinguished by different cost functions, the issue of more carriers is also addressed in the article (Chu, 2005). A carrier is characterized by a fleet of vehicles (with either bounded or infinite capacities) and a cost function. The cost function of a carrier can depend on the distance traveled and on the weights of items to be delivered. Even further factors can play a role in practice, such as the ordering of edges visited or arrival times to nodes, but these are not taken into account in this study (e.g. the problem of the VRP with time windows is discussed in (Baldacci, Mingozzi and Roberti, 2012; Jiang et al., 2014)).

This paper is focused on the distinction of carriers according to their cost functions. We consider three types: costs depending only on the distance traveled, costs depending only on the weights loaded in the depot and on a combination of both. This distinction does not play a crucial role in the ILP formulation of the problem (simply, the objective is the sum of the corresponding cost functions), but plays a role in the design of heuristics.

To formalize the problem more precisely, assume that we are given a family of indivisible batches with known weights. The batches are to be delivered into known nodes on a given graph with weighted edges, where weights correspond to distances. There are no assumptions on the distances. In the same node three different carriers are also located. There are no restrictions on delivery times. All batches are located in a City 1 (depot node).
Carrier I has a fleet of vehicles with known capacities. This fleet can be used in the TSP-style (i.e., a vehicle can pass through a sequence of nodes and return to City 1) and the price of the delivery is proportional to the sum of distances traveled by the vehicles. Carrier II works in a similar way, except for that he has a single vehicle with an unlimited capacity and the price is proportional to the distance traveled plus total weight loaded in the depot. Carrier III is a kind of `post-office' - the costs are proportional to the weight transported only. The goal is to minimize the overall costs.

2 ILP MODEL

Let $G = \{V;E\}$ be a complete undirected graph with $V = \{1; 2; \dots ; n\}$, where node 1 is the depot.

Input data:

- m = number of batches,
- l = number of vehicles of Carrier I,
- d_{ij} =distance between nodes i, j (generally, the matrix (d_{ij}) need not be symmetric),
- w_k = weight of batch k,
- c_q = capacity of the q-th vehicle of Carrier I,
- p_q = costs per km of the q-th vehicle of Carrier I,
- p_{Carr_II} = costs per km of the vehicle of Carrier II,
- r_{Carr_II} = costs per ton loaded by Carrier II,
- r_{Carr_III} = costs per ton transported by Carrier III,
- $z_{ki} \in \{0; 1\}$, $z_{ki} = 1$ iff node i is the destination for batch k.

Model variables:

- $y_k^q \in \{0; 1\}$, $y_k^q = 1$ iff batch k is transported by vehicle q of Carrier I,
- $x_{i,j}^q \in \{0; 1\}$, $x_{i,j}^q = 1$ iff vehicle q of Carrier I travels from node i to node j,
- $y_k^s \in \{0; 1\}$, $y_k^s = 1$ iff batch k is transported by Carrier II,
- $x_{i,j}^s \in \{0; 1\}$, $x_{i,j}^s = 1$ iff vehicle of Carrier II travels from node i to node j,
- $y_k^t \in \{0; 1\}$, $y_k^t = 1$ iff batch k is transported by Carrier III,
- u_i^q are auxiliary variables in anti-cyclic constraints for Carrier I,
- u_i^s are auxiliary variables in anti-cyclic constraints for Carrier II.

Objective function:

$$\min \sum_{q,i,j} p_q d_{ij}^q x_{ij}^q + p_{Carr\_II} \sum_{i,j} d_{ij}^s x_{ij}^s + r_{Carr\_II} \sum_k w_k y_k^s + r_{Carr\_III} \sum_k w_k y_k^t \quad (1)$$

Constraints:

$$\sum_i x_{ij}^q = \sum_i x_{ji}^q \qquad \forall q, j, \qquad (2)$$

$$\sum_i x_{ij}^s = \sum_i x_{ji}^s \qquad \forall j, \qquad (3)$$

$$\sum_k w_k y_k^q \le c_q \qquad \forall q, \qquad (4)$$

$$\sum_q y_k^q + y_k^s + y_k^t = 1 \qquad \forall k, \qquad (5)$$

$$\sum_{j,j \neq i} x_{ji}^q \le \sum_k y_k^q z_{ki} \qquad \forall q, \forall i \ge 2, \qquad (6)$$

$$\sum_{j,j \neq i} x_{ji}^s \le \sum_k y_k^s z_{ki} \qquad \forall i \ge 2, \qquad (7)$$

$$\sum_{j,k,j \neq i} z_{ki} x_{ji}^q \ge \sum_k y_k^q z_{ki} \qquad \forall q, \forall i \ge 2, \qquad (8)$$

$$\sum_{j,k,j \neq i} z_{ki} x_{ji}^s \ge \sum_k y_k^s z_{ki} \qquad \forall i \ge 2, \qquad (9)$$

$$\sum_j x_{1j}^q \le 1 \qquad \forall q, \qquad (10)$$

$$u_i^q - u_j^q + n x_{ij}^q \le n - 1 \qquad \forall q, i, \forall j \ge 2, i \neq j, \qquad (11)$$

$$u_i^s - u_j^s + n x_{ij}^s \le n - 1 \qquad \forall i, \forall j \ge 2, i \neq j, \qquad (12)$$

Constraint (2) states that a vehicle of Carrier I must enter and leave a node. Constraint (3) states same for the vehicle the Carrier II. Inequality (4) assures that capacity of q-th vehicle of Carrier I is not exceeded. Equation (5) tells that all batches must be delivered. Constraint (6) prevents the q-th vehicle of Carrier I from entering a node if the node is not served by the vehicle. Similarly, constraint (7) states the same for Carrier II. Conversely, constraints (8) and (9) require entry to node by vehicle if this vehicle serves the node for both carriers I and II. At least one leave the City 1 is possible for each vehicle of the carriers I state (10). Anti-cyclic conditions are in (11) and (12).

3 HEURISTICS

In this section, we design two heuristics for the problem. Both the heuristics take batches one by one and consecutively assign them to carriers and vehicles, while constructing the routes for vehicles of the distance-aware carriers. In each iteration, both the heuristics assign at least one batch unassigned so far and update the so-far-constructed routes. The difference is that the first heuristic select the first undelivered batch from the list to be assigned, while the second takes all unassigned batches into account in every iteration.

For Carrier I and Carrier II, a kind of the Insert Heuristic is used to update the route. Carrier III need not care about the vehicle position: we can assume that his batches are delivered directly from the depot.
For the description of the heuristics, the following notation will be useful.

Route: The route is a finite sequence of nodes from V where the first and last node is the depot.

- $(a_1^q, a_2^q, \cdots, a_{N_q}^q \equiv a_1^q)$ = route of the q-th vehicle of Carrier I, and,
- $(b_1, b_2, \cdots, b_M \equiv b_1)$ = route of the vehicle of Carrier II.

Delivered batches: The delivered batches is a set of batches which are delivered by a vehicle.

- Y^q = delivered batches by the q-th vehicle of Carrier I,
- Y^s = delivered batches by the vehicle of Carrier II, and
- Y^t = delivered batches by the vehicle of Carrier III.

Configuration: The configuration is a set of current routes and current assigned batches of all vehicles.

Cost: The cost is the lowest cost of delivering selected batch k with respect the current configuration.

- μ_k^q = the lowest cost of inserting batch k to the route of q-th vehicle of Carrier I,
- υ_k = the lowest cost of inserting batch k to the route of the vehicle of Carrier II,
- ω_k = the cost of using the vehicle of Carrier III for delivering batch k, and
- σ = the lowest cost of delivering at a given configurations.

The actual $\mu_k^q, \upsilon_k, \omega_k$ *and* σ are to be specified for each of the heuristics individually.

3.1 Heuristic I

The batches are processed one-by-one in an initial ordering. For each batch, the cheapest way of shipping is found based on the routes of individual carrier vehicles planned so far. In case that the batch is transported by vehicle of the Carrier I, if possible, additional batches that have not yet been delivered to the destination are added to the car. The heuristic ends after processing the last undelivered batch from the list.

Step 1: Initialize start parameters. All routes are initialized to the value (1,1), i.e. start and stop node must be City 1 (depot).

Step 2: Calculate delivery costs of batch *k* for every vehicle. A batch k is selected from the list of the unassigned batches. The costs for each vehicle are calculated as follows:

- $\mu_k^q = \infty \; if \; w_k > c_q - \sum_{l, l \epsilon Y^q} w_l, else \; \mu_k^q = \min_{a_j}\left\{p_q\left(d_{a_j,i_k} + d_{i_k,a_{j+1}} - d_{a_j,a_{j+1}}\right)\right\}$ where $a_j \in \left(a_j^q\right) \wedge j < N_q, i_k$ = the destination node for batch k,
- $\upsilon_k = \min_{b_j}\left\{p_{Carr\_II}\left(d_{b_j,i_k} + d_{i_k,b_{j+1}} - d_{b_j,b_{j+1}}\right) + w_k r_{Carr\_II}\right\}$ where $b_j \in \left(b_j\right) \wedge j <$ M, i_k = the destination node for batch k,
- $\omega_k = w_k r_{Carr\_III}$.

Step 3: Choose the carrier's vehicle with minimal costs. The costs are calculated as follows:

- $\sigma = \min_q\{\mu_k^q, \upsilon_k, \omega_k\}$.

For the minimizing carrier and vehicle, do the following:

Carrier I: add additional undelivered batches to the vehicle for the destination (if it can carry), add another node to the vehicle route $\left(a_j^q\right)$, mark laden batches as delivered. Update the set of delivered batches Y^q for the vehicle q with laden batches.

Carrier II: add another node to the vehicle route $\left(b_j\right)$, mark laden batch as delivered. Update the set of delivered batches Y^s with laden batches.

Carrier III: mark laden batch as delivered. Update the set of delivered batches Y^t with laden batches.

If there is still an undelivered batch, take the first one and proceed to Step 2, otherwise the processing will stop.

3.2 Heuristic II

This heuristic looks for the batch with the cheapest possible transport among the batches unallocated so far. The calculation is based on the actual configuration. In case the batch is transported by a vehicle of Carrier I, additional batches that have not yet been delivered (if any) to the destination are added to the vehicle. In case the batch is transported by a vehicle of Carrier II, additional batches with the same destination are added to the vehicle under the condition that their transport is cheaper than if they were transported by Carrier III. The heuristics ends after processing the last batch.

Step 1: Initialize start parameters. All routes are initialized to the value (1,1), i.e start and stop node must be City 1 (depot).

Step 2: Calculate delivery costs for every vehicle. The costs for each vehicle and each undelivered batch are calculated as follows:

- $\mu_k^q = \infty \; if \; w_k > c_q - \sum_{l,l\epsilon Y^q} w_l, else \; \mu_k^q = \min_{a_j}\left\{p_q\left(d_{a_j,i_k} + d_{i_k,a_{j+1}} - d_{a_j,a_{j+1}}\right)\right\}$ where $a_j \in \left(a_j^q\right) \wedge j < N_q, i_k$ = the destination node for batch *k*,
- $\upsilon_k = \min_{b_j}\left\{p_{Carr\_II}\left(d_{b_j,i_k} + d_{i_k,b_{j+1}} - d_{b_j,b_{j+1}}\right) + w_k r_{Carr\_II}\right\}$ where $b_j \in \left(b_j\right) \wedge j <$ M, i_k = the destination node for batch *k*,
- $\omega_k = w_k r_{Carr\_III}$.

For each vehicle minimal costs are selected over all undelivered batches:

- $\mu_{k_I}^q = \min_k\{\mu_k^q\}$ where the decision on the minimum is made on the basis of expression μ_k^q / w_k,
- $\upsilon_{k_{II}} = \min_k\{\upsilon_k\}$,
- $\omega_{k_{III}} = \min_k\{\omega_k\}$.

Step 3: Choose the carrier's vehicle with minimal costs. The costs are calculated as follows:

- $\sigma = \min_q\{\mu_{k_I}^q, \upsilon_{k_{II}}, \omega_{k_{III}}\}$.

For the minimizing carrier and vehicle, do the following:

Carrier I: add additional undelivered batches to the vehicle for the destination (if it can carry), add another node to the vehicle route $\left(a_j^q\right)$, mark laden batches as delivered. Update the set of delivered batches Y^q for the vehicle q with laden batches.

Carrier II: add additional undelivered batches to the destination on the condition that their delivery is cheaper than if they were delivered by Carrier III, add another node to the vehicle route (b_j), mark laden batch as delivered. Update the set of delivered batches Y^s with laden batches.

Carrier III: mark laden batch as delivered. Update the set of delivered batches Y^t with laden batches.

If there is still an undelivered batch, take the first one and proceed to Step 2, otherwise the processing will stop.

4 PERFORMANCE OF THE HEURISTICS

The model and both heuristics were tested on six datasets. Their parameters are described in Table 1. The samples are intentionally small so that we can get either an optimum solution or a tight lower bound from a solver (we used CPLEX here) and compare it to the output of the heuristic methods.

We used a computer with the following parameters: Intel Core i5-4200U CPU @ 1.60 GHz 2.30 GHz, 4 GB RAM, Windows 10 (x64). The ILP model is solved by CPLEX 12.6 and both heuristics are implemented in Python 3.4.

Results summarized in Table 2. Note that Heuristic I is sensitive to the initial ordering of batches, which also affects the search for the routes of vehicles.

Table 1. Summary data for test instances. The columns stand for the number of nodes, number of batches, number of vehicles of Carrier I, number of variables and number of constraints of the ILP model. In Ex. 1, batches are sorted at random, in Ex. 1-v2, batches are sorted in the descending order according to the weights, and in Ex. 1-v3, batches are sorted by distances to the target node. Ex. 2 v2 is a modification of Ex. 2 with lower per-km costs of Carrier III.

	vert's	batches	No. I	var's	constr's
Example 1	5	10	2	115	96
Example 1 v2	5	10	2	115	96
Example 1 v3	5	10	2	115	96
Example 2	10	20	5	740	693
Example 2 v2	10	20	5	740	693
Example 3	30	40	8	8 500	8 446

Table 2. Results of experiments: the optimal values of objective function found by CPLEX (in Ex. 3, only a lower bound was available due to excessive memory requirements), the values found by Heuristics I and II and the corresponding ratios (HI/C, HII/C, HI/HII).

	Cplex	Heur. I	HI/C	Heur. II	HII/C	HII/HI
Example 1	68 100	71 343	1.05	70 283	1.03	0.99
Example 1 v2	68 100	74 833	1.10	70 283	1.03	0.94
Example 1 v3	68 100	73 221	1.07	70 283	1.03	0.96
Example 2	128 187	162 366	1.27	133 790	1.04	0.82
Example 2 v2	67 988	78 581	1.16	70 928	1.04	0.90
Example 3	≈334 237	360 311	1.08	346 349	1.04	0.96

5 EXAMPLE 1 IN DETAIL

In Example 1 we have $n = 5$ (number of nodes), $m = 10$ (number of batches), $l = 2$ (number of vehicles of Carrier I), the distance matrix $D = (d_{ij})$ has the form

$$D = \begin{pmatrix} 0 & 13 & 9 & 23 & 17 \\ 13 & 0 & 16 & 8 & 25 \\ 9 & 16 & 0 & 30 & 28 \\ 23 & 8 & 30 & 0 & 15 \\ 17 & 25 & 28 & 15 & 0 \end{pmatrix},$$

and parameters of batches are as follows:

Batch No.	1	2	3	4	5	6	7	8	9	10
Weight	20	35	19	57	74	10	8	11	39	52
Target node	2	4	3	4	3	4	5	5	3	2

Capacities c_1; c_2 of the vehicles of Carrier I are 50; 100, and their per-km costs p_1; p_2 are 6; 12, respectively. Further we have $p_{Carr\_II} = 11$ (per-km costs of Carrier II), $r_{Carr\_II} = 380$ (per-ton costs of Carrier II) and $r_{Carr\_III} = 450$ (per-ton costs of Carrier III).

The results of Example 1 are presented in Table 3 and in Figure 1.

Table 3. Results (Example 1): The ratio of the weight of batches transported by vehicles of Carrier I to its capacity found by CPLEX and Heuristics I and II.

Filling	**Cplex**	**Heur. I**	**Heur. II**
Carrier I - Vehicle 1	50/50	49/50	50/50
Carrier I - Vehicle 2	100/100	100/100	93/100

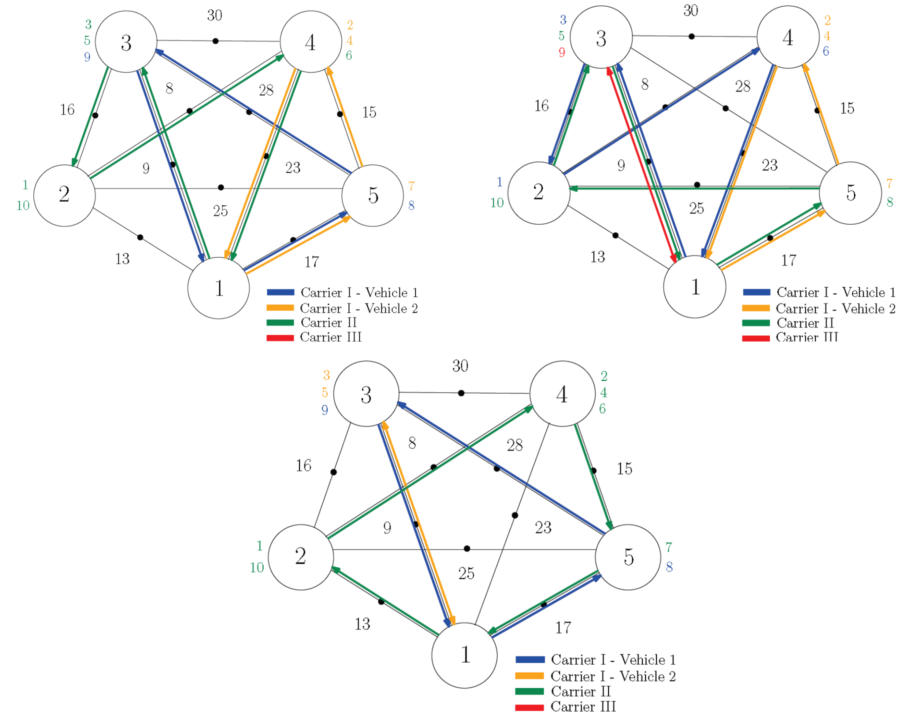


Figure 1. Results (Example 1): The graph with solution founded by CPLEX (top left), Heuristic I (top right) and Heuristic II (bottom). Color edges represent vehicle's routes. The numbers next to the nodes identify the batches.

6 CONCLUSIONS

We considered the case of vehicle routing where carriers differ by their cost functions. We considered three types of carriers, whose costs depend only on the distance traveled, only on the weight transported and on both. We formulated the problem as an ILP and showed that even small instances are not solvable exactly (which is not surprising); therefore we designed heuristic approaches. The heuristics utilize the particular forms of cost functions of the carriers (but note that they can be useful for a broader class of cost functions). It turns out that Heuristic II outperforms Heuristic I and, on average, Heuristic II constructs feasible solution just about 3 % - 4 % worse compared to the true optimal value (at least with the tested instances where the optimum solution could be found).

Acknowledgements: The work was supported by the Czech Science Foundation under grant 16-00408S.

References

[1] Baldacci, R., Mingozzi, A. and Roberti, R. (2012). Recent exact algorithms for solving the vehicle routing problem under capacity and time window constraints. *European Journal of Operations Research*, 218(1), pp. 1-6.

[2] Braekers, K., Ramaekers, K. and Van Nieuwenhuyse, I. (2016). The vehicle routing problem: State of the art classification and review. *Computers & Industrial Engineering*, 99, pp. 300-313.

[3] Campbell, A.M. and Savelsbergh, M. (2004). Efficient insertion heuristics for vehicle routing and scheduling problems. *Transportation Science*, 38(3), pp. 369-378.

[4] Chu, Ch-W. (2005). A heuristic algorithm for the truckload and less-than-truckload problem. *European Journal of Operations Research*, 165(3), pp. 657-667.

[5] Clarke, G. and Wright, J.W. (1964). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operational research*, pp. 568-581.

[6] Eksioglu, B., Vural, A. V. and Reisman, A. (2009). The vehicle routing problem: A taxonomic review. *Computers & Industrial Engineering*, 57, pp. 1472-1483.

[7] Jiang, J., Ng, K. M., Poh, K. L. and Teo, K. M. (2014). Vehicle routing problem with a heterogeneous fleet and time windows. *Expert Systems with Applications*, 41(8), pp. 3748-3760.

[8] Lin, S. and Kernighan, B. W. (1973). An Effective Heuristic Algorithm for the Traveling Salesman Problem. *Operations Research*, 21(2), pp. 498-516.

Author's address

RNDr. Petr Štourač
University of Economics, Faculty of Informatics and Statistics, Department of Econometrics
nám. W. Churchilla 1938/4, 130 67 Praha 3
Czech Republic
email: stop02@vse.cz

prof. RNDr. Ing. Michal Černý, Ph.D.
University of Economics, Faculty of Informatics and Statistics, Department of Econometrics
nám. W. Churchilla 1938/4, 130 67 Praha 3
Czech Republic
email: cernym@vse.cz

Ing. Miroslav Rada, Ph.D.
University of Economics, Faculty of Informatics and Statistics, Department of Econometrics
nám. W. Churchilla 1938/4, 130 67 Praha 3
Czech Republic
email: miroslav.rada@vse.cz

CONVERGENCE OF UNEMPLOYMENT IN THE SLOVAK REPUBLIC

Kvetoslava Surmanová, Marian Reiff, University of Economics in Bratislava

Abstract
Economic growth and regional disparities increment is a major problem in several countries that require a resolution. The starting point for solving this issue is to know how large disparities are. Often the convergence method is used to assess region comparisons, respectively, the divergence method. In empirical studies various convergence methods are applied to study the regional differences. The choice of the appropriate method depends on the objective pursued. Acquired knowledges and understanding of the persistence of regional unemployment policies. Ultimately it is a tool used by economic policy makers to address various economic issues. In our analysis, we focused on the β- convergence of the unemployment rate at NUTS 4 level in 79 districts of the Slovak Republic over the years 2010-2017. The results of the analysis point to a faster convergence in the period 2012-2017 compared to the convergence for the period 2010-2017.

Keywords: *regional disparities, unemployment, classification, β-convergence*

JEL Classification: C21, J64
AMS Classification: 62J05

1 INTRODUCTION

Convergence is a concept for which there are several explanations. Real convergence is the approximation of one country to another, more advanced country. As a rule, they are measured by the GDP per capita indicator. Analytically, the definition of convergence at time *t* can be written using the following relationship (Slavík, 2007):

$$|y_{1,t} - y_{2,t}| > |y_{1,t+1} - y_{2,t+1}|, \tag{1}$$

where $y_{1,t}$ and $y_{2,t}$ are the relative economic quantities of two countries (regions) at time *t*. If the relationship (1) is met as inequality of opposite sign, it would be divergence.

In empirical literature, we can meet two directions of convergence research. First group of authors examines the convergence of economic growth and income. Barro and Sala-i-Martin (1992, 1995) are the most famous works on convergence analysis based on cross-sectional data. Other work on income convergence is, for example, Nevima and Melecký (2011) or Szomolányi, Lukáčiková and Lukáčik (2011). Boldrin and Canova (2001) have focused on the EU countries and Egger et al. (2005) on convergence of Central and Eastern Europe. countries.

Second group of authors examines the convergence of the labor market, focusing on wages and the unemployment rate. In this area it is possible to find rather work of a regional character, for example Pehkonen and Tervo (1998) or Boeri and Terrell (2002). Katrencik, Wójcik and Tyrowicz (2008) focused on convergence models applicable to the transition economies of the Czech Republic, Poland and Slovakia. Chocholatá and Furková (2016) analyzed the regional disparities of net disposable income of households across 82 NUTS 2 regions of Austria, Czech Republic, Slovakia, Poland, Hungary and Germany over the period 2000-2013.

The main aim of the paper is to analyze regional disparities on labor market base on the concept of beta convergence across 79 NUTS4 regions of Slovakia during the period 2010 – 2017.

The structure of the paper is as follows. After the introductory section, the second section introduces the methodological issues to beta convergence in the context of across data, contains the data description and preliminary evidence. The empirical results of beta-convergence testing are given in the third section. The fourth section concludes with some challenges for future research.

2 METHODOLOGY AND DATA DESCRIPTION

The first empirical analysis was carried out by Baumol (1986). The basic regression was linear modeling the per capita income depending on its initial level. The formal convergence model was created by Barro and Sala-i-Martin (1992), deriving linear regression from the neoclassical model. Other variables have been considered, based on theory of growth. This model has the following form:

$$\frac{1}{T} ln\left(\frac{y_{iT}}{y_{i0}}\right) = \alpha_0 + \alpha_1 ln(y_{i0}) + \sum_{k=1}^{K} \emptyset_k X_{k,iT} + \varepsilon_i, \tag{2}$$

where

y_{i0} and y_{iT} is per capita income in the start and end times of the i-th cross-section unit,
α_1 is the parameter used to draw conclusions about the existence of convergence,
$X_{k,iT}$ is set of control variables ($k = 1, \dots, K$)
ε_i is an error term.

The left-hand side of (2) represents the rate of economic growth. For economic growth, the catching-up process takes place if α_1 is negative and statistically significant. The convergence in relation (2) is conditional on the factors of economic growth contracted by the set of control variables $X_{k,iT}$. The speed of convergence expresses what portion of the gap between the current state and the baseline condition will manage to fade in each year and we can express it from relationship (2) as follows:

$$\beta = \frac{-\ln(1+\alpha_1 T)}{T} \tag{3}$$

Half-life of convergence can be calculated from conversion rate:

$$t^* = -\frac{\ln(2)}{\beta} \tag{4}$$

and means the number of years for which the original gap between regions is reduced by half. Absolute (unconditional) convergence means that countries are approaching the same steady state. This convergence will also be addressed in our empirical part, but we will use the slightly transformed convergence model of Próchniak and Witkowski (2012).The growth rate in (2) is not divided by T:

$$ln\left(\frac{u_{it}}{u_{i0}}\right) = \alpha_0 + \alpha_1 ln(u_{i0}) + \varepsilon_i, \tag{5}$$

where u_{i0} and u_{iT} are indicators of unemployment in the start and ending periods.
And estimates of β is in this analysis are calculated as follows:

$$\beta = \frac{\ln(1+\alpha_1)}{T} \tag{6}$$

The only difference in the convergence of unemployment compared to the income convergence is the expectation of sign of the parameter α_1. In this case, it should be positive, because in the convergence of regions we expect to reduce the differences, a region with high unemployment should reduce the number of unemployed faster than a region with low unemployment

In the analysis we use the following data: unemployment rate (in %), the number of economically active population (EAP) in persons, and the number of job seekers (NJS) in persons from sources Central Office of Labour, Social Affairs and Family [15] for the years 2010, 2012 and 2017. The data are always relevant for December of the year in question. The data were retrieved for the 79 NUTS 4 regions of Slovak republic. In the analysis, we used the ratio, expressed as a proportion NJS and EAP. The one part of analysis was performed using the software GeoDa [17] and second part of analysis by Eviews. From the shape file (.shp) of Slovakia [16], the 79 NUTS 4 Slovakian regions were selected in GeoDa.

3 EMPIRICAL RESULTS

Based on the graphical representation of the spatial distribution of the unemployment rate in Slovakia (Figures 1 and 2) it is clear that the high unemployment rate is in eastern Slovakia (with the exception of Košice only), in the southern districts of central and partly south-western Slovakia. The highest level is in Rimavská Sobota, while the first level remains in 2017 too. Districts with a high rate of unemployment are characterized by a low level of wages, and in the case of Eastern Slovakia by high level of Roma population (Mušinka et al., 2014) Unemployment rates in 2010 and 2017 in Eastern Slovakia did not show any significant redistribution, as weak job descriptions are of a fixed nature. Compared to 2017, lower unemployment is concentrated mainly in western Slovakia and northern districts of central Slovakia.

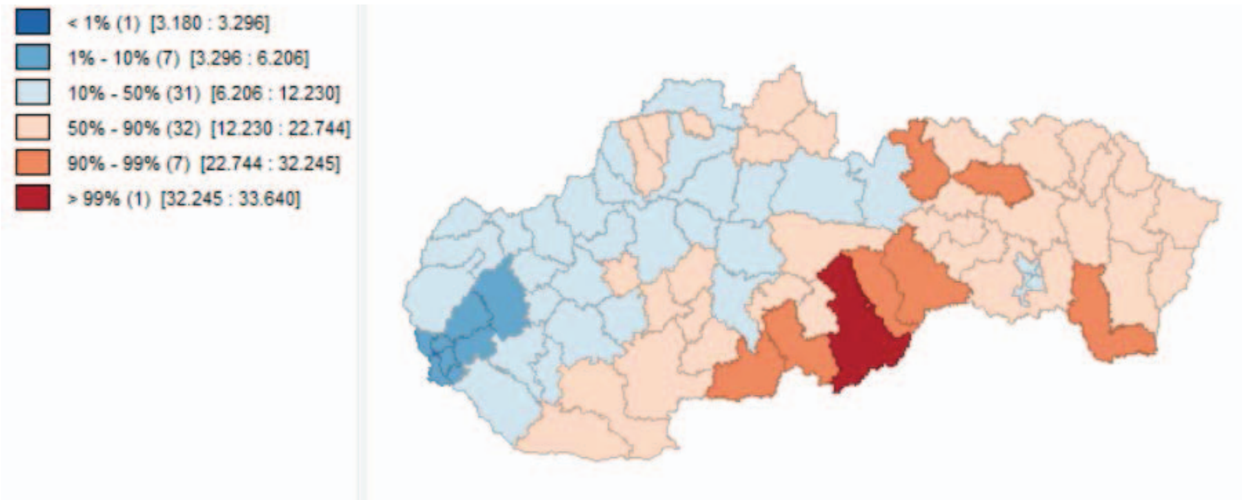


Figure 1 Percentile Map of Unemployment Rate in 2010 [1]
Source: Own calculation

[1] All figures presented in the paper are available in the colour version on the web-site: http://www.fhi.sk/kove/konferencie

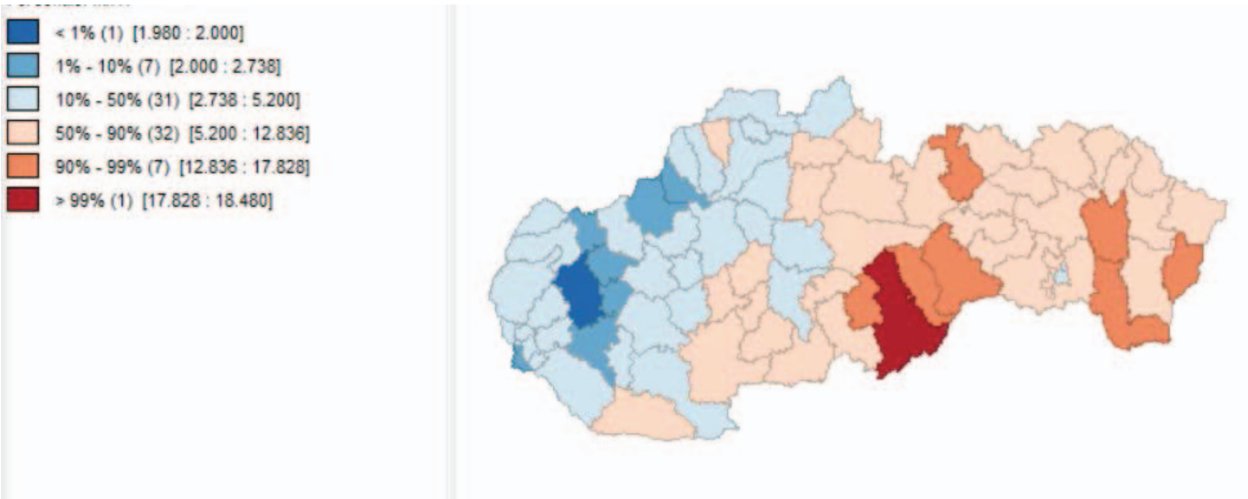


Figure 2 Percentile Map of Unemployment Rate in 2017
Source: Own calculation

The essence of spatial analysis is the finding that what is happening in one region is influenced by events in other regions, the Closter map captures spatial autocorrelation through the color range. The map is generated when calculating the Moran coefficient. The districts marked with a gray color are statistically insignificant. The red and the dark blue are positively autocorrelated and light blue, respectively pink are negatively autocorrelated. Positive dependence means that the neighboring values are similar and vice versa, the negative dependence means that the adjacent values are different.

In the Figures 3 and 4 below, LISA cluster map (Local Indicators of Spatial Association) that captures the weighted value of a variable in nearby units. By comparing the two images we can see how the correlation between districts has changed.

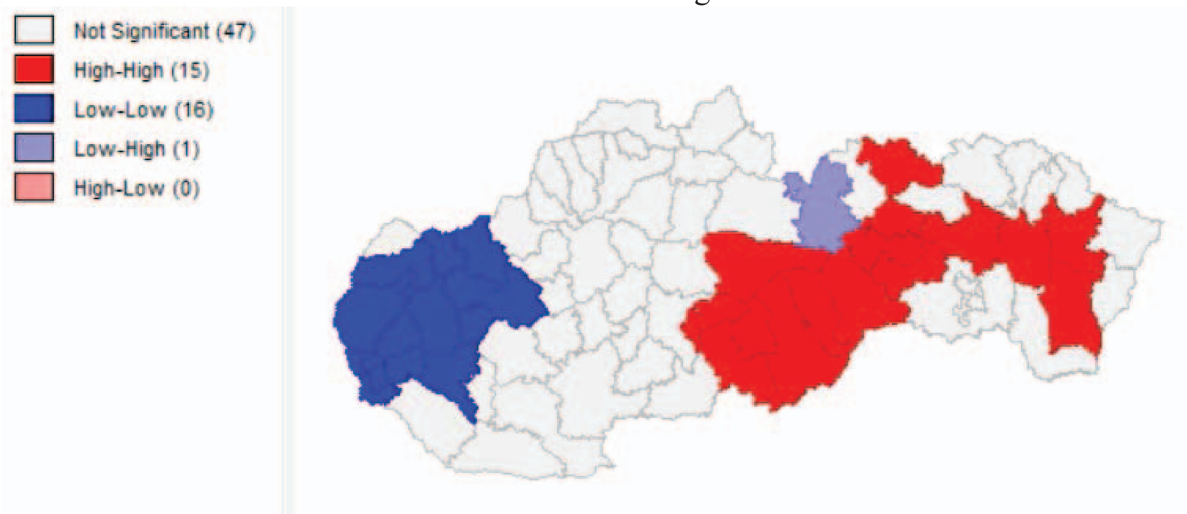


Figure 3 LISA Cluster map, year 2010
Source: Own calculation

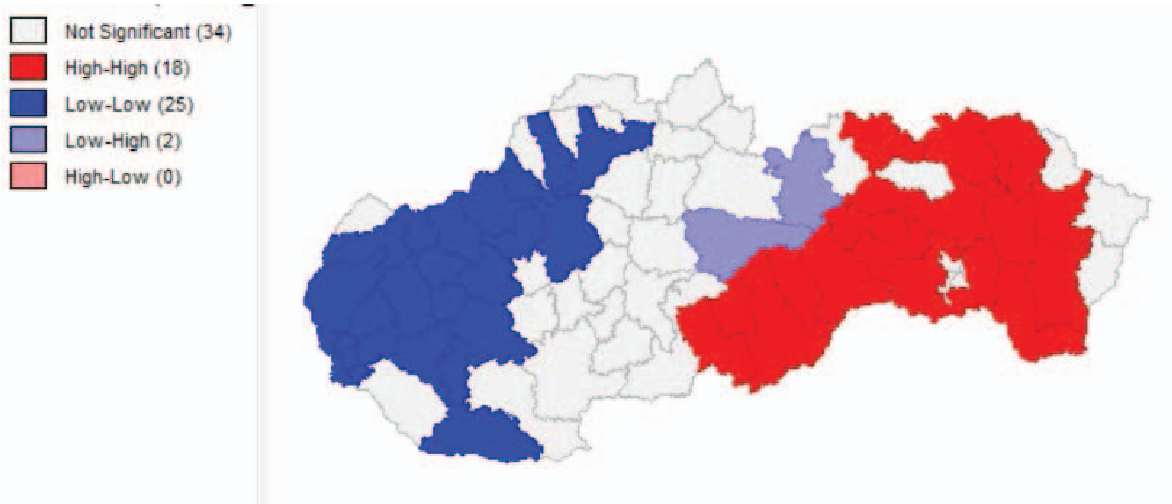


Figure 4 LISA Cluster map, year 2017
Source: Own calculation

By comparing the two sessions, there is an increasing proportion of districts in which there is a positive autocorrelation at the expense of districts that are statistically insignificant. While only 31 districts were positively correlated in 2010, this number increased to 43 (18 districts with *high - high* and 25 regions with *low - low* association) in 2017. The strongly colored regions are therefore those that contribute significantly to a positive global spatial autocorrelation outcome, while paler colors contribute significantly to a negative autocorrelation outcome. Based on the above figures it can be stated that in the last 8 years in the Slovakia there was tendency of the districts clustering in space, although the unemployment rate varies from one district to another, there is usually a tendency for adaptation, convergence of development to the same value.

The value of the Univariate Moran's coefficient[2] is 0.6493 (2010) and 0.6507 (2017), which also expresses positive spatial autocorrelation.

In this section we used non-spatial convergence econometric models briefly described in part 2. Regression analysis is based on relationship 5 and 6. As an indicator of unemployment in the initial period, we chose the proportion of job seekers (NJS) and economically active population (EAP), $u_{i0} = NJS_{i0}/EAP_{i0}$. We have also calculated the unemployment rate over time T.

The results of the convergence analysis for all districts were not sufficiently significant, so we have estimated two more versions of the convergence model. Model with 67 regions, since in the regions of Western Slovakia (Bratislava 1 – Bratislava V, Malacky, Pezinok and Senec) and Košice districts (KO1 - KO IV) the values are significantly different from the rest of Slovakia. Another version was made up of 71 districts and we did not consider data for 8 districts (Bratislava 1 – Bratislava 5, Malacky, Pezinok, Senec).

Convergence tendencies were followed for two periods 2010 - 2017 and 2012 - 2017. The second time period 2012-2017 was chosen because of a greater distance from the outbreak of the crisis in 2008-2009.

The results of the estimated models are presented in Table 1.

[2] Moran´s coefficient expresses the measure of spatial autocorrelation.

Table 1 Results of convergence models

Estimation	Convergence 2010-2017 (OLS)			Convergence 2012-2017 (OLS)		
	67 regions	71 regions	79 regions	67 regions	71 regions	79 regions
α_0	-0,170**	-0,258***	-0,728***	-0,173**	-0,234***	-0,672***
α_1	0,319***	0,261***	-0,008	0,396***	0,351***	0,084*
R^2	0,452	0,296	0,003	0,553	0,444	0,034
	Convergence characteristics					
Speed of convergence	0,039 (3,9 %)	0,033 (3,3 %)	-	0,047 (4,7%)	0,060 (6,0 %)	0,016 (1,6 %)
Half-life	17,531	20,943	-	14,557	11,522	42,874

Note: Symbol ***, **, * indicates statistical significance at 1, 5 and 10 % level of significance.
Source: Own calculations

The convergence process was most pronounced in the results in 2012-2017 of the 71-district model (highest coefficient of determination). In this case, the convergence is at a rate of 6%. For the same reporting period, we missed out on an analysis of district Košice. The result was a smaller convergence speed, namely 4.7 percent. The speed of convergence 1,6% reflects a small reduction in regional disparities in all districts of the SR. The result can be considered relevant, as we are also considering the districts of Western Slovakia (variance is higher in this case).
The results for the period 2010-2017 show minimal differences, as the unemployment rate was rising at the beginning of the reference period (the average rate in 2010 was 13.52% and in 2012 increased to 15.65%). The convergence tendencies in 2010-2017 have not been confirmed at the level of all districts in Slovakia, so we do not report convergence characteristics in Table 1.

4 CONCLUSIONS

Better results - for faster convergence are shown by the models in the period 2012 – 2017. A lower speed of convergence from the analysis of the 2010-2017 period can be attributed to a stronger impact of the crisis on the labor market. The convergence process at the level of the districts of the Slovak Republic appears to be more pronounced in cases when the Bratislava districts (5), Kosice districts (4) and 3 districts in the vicinity of Bratislava were not used in analysis. When we focus solely on the results of the convergence process itself in 2012 and 2017, the fastest (6.0%) convergence is present in 71-district model. The most remarkable decline in unemployment was in the Revúca (17.67%), Rimanská Sobota (17.11%), Veľký Krtíš (16.69%), Sabinov (16.24%). While the smallest change (5-6%) is recorded the districts of Western and Central Slovakia, Galanta, Skalica, Trnava, Myjava, Žilina, B. Bystrica and Košice 1 - Košice 4. The results of the analysis for all districts for the period 2012-2017, despite the significance, have low reporting ability (low R^2). At a low convergence rate, the half-life is estimated to almost 43 years. Based on the results of spatial analysis, it can be stated that the data show spatial autocorrelation, so in the next research it is necessary to consider the spatial aspect in models.

The presence of this dependence requires a much more complex model to be formulated, i.e. spatial model (SAR, SEM, or some of the panel data model models).

It would be interesting to compare the results of this analysis with the results of the regional convergence incomes.

Acknowledgements: This work was supported by the Grant Agency of Slovak Republic – VEGA grant No. 1/0248/17 "Analysis of Regional Disparities in the EU based on Spatial

Econometric Approaches" and VEGA 1/0368/18 “Pricing Strategies in the Environment of an Effective Regulated Mechanism in the Network Industries Markets of the Slovak Economy“.

References

[1] Barro, R. and Salla-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2), pp. 223-251.

[2] Barro, R.J. and Sala-i-Martin, X. (1995). *Economic Growth.* New York: McGraw-Hill.

[3] Baumol, W. J. (1986). Productivity Growth, Convergence and Welfare: What the Long-Run Data Show. *The American Economic Review*, 76(5), pp. 1072-1085.

[4] Boeri, T. and Terrell, K. (2002). Institutional Determinants of Labor Reallocation in Transition. *Journal of Economic Perspectives* 16(1), pp. 51-76.

[5] Boldrin, M. and Canova, F. (2001). Inequalities and convergence in Europe’s regions. Reconsidering European Regional Policies. *Economic Policy*, 16(32), pp. 206-253.

[6] Egger, P., Huber, P. and Pfaffermayr, M. (2005). A note on export openness and regional wage disparity in Central and Eastern Europe. *The Annals of Regional Science*, 39(1), 63-71.

[7] Chocholatá, M. and Furková, A. (2016). Income disparities and convergence across regions of Central Europe. *Croatian Operational Research Review,* [online] 7(2), pp. 303-318. Available at: file:///C:/Users/ssov/Downloads/4539-Article%20Text-14448-1-10-20170130.pdf [Accessed 25 Feb. 2018].

[8] Katrencik, D., Wójcik, P. and Tyrowicz, J. (2008). Unemployment convergence in transition. Available at: http://www.iu.edu/~econdept/workshops/Spring_2008_Papers/Tyrowicz_Convergence_in_Transition.pdf [Accessed 25 Feb. 2018].

[9] Mušinka, A. et al. (2014). *Atlas rómskych komunít na Slovensku 2013*. Bratislava: Regionálne centrum Rozvojového program OSN.

[10] Pehkonen, J. and Tervo, H. (1998). Persistence and Turnover in Regional Unemployment Disparities. *Regional Studies* 32(5), pp. 445-458.

[11] Próchniak, M. and Witkowski, B. (2012). *β-Convergence Stability Among “Old” and “New” EU Countries: The Bayesian Model Avaraging Perspective*. Warsaw School of Economics. Available at: http://coin.wne.uw.edu.pl/wiem/papers_2012/Witkowski [Accessed 25 Feb. 2018].

[12] Nevima, J., and Melecký, L. (2011). The β-convergence analysis of the visegrad four NUTS 2 regions. *Mathematical Models and Methods in Modern Science*, pp.48-53.

[13] Slavík, C. (2007), Reálná konvergence České republiky k Evropské unii v porovnání s ostatními novými členskými zeměmi. *Politická ekonomie*, 55(1), pp.23-40

[14] Szomolányi, K., Lukáčiková, A. and Lukáčik, M. (2011). The β-convergence of the EU27 Countries. In: *Mathematical methods in Economics 2011*. Praha: Professional Publishing, pp. 287-291.

[15] Central Office of Labour, Social Affairs and Family, Statistics/ Available at: http://www.upsvar.sk/statistiky.html?page_id=1247 [Accessed 05/02/2018]

[16] Geoportal, Available at: https://www.geoportal.sk/sk/zbgis_smd/na-stiahnutie/ [Accessed 05/02/2018].

[17] The Centre for Spatial Data Science, University of Chicago, Software downloads, Available at: https://geodacenter.asu.edu/software/downloads [Accessed 05/02/2018].

Author's address
Ing. Kvetoslava Surmanová, PhD.
University of Economics in Bratislava, Faculty of business informatics, Department of Operations research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
email: kvetoslava.surmanova@euba.sk

Ing. Marian Reiff, PhD.
University of Economics in Bratislava, Faculty of business informatics, Department of Operations research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
email: marian.reiff@euba.sk

ESTIMATION OF PRODUCT SUBSTITUTABILITY AND COMPLEMENTARITY

Michal Švarc, University of Economics, Prague, Czech Republic

Abstract

Deep knowledge of product relationships represents a strategic advantage for retailers. This work proposes a methodology for classifying relationships between two products based on transactional data. Firstly, data mining framework }alled association rules is used for identification of substitutes and complements. The next step is verification of the product pair relationship via estimation of products' price cross-elasticity. The methodology was applied on real dataset with retailer's transactions. The empirical results clearly show that stand-alone usage of association rules provides only pure classification for product pairs while incorporation of product price cross-elasticity helps to increase the accuracy of product pairs categorization.

Keywords: *association rules, cross-price elasticity, substitutability, complementarity*

JEL Classification: *C55, D12*
AMS Classification: *62H20, 62P20*

1 INTRODUCTION

In the center of any successful retailer stands the effective product range management. This discipline is also known as "assortment" or "ranging" and is strongly connected to category management questions like: What is the right depth and width of the product range? How does a proper pricing strategy look like? The ability to provide answers to above mentioned questions requires deeper understanding of product range and relationships between products. One of the most important type of relationships between products are their substitutability and complementarity.

The knowledge of these product properties can also help retailers in many other fields like optimization of resources fixed in stocked products, replacement of unavailable products with best possible substitutes (Corsten and Gruen, 2003) or shelf assortment optimization. (Yang and Hao, 2011) Norm Borin proposed a model which helps retailers to decide which products should be listed and how much shelf space should be allocated to those products, (Borin et al., 1994). Good knowledge of products relationships is an integral part of an e-shop recommendation system. (Isinkaye et al., 2015)

Finally, the information about products relationships can be used for evaluation of retailer's promotions activity. Temporary product price reduction is understood under the term "product promotion". One of the commonly used approach for identification of the impact of product promotion is distinguishing between four independent effects: lift in promoted item sales, halo, cannibalization, pull-forward.

The calculation of the lift in promoted product sales is a difference between the sales when the product is promoted and time frame when the product is sold for regular price. The Halo effect occurs when product promotion causes additional purchase of product complements. On the opposite side stands cannibalization. Cannibalization is a sale decrease of products which can be considered as substitutes. The so-called pull-forward effect occurs especially for products with long shelf life and is connected to temporary sales drop down of promoted products after the promotion end. (Kassam, 2012)

The main contribution of this paper lies in proposed framework for scoring products relationships like substitutes or complements based on information from transactional data. The key features for obtaining these links between products are products occurrences in the shopping baskets and estimated cross-price elasticity of demand. The biggest advantage of this approach is the fact that no additional product attributes are required and relationship between products is derived from customer's behavior data. This puts the customer's opinion in the middle of retailer decision making process and is perfectly aligned with successful customer centric business model. (Kindström, 2010)

2 LITERATURE REVIEW

Customer theory in economics says that two goods which can be described with the word substitutes are products or services that customer sees as similar to each other and could be used for the same purpose. On the opposite side stand complementary products. These products or services are usually consumed together because its separate consumption has a little or no value for customer. This fact is also described with the term "joint demand".

Above mentioned characteristics of substitutes and complements lead data-mining community to following idea. The product relationships can be derived from market basket data analysis, because the complements will be more likely together in one basket whereas two substitutes are less likely present in one basket. This approach is always described in the literature as Market basket analysis. One of the first proposals of this approach occurs in Sweden and was applied on supermarket data analysis (Julander, 1992). However, many authors cite the association rules framework for discovery connectedness of products, which was firstly presented by (Agrawal et al., 1993). This framework becomes a ground for marketing strategies focusing on cross-selling (Anand et al., 1998) as well as for product assortment optimization in retail environment (Bell et al., 1999).

One of the biggest challenges related to the classification of products basket on market basket analysis is related to the phenomenon called "horizontal variety seeking". For instance, customers' willingness to switch their consumption of some products between several alternative flavors was documented already in year 1982 (McAlister, 1982). Explanation for this variety seeking behavior may lie in the uncertainty of their taste in the future time of consumption. Another reason for this type of behavior could be varying tastes of consumers of shopping basket content. (Dubé, 2004) This assumption seems to be valid especially for large shopping baskets intended for household consumption.

Above mentioned customer's behavior can cause a situation, when products that should be considered as substitutes are classified via market basket analysis as complements.

However, economic theory defines other properties of substitutes and complements that are based on demand cross-elasticities. Formal language of economics says, that A and B are substitutes if the demand for A increases when the price of B increases, or if there is a positive cross elasticity of demand. In contrast to substitutes, goods with a negative cross elasticity of demand are considered as complements. If products A and B are complements than A's demand increases when the price of B decreases.

Definition of demand cross-elasticity value for substitutes and complements can be either used as validation tool for outputs from market basket analysis or may serve as additional condition which must be met if couple of product should be classified as substitutes or complements. The research of (Vindevogel et al., 2005) used this additional property for validation of product pairs created via association rules and their findings show that the market basket analysis cannot automatically identifies complements and substitutes. This result seems to be very important especially if identified product links are used for promotions planning.

The last important note regarding to products' complementarity and substantiality is the fact that these relationships are not necessarily symmetric. Hicks and Allen had already mentioned the possibility of cross-price effects with opposite sign in the case of two goods. (Hicks and Allen, 1934) Also the asymmetric behavior of goods can be represented by following example: product A may be a good substitute for product B if product A is similar to product B and is cheaper. In this case, product B would be a poor substitute for product A.

3 METHODOLOGY

As was already pointed out in the introduction the main aim of this paper is to provide a framework for classification of product pairs into four categories: substitutes, complements, independent products and unclassified relationship. Product pairs are considered to be an ordered arrangements of 2 products (2-permutations) due to the fact that the product A can be a complement of product B, but the product B cannot be a complement of product A.

The pool of products used for creating product pairs is represented by retailer's product range with several exclusions. Firstly, all products that are in range less than 90 days are excluded from the pool. The aim of this condition is to provide enough space for the product to interact with other products. Also, products like shopping bags, bottles deposits and beer containers are removed from the product pool either. Finally, low penetrated products are excluded from the classification framework, too. Low penetrated product means product which occurs in less than 0.001‰of all shopping baskets. Table 1 shows how the sales coverage is distributed by different threshold for product basket penetration. Chosen threshold for low penetrated product covers approximately 70% of retailer's sales. This decision also significantly reduces the number of product that are in the pool and corresponds with the first step of apriority algorithm proposed for mining of association rules. (Agrawal and Srikant, 1994)

product penetration threshold	approx. number of products (in thousands)	approx. sales coverage	theoretical count of 2-permutations
$\geq$ 0.00000000	735 k	100%	540 224 265 k
$\geq$ 0.00001000	31 k	85%	960 969 k
$\geq$ 0.00010000	10 k	70%	99 990 k
$\geq$ 0.00100000	1 k	40%	999 k

Table 1: Sales coverage by product basket penetration threshold

The entire framework for identification products' relationships can be divided into three stages which are repeatedly executed for all pairs of products. First stage crunches data from retailer's transactional database. Second stage focuses on getting association rules characteristics and the final stage analyses products' price cross-elasticities for assessing if association rules provided correct/incorrect classification of product pair.

3.1 Data preparation

As was outlined, the first step toward product classification involves mining data from retailer's transactional database. At this stage is a set of conditions which a pair of products needs to meet to be involved into the pool of classified pair. In the first place, products in pair need to be in stores assortment together at least for 8 weeks. The 8 weeks threshold was chosen for ensuring high probability of unit price change occurrence for product due to the retailer's price policy which runs promotion in weekly, fortnightly and monthly periods.

The ability to classify products' price cross elasticity fully dependends on occurrence price changes. This fact resulted in second condition related to price changes. Product pair is classified only if both products in pairs have at least two different prices (regular price and promo price). Promo price is defined as a unit price which at least 5 % percent lower than the regular unit price and the absolute change in unit price (compared to regular unit price) needs to be greater than 1 CZK. At the same time, basket penetration of the product during promo period needs to be higher than the basket penetration during the time when the product is sold for regular price. Alternatively, the regular unit price needs to be higher than the product promo price and basket penetration over the regular price period should be below the basket penetration in promo period.

Last condition requires price changes do not occur concurrently. It means, there is required presence of time periods when product A is in promotion and product B is sold for regular price or otherwise. This condition ensures that one of the products from product pair will be cheaper/more expensive for certain period for retailer's customers that may cause changes in their shopping behavior and helps to identify products' relationships.

Due to the fact that the retailer runs a customer loyalty program, there is a possibility to identify behavior of a particular customer. This helps to split the analysis into two areas. One possibility is to run the analysis on all available transaction which meets defined condition. This option has the advantage of quite robust based of transactions but there are also transactions which are made by random buyers or trialists which can bias results towards substitutes. This issue can be eliminated via inclusion of transactions made by products' heavy users (Twedt, 1964). For this purpose, the definition of heavy users of product A is based on total volume of product A bought by customers/users. The amount of products is calculated for each customer and customers are sorted by this amount in descending order. Customers labeled as heavy users meet at least one of these two conditions:

- comprises at top 1/3 of all customers of product A,
- accounts for at least 2/3 of product A's total sales volume.

3.2 Association rules

One of the tools for market basket analysis is the association rules framework that focuses on composition of products in customers' baskets. Measures like support, confidence and lift from association rules framework are very suitable for product relationships classification. The support for a product A is the fraction of transactions in dataset that contain the product A. The support measure can be also described as a probability that product A will be present in a basket. The measure called confidence describes the likelihood of buying product B if product A is in basket. The main drawback of this measure is the absence of control for popularity of product B. With the increasing popularity of product B, the value of confidence (A, B) is also increasing. This disadvantage is overcome in measure called lift. Lift says, how likely product B is present in basket when product A is in baskets, while controlling how popular product B is. Lift values around 1 would imply that product A and product B occur randomly in baskets. Lift values below 1 and close to 0 says that products A and B will be unlikely in one basket, which can lead to the assumption that these products are substitutes. Lift values over 1 may indicate that product

A and B are more likely together in one basket. This is sometimes rated as a complimentary relationship.

3.3 Price cross-elasticity evaluation

Although there are many studies that are using multivariate time-series techniques like vector autoregressive models for exploring effects of product price promotions on products sales, this paper keeps a simple approach to examining product price elasticity and product cross-elasticity. This simple approach tries to estimate relationship between sold amount of product A and unit price of product A as well as unit price of product B (described in equation 1a) or alternatively relationship between number of sold products B and unit prices of product A and B (equation 1b).

$$q_i^A = c_0 + c_1^A p_i^A + c_1^B p_i^B + u_i, \tag{1a}$$

$$q_i^B = c_0 + c_1^A p_i^A + c_1^B p_i^B + u_i \tag{1b}$$

Variables q_i^A and q_i^B stands for amount of sold products of product A and B in particular day i. Equations 1a, 1b incorporates unit prices for product A and B as p_i^A and p_i^B as modal unit price for product A and B, if the modal unit price covers at least 75% of all transactions for day i. If this condition is not met, the modal unit price is replaced by average unit price. The reason for this switching lies in retailer's and suppliers' activities like targeted discounts or different types of coupons. Usual way, how the product price elasticity and cross-elasticity of products are interpreted, is using percentages. More precisely, elasticity says what is a reaction of consumers on percental change of product price where the reaction is measured as percental change of demanded number of products. There is also a need to control other factors which may influence the demand for a product. The most important factor is the total number of visits for particular day v_i. Also the day of the week impacts the product's demand because during the working days (except Friday) customers concentrate mostly on top-up goods when they runs out of stock but during the Friday and weekend they do fuller buying which is characterized by bigger and heavier baskets. This controls are added into the model via binary variables $d_i^1 \dots d_i^6$ which indicates Tuesday till Sunday. Customers behavior is strongly biased by events like Christmas and Eastern. Both events' periods are distinguished from regular trading period by indicator variables o_i^X for Christmas period and o_i^E for Eastern. As long as the time period is continuing and longer than 90 days, there is also a control for a linear or polynomial trend $f(t)$. If the period length goes over 2 years, the seasonality controlling is enabled via quarters indicators (qt_i^1 for winter, qt_i^2 for spring and qt_i^3 for summer). Finally, products with longer expiration period, where customers tend to pull-forward, are controlled via dummy variables for a promotion weeks ($pw_i^1, pw_i^2, ..pw_i^{n(pw)}$). Index $n(pw)$ denotes the number of consecutive promotion weeks. It is expected that the first week of promotion will be stronger in terms of quantity of sold products then the following weeks.

All above mentioned factors are incorporated into equations 2a and 2b for estimating price elasticity and price cross-elasticity:

$$\begin{aligned}
log_e(q_i^A) &= c_0 + c_1^A log_e(p_i^A) + c_1^B log_e(p_i^B) + c_2 v_i + f(t) + c_4^X o_i^X + c_4^E o_i^E + \\
&+ \sum_{j=1}^{6} (c_3^j d_i^j) + \sum_{k=1}^{3} (c_5^k qt_i^k) + \sum_{l=1}^{n(pw)} (c_6^l pw_i^l) + u_i,
\end{aligned} \tag{2a}$$

$$\begin{aligned}
log_e(q_i^B) &= c_0 + c_1^A log_e(p_i^A) + c_1^B log_e(p_i^B) + c_2 v_i + f(t) + c_4^X o_i^X + c_4^E o_i^E + \\
&+ \sum_{j=1}^{6} (c_3^j d_i^j) + \sum_{k=1}^{3} (c_5^k qt_i^k) + \sum_{l=1}^{n(pw)} (c_6^l pw_i^l) + u_i
\end{aligned} \tag{2b}$$

Automated model self-diagnosis detected a presence of autocorallation in majority of models. Due to this fact, presented results are obtained Prais-Winsten estimation procedure.

4 EMPIRICAL RESULTS

The methodological proposal was verified on 2.5 years history of one Czech retailer‘s transactions which were done in stores with the size corresponding to hypermarkets. Total data volumen is approximately 235 000 k of baskets and 2 030 000 k of product items. Restriction of transactions only to retailer's hypermarkets ensures the product prices consistency, because the prices may differ by store formats. As was already mentioned in previous section, the product range was subseted only for products with basket penetration higher or equal to 0.001 ‰. This decision significantly reduced the number of product (from 735 k to 10 k) with keeping sufficient coverage of total retailer sales (approx. 70%).

basket_id	date_id	customer_id	product_id	net_spend	item_qty	unit_price
3287513	02-Apr-17	193516	4814	54.2	1	54.2
3287513	02-Apr-17	193516	4818	97.5	3	32.5
3287514	03-Apr-17	null	4251	14.5	5	2.9
3287514	03-Apr-17	null	3951	37.4	3	18.7
3287515	03-Apr-17	193516	4818	195.0	6	32.5

Table 2: Data base table structure

Base table for the entire analysis consists of these columns: basket_id (unique identifier of the basket/transaction), date_id (date of the transaction), customer_id (unique identifier of customers with retailer's loyalty card), product_id (unique id of product), net_spend (spend for all product's units in basket), item_qty (number of product's units in basket) and calculated column unit_price with product unit price. Also, baskets with some assigned discount coupons were excluded from the analysis. Table 2 shows the base table structure.

Once all conditions defined in methodological part focused on data preparation was applied on the high penetrated products pool, the number of products pairs was substantially reduced to 55.5 % of all original product pairs. The total number of product pairs which become inputs for association rules analysis is approx. 27 000 K.

The most impactful constraint is the requirement on common occurrence of products for at least 8 weeks. This condition removes product pairs with seasonal goods especially. The requirement for nonparallel price changes removes pairs of products which are usually promoted together.

These products are usually part of one product family. The figure 1 shows the evolution of product pairs by applying defined conditions.

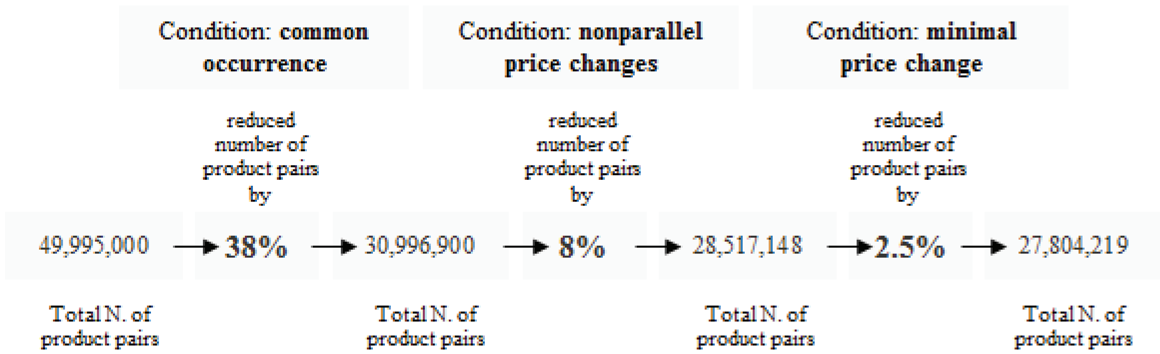


Figure 1: Effect of conditions on count of product pairs

As can be seen on figure 2, the lift values distribution is trimodal. The major number of products (63.4%) pairs lies between lift values 1 and 4. This is common for calculated lift values based on all transaction as well as for values based on product heavy users shopping behavior. The lift values obtained from the transactions restricted to heavy users tend to be more distinguishable in terms of product substitutability and complementarity, under the assumption that lift values close to null should be associated with substitutes and high lift values stand for complements.

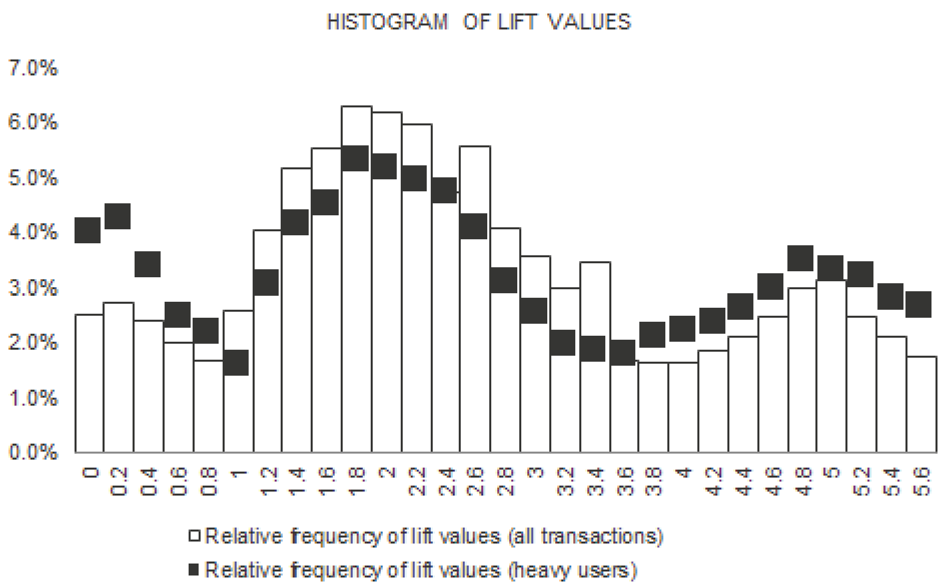


Figure 2: Histogram of lift values for all transactions and heave users

Although the association rules seem to provide an effective way for segregation, the outputs of the model for product price cross-elasticity shows different results. These results are summarized in table 3.

Products pairs were firstly categorized with the help of distribution of lift values into three buckets:

- substitute (lift values < 1)
- complements (lift values > 4)

	Percentage of products pairs in buckets		
	substitutes	no relationships	complements
Classified based on lift (calculated on all transactions)	9.6%	63.9%	26.5%
Classified based on lift (calculated on heavy users)	14.1%	50.7%	35.2%
Classified based price cross-elasticity	5.2%	90.1%	4.7%

Table 3: Percentage of products in product pairs category

- products with no relationships (lift values: <1;4>).

The product categorization based on estimated values of product price cross-elasticity coefficient and their p-value takes place in the last row of table 3. The table 3 clearly shows that the number of product pairs based on price cross-elasticity is significantly lower than via association rules and the lift measure. Moreover, only 74.3% of all product pairs categorized as complements were symmetric, it means product A is a complement for product B and product B is complement for product A. The substantiality of products reports higher symmetry, approx. 92.7%.

The next interesting finding relates to the consistency of classification is that 21.2% of all substitutes classified by product price cross-elasticity was found within product pairs classified by association rules as complements. Also, 10% of all complement sourced by cross-elasticity model was presented in substitutes from association rules analysis.

5 CONCLUSIONS

Outputs of the research proved that association rules as stand-alone technique is not able to provide reliable insight related to product complementarity and substitutability. Moreover, the association rules methodology fully ignores the fact that the relationship between products can be asymmetric. Due to both reasons, the product pairs relationships (complements/substitutes) should be classified not only with the help of presence or absence of products in baskets but also there is a need to incorporate information about price changes and changes of customers' behavior caused by these price changes. This work provides a guidance on how to estimate product pairs relationships strictly based on transactional data via applying framework for finding association rules and verifying results through the price cross-elasticity.

The space for further investigation lies in optimization of the process of generation especially, validation and classification product pairs, because the procedure for generating information about one product pair took about 1.2 seconds and the duration for getting information about 28 million of product pairs was 5 days with the usage of one hundred concurrent workers. Also, incorporation of product life cycle can contribute to the better results of product pairs classification, especially via removing launching and decommission stage of product. The same impact could have controlling of the product availability.

References

[1] R. Agrawal and R. Srikant. Fast Algorithms for Mining Association Rules in Large Databases. In *Proceedings of the 20th International Conference on Very Large Data Bases*, VLDB '94, pages 487–499, San Francisco, CA, USA, 1994. Morgan Kaufmann Publishers Inc. ISBN 978-1-55860-153-6. URL http://dl.acm.org/citation.cfm?id=645920.672836.

[2] R. Agrawal, T. Imieliński, and A. Swami. Mining Association Rules Between Sets of Items in Large Databases. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, SIGMOD '93, pages 207–216, New York, NY, USA, 1993. ACM. ISBN 978-0-89791-592-2. doi: 10.1145/170035.170072. URL http://doi.acm.org/10.1145/170035.170072.

[3] S. S. Anand, A. R. Patrick, J. G. Hughes, and D. A. Bell. A Data Mining methodology for cross-sales. *Knowledge-Based Systems*, 10(7):449–461, May 1998. ISSN 0950-7051. doi: 10.1016/S0950-7051(98)00035-5. URL http://www.sciencedirect.com/science/article/pii/S0950705198000355.

[4] D. R. Bell, J. Chiang, and V. Padmanabhan. The Decomposition of Promotional Response: An Empirical Generalization. *Marketing Science*, 18(4):504–526, Nov. 1999. ISSN 0732-2399. doi: 10.1287/mksc.18.4.504. URL http://pubsonline.informs.org/doi/abs/10.1287/mksc.18.4.504.

[5] N. Borin, P. Farris, and J. Freeland. A Model for Determining Retail Product Category Assortment and Shelf Space Allocation. *Decision Sciences*, 25(3):359–384, June 1994. URL http://digitalcommons.calpoly.edu/mkt_fac/7.

[6] D. Corsten and T. W. Gruen. Desperately seeking shelf availability: an examination of the extent, the causes, and the efforts to address retail out-of-stocks. *International journal of retail & distribution management*, 31(12):605–617, 2003. ISSN 0959-0552. URL https://www.alexandria.unisg.ch/16824/.

[7] J.-P. Dubé. Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks. *Marketing Science*, 23(1):66–81, Feb. 2004. ISSN 0732-2399. doi: 10.1287/mksc.1030.0041. URL http://pubsonline.informs.org/doi/abs/10.1287/mksc.1030.0041.

[8] J. R. Hicks and R. G. D. Allen. A reconsideration of the theory of value: part I, part II. *Economica*, 1(1):52–76, 1934.

[9] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh. Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 16(3):261–273, Nov. 2015. ISSN 1110-8665. doi: 10.1016/j.eij.2015.06.005. URL http://www.sciencedirect.com/science/article/pii/S1110866515000341.

[10] C. R. Julander. Basket analysis: a new way of analysing scanner data. *International Journal of Retail & Distribution Management*, 20(7), July 1992. ISSN 0959-0552. doi: 10.1108/09590559210022362. URL http://www.emeraldinsight.com/doi/abs/10.1108/09590559210022362.

[11] R. Kassam. Guide To Analyzing The Overall Lift Of A Retail Promotion, Nov. 2012. URL https://www.crosscap.com/retail-marketing-blog/guide-to-analyzing-the-overall-lift-of-a-retail-promotion.

[12] D. Kindström. Towards a service-based business model – Key aspects for future competitive advantage. *European Management Journal*, 28(6):479–490, Dec. 2010. ISSN 0263-2373. doi: 10.1016/j.emj.2010.07.002. URL http://www.sciencedirect.com/science/article/pii/S0263237310000538.

[13] L. McAlister. A Dynamic Attribute Satiation Model of Variety-Seeking Behavior. *Journal of Consumer Research*, 9(2):141–150, Sept. 1982. ISSN 0093-5301. doi: 10.1086/208907. URL https://academic.oup.com/jcr/article/9/2/141/1813215/A-Dynamic-Attribute-Satiation-Model-of-Variety.

[14] D. W. Twedt. How important to marketing strategy is the" heavy user"? *Journal of Marketing (pre-1986)*, 28(000001):71, 1964.

[15] B. Vindevogel, D. Van den Poel, and G. Wets. Why promotion strategies based on market basket analysis do not work (PDF Download Available). *Expert Systems with Applications*, 28(3):583–590, Apr. 2005. ISSN 0957-4174. doi: https://doi.org/10.1016/j.eswa.2004.12.019. URL https://www.researchgate.net/publication/223341118_Why_promotion_strategies_based_on_market_basket_analysis_do_not_work. DOI: http://dx.doi.org/10.1016/j.eswa.2004.12.019.

[16] Y. Yang and C. Hao. Product selection for promotion planning. *Knowledge and Information Systems; London*, 29(1):223–236, Oct. 2011. ISSN 02191377. doi: http://dx.doi.org.zdroje.vse.cz/10.1007/s10115-010-0326-8. URL https://search-proquest-com.zdroje.vse.cz/docview/895729710/abstract/E4AF6538FB4C43E6PQ/1.

Author's address
Ing. Michal Švarc
University of Economics, Faculty of Informatics and Statistics, Department of Econometrics
nám. W. Churchilla 1938/4, 130 67 Praha 3
Czech Republic
email: xsvam00@vse.cz

QUEUING SYSTEM AS A MODEL OF A SPECIFIC SMALL BUSINESS PROBLEM AND ITS SOLUTION BY SIMULATION TECHNIQUES

Milan Svoboda, University of West Bohemia in Pilsen,
Miroslav Plevný, University of West Bohemia in Pisen,
Pavla Říhová, University of West Bohemia in Pilsen

Abstract
The study deals with an analysis and optimization of a queuing system in a small enterprise operating in the service sector using simulation method. The company under review is a traditional men's hairdresser (barber) with a long history, whose business model is based on limited offer of services (dry hair cut) with the emphasis on short waiting time and no orders in advance. The aim of the study is to analyze the current model of service process, i.e. a given number of channels (operator lines) for the whole opening hours, compared to models with variable (lower than the original) number of channels. The service process parameters (customers waiting times) should not be deteriorated significantly. Random distributions of system input parameters are modeled on the basis of estimates of the number of arriving customers and the time of service provided by the enterprise. The simulation was performed by MS Excel using VBA (Visual Basic for Applications). The simulation results show that the business operations can be streamlined without significant deterioration in the service.

Keywords: *queuing system, simulation methods, MS Excel, service quality*

JEL Classification: C630, L840
AMS Classification: 68U20

1 INTRODUCTION

Queuing theory is a widespread discipline used in mathematical analysis for modelling and determining characteristics of processes in which there are units (requirements) entering the system and queuing until they are served. The occurrence of these units is random. As a result of a limited number of operator lines (service channels) there may be a problem consisting in accumulation of requirements and formation of queues. These processes may be modelled by means of queuing models. It is usually necessary to find such a state of a given model as seen from the point of the number of operator lines in which the requirements for the service do not wait too long and, at the same time, the operator lines are sufficiently busy.

Simpler models of queuing, where the entry of units into the system and the length of the duration of their operation can be described by means of some basic random distributions, can be solved analytically. This means that, on the basis of the known values of the input parameters of a particular model of queuing, it is possible to determine the output characteristics which describe the behaviour of such a model. Analytical approaches to the systems of queuing are described, for example, in Ross (2010).

For more complex models or for models described by means of less common random distributions it is more suitable to estimate the output characteristics on the basis of a simulated passage of units through the modelled system. The description of some selected simulation techniques used in the queuing systems is presented, for example, by Stewart (2009).

A number of studies related to the application of queuing models for solving diverse issues have been published. As examples of the above applications it is possible to mention their use in the field of telecommunication networks (De Boer 2000) or in banking (Madani 2013). A number of studies use queuing models for solving issues related to improving customer service in the field of retailing and public services (Xian et al. 2016), (Xing et al. 2015).

The aim of the study presented by our team is to continue in the already presented works and to show possible economic benefits of the application of the above mentioned techniques even in case of a relatively small enterprise with a very simple system of operation where it might seem that the already well established management system cannot be improved (Vacek et al. 2011). The subject under review is an analysis of the effectiveness of operation in a small enterprise operating in the service sector, particularly in a men´s hairdresser shop, with the benefit of applying the solution of a queuing model by means of simulation.

2 THE ANALYSED ENTERPRISE

The enterprise for which the queuing model is used in this study is a traditional men´s hairdresser´s with a long history in the centre of town. The main advantage of its central location is a big number of potential customers moving in its proximity on a daily basis. These potential customers either work directly in the town centre, possibly in one of a big number of shops, municipal authorities, banks or offices or they belong to their customers. A lot of stops of numerous lines of municipal transport can also be found near the hairdresser´s. The disadvantage of the location of the shop in the centre is, on the contrary, the lack of car parks in the close vicinity and related high parking rates. The business model of the hairdresser´s is based on a limited offer of services (dry hair cut), on the system of no orders in advance, lower prices and short waiting times. The hairdresser´s is open from 7 a.m. to 8 p.m. only on weekdays, which is the fact related to its location.

2.1 Characteristic of the service

In the hairdresser´s there are five hairdressing chairs. These chairs are served in two shifts, always by five hairdressers. The working hours of the hairdresser´s are seven hours, the beginning is 15 minutes before the start of the operation and the end is 15 minutes after the end of the operation. The first shift works from 7 a.m. to 1:30 p.m. and the second shift from 1:30 p.m. to 8 p.m. The time of the operation of each shift therefore amounts to 6.5 hours. The time of serving one customer is understood as the total time which a hairdresser devotes to that particular customer. The time consists of the following activities: calling on a customer whose turn it is, accommodating the customer in the chair, protecting customer´s clothes, agreeing on the type of haircut, the haircut itself and potential corrections in case the customer is not satisfied, removing the protection of clothes and removing the cut off hairs (brushing down), issuing the receipt and receiving the payment (only cash payments are received) and sweeping the work area. The time of serving one customer depends on a lot of factors: the chosen haircut, the density and length of hair, the technique of cutting (scissors or electric trimmer), the speed of the particular hairdresser, corrections of the haircut and possibly also on additional requirements, for example beard trimming. The service provider estimates that the time of serving one customer is most often 12 minutes, the shortest estimated service time is 9 minutes and the maximum service time is 18 minutes.

2.2 Customers´ entries

There is no electronic system in use in the shop that would allocate the customer an order number. The exact times of customers´ entries in the shop are not known. A daily number of the served customers is known from the accounting documents. According to the provider the number of the served customers may vary from day to day. The number of customers served in one day may vary from 140 to 200, the average number being around 170 customers. The factors influencing the number of customers are as follows: the weather - if the weather is nice, there are more customers than on rainy days. The distribution of the number of customers during the day is not even. There are three local peak periods during the day: the first one approximately between 7:30 a.m. and 8:30 a.m., the second one between 11:30 a.m. and 12:30 p.m., and the third one between 3 p.m. and 5 p.m. The above local peak periods are related to the type of customers of the hairdresser´s. The first type represents customers working on clerical positions, where they start at 9 o´clock in the morning, or possibly customers waiting for the nearby shops to open. The second peak period comes in the time of the lunch break and some customers may use it for visiting the hairdresser´s. The last peak period relates to the end of the opening hours or the closing of the nearby shops. On the basis of the empirically gained data the below Figure 1 can characterize the average number of entering units in the individual 30 minute intervals.

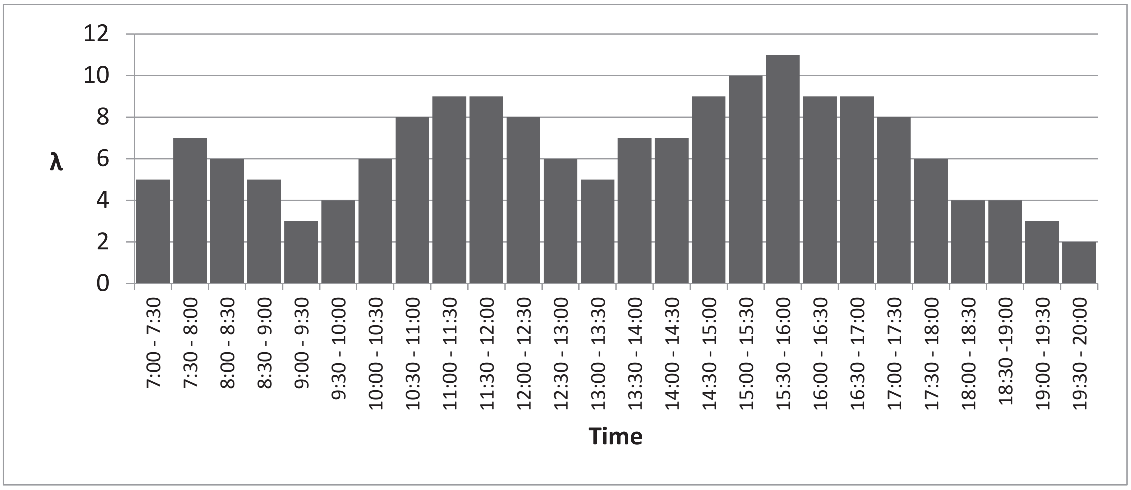


Figure 1: Intensity of customers´ arrivals

3 MODEL FORMATION

Based on the provided information we will model the operating time by a triangular distribution. The intervals between the arrivals of two customers will be modelled by means of exponential distribution with a variable mean value, which is a standard way of modelling in similar cases. Both the distributions, including their characteristics, are described, for example in (Dlouhý et al. 2007).

The triangular distribution *Tri(a,b,c)* with parameters *a, b, c* is recommended in cases where there are no particular data available but we know the minimum value (parameter *a*), the most frequent value (parameter *b*) and the maximum value (parameter *c*). In our case we have a distribution with parameters *Tri*(9, 12, 18). The mean value of the triangular distribution is given by the relation (1):

$$E(X) = \frac{a+b+c}{3} \qquad (1)$$

By substituting to relation (1) we will gain the mean value of customer service time which equals 13 min. One operating channel thus serves 4.6 customers in one hour on average. It can be easily calculated that one operating channel could serve, on average, 59.8 customers in a thirteen hours long shift.

The exponential distribution has the only parameter λ which is interpreted as the intensity of entries, i.e. the number of entries over a time unit. The reciprocal value $1/\lambda$ is the mean value of this distribution, i.e. the average time between the arrivals of two customers. In our model the value λ is variable and dependent on time, according to Figure 1. The waiting space will be considered as unlimited and all the customers as patient, i.e. they will be waiting even if the queue is long. Here we had to simplify a little, as in the real life some customers would most probably not enter the shop if they saw a long queue.

Owing to the above mentioned characteristic of the variability of the intensity of inputs we will solve the task by means of simulation. The simulation model was programmed in the environment of MS Excel by means of VBA language. One simulation run models (simulates) the course of one working day in the shop. It is a simulation with a final time horizon. The simulation run is started at 7 a.m. and finished at 8 p.m. In the serving system we are interested in the following characteristics: number of entries, average queue size, total waiting time in the queue, average waiting time in the queue, total waiting time for the customer, waiting time for the customer in percentage.

4 EXPERIMENTAL RESULTS

Overall 3 models of the above described system of queuing were analysed. For each model 200 repetitions of a simulation run were carried out. On the basis of these repetitions we determined the average value, the standard deviation, and the 5^{th} and 95^{th} percentile.

4.1 Model 1 - The current model of service

First we will analyse the current state of service, i.e. the 5 channel system of service for all the time of opening hours. The results of the simulation are given in Table 1.

Table 1

	Average	St. dev.	5^{th} percentile	95^{th} percentile
Number of entries	170.78	13.80	147.00	192.10
Average queue size	0.42	0.36	0.09	1.27
Total waiting time in the queue	5:37:36	4:51:53	1:13:41	17:12:36
Average waiting time in the queue	0:01:55	0:01:33	0:00:28	0:05:23
Total waiting time for the customer	28:00:34	2:56:48	23:10:21	33:04:49
Wait time for the customer %	43.1%	4.5%	35.7%	50.9%

As is obvious from the results, the current system of service is very helpful to customers. The overall time spent by customers waiting to be served is 5.62 hours on average, and it is much shorter than the total time of waiting for the customer, which is 28.01 hours on average. The average time spent by customer in the queue is 1.92 minutes with the standard deviation of 1.55. A possible course of the development of the waiting time and the queue size is shown in Figure 2 and Figure 3.

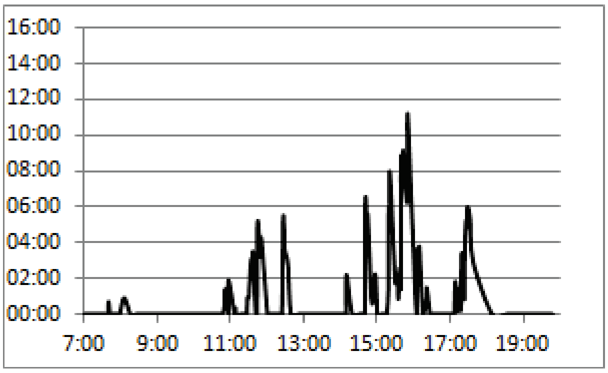


Figure 2 Waiting time (minutes)

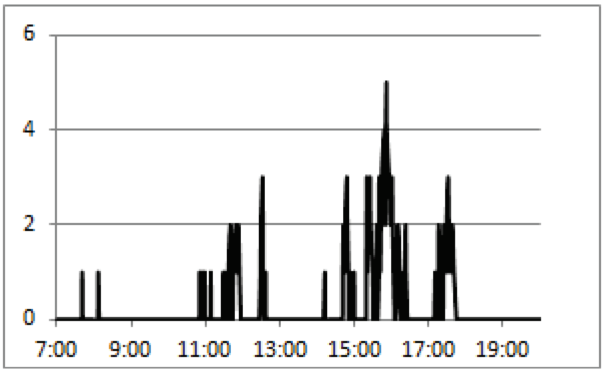


Figure 3 Queue size

The graphs in Figures 2 and 3 can be considered as typical because the longest queue and waiting time is between 3:00 p.m. and 4:00 p.m. when the biggest medium intensity of operation is 0.91 and it is calculated as (10 + 11)/(5 * 4.6). The high intensity of operation further occurs in the time periods from 10:30 a.m. to 12:00 and from 2:30 p.m. to 5:30 p.m.

4.2 Model 2

The second model is based on the analysis of the first model (current system of service) with a high level of ineffectiveness (idle time) of the service staff and at the same time with a high medium intensity of operation in the time between 3:00 p.m. and 4:00 p.m. which does not allow to reduce the number of the hairdressing chairs in this time period. If we, at that moment, reduce the number of the hairdressing chairs by one, the medium operation intensity would be 1.14, which would cause a very significant growth of queues. Therefore, there is a possibility of using the variable service system. Overall 4 channels will be available for all the opening hours. The fifth channel will only be available in the peak time, from 11:00 a.m. to 5:30 p.m. The results of the simulation are given in Table 2.

Table 2

	Average	St. dev.	5th percentile	95th percentile
Number of entries	170.13	14.45	147.00	194.05
Average queue size	0.55	0.51	0.15	1.23
Total waiting time in the queue	7:06:38	6:37:20	2:00:35	16:03:13
Average waiting time in the queue	0:02:25	0:01:58	0:00:49	0:05:02
Total waiting time for the customer	21:37:14	3:12:25	16:03:25	26:43:26
Waiting time for the customer %	37.0%	5.5%	27.4%	45.7%

As is obvious from the results, the average waiting time in the queue increased to 2.42 minutes, with the standard deviation of 1.96 minutes, which can still be considered a satisfactory time of waiting. The time of waiting for the customer also decreased but 37% of staff´s idle time is still high. A possible course of the development of the waiting time in the queue and the queue size is shown in Figure 4 and Figure 5.

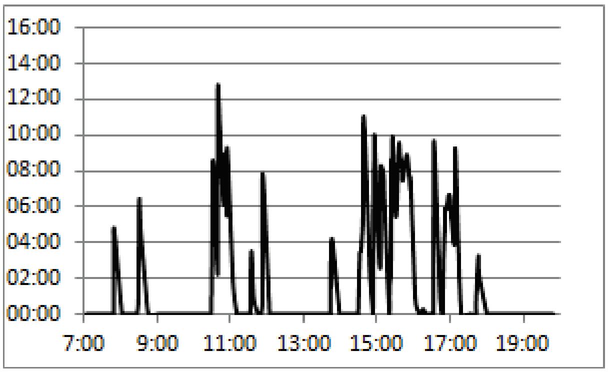


Figure 5 Waiting time (minutes)

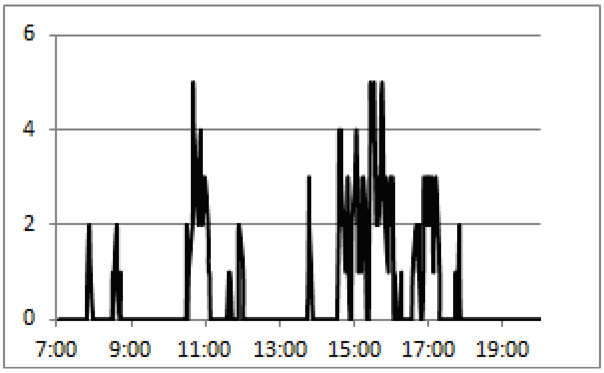


Figure 6 Queue size

The figures can be considered typical again. The waiting time in the period from 7:00 a.m. to 11:00 a.m. increased, which is the result of the decrease by one hairdressing chair. The decrease by one service channel made itself felt more in the time from 7:00 a.m. to 11:00 a.m. than in the time between 6:00 p.m. and 8:00 p.m., which can be explained by a higher intensity of the customers arriving in the morning time as is obvious from Figure 1. Despite the decrease of the service by one hairdressing chair the medium operation intensity in the morning peak time (from 7:30 a.m. to 8:30 a.m.) is at an acceptable level of 0.71, which, together with the course of the development of the waiting time, makes us think of reducing the number of the hairdressing chairs by another one. The third model is based on these considerations.

4.3 Model 3

In this model there will be three operation channels available for all the opening hours. The fourth channel will be available in the time from 10:30 a.m. to 5:00 p.m. and the fifth one will be available from 11:00 a.m. to 5:30 p.m. The total service time will then be 52 hours, which, opposed to model 1 (the existing service system), means a decrease of 13 hours, i.e. by 20%. The results of the simulation are shown in Table 3.

Table 3

	Average	St. dev.	5th percentile	95th percentile
Number of entries	170.48	11.87	151.95	188.00
Average queue size	0.84	0.59	0.27	2.10
Total waiting time in the queue	10:58:16	7:56:21	3:26:31	28:32:10
Average waiting time in the queue	0:03:46	0:02:31	0:01:20	0:09:37
Total waiting time for the customer	15:00:00	2:37:20	10:44:05	19:03:51
Waiting time for the customer %	28.8%	5.0%	20.6%	36.7%

As is obvious from the results, the average waiting time in the queue increased to 3.77 minutes, with the standard deviation of 2.52 minutes, which can still be considered a satisfactory time of waiting. The overall time of waiting for the customer decreased to 15 hours on average, from which it is obvious that the staff´s idle time decreased to 28.8%. A possible course of the development of the waiting time in queue and the queue size is shown in Figure 6 and Figure 7.

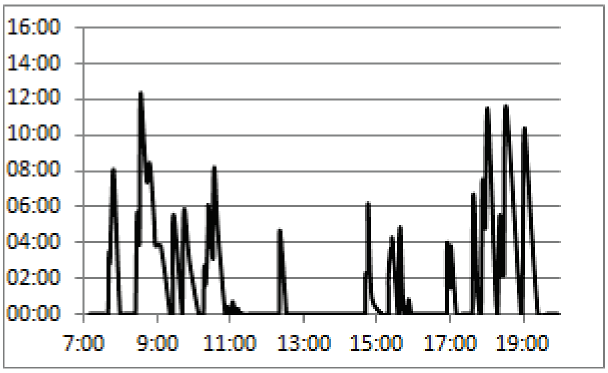


Figure 6 Waiting time (minutes)

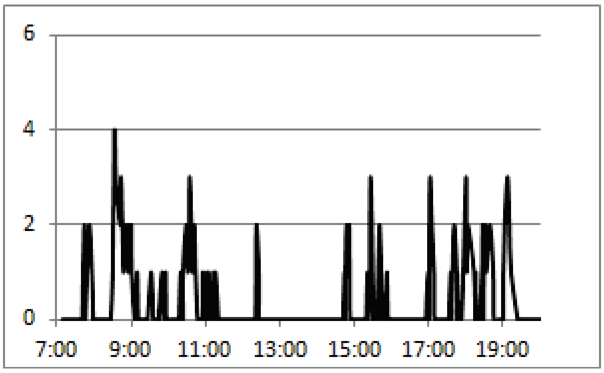


Figure 7 Queue size

The above stated figures show a possible increase of the waiting time and the queue size as opposed to the previous models in the time periods from 7:00 a.m. to 10:30 a.m. and from 5:00 p.m. to 8:00 p.m. The morning peak time between 7:30 a.m. and 8:30 a.m. may be critical, as the medium operation intensity reaches the value of 0.94. In this situation it may be assumed that the group of customers arriving in the morning peak to be served at the hairdresser´s on their way to work may not be willing to wait to be served.

5 CONCLUSIONS

The presented study showed that by the change of the service system it is possible to reduce the service costs without a significant deterioration of the customer service parameters. In case of the existing service system (Model 1) the overall average waiting time of both the service staff and the customers amounts to 33.63 hours. In case that the average waiting time of the customer is 1.92 minutes, the average idle time of the staff amounts to 43.1%. In case of the service system which was analysed by means of Model 3 there would be a 20% reduction of the service staff cost (the overall capacity of the service staff decreased from 65 hours to 52 hours) and the average idle time of the opening hours would decrease to 28.8%. At the same time the average waiting time in the queue would increase mildly by almost two minutes (from 1.92 to 3.77 minutes). The final assessment of the advantageousness of the individual variants of operation is, of course, at the discretion of the shop management. The increase of the average time of customers waiting in the queue by half a minute when using Model 2 is practically almost negligible. Even in case of using Model 3 it is still possible to assess the state as acceptable in comparison with the generally accepted waiting times of customers in similar systems. On the other hand, the economic effects as seen from the point of view of reducing costs for the shop operation may be considered as significant.

Acknowledgements: This work was supported by Grant SGS-2018-042 of the University of West Bohemia in Pilsen, Faculty of Economics.

References

[1] De Boer, P. T. (2000). *Analysis and Efficient Simulation of Queueing Models of Telecommunication Systems*. Enschede: Centre for Telematics and Information Technology.

[2] Dlouhý. M., Fábry, J., Hladík, T. and Kuncová, M. (2007). *Simulace podnikových procesů.* Brno: Computer Press.

[3] Madani, N., Roudsari, A. H., Wong, K. Y. and Galankashi, M. R. (2013). Modeling and Simulation of a Bank Queuing System. In*: 5th International Conference on Computational Intelligence, Modelling and Simulation*. [online] IEEE: CPS, pp. 209-

215. Available at: https://ieeexplore.ieee.org/document/6663187/ [Accessed 26 Feb. 2018].
[4] Ross, S. M. (2014). *Introduction to probability models*. 11th ed. Oxford: Elsevier.
[5] Stewart, W. R. (2009). *Probability, Markov Chains, queues and simulation: the mathematical basis of performance modeling*. Princeton: Princeton University Press.
[6] Vacek, J., Egerová, D. and Plevný, M. (2011). Innoskills: Innovation Guide for Small and Medium-Size Enterprises. In *10th International Conference of Liberec Economic Forum 2011*. Liberec: Technical University, pp. 538-546.
[7] Xian, T. C., Hong, C. W. and Hawari, N. N. (2016). Modeling and Simulation of Queuing System for Customer Service Improvement: A Case Study. In: *AIP Conference Proceedings*. [online] AIP Publishing, 1782(1), Article Number: 040020. Available at: https://aip.scitation.org/doi/abs/10.1063/1.4966087 [Accessed 26 Feb. 2018]
[8] Xing, W. J., Li, S. B. and He, L. H. (2015). Simulation Model of Supermarket Queuing System. In: *34th Chinese Control Conference*. [online] IEEE, pp. 8819-8823. Available at: https://ieeexplore.ieee.org/document/7261032/ [Accessed 26 Feb. 2018]

Author's address

Ing. Mgr. Milan Svoboda, PhD.
University of West Bohemia in Pilsen, Faculty of Economics, Department of Economics and Quantitative Methods
Univerzitní 22, 306 14 Plzeň
Czech Republic
email: svobodm@kem.zcu.cz

doc. Dr. Ing. Miroslav Plevný
University of West Bohemia in Pilsen, Faculty of Economics, Department of Economics and Quantitative Methods
Univerzitní 22, 306 14 Plzeň
Czech Republic
email: plevny@kem.zcu.cz

Ing. Pavla Říhová
University of West Bohemia in Pilsen, Faculty of Economics, Department of Economics and Quantitative Methods
Univerzitní 22, 306 14 Plzeň
Czech Republic
email: divisova@kem.zcu.cz

OUTPUT DYNAMIC CHARACTERISTICS IN CHOSEN ECONOMIES

Karol Szomolányi, University of Economics in Bratislava,
Martin Lukáčik, University of Economics in Bratislava,
Adriana Lukáčiková, University of Economics in Bratislava

Abstract
In comparison with developed economies such as USA, European Union or Euro Area, the relatively high proportion of output volatility in Czech and Slovak economies is driven by persistent shocks. It follows from the unobserved components model of given countries. Using the model, the output is decomposed into stochastic cyclical and stochastic trend components and their variabilities are estimated with the maximum likelihood estimator. Data series of quarterly seasonally adjusted real GDP for the period 1996 to 2017 are used. The observed results correspond to the business-cycle characteristics of post-communist countries, such as Czech Republic and Slovakia, analysed in recent studies.

Keywords: *unobserved components model, post-communist economies, cyclical and trend components of output*

JEL Classification: A10, C20, E32
AMS Classification: 62P20

1 INTRODUCTION

Analysing an economic development of post-communist countries integrated into the European Union we consider a hypothesis that long-run real shocks have been a significant source of business cycles in these economies during last two decades. Szomolányi, Lukáčik and Lukáčiková (2017) measured business cycles of the countries and measured business cycle characteristics are compared with the same characteristics in countries around the world. Authors state that even if European post-communist countries are emerging market economies, their short-run economic performance do not correspond to the observed business cycles of other global emerging markets. The business cycles of the studied countries are longer, and their recessions are more pronounced. Moreover, economic activity in the studied countries is volatile, and the trade balance and government purchases have a relatively significant countercyclical character. On the other hand, the consumption is relatively more volatile in European post-communist countries than in developed countries, even if it is not as much volatile in European post-communist countries as it is in emerging countries.

A possible explanation is provided by the real business cycle theory. Uribe and Schmitt-Grohé (2017) highlight a theoretical principle according to which the more persistent the productivity shocks are, the more volatile consumption is and thus the more counter-cyclical the trade balance is. The hypothesis of long-run real shocks in European post-communist countries matches with business cycles observations in these countries.

Szomolányi, Lukáčiková and Lukáčik (2011) as well as Furková and Chocholatá (2016) show that post-communist countries were in a transitional state in the period 1997-2009 and that they were converging to a higher level expressed by steady state GDP per capita. The steady state was similar for European Union countries. In other words, post-communist European Union members may be as rich as Germany several decades from now. It follows from neoclassical growth theory that as post-communist countries have been in a transitional state, their steady state has changed (risen) dramatically. This change stems from events such as economic system transformation, opening up to international markets, price liberalization, privatization, tax reformation and legal system transformation, among others. All these events have been present in some form in all post-communist countries during the period studied. Moreover, there were currency transformations in many countries. As explained by neoclassical growth theory, events causing changes in steady states are themselves mostly caused by permanent real shocks. Indeed, the afore-mentioned events have the character of permanent real shocks, which may explain why post-communist business cycles last longer than those of other countries. König and Radvanský (2015) forecasted a short-run Slovak economic development.

The aim of the paper is to verify the hypothesis of an importance of persisted real shocks in the business cycles of Slovakia and Czech Republic, two of the European post-communist country group.

The long-term effects on business cycles can be measured by the unobserved components model of Clark (1987). Čížek (2016) realised a similar study using Czech data. Author measures relatively large variability of long-run output in Czech Republic during last two decades. Surmanová and Furková (2011) summarised many methodologic tools decomposing data series into the cyclical and trend components. However, not many of them allow measuring the variability of trend component.

In the paper the unobserved components model will be used to measure long-term effects on business cycles in Slovakia, Czech Republic, European Union and USA during last two decades. The variability of long-run output will be compared in studied countries. If the hypothesis of long-run real shocks in Slovakia and Czech republic is relevant than we predict relatively high variability of long-run output in these countries in comparison with the same measure in steady state countries as USA or European Union.

2 MODEL

Analysis of univariate time series starts from the assumption of independence its components. It can be shown that time series y_t as sum of its components y_t^i is described by dynamic linear model with normally distributed disturbances:

$$y_t = \mathbf{F}\boldsymbol{\theta}_\mathbf{t} + v_t, \quad v_t : N(0,V) \tag{1}$$

$$\boldsymbol{\theta}_\mathbf{t} = \mathbf{G}\boldsymbol{\theta}_\mathbf{t-1} + \mathbf{w}_\mathbf{t}, \quad \mathbf{w}_\mathbf{t} : N(\mathbf{0},\mathbf{W}) \tag{2}$$

where $\boldsymbol{\theta}_\mathbf{t}$ is vector, $\mathbf{F}$ is row vector and $\mathbf{G}$ and $\mathbf{W}$ are block diagonal matrices.

The unobserved components model equations by Clark (1987) or Petris, Petrone and Campagnoli (2009) are:

$$y_t = y_t^{LR} + y_t^C \tag{3}$$

$$y_t^{LR} = y_{t-1}^{LR} + \delta_t + \varepsilon_t \tag{4}$$

$$\delta_t = \delta_{t-1} + z_t \tag{5}$$

$$\phi(L)\, y_t^C = u_t \tag{6}$$

where y_t is the logarithm of the output, y_t^{LR} represents the long-term component and y_t^C is the cyclical component and ϕ(L) is a finite polynomial in the lag operator L.

The component y_t^{LR} follows a linear growth model observed without error. The state vector is $\boldsymbol{\theta}_\mathbf{t}^\mathbf{LR} = (y_t^{LR}, \delta_t)^\mathrm{T}$, with innovation vector $\mathbf{w}_\mathbf{t}^\mathbf{LR} = (\varepsilon_t, z_t)^\mathrm{T}$. The observation matrix and observation variance are $\mathbf{F^{LR}} = [1\ 0]$ and $V^{LR} = 0$, while the system evolution matrix and innovation variance are $\mathbf{G^{LR}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{W^{LR}} = \text{diag}[\sigma_\varepsilon^2, \sigma_z^2]$. The error terms ε_t and z_t are interpreted as shocks to the output level and to the output growth rate, respectively.

The component y_t^C is described by an autoregressive model of the p-th order with the state vector $\boldsymbol{\theta}_\mathbf{t}^\mathbf{C} = (y_t^C, \theta_{t,2}, \dots, \theta_{t,p})^T$ and innovation vector $\mathbf{w_t^C} = (u_t, 0, \dots, 0)^T$. For example, if we choose AR(2) process, the observation matrix and observation variance are $\mathbf{F^C} = [1\ 0]$ and $V^C = 0$ and evolution matrix innovation variance are $\mathbf{G^C} = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix}$ and $\mathbf{W^C} = \text{diag}[\sigma_u^2, 0]$.

The higher order of the autoregressive process allows the departures from the trend, to have a dumped cyclic autocorrelation function, which is often observed in economic time series.

The matrices of the resulting model (in case of AR(2) process) are:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix},\ \mathbf{G} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \phi_1 & 1 \\ 0 & 0 & \phi_2 & 0 \end{bmatrix},\ V = 0,\ \mathbf{W} = diag\left(\sigma_\varepsilon^2, \sigma_z^2, \sigma_u^2, 0\right) \qquad (7)$$

The resulting model has five unknown parameters: σ_ε^2 and σ_z^2 for the trend component, and ϕ_1, ϕ_2 and σ_u^2 for the AR(2) component. The maximum likelihood estimator is used to estimate the parameters.

The parameters of our interest are variabilities σ_ε^2, σ_z^2 and σ_u^2. The first two of them are variabilities in the trend component, while the third one is variability in the autoregressive component interpreted as the variability of temporary cyclical component.

3 DATA

We use the seasonally adjusted time series of natural logarithms of real GDP in Czech Republic, Slovakia, European Union (28 countries), Euro Area and USA. Quarterly data for the period 1996 to 2017 are gathered from EUROSTAT database and Federal Reserve Economic Data (real seasonally adjusted U.S. GDP). The last published observation of the U.S. GDP is in the 3rd quarter, 2017. Data series gathered from the EUROSTAT were seasonally adjusted with the X12 procedure.

4 RESULTS

We choose convenient values for the initial mean of the states of the trend component, with a fairly large variance, and use the default initial values in the used software function for the autoregressive component. We also take a very small value for V, which can be considered zero for all practical purposes.

The MLE estimates for the United States (USA) are $\hat{\sigma}_\varepsilon = 0,00135$, $\hat{\sigma}_z = 0,00265$, $\hat{\sigma}_u = 0,00348$. The estimated autoregressive parameters satisfy the stationarity conditions.

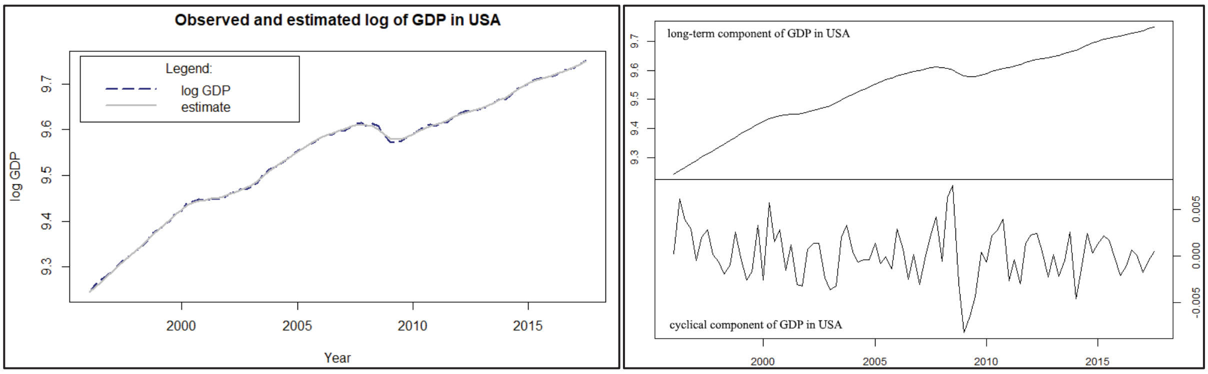


Figure 1: Observed and estimated log of GDP in USA and both components of log of GDP

The MLE estimates for the European Union (EU) are $\hat{\sigma}_{\varepsilon} = 0,00066$, $\hat{\sigma}_{z} = 0,00378$, $\hat{\sigma}_{u} = 0,00412$. The estimated autoregressive parameters satisfy the stationarity condition.

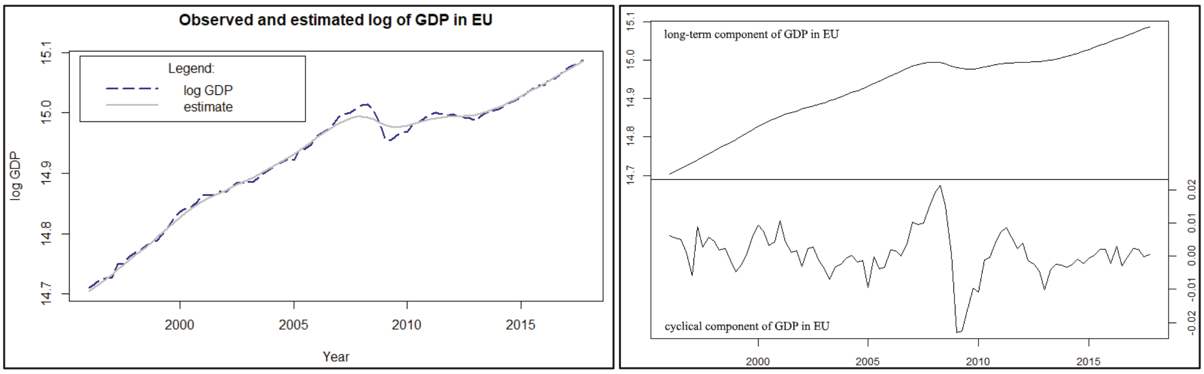


Figure 2: Observed and estimated log of GDP in EU and both components of log of GDP

The MLE estimates for the Euro Area (EA) are $\hat{\sigma}_{\varepsilon} = 0,00200$, $\hat{\sigma}_{z} = 0,00343$, $\hat{\sigma}_{u} = 0,00580$. The estimated autoregressive parameters satisfy the stationarity condition.

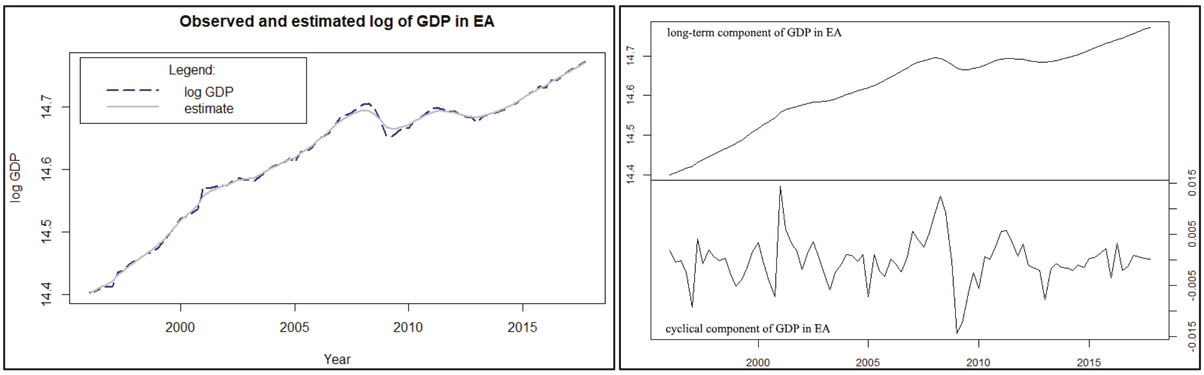


Figure 3: Observed and estimated log of GDP in EA and both components of log of GDP

The MLE estimates for the Czech Republic (CZE) are $\hat{\sigma}_{\varepsilon} = 0,00801$, $\hat{\sigma}_{z} = 0,00291$, $\hat{\sigma}_{u} = 0,00127$. The estimated autoregressive parameters satisfy the stationarity condition.

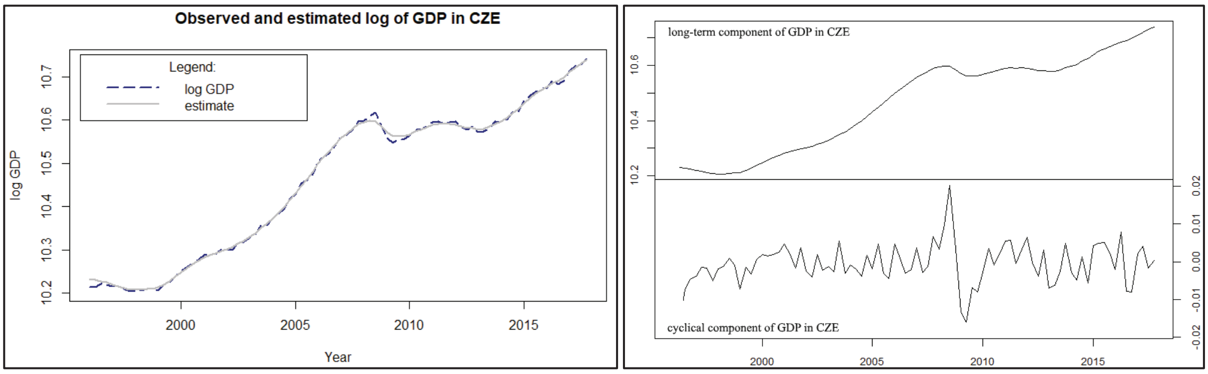


Figure 4: Observed and estimated log of GDP in CZE and both components of log of GDP

The MLE estimates for the Slovakia (SVK) are $\hat{\sigma}_{\varepsilon} = 0,01714$, $\hat{\sigma}_{z} = 0,0007$, $\hat{\sigma}_{u} = 0,00339$. The estimated autoregressive parameters satisfy the stationarity condition.

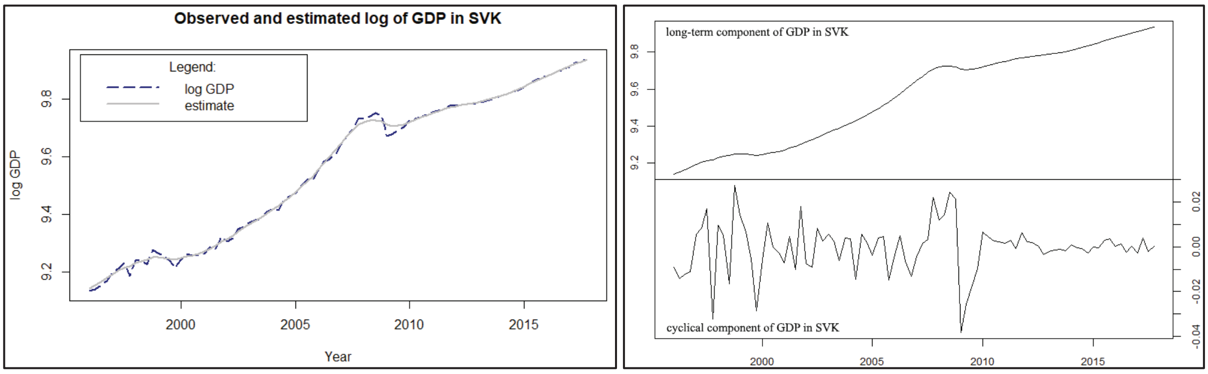


Figure 5: Observed and estimated log of GDP in SVK and both components of log of GDP

The Table 1 sums up the results. In the table, the volatilities of the output level ($\hat{\sigma}_{\varepsilon}$), output growth ($\hat{\sigma}_{z}$) and temporary cycle shocks ($\hat{\sigma}_{u}$) are displayed. In the next column the whole output volatility ($\hat{\sigma}$) is computed as the sum of the three volatilities. In the last column are the shares of temporary cycles on the whole volatility in each country (*sh*). The results are in percents (as we use the log specification).

	$\hat{\sigma}_{\varepsilon}$	$\hat{\sigma}_{z}$	$\hat{\sigma}_{u}$	$\hat{\sigma}$	*sh*
CZE	0,801	0,291	0,127	1,220	10,421
SVK	1,714	0,070	0,339	2,122	15,958
EU	0,066	0,378	0,412	0,856	48,127
EA	0,200	0,343	0,580	1,123	51,647
USA	0,135	0,265	0,348	0,749	46,533

Table 1: Volatility decomposition of the output

By the results, we confirm that Slovak and Czech economy are rather more volatile supporting the openness of both countries. The persistent shocks participate in the whole output volatility in a higher degree absolutely as well as relatively in both countries. The share of temporary cycles on the whole volatility is approximately one half in big developed countries (USA, EA, and EU). This result corresponds to Clark's (1987) observation, even if we estimate slightly higher importance of the growth output component in it whole volatility. This result may be caused by the studied, more turbulent period including financial crisis, monetary expansions, American health reforms and European debt crisis influencing the steady state and so growth of developed countries. Interestingly, the Euro Area is rather volatile, even if the temporary cycles have a large share on the output volatility in the Euro Area.

5 CONCLUSIONS

In our paper, using the unobserved data model, we confirmed that Czech and Slovak output has been influenced by the long-run shocks.

In comparison with USA and European Union, Slovak and Czech economies are highly volatile and the volatility is driven by rather more persistent output shocks.

However, the methodology used is not without problem. As Clark (1987) himself states, innovations in trend and cycle components are correlated. Weber (2011) discussed the problem citing the relevant literature and suggested the possible solutions. Our study is a first-look use of the unobserved data model to decompose the output into the cycle and trend variability. Possible improvements of the study will be considered in the next research.

Acknowledgements: This work was supported by The Grant Agency of Slovak Republic - VEGA, grant no. 1/0294/18, "Short-run and long-run dynamic analysis of economic development of European post-communist countries and regions".

References

[1] Clark, P. K. (1987). The Cyclical Component of U.S. Economic Activity. *The Quarterly Journal of Economics*, 102(4), pp. 797-814.

[2] Čížek, O. (2016). Long-Run Growth in the Czech Republic. In: *34th International Conference on Mathematical Methods in Economics 2016*. Liberec, Czech Republic: Technical University of Liberec, pp. 143-148.

[3] Furková, A. and Chocholatá, M. (2016). Spatial econometric modelling of regional club convergence in the European Union. *Ekonomický časopis*, 64(4), pp. 367-386.

[4] König, B. and Radvanský, M. (2015). Strednodobá prognóza vývoja ekonomiky SR v rokoch 2015-2018. In: *Pohľady na ekonomiku Slovenska 2015.* Bratislava: Slovenská štatistická a demografická spoločnosť, pp. 45-55.

[5] Petris, G., Petrone, S. and Campagnoli, P. (2009). *Dynamic Linear Models with R*. New York: Springer.

[6] Szomolányi, K., Lukáčik, M. and Lukáčiková, A. (2017). Business Cycles in European Post-Communist Countries. *Contemporary Economics*, 11(2), pp. 171-186.

[7] Szomolányi, K., Lukáčiková, A. and Lukáčik, M. (2011). The β-convergence of the EU27 countries. In*: 29th International Conference on Mathematical Methods in Economics 2011*. Jánska Dolina, Slovakia, Praha: Professional Publishing, pp. 287-291.

[8] Surmanová, K. and Furková, A. (2011). Methods for output gap detection. In*: 29th International Conference on Mathematical Methods in Economics 2011*. Jánska Dolina, Slovakia, Praha: Professional Publishing, pp. 275-280.

[9] Uribe, M. and Schmitt-Grohé, S. (2017). *Open economy macroeconomics*. Princeton: Princeton University Press.

[10] Weber, E. (2011). Analyzing U.S. Output and the Great Moderation by Simultaneous Unobserved Components. *Journal of Money, Credit and Banking*, 43(8), pp.1579-1597.

Author's address

doc. Ing. Karol Szomolányi, PhD.
University of Economics in Bratislava, Faculty of Economic Informatics, Department of Operations Research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
email: karol.szomolanyi@euba.sk

doc. Ing. Martin Lukáčik, PhD.
University of Economics in Bratislava, Faculty of Economic Informatics, Department of Operations Research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
email: martin.lukacik@euba.sk

Ing. Adriana Lukáčiková, PhD.
University of Economics in Bratislava, Faculty of Economic Informatics, Department of Operations Research and Econometrics
Dolnozemská 1/b, 852 35 Bratislava
Slovak Republic
email: adriana.lukacikova@euba.sk

FACILITY LOCATION PROBLEMS WITH SEMI-FIXED COSTS

Dušan Teichmann, VŠB – Technical University of Ostrava
Alexandra Dvořáčková, Czech Technical University in Prague
Denisa Mocková, Czech Technical University in Prague

Abstract
Facility location problems belong to a group of basic tasks that are solved within distribution logistics. Previously published models consider investment costs that are incorporated into the total costs of the distribution system but do not take into account the volume of goods that are distributed via the warehouse. However, in practice this simplification is incorrect. It is usual for investment costs to be stepped and linked to the volume of goods that go through the warehouse. In this article, a model of such a distribution system is presented and selected results of calculation experiments carried out in order to test the functionality of the proposed mathematical model are demonstrated.

Keywords: *mathematical programming, linear programming, facility location problem*

JEL Classification: C61
AMS Classification: 90C05

1 INTRODUCTION – MOTIVATION TO SOLVE THE PROBLEM

Facility location problems play an important role in distribution logistics. Their objective is to optimize positions of warehouses, which are to be placed in predefined locations in a given territory and to create attraction zones for them. The most frequently selected optimization criterion is the total running costs of the system.

Usually, these costs comprise of two parts. The first part includes fixed costs – the cost of building the storage space or renting it (see, for example, (Mlčák, 2017)) while the second part includes transport costs. Transport costs are the costs of transporting goods from the warehouse to the customers. In principle, the models of facility location problems do not allow the changing of the fixed costs for individual locations (the costs may vary for different locations but they do not change for one particular location). This theoretical assumption can be very restrictive in practice.

As a matter of fact, there are two possibilities when deciding about new warehouses. A new warehouse can either be built (or leased) in various locations or an existing warehouse can be extended and remodelled, see the study (Novotný, 2017), etc. For that reason, it is advisable to focus on adjusting the model of the facility location problem in such a way as to make the fixed costs variable at certain times. Economic literature, e.g. (Taylor, Garret and Iseki, 2000), describes the fixed costs with the possibility of a step change as the so-called semi-fixed costs. For the facility location problem, semi-fixed costs can be characterized as the costs of construction or lease of the warehouses in possible locations that are likely to change. Last but not least, the factor that affects the amount of fixed costs should also be mentioned. For instance, the fixed costs may vary depending on the volume of goods that are to be distributed through the given warehouse.

The presented article contains a proposal of a mathematical model that incorporates the semi-fixed costs, which will be stepped when certain flows of goods through the said warehouse are exceeded.

2 STATE OF THE ART

The first attempts to characterize a discrete facility location problem, as it is known today, and as characterized, for example, in (Ulukan and Demircioğlu, 2015) were made between 1960 and 1970. The facility location problems can be classified according to several criteria.

The first criterion is the continuity of a set of potential locations for warehouses. From the perspective of this criterion the following problems are identified:

- discrete facility location problems – the number of possible locations of warehouses within the proposed transport network is finite,
- continuous facility location problems – the number of possible locations of warehouses within the proposed transport network is infinite.

Other criterion is the capacity limit of the warehouses. From the perspective of this criterion the following problems are identified:

- uncapacitated facility location problems,
- capacitated facility location problems.

The third criterion can be the nature of the facilities to be located. Based on this criterion the following problems are identified:

- facility location problems – these problems are focused on the location of facilities with a positive impact on users (supply problems, etc.),
- obnoxious facility location problems – problems that are focused on the location of facilities with a negative impact on users (location of dump sites, etc.).

Summaries of findings obtained in the field of the locational analysis have been covered several times by articles published in international journals. In their research works, the vast majority of authors focus on capacitated or uncapacitated discrete location problems with positive impact of facilities on users. Out of the reference literature the works (Verter, 2011) and (Farahani and Hekmatfar, 2010) should be noted. The work Farahani, SteadieSeifi and Asgari, 2010) is focused on multi-criterial approaches to location problems. Only a small number of authors deal with location problems of facilities with a negative impact on users. For instance, this problem is addressed in (Cappanera, 1999).

Not only are the facility location problems used for supply systems but they are also used in emergency service systems as specified in (Janáček and Kvet, 2017).

3 PROBLEM FORMULATION – DEFINITION OF SETS, INPUT DATA AND VARIABLES

A set of locations I into which warehouses can be located is defined in the network together with a set of users J. For each location $i \in I$ the potential storage capacity for a certain planning period a_i is defined (provided the warehouse is to be operated there), and also the requirement b_j of each user $j \in J$ for the same planning period is defined. In addition, the unit costs of supply to the user $j \in J$ from the warehouse in the location $i \in I$ are known; these are referred to as c_{ij} further in the text. For each location $i \in I$ a set of intervals K_i is defined; these intervals represent the volume of goods distributed through this location where there are different fixed costs for the operation of the warehouse. A number of elements in the set K_i is p_i, it means $|K_i| = p_i$. It is assumed for the previously mentioned intervals that they continuously follow each other; i.e. $\langle 0; h_{i1} \rangle \cup (h_{i1}; h_{i2} \rangle \cup \ldots \cup \left(h_{p_{i-1}}; h_{p_i}\right\rangle$ and then let $h_{p_i} = a_i$. For each location $i \in I$ and interval $k \in K_i$ the semi-fixed costs f_{ik} are known. Let us assume that $f_{ik-1} < f_{ik}$ for $i \in I$ and $k \in K_i \setminus \{1\}$ (this assumption defines that that fixed costs will increase with the increasing volume of goods being distributed through the warehouse). The task is to determine a way of clearly assigning the users to the operational warehouses in order to minimize the amount of total costs of supplying users from the sources. Let us assume that the user cannot be supplied from several warehouses at the same time. In addition, let us assume that each user can be supplied from any operational warehouse.

Two groups of bivalent variables are entered into the problem. The first group includes variables that model decisions on operation of warehouses in individual locations and which will generate fixed costs for the operation of the warehouse in the location. Let us designate y_{ik} a variable modelling a decision on operating a warehouse in a location $i \in I$ with a capacity from the interval $k \in K_i$. If $y_{ik} = 1$ after completing the optimization calculation, then the warehouse in the location $i \in I$ will be operated and the volume of goods distributed through this warehouse will be from the interval $k \in K_i$. If $y_{ik} = 0$ after completing the optimization calculation, then the warehouse in the location $i \in I$ will not be operated in the interval $k \in K_i$. The second group of variables includes variables modelling the decision on assigning users to attraction zones of the operated warehouses (based on their values it will be possible to identify the attraction zones). Let us designate a variable modelling the said decision by the symbol x_{ij}. If $x_{ij} = 1$ after completing the optimization calculation, then the user $j \in J$ will be supplied from the operated warehouse in the location $i \in I$. If $x_{ij} = 0$ after completing the optimization calculation, then the user $j \in J$ will not be supplied through the location $i \in I$.

4 MATHEMATICAL MODEL

Upon introducing symbols to denote the constants and variables in the problem, the mathematical model can be described. The mathematical model is to be expressed as a linear model as follows:

$$\min f(x, y) = \sum_{i \in I} \sum_{k \in K_i} f_{ik} y_{ik} + \sum_{i \in I} \sum_{j \in J} \sum_{k \in K_i} c_{ij} b_j x_{ij} \tag{1}$$

subject to:

$$\sum_{i \in I} x_{ij} = 1 \qquad \text{for } j \in J \tag{2}$$

$$\sum_{j \in J} b_j x_{ij} \leq \sum_{k \in K_i} h_{ik} \, y_{ik} \qquad \text{for } i \in I \tag{3}$$

$$\sum_{k \in K_i} y_{ik} \leq 1 \qquad \text{for } i \in I \tag{4}$$

$$x_{ij} \in \{0; 1\} \qquad \text{for } i \in I, j \in J \tag{5}$$

$$y_{ik} \in \{0; 1\} \qquad \text{for } i \in I, k \in K_i \tag{6}$$

Function (1) represents the optimization criterion. The group of constraints (2) ensures that each user is supplied. The number of the constraints corresponds to the number of the users. Constraints (3) ensure that the amounts of goods distributed through the individual warehouses will correspond to the amount of the fixed costs of operating the warehouses. The group of constrains (3) further ensures that the storage capacities will be accepted. In addition, the group of constraints (3) provide logical links among variables from various groups. The group of constraints (3) is created for every location. Therefore, their number corresponds to the number of locations. Constraints (4) will ensure that a maximum of one fixed cost rate will be selected for each location suitable for the operation of the warehouse. The number of the constraints in this group corresponds to the number of the locations. Groups of constraints (5) and (6) define the domains of definition of individual variables in the model.

5 CALCULATION EXPERIMENT

In order to verify the functionality of the proposed model the optimization calculations were made. The calculations were performed by optimization software Xpress-IVE (FICO®, 2018).

Let us consider there are 3 locations suitable for warehouses ($|I| = 3$) and 10 users to be supplied ($|J| = 10$). Let us assume that there are four values of the fixed costs related to each location suitable for setting up a warehouse, i.e. $p_i = 4$ for $i \in I$. It can be noted that with the increasing capacity of the warehouse, the fixed costs also increase. These fixed costs are stepped when the values of capacities $h_{i1}, h_{i2}, \dots, h_{ip_i} = a_i$ are exceeded. Specific information related to the locations is given in Table 1.

Table 1: Information related to the boundary values at which the fixed costs of warehouse operation are stepped and the values of the fixed costs

	h_{i1}	h_{i2}	h_{i3}	$h_{i4} = a_i$	f_{i1}	f_{i2}	f_{i3}	f_{i4}
1	110	120	130	170	1 100	1 200	1 300	1 700
2	40	50	65	70	4 000	5 000	6 500	7 000
3	30	35	40	45	3 000	3 500	4 000	4 500

Specific information related to the unit costs of transport between origin and destination, and requirements of users are stated in Table 2.

Table 2: Unit costs of transport and requirements of users

Location \ User	1	2	3	4	5	6	7	8	9	10
1	10	2	3	4	5	6	7	8	9	7
2	11	3	5	7	9	4	6	5	3	8
3	8	9	5	6	4	3	2	9	8	16
User requirements	30	35	40	45	10	15	30	20	15	30

Values from Tables 1 and 2 were entered into the mathematical model (1) – (6). Upon solution the following results were obtained, see Tables 3 and 4. Table 3 contains the resulting values of variables x_{ij} in the optimal solution.

Table 3: Values of variables x_{ij}

Location \ User	1	2	3	4	5	6	7	8	9	10
1	1	0	1	1	1	1	0	0	0	1
2	0	1	0	0	0	0	0	1	1	0
3	0	0	0	0	0	0	1	0	0	0

Table 4 contains the resulting values of variables y_{jk} in the optimal solution.

Table 4: Values of variables y_{ik}

Location \ Variable y_{ik}	y_{i1}	y_{i2}	y_{i3}	y_{i4}
1	0	0	0	1
2	0	0	0	1
3	1	0	0	0

Table 5 contains the resulting flows of goods between warehouses and users in the optimal solution.

Table 5: Flows of goods between warehouses and users

Location \ User	1	2	3	4	5	6	7	8	9	10	Capacity of the location	% use
1	30	0	40	45	10	15	0	0	0	30	170	100
2	0	35	0	0	0	0	0	20	15	0	70	100
3	0	0	0	0	0	0	30	0	0	0	45	66.67

Figure 1 illustrates the status report on achieving the optimal solution.

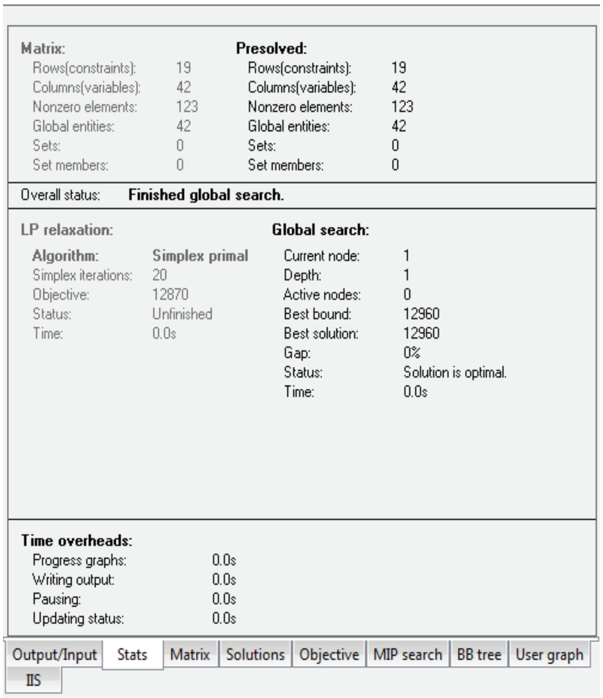


Figure 1: Status report on achieving the optimal solution

The last two columns show the maximum storage capacities of warehouses in individual locations and the percentage utilization of these capacities. Line totals indicate the numbers of units of goods distributed through individual warehouses. The proposed solution expects the warehouses to be operated in all localities while maintaining maximum utilization of their storage capacity.

In addition, the total costs equal to 12,960 monetary units in the optimal solution. Out of this amount, 11,700 monetary units are to be spent on the operation of the warehouses and 1,260 units are the transport costs.

As can be seen in Table 5, the requirements of all users will be met. The maximum capacity of any resource will not be exceeded.

In view of the fact that the storage capacities in locations 1 and 2 are used to the maximum extent possible, the operating of the warehouses in these locations generates the highest possible costs.

The optimization calculation was performed on the computer with processor 3.3 GHz, AMD FX – 8300 Eight-Core and 8GB RAM and the time of optimization calculation was 0.03 s. Its course can be seen in Figure 2.

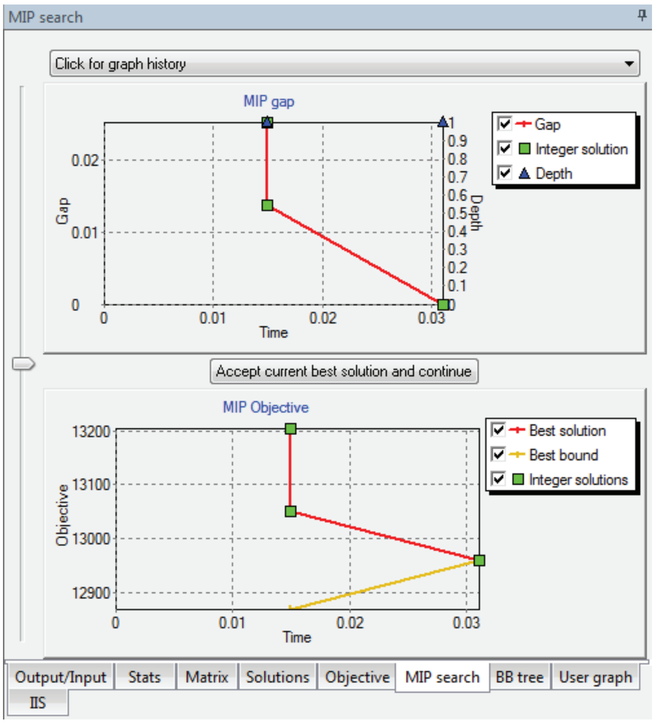


Figure 2: Course of the optimization calculation in time

6 CONCLUSIONS

The presented article is focused on location problems with respect to semi-fixed costs of the operation of warehouses. The models used are extended versions of traditional models applied in location problems for which only one fixed cost value is defined for each location suitable for operating a warehouse. Unlike traditional models of location problems, there is a higher number of bivalent variables related to the warehouse operation. There is no longer just one variable generating the fixed cost defined for each location but there are several variables that correspond to the rates of the fixed costs for the given warehouse.

The proposed model is expected to increase the potential of location facility problems so that they can be used in real-life logistic practice. Other uses of the proposed model are, for example, in a situation when it is necessary to determine whether it is more efficient to expand the capacity of an existing warehouse or to establish a new warehouse in another location.

In the future, the authors plan to focus on more extensive problems with respect to the obtained calculation times that are to be compared with the calculation times needed for solving facility location problems without semi-fixed costs. Since the number of bivalent variables in the proposed model is higher, the extended calculation times are expected due to the higher complexity of calculations.

References

[1] Cappanera, P. (1999). *A Survey on Obnoxious Facility Location Problems: Technical Report*. Pisa: University of Pisa.

[2] Farahani, R. and Hekmatfar, M. (2009). *Facility location: concepts, models, algorithms and case studies*. Springer Science & Business Media.

[3] Farahani, R., SteadieSeifi, M. and Asgari, N. (2010). Multiple criteria facility location problems: A survey. *Applied Mathematical Modelling*, [online] 34(7), pp.1689-1709. Available at: http://www.sciencedirect.com/science/article/pii/S0307904X09003242 [Accessed 16 Feb 2018].

[4] FICO®. (2018). *FICO® Xpress Optimization*. [online] Available at: http://www.fico.com/en/products/fico-xpress-optimization [Accessed 8 Feb 2018].

[5] Janáček, J. and Kvet, M. (2017). Semi-fair design of emergency service system with failing centers. *Central European Journal of Operational Research*, 25(3), pp.665-677.

[6] Mlčák, M. (2017). Logistické centrum Ostrava - Hrabová. *Stavby v MS kraji*. [online] Available at: http://www.msstavby.cz/logisticke-centrum-ostrava-hrabova-13-07-2017/ [Accessed 10 Feb 2018].

[7] Novotný, R. (2017). Výrobce textilu Adler rozšíří své distribuční centrum v Ostravě. Kapacita vzroste o 150 procent. *Hospodářské noviny*, [online] 2017(13.10.). Available at: https://archiv.ihned.cz/c1-65914750-vyrobce-textilu-adler-rozsiri-sve-distribucni-centrum-v-ostrave-kapacita-vzroste-o-150-procent [Accessed 10 Feb 2018].

[8] Taylor, B., Garrett, M. and Iseki, H. (2000). Measuring Cost Variability in Provision of Transit Service. *Transportation Research Record: Journal of the Transportation Research Board*, 1735, pp.101-112.

[9] Ulukan, Z. and Demircioğlu, E. (2015). A survey of discrete facility location problems. *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, 9(7), pp.2487-2492.

[10] Verter, V. (2011). Uncapacitated and Capacitated Facility Location Problems. In: Eiselt H., Marianov V. (eds) *Foundations of Location Analysis*. Boston, MA: Springer, pp.25-37.

Author's address

doc. Ing. Dušan Teichmann, Ph.D.
VŠB – Technical University of Ostrava, Faculty of Mechanical Engineering, Institute of Transport
17. listopadu 15/2172, 708 33 Ostrava-Poruba
Czech Republic
email: dusan.teichmann@vsb.cz

Ing. Alexandra Dvořáčková, doc. Ing. Denisa Mocková, Ph.D.
Czech Technical University in Prague, Faculty of Transportation Sciences, Department of Logistics and Management
Konviktská 20, 110 00 Praha 1
Czech Republic
email: dvoraale@fd.cvut.cz, mockova@fd.cvut.cz

THE FUZZY APPROACH TO RE-OPTIMIZATION OF WASTE COLLECTION ROUTES BASED ON REAL-TIME BIN STATUS DATA

Milorad Vidović, Gordana Radivojević, Branislava Ratković, Nenad Bjelić, Dražen Popović,

University of Belgrade, Faculty of Transport and Traffic Engineering

Abstract

Challenges associated with municipal solid waste management are constantly increasing due rapid urbanization, technological and economic development. Collection of waste is one of the most difficult operational problems that have a number of effects on collection efficiency. In this paper we propose fuzzy approach for dynamic re-optimization of collection vehicles routes in the case when waste bins are equipped with fill level sensors. The idea is to analyse waste collection network served by fleet of vehicles which simultaneously possesses bin status information, and the other vehicles status. We consider collection process realized in dynamic environment where vehicles visit waste collection bins accordingly to predefined routes with opportunity of dynamically changing routes accordingly to the current bin status. We analyse effects of possible route changes which are determined by the proposed approach. Performances of the proposed algorithm, as well as the effects of the routes change strategies are illustrated on numerical example.

Keywords: *waste collection, dynamic routing, fuzzy theory*

JEL Classification: C63, R41
AMS Classification: 90B06

1 INTRODUCTION

Solid waste analyzed in this paper represents the daily generated materials from households, and therefore correspond to waste and recyclables. Those materials are transported to a location for further processing, transfer, or disposal. Collection of solid waste is of special importance because it includes loading-unloading and hauling operations, which are, in average, responsible for 50 to 70 percent of total waste management costs. Because of that, even a small percentage improvement in the collection operation can affect a significant savings in the overall system cost (Tchobanoglous and Kreith, 2002). Since the collection process efficiency is directly related to the routing of collection vehicles, during the last decades, lot of effort has been directed to improvements of routing decisions, both in theory and practice. As an example can serve booklet (Shuster and Schur, 1974) which presents heuristics and practical instructions for routing of collection vehicles. Collection vehicle routes are usually fixed, based on long term expected waste quantities and, as such, do not respect uncertainty which is, de facto, immanent and particularly related to the waste quantity. In turn, this fact results with inabilities of finalizing predefined vehicle routes, i.e. braking off the route realization when waste quantity is larger than expected, which means that some of network nodes belonging to the route are not visited. Also, for the case when generated quantities of waste are smaller than expected, partially loaded vehicles decreases collection efficiency.

There are literally thousands of papers dealing with vehicle routing, but among them, much lesser is the number of those that respect uncertainty, and particularly uncertain demand. Teodorović and Pavković, (1996), were among the first that proposed fuzzy logic approach to solving vehicle routing problems with uncertain demand. On the other hand, recent advances in information and communication technologies (ICT), particularly waste bins fill level sensors, which provide real-time data on the waste quantity, may potentially increase collection efficiency through providing opportunity for dynamic routing of vehicles, dependent on the actual quantity of the waste. This challenging research area is relatively new, more heightened in last decade, as a result of advances in ICT, and Internet of things, although dynamic routing itself has relatively long history, which can be seen in Psaraftis et al. (2016). The most recently, Ramos et al. (2018), proposed operational management approach based on real time information to solve "The Smart waste collection routing problem", and concluded that several aspects of the problem should be further addressed.

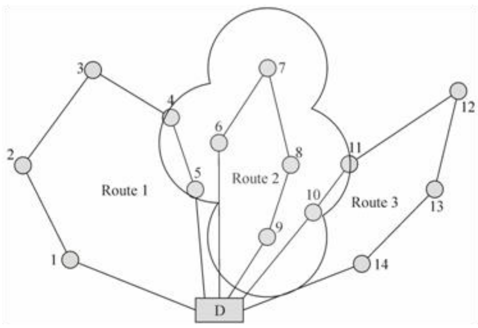
In this research we propose fuzzy approach to dynamic re-routing of collection vehicles, respecting uncertainty which is immanent in the waste collection process, even when real-time data about the bin quantities exist, because changes may happen after the route started. Remaining part of the paper is organized in following way. Section 2 describes the problem, while the modelling approach is presented in the section 3. Illustrative numerical example is given in the section 4, while the section 5 concludes the paper.

2 PROBLEM DESCRIPTION

The idea of this research is to analyse waste collection network served by fleet of vehicles which simultaneously possesses bin status information, as well as status of all vehicles which participate in collection routes realization. Collection process is considered as dynamic, with opportunity for re-routing of vehicles in accordance to the current bin status. Namely, depending on the actual quantities of the waste, generated in networks nodes, following situations, and resulting actions are possible:

- **S-I:** Actual quantities of generated waste are close to expected in all of preliminary defined routes so that collection process can be realized in accordance to the previous plan
- **S-II:** Actual quantities of generated waste are smaller than expected in all of preliminary routes, so it is possible to analyze re-routing possibilities with the objective of decreasing number of routes, i.e. used vehicles
- **S-III:** Actual quantities of generated waste are bigger than expected in all of preliminary routes, so it is necessary to exclude some network nodes from preliminary defined routes, and to construct additional routes which visits excess quantity nodes
- **S-IV:** Actual quantities of generated waste in some of network nodes are smaller, while in some other are bigger than expected, so it is possible to interchange nodes between routes, re-route vehicles, and in case of excess quantities in certain nodes, even after interchanges are performed, to construct additional routes which visits excess quantity nodes

Let N be the set of network nodes, representing collection bins, and q_i quantities of waste in bins at the moment $t=0$, when vehicles start from the depot $i=0$. Because the quantities of generated waste may vary, even during the collection process realization, in the proposed approach we use the following strategy. If q_i^T is the quantity of waste in the bin i, T time units before the time when vehicles start, then the trend coefficient of waste quantities change γ, during the collection process realization, is estimated as $\gamma = \sum_{i \in N} q_i \Big/ \sum_{i \in N} q_i^T$. The period of the T time units corresponds to the half of average time of collection routes realization, so that γq_i represents expected quantity of collected waste in the node $i \in N$. Note that we use averaged trend coefficient γ for all network nodes because it is intuitively clear that the re-routing possibilities can be utilized only in smaller collection zones with relatively close network nodes where, consequently, similar general trend in waste quantities changes can be expected. The reason of limiting collection zones in analysis of dynamic re-routing improvements potentials lies in fact that the re-routing of collection vehicles, i.e. exclusion or inclusion certain nodes in preliminary defined routes, requires drivers familiarity with the collection area, streets, traffic conditions and bins positions. Averaged trend coefficient γ is then used to define fuzzy estimates of uncertain waste quantities $\tilde{q}_i$ in bins, represented by triangular fuzzy numbers given by $\tilde{q}_i = ((1-\alpha)\gamma q_i, \gamma q_i, (1+\alpha)\gamma q_i)$, where $0 < \alpha < 1$ represent coefficient of bins quantity spread, where bigger spread means larger uncertainty. Because the waste quantities are uncertain until the moment when the bin's content is loaded to the collection vehicle, we introduced additional measure q_i^C which represents actual quantity in the node $i \in N$ found at the moment of collection, which can be used to asses re-routing effects. Set of preliminary routes K, based on long term expected waste quantities $\bar{q}_i$, is known in advance. Each route $k \in K$ is defined with the sequence R_k of nodes $i \in N$, arranged by order of visit. To each preliminary route $k \in K$ is assigned set N_k of neighbour nodes $j \in N_k$, which belong to neighbouring routes $u \in K, u \neq k$. Sets of neighbour nodes include nodes from routes $u \in K$ which can be inserted in a routes $k \in K$, satisfying drivers' familiarity criterion. An example of the neighbour nodes set is given in the Figure 1.



Set of route 2 neighbor nodes
$N_2 = \{4,5,10,11\}$

Figure 1. Example of route's neighbouring nodes

In order to keep track on the nodes are excluded from preliminary defined routes because of excess quantities, and on the routes with free vehicle capacities, which are candidates for re-routing improvements, following two sets are also defined. Set of network nodes D_k, excluded from preliminary route $k \in K$, because of excess quantity, is defined by expression (1).

$$D_k = \left\{ w \in R_k \mid \sum_{i \in R_k} \tilde{q}_i - \sum_{w \in R_k} \tilde{q}_w \leq q_v \right\} \tag{1}$$

Obviously, set of all network nodes with excess quantity is obtained as $D = \bigcup_{k \in K} D_k$. Set of routes with free vehicle capacity F is defined by expression (2).

$$F = \left\{ R_k \mid \sum_{i \in R_k} \tilde{q}_i \leq q_v \right\} \tag{2}$$

Based on introduced concept, possibilities and effects of collection vehicles dynamic re-routing are determined by applying proposed fuzzy modelling approach.

3 MODELING APPROACH

To solve the problem, two-phase fuzzy arithmetic approach has been proposed. In the first phase, based on information obtained from collection bin sensors, is necessary to calculate trend coefficient *γ*. Then, sets *D* and *F* are created, and depending of its characteristics all possible dynamic re-routing actions (S-I, S-II, S-III, S-IV) are defined. Re-routing algorithms are performed in the second phase of the approach. Finally, after proposed algorithms are applied, we calculate effects of performed re-routing actions.

3.1 The first phase of the algorithm

Algorithm for the first phase of the proposed approach is given in the form of pseudo code in Figure 2.

THE FIRST PHASE ALGORITHM

1 ***Set*** $D_k = \varnothing \quad \forall k \in K$

2 ***Set*** $F = \varnothing$

3 **for** *i* **in** *N:*

4 $\qquad \gamma = \dfrac{\sum_{i \in N} q_i}{\sum_{i \in N} q_i^T}$: # trend coefficient calculation

5 $\qquad \tilde{q}_i = ((1-\alpha)\gamma q_i, \gamma q_i, (1+\alpha)\gamma q_i)$:# fuzzy estimates of the waste quantities in bins

6 **for** *k* **in** *K:*

7 $\qquad$ **if** $\sum_{i \in R_k} \tilde{q}_i > q_v$:

8 $\qquad\qquad D_k = \left\{ w \in R_k \mid \sum_{i \in R_k} \tilde{q}_i - \sum_{w \in R_k} \tilde{q}_w \le q_v \right\}$: # nodes to be excluded from the route

9 $\qquad\qquad R_k^U = R_k \setminus D_k$: # updating routes R_k by excluding excess quantity nodes

10 $\qquad$ **else**

11 $\qquad\qquad R_k^U = R_k$

12 $D = \bigcup_{k \in K} D_k$

13 $F = \bigcup_{k \in K} R_k^U$

Figure 2. Algorithm for the first phase of the proposed approach

Note that sets D_k contain one or more nodes whose exclusion provide routes realization without exceeding vehicle capacity. The idea is to find nodes with total waste quantity as close as possible to the excess quantity $\sum_{i \in R_k} \tilde{q}_i - q_v$, which can be done in different ways, by using optimization, heuristics, or even enumeration since the dimension of that sub-problem is not so large. Set *F* includes routes with free vehicle capacity. When, for the arbitrary route R_k, corresponding set D_k is empty, which means that capacities of its nodes are lesser or equal to the vehicle capacity, then the updated route R_k^U is identical to the route R_k . Fuzzy comparison in block 7 may be done by different methods, and here we proposed center of gravity defuzzification approach (COG), because of its simplicity, one of the most popular defuzzification techniques of fuzzy mathematics. More details about defuzzification approaches can be found in Skalna et al., (2015).

3.2 Fuzzy arithmetic method for the second phase of the approach

Algorithm for the second phase of the proposed approach, based on fuzzy arithmetic, comprises two main steps. Based on sets D and F, the first step proposes application one of defined re-routing algorithms: S-I, S-II, S-III, S-IV. The second step comprises three algorithms corresponding to one of defined re-routing actions (S-II, S-III, S-IV), while the first (S-I) assumes realization of preliminary defined routes. Pseudo code of the algorithm's first step is given in Figure 3.

THE SECOND PHASE ALGORITHM

- THE FIRST STEP -

if $D = \varnothing$ **and** $\sum_{k\in F}\sum_{i\in R_k}\tilde{q}_i \cong |K|\, q_v$:

  S-I: # preliminary routes realization

if $D = \varnothing$ **and** $\sum_{k\in F}\sum_{i\in R_k}\tilde{q}_i < |F|\, q_v$:

  S-II: # re-routing with the objective of decreasing number of routes

if $D \neq \varnothing$ **and** $\sum_{k\in F}\sum_{i\in R_k}\tilde{q}_i \cong |F|\, q_v$:

  S-III: # re-routing with the objective to construct additional routes

if $D \neq \varnothing$ **and** $\sum_{k\in F}\sum_{i\in R_k}\tilde{q}_i < |F|\, q_v$:

  S-IV: # re-routing with the nodes interchange and additional routes construction

Figure 3. Algorithm for the first step in the second phase of the proposed approach

S-II algorithm, given in Figure 4, tries to choose route R_k^*, which is candidate for removal, whose nodes will be inserted into the neighbour routes, which in turn decreases number of routes and, as a consequence, number of vehicles used. To do that, in the block 2 we find routes with smallest quantity of waste, and in blocks 3-10 search for best insertion of its nodes. The algorithm analyses effects of simultaneous multiple as well as single nodes insertion, considering all subsets of the power sets $\wp(R^* \cap N_z)$ which are determined as intersection of the set R^* and all sets of neighbouring nodes $N_z, \forall z \in K$, whose elements $j \in N_z$ are, also, elements of the subsets $e \in R^*$, and, therefore, give opportunity for insertion of nodes from the subsets $e \in R^*$ into the route R_z. Effects ξ_{ez} of the insertion of the nodes in the subset $e \in R^*$ into the route R_z are estimated by the calculation of route length extension, and vehicle use improvements will resulted from the insertion. The best insertion has minimum route length extension, and maximum vehicle capacity use improvement.

S-III algorithm intends to construct additional routes which visits surplus quantity nodes from the set D, i.e. nodes which are excluded from the preliminary routes. The algorithm is inspired by the "sweep" algorithm based on fuzzy logic approach (Teodorović and Pavković, 1996 but our idea differs, since it is based on fuzzy arithmetic only. The solution approach is similar to application of original "sweep" algorithm proposed by Gillet and Miller (1974), but here, ability of the vehicle to serve the next node after it has served several nodes before is evaluated by fuzzy comparison which is done by defuzzification approach. Pseudo code of the S-III algorithm is given in Figure 5.

THE SECOND PHASE ALGORITHM

- THE SECOND STEP -

*** S-II algorithm***

if $\sum_{k\in F}\sum_{i\in R_k}\tilde{q}_i \le (|F|-1)q_v$: # possibility for decreasing number of routes
    # determination of routes which are candidates for removal R^* by inserting its nodes into other routes with free vehicle capacity
    $\arg\min_{R^*\in F} Q_k = (R^* \in F \mid Q_k = \sum_{R_k\in F}\sum_{i\in R_k}\tilde{q}_i < \sum_{R_s\in F}\sum_{i\in R_s}\tilde{q}_i \quad \forall R_s, R_k \in F, R_s \ne R_k)$
    for z in K:
        $INS^* = \wp(R^* \cap N_z)$: # power set of the sets intersection
        if $INS^* \ne \varnothing$:
            for e **in** INS^*:
                if $\sum_{i\in e}\tilde{q}_i \le \sum_{j\in R_z}\tilde{q}_i$: # possibility of insertion $e \in INS^*$ into the route R_z
                    $\xi_{ez} = \xi(L_{ez}, Q_{ez})$ # effects of the insertion $e \in INS^*$ into the route R_z
    $\xi^*_{ez} = \min\{\xi_{ez}\}$: # best insertion of $e^* \in INS^*$ into the route R_z
    $R^* = R^* \setminus e^*$
    $R_z^U = R_z \cup e^*$
    $F = F \cup R_z^U$
else
    run S-I: # preliminary routes realization

Figure 4. Pseudo code of the S-II algorithm

- THE SECOND STEP -

*** S-III algorithm***

Arrange nodes in the set D in nondecreasing order by its polar coordinates
while $D \ne \varnothing$:
    $\tilde{Q}_w$ =0: # quantity of waste loaded on the vehicle after visiting all previous nodes in the route
    $k = k + 1 : R_k = \varnothing$: # new route creating
    for i **in** D:
        if $\tilde{Q}_w + \tilde{q}_i \le q_v$: # checking available capacity of the collection vehicle
            $R_k \leftarrow i$: # insertion of the node into the route
            $\tilde{Q}_w = \tilde{Q}_w + \tilde{q}_i$: # updating quantity of waste loaded on the vehicle
            $D = D \setminus \{i\}$: # updating set

Figure 5. Pseudo code of the S-III algorithm

S-IV algorithm deals with the situations when some routes with a free vehicle capacity exist, while some other require nodes exclusion to prevent exceeding a vehicle capacity. Therefore, this algorithm can be seen as a tool that can solve both, re-routing problems with objective to insert excess waste quantity nodes into routes with free vehicle capacity, and to provide additional routes creation if some nodes cannot be inserted into the existing routes. Because of that S-IV algorithm is based on successive application of the S-II, and the S-III algorithms. Only difference is in defining the sets which are used during the algorithm application. Namely, instead of searching for the set $R^* \in F$ whose nodes $i \in R^*$ will be inserted into the another routes, application of the S-IV algorithm tries to find best insertions of nodes $i \in D$ into the available routes $R_k^U \in F$ which are updated accordingly to the block 9 of the algorithm in Figure 2. Since we try to insert all nodes $i \in D$ into the existing routes from the set $F = \bigcup_{k \in K} R_k^U$, the set R^* defined in the algorithm S-II in Figure 4, when applying algorithm S-IV, becomes $R^* = D$. After creation of the set R^*, all remaining blocks defined in the algorithm S-II should be performed in the same way.

4 ILLUSTRATIVE NUMERICAL EXAMPLE

The waste collection network is based on Solomon's instance R101 with *N=25* nodes, whose coordinates were divided by 10. Quantity of waste - node demands, were simulated by Uniform distributions, and represent tones of waste:

- expected long term waste quantities $\bar{q}_i$ ~ U(0.5, 1.5)
- quantity of waste, T time units before vehicles started q_i^T ~ U(0.8 $\bar{q}_i$, 1.2 $\bar{q}_i$)
- quantity of waste at the moment when vehicles started q_i ~ U(q_i^T, 1.2 q_i^T)
- quantity of waste at the moment of collection q_i^C ~ U(q_i, 1.3 q_i)

Based on simulated values trend coefficient of waste quantities change *γ=1.10*. Fuzzification of uncertain waste quantities in bins $\tilde{q}_i$ was based on coefficient of bins quantity spread *α=0.1*. Table 1 summarizes simulated values.

Table 1. Simulated parameters' values

node	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
$\bar{q}_i$	0.86	0.99	0.66	0.97	0.76	1.13	1.04	0.66	1.44	1.15	1.01	0.89	0.91
q_i^T	0.94	0.77	1.37	1.25	0.77	1.17	0.76	0.59	0.53	0.82	1.29	0.8	0.74
γq_i	1.06	0.96	1.63	1.64	0.87	1.53	0.94	0.69	0.6	0.99	1.48	1.05	0.83
q_i^C	1.04	1.03	1.77	1.61	0.86	1.74	1.06	0.74	0.71	1.15	1.44	1.16	0.98
node	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	
$\bar{q}_i$	1.28	0.96	1.25	1.1	1.33	0.52	0.71	0.57	0.61	0.83	1.28	0.98	
q_i^T	0.98	0.75	0.84	0.54	0.98	0.71	1.36	1.09	1.25	1.43	0.83	1.04	
γq_i	1.08	0.89	1.03	0.71	1.2	0.84	1.76	1.4	1.57	1.8	1.1	1.23	
q_i^C	1.06	0.94	0.96	0.84	1.31	0.77	1.87	1.31	1.47	2.03	1.08	1.13	

Seven preliminary defined routes, based on expected long term waste quantities $\bar{q}_i$, were constructed for the vehicles' capacities of 4 tonnes, are shown in the Figure 2 (left), an in the Table 2, together with its neighbouring sets N_k. Using the proposed approach, first we defined sets D_k, presented in the Table 2. Set $D = \bigcup_{k=K} D_k = \{1,19,23,4\}$ includes nodes which are excluded from the routes R_1, R_2 and R_6, to prevent exceeding a vehicle capacity. After those nodes are excluded we updated routes R_k, by $R_k^U = R_k \setminus D_k$. After the update, all route sets have possible free vehicle capacity, while nodes with excess quantity exists in set *D*. Possible quantities of waste which have the maximum membership function are also shown in Table 2, for preliminary $\sum_{i \in R_k} \gamma q_i$, and updated routes $\sum_{i \in R_k^U} \gamma q_i$. Then, we apply algorithm S-IV by performing procedure defined in algorithm S-II, starting from the block 3, but, before that, defining $R^* = D$. Intersection of the set R^* and all neighbouring sets $(R^* \cap N_z)$, on which is applied power set operator $\wp$, requires $|R_k|^2=49$ iterations which show that only two of four excluded nodes can be inserted in updated routes.

Table 2. Sets R_k and N_k

Route k,z	Sets R_k	$\sum_{i \in R_k} \gamma q_i$	Sets D_k	$\sum_{i \in R_k^U} \gamma q_i$	Sets N_k	$(R^* \cap N_z)$
1	$R_1 = \{3,9,20,1\}$	5.05	$D_1 = \{1\}$	3.99	$N_1 = \{10,12,24\}$	$\{\ \}$
2	$R_2 = \{10,11,19,7\}$	4.25	$D_2 = \{19\}$	3.41	$N_2 = \{1,8,18\}$	$\{1\}$
3	$R_3 = \{18,8,17\}$	2.6	$D_3 = \varnothing$	2.60	$N_3 = \{5,6,7,16\}$	$\{\ \}$
4	$R_4 = \{5,16,6\}$	3.43	$D_4 = \varnothing$	3.43	$N_4 = \{8,13,14,17,18\}$	$\{\ \}$
5	$R_5 = \{14,15,13\}$	2.8	$D_5 = \varnothing$	2.80	$N_5 = \{2,6,16\}$	$\{\ \}$
6	$R_6 = \{2,22,23,4,21\}$	7.37	$D_6 = \{23,4\}$	3.93	$N_6 = \{12,13,15,25\}$	$\{\ \}$
7	$R_7 = \{25,24,12\}$	3.38	$D_7 = \varnothing$	3.38	$N_7 = \{3,4,21\}$	$\{4\}$

The node 1, which is excluded from the route 1, can be only inserted in updated route 2 (preliminary route 2, after the node 19 has excluded), and the node 4 which is excluded from the route 6, can be only inserted in route 7. After the nodes insertions and removals, sets $R_1 = \{3,9,20\}$, $R_2 = \{1,10,11,7\}$, $R_6 = \{2,22,21\}$, and $R_7 = \{4,25,24,12\}$ are changed, while the sets R_3, R_4 and R_5 remain the same as is was preliminary defined. Nodes in updated set $R^* = \{23,19\}$, which corresponds to updated set D, should be visited by additionally constructed routes $R_8 = \{7,19,20\}$, and $R_9 = \{12,23\}$. Similarly, additional routes, for the case when only preliminary defined routes are used should be $R_8^P = \{7,20,1\}$, $R_9^P = \{4,21\}$ and $R_{10}^P = \{23\}$. Figure 6 shows preliminary defined and realized routes (left), as well as routes after the algorithm application (right).

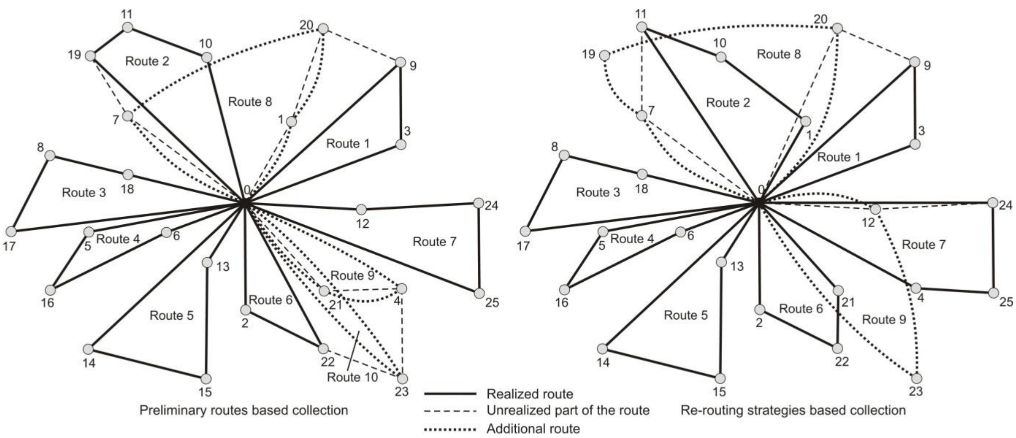


Figure 6. Preliminary defined routes and dynamic re-routing

The algorithm's effects estimation, shown in Table 3, was based on the actual collected quantities q_i^C, where both, preliminary and routes determined by the algorithm, include additional routes. When apply re-routing algorithm two additional routes are needed, R_8, R_9, whose capacities are 3.70 and 3.19, respectively - in total 6.89. Preliminary routes based concept requires three additional routes R_8^P, R_9^P, R_{10}^P, whose capacities are 3.97, 2.92, 2.03, respectively - in total 8.92. Total lengths of additional routes are 20.72 and 17.59 respectively, for the preliminary and the re-routing concepts.

Table 3. Estimated effects of the proposed algorithm

	Number of nodes visited (based on actual quantity q_i^C without additional routes)	Total collected quantity (based on actual quantity q_i^C without additional routes)	Number of routes (including additional routes)	Total length of routes (including additional routes)	Average vehicles utilization
Preliminary defined routes	19	21.12	10	69.86	0.75
Routes based on re-routing algorithm	20	25.95	9	66.96	0.83

5 CONCLUSIONS

In this paper we presented the idea for introducing re-routing strategies which can handle immanent process uncertainty, even in case when real-time information are available. Those uncertainties result from the fact that the transportation process characteristics may change during its realization. We propose modelling approach in which re-routing actions rely on prediction which is based on real-time data from sensors. In this way is tried to handle uncertainty by predefined strategies, instead of directly following real-time data by certain on-line dispatching rules. Proposed re-routing strategies were directed to waste collection process improvement, since waste bins sensors provide real-time information of current waste quantity. The algorithms proposed are based on fuzzy theory approach which in its essence handle uncertainty. We used triangular fuzzy numbers, because of simply use and easiness in presenting the concept, but instead, the model can be based on any other shape of fuzzy number which describes waste quantity uncertainty in the best way. We define four strategies and developed three fuzzy approach based algorithms to provide re-routing decisions in waste collection process. In this research we formulated only the main idea, but we plan to further extend this approach to be closer to an expert system. This will include not only appropriate software tools development, but also possible application of other fuzzy theory concepts like fuzzy logic, considering sets *D* and *F* as fuzzy sets, etc., as well as further adaptation of the concept to real world processes. Also, one important direction is to make comparison of the proposed and other possible approaches. Since the results presented here are very promising, we do hope that our intentions will be successfully implemented in the near future.

Acknowledgements: This work was supported by the Ministry of Education, Science and Technical development of the Government of the Republic of Serbia through the projects TR36005 and TR36006, for the period 2011-2018.

References

[1] Gillet, B. and Miller, L. (1974). A heuristic algorithm for the vehicle dispatch problem, *Operations Research*, 22(2), pp. 340-349.

[2] Tchobanoglous, G. and Kreith, F. (2002). *Handbook of Solid Waste Management*. 2nd Edition, McGraw Hill Handbooks. New York.

[3] Shuster, K.A. and Schur, D.A. (1974). *Heuristic routing for solid waste collection vehicles*. U.S. Environmental Protection Agency.

[4] Teodorović, D. and Pavković, G. (1996). The fuzzy set theory approach to the vehicle routing problem when demand at nodes is uncertain, *Fuzzy Sets and Systems*, 82(3), pp. 307-317.

[5] Psaraftis, H.N., Wen, M. and Kontovas, C.A. (2016). Dynamic Vehicle Routing Problems: Three Decades and Counting, *Networks*, 67(1), pp. 3–31.

[6] Ramos, T.R.P., Morais, C.S. and Barbosa-Póvoa, A.P. (2018). The Smart Waste Collection Routing Problem: alternative operational management approaches, *Expert Systems with Applications*, 103, pp. 146-158.

[7] Skalna, I., Rębiasz, B., Gaweł, B., Basiura, B., Duda, J., Opiła, J. and Pełech-Pilichowski, T. (2015). *Advances in Fuzzy Decision Making: Theory and Practice (Studies in Fuzziness and Soft Computing)*. Springer International Publishing Switzerland.

Author's address

Milorad Vidović, Professor
University of Belgrade, Faculty of Transport and Traffic Engineering, Department of Logistics
Vojvode Stepe 305, Belgrade
Republic of Serbia
email: m.vidovic@sf.bg.ac.rs

Gordana Radivojević, Professor
University of Belgrade, Faculty of Transport and Traffic Engineering, Department of Logistics
Vojvode Stepe 305, Belgrade
Republic of Serbia
email: g.radivojevic@sf.bg.ac.rs

Branislava Ratković, Assistant Professor
University of Belgrade, Faculty of Transport and Traffic Engineering, Department of Logistics
Vojvode Stepe 305, Belgrade
Republic of Serbia
email: b.ratkovic@sf.bg.ac.rs

Nenad Bjelić, Assistant Professor
University of Belgrade, Faculty of Transport and Traffic Engineering, Department of Logistics
Vojvode Stepe 305, Belgrade
Republic of Serbia
email: n.bjelic@sf.bg.ac.rs

Dražen Popović, Assistant Professor
University of Belgrade, Faculty of Transport and Traffic Engineering, Department of Logistics
Vojvode Stepe 305, Belgrade
Republic of Serbia
email: d.popovic@sf.bg.ac.rs

THE USE OF BACKTESTING BINOMINAL TESTS IN ASSESMENT OF THE VALUE-AT-RISK ON GERMAN AND POLISH CAPITAL MARKETS

Marek Zinecker, Brno University of Technology,
Tomáš Meluzín, Brno University of Technology,
Marcin Fałdziński, Nicolaus Copernicus University in Toruń,
Adam P. Balcerzak, Nicolaus Copernicus University in Toruń,
Michał Bernard Pietrzak, Nicolaus Copernicus University in Toruń

Abstract
The article concerns the problem of risk measurement related to the functioning of capital markets in Germany and Poland. This risk is related to the constantly progressing phenomenon of globalization of economies, which in turn results in a dynamic growth of links between financial markets. In connection with the issue of risk measurement, the attention of the authors is focused on determining the value of Value-at-Risk measure and testing the quality of the received estimates. The VaR measure is a useful risk assessment tool on the financial market due to its universality and methodology providing many tools for its assessment and testing. The purpose of the proposed article is to assess the quality of VaR measure estimates used to determine the market risk for DAX and WIG indices. The research was done for the years 2000-2012. The period was chosen deliberately as it covered the time of global financial crisis. For the quality evaluation, backtesting was used, where binominal tests LRuc, LRind, LRcc were applied.

Keywords: *Value-at-Risk, backtesting, LRuc test, LRind test, LRcc test*

JEL Classification: G15, C58
AMS Classification: 91B30, 91B84

1 INTRODUCTION

The article focuses on measuring the risk related to the operation of capital markets, which concerns individual assets as well as to capital markets as a whole. The increasing globalization of the world economy contributes to the increase in the number of linkages between capital markets, which extends the sources of market risk significantly (Osińska, Fałdziński and Zdanowicz, 2009; Vychytilová, 2015; Heryán and Ziegelbauer, 2016; Fałdziński et al., 2016; Balcerzak, 2016; Cevik, Korkmaz, and Cevik, 2017; Śliwiński and Łobza, 2017; Navid and Shabantaheri, 2017; Kubiszewska, 2017; Strielkowski, Tcukanova, and Zarubina, 2017; Balcerzak et al., 2017). Also, the dynamic development of the digital economy, through increasingly wider and faster access to information, strengthens this phenomenon (Balcerzak and Pietrzak, 2017a). Moreover, the global value of capital markets is constantly increasing due to regular introductions of new joint-stock companies to the market (Meluzin and Zinecker, 2014; 2016; Draženović and Kusanović, 2016; Meluzín et al., 2017; 2018; Skalická Dušátková, Zinecker and Meluzin, 2017; Sosnowski, 2017). The situation on the financial markets and the risks related to them are a significant factor affecting the level of socio-economic development of national economies, including the entrepreneurship potential and their international competitiveness (Lisowska, 2016; Balcerzak and Pietrzak, 2016; Dvorsky, Sopkova and Janoskova 2017; Cheba and Szopik-Depczyńska, 2017; Kruk and Waśniewska, 2017; Zygmunt, 2017; Pietrzak et al., 2017a). Therefore, the identification of linkages between the markets and the measurement of the resulting risk is an important element of ongoing research in the field of financial market analysis (Zinecker et al., 2016; Pietrzak et al., 2017b; Vukovic, Hanic and Hanic, 2017). Many measures have been developed for the assessment of financial risk, of which the Value at Risk measure has been applied to the largest extent.

In connection with the undertaken subject of risk measurement, the article focuses on the issues related to the determination of the Value-at-Risk measure and testing the quality of the received measurement estimates. Due to the abundant methodology of VaR calculation methods and the properties of statistical tests, the VaR measure is a useful tool for risk assessment in the financial market (Fałdziński, 2011; Fałdziński, 2017). The research objective of the proposed article is to assess the quality of VaR estimates used to determine the risk of the major German and Polish market indices. The objective was achieved through the use of backtesting. Binominal tests were made to evaluate the received values of the measure, which allowed us to assess the quality of the measure in terms of the coverage and independence properties. In order to calculate VaR we used a GARCH class model with the conditional Student's *t* distribution. Then, the quality of the VaR obtained was checked by means of the LRuc, LRind, LRc tests (Christoffersen, 1998; Fałdziński, 2017). The analysis was carried out in the 2000-2012 period, which allowed a comprehensive assessment of the risk level for both market indices to be made.

2 THE GARCH MODEL AND VALUE AT RISK

VaR is used to measure the market risk of securities, most often these are individual shares or an investment portfolio composed of them. Value at risk expresses risk in a way that is easy to interpret. It calculates the potential maximum losses. The key issue when measuring VaR is to determine the distribution of returns for a specific asset (or for an investment portfolio). The research has confirmed many times the variance of variability over time (Fałdziński and Pietrzak, 2015) for returns. For such purposes, GARCH class models are commonly used. They allow modelling of the conditional variance for individual assets or stock indices (Szumilo et al., 2018; Smolović, Lipovina-Božović, and Vujošević, 2017).

We use the model ARMA(p,q)-PARCH(1,1) (The Power ARCH model) (Ding, Granger, and Engle, 1993), which can be defined as follows:

$$y_t = c + \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \varepsilon_t | F_{t-1} \sim t(0, h_t, \mathrm{v}) \tag{1}$$

$$h_t^\delta = \alpha_0 + \alpha_1(|\varepsilon_{t-1}| - \gamma_1 \varepsilon_{t-1})^\delta + \beta_1 h_{t-1}^\delta, \tag{2}$$

where $\mathrm{y_t}$ is the process of returns, $\mathrm{h_t}$ is the conditional standard deviation equation, $\varphi_1, \varphi_p, \theta_1, \theta_p, \alpha_0, \alpha_1, \gamma_1, \beta_1, \delta$ are the parameters of both equations, the conditional expected value equation and the conditional variance equation, ε_t is the noise term following the t-distribution with v degrees of freedom and $\alpha_0 > 0, \delta \geq 0, \alpha_1 \geq 0, -1 < \gamma_1 < 1, \beta_1 \geq 0$.

Value at Risk is a measure allowing to determine the loss of the market value (of a specific asset or of an investment portfolio). The measure determines the loss with the probability of its occurrence or exceedance equal to the tolerance level set. The VaR measure is expressed in the following form:

$$P(C_1 \leq C_0 - VaR) = \alpha \tag{3}$$

where C_0 is the initial value of the asset (portfolio), C_1 is the final value of the asset (portfolio) and α represents the tolerance level.

Measuring VaR consists in determining the value of the measure based on the distribution of returns, where the distribution characteristics are assumed based on the obtained results of the ARMA(p,q)-PARCH(1,1) model estimation. The formula for determining the one-day VaR measure is defined as follows:

$$VaR_\alpha^t = \mu_{t+1} + h_{t+1} z_q \tag{4}$$

where μ_{t+1} is the prediction of the conditional mean of returns, h_{t+1} is the prediction of the conditional volatility and z_q is q-quantile of the conditional distribution, and α is known as coverage level.

A significant element of the application of the VaR measure is the assessment of its quality in terms of unconditional coverage and independence. For this purpose, backtesting is used, where tests are performed to check specific properties. The first one assume that the expected number of hits (cases when returns are greater than estimated VaR in absolute term) equals the given coverage level. The second one state that the hit process is independent. The article uses binominal tests such as, for instance, the unconditional coverage test LRuc, the test of independent LRind and the conditional coverage test LRcc in the backtesting procedure (Christoffersen, 1998). The use of the tests allows not only the assessment of the quality of the VaR measure applied to a specific asset or portfolio, but also the assessment of different methods of VaR determination (Fałdziński, 2017).

3 APPLICATION OF THE LRUC, LRIND, AND LRCC BACKTESTS FOR MEASURING VALUE AT RISK

With regard to risk assessment on the capital markets of Germany and Poland, time series of the WIG and DAX stock indices were used. For each of the indexes, the logarithmic daily returns were determined in the period from 3 January 2000 to 3 January 2012, and the data was obtained from the www.finance.yahoo.com website. In the first stage of the study, the ARMA(p,q)-PARCH(1,1) model parameters were estimated using the maximum likelihood method with the *t*-distribution. The ARMA(p,q) lag order was selected based on the Akaike information criterion. Therefore, for each of the indices, the specification of the model defined by the formulas (1-2) was adopted. The obtained results of the estimation of the model parameters are presented in Table 1. In the case of the conditional mean equation not all the parameters are significant which is due to a selection process based on the AIC However, in the case of the conditional equation of variance (formula 2), all parameters proved to be statistically significant at the 5% level of significance. The estimated parameter *ν* of the *t*-distribution indicates the occurrence of thick tails in the distribution of returns, which justifies the use of this distribution in the estimation procedure.

Table 1. The estimation of the ARMA(p,q) -PARCH (1,1) model parameters

Parameter	**Estimate**	**Std. error**	**p-value**	**Estimate**	**Std. error**	**p-value**
	WIG			**DAX**		
c	0.0507	0.0209	0.0155	0.0181	0.0207	0.3818
φ_1	-0.0192	0.2243	0.9314	0.2417	0.0275	0.0000
φ_2	-0.7506	0.1376	0.0000	-0.8852	0.0297	0.0000
φ_3	-0.3969	0.2092	0.0578	-	-	-
θ_1	0.0709	0.2204	0.7474	-0.2495	0.0331	0.0000
θ_2	0.7356	0.1381	0.0000	0.8888	0.0301	0.0000
θ_3	0.4351	0.2042	0.0332	0.0157	0.0201	0.4351
α_0	0.0216	0.0061	0.0005	0.0254	0.0037	0.0000
α_1	0.0624	0.0110	0.0000	0.0660	0.0059	0.0000
β_1	0.1936	0.0655	0.0031	0.9999	0.0000	0.0000
γ_1	0.9270	0.0098	0.0000	0.9245	0.0072	0.0000
δ	1.8440	0.3549	0.0000	1.1393	0.1485	0.0000
v	8.0382	1.1257	0.0000	20.726	6.1892	0.0000

Source: own calculations.

The estimation of the ARMA(p,q)-PARCH(1,1) model to the empirical data allowed us to go on to the procedure of the estimation of the Value at Risk measure. The procedure consisted in estimating the PARCH model for the first 1000 observations and determining the forecast of one-day VaR in accordance with formula (4). Then, the observation window was rolling by a one-day session, the parameters of the PARCH model were estimated again and a one-day VaR forecast was determined. In this way, a set of 2000 VaR values were obtained. The forecasted VaR values were used to carry out the LRuc, Lrind, and LRcc backtests. The results regarding the use of selected tests are presented in Table 2. Only for the 99% coverage level unconditional coverage property is not met for both indices. It should be noted that for both market indexes (DAX and WIG) the properties of the VaR measure can be considered as correct for 95% and 90% coverage level. The results obtained confirmed the correctness of the VaR measure estimation in the case of both market indices.

Table 2. Backtesting results

Test	LRuc			LRind			LRcc		
coverage level	99%	95%	90%	99%	95%	90%	99%	95%	90%
DAX index									
Statistics	8.1819	8.1426	2.3776	0.2627	3.1029	0.1251	8.4447	11.2455	2.5027
p-value	**0.0042**	**0.0043**	0.1231	0.6083	0.0782	0.7236	**0.0147**	**0.0036**	0.2861
WIG index									
Statistics	5.2330	0.0418	0.0884	0.4442	3.9660	0.5806	5.6773	4.0078	0.6690
p-value	**0.0222**	0.8379	0.7663	0.5051	**0.0464**	0.4461	0.0585	0.1348	0.7157

Source: own calculations.

4 CONCLUSIONS

The subject of the article is measuring risk related to the functioning of financial markets. This is the risk occurring within the constantly developing network of linkages between capital markets whose strength and range keep growing along with the increasing globalization. Risk identification and risk measurement make a challenging research problem, since they allow to determine the probability of a significant loss under a crisis situation. VaR estimates were made to assess the level of risk for Germany's major market indices. In order to estimate VaR a GARCH model with the conditional *t* distribution was used and the quality of the values obtained was checked using the LRuc, LRind, and LRc tests. The results were obtained based on the use of selected tests and allowed to state that for the two indices VaR properties can be considered as correct. This means that in the case of the DAX and WIG indices, the VaR measure is a useful risk assessment tool for the capital market.

References

[1] Balcerzak, A. P. (2016). Multiple-criteria Evaluation of Quality of Human Capital in the European Union Countries. *Economics & Sociology*, 9(2), pp. 11-27.

[2] Balcerzak, A. P., Kliestik, T., Streimikiene, D. and Smrčka L. (2017). Non-parametric approach to measuring the efficiency of banking sectors in European Union Countries. *Acta Polytechnica Hungarica*, 14(7), pp. 51-70.

[3] Balcerzak, A. P. and Pietrzak, M.B. (2016). Quality of Institutions for Knowledge-based Economy within New Institutional Economics Framework. Multiple Criteria Decision Analysis for European Countries in the Years 2000–2013. *Economics & Sociology*, 9(4), pp. 66-81.

[4] Balcerzak, A. P. and Pietrzak, M. B. (2017). Digital Economy in Visegrad Coutnries. Multiple-criteria decision analysis at Regional Level in the Years 2012 and 2015. *Journal of Competitiveness*, 9(2), pp. 5-18.

[5] Cevik, E. I., Korkmaz, T. and Cevik, E. (2017). Testing causal relation among central and eastern European equity markets: evidence from asymmetric causality test. *Economic Research-Ekonomska Istraživanja*, 30(1), pp. 381-393.

[6] Cheba, K. and Szopik-Depczyńska, K. (2017). Multidimensional comparative analysis of the competitive capacity of the European Union countries and geographical regions. *Oeconomia Copernicana*, 8(4), pp. 487-504.

[7] Christoffersen, P. (1998). Evaluating Interval Forecasts. *International Economic Review*, 39(4), pp. 841-862.

[8] Ding, Z., Granger, C.W.J. and Engle R.F. (1993). A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, 1, pp. 83–106.

[9] Draženović, B.O. and Kusanović, T. (2016). Determinants of capital market in the new member EU countries. *Economic Research-Ekonomska Istraživanja*, 29(1), pp. 758-769.

[10] Dvorsky, J., Sopkova, G. and Janoskova, M. (2017). Evaluation of Social Environment and Access to Financial Resources for Business: Case Study of the Czech and Slovak Republic. *Economicko-manazerskie spektrum/Economic and Managerial Spectrum*, 11(1), pp. 62-73.

[11] Fałdziński, M. (2011). On The Empirical Importance of The Spectral Risk Measure with Extreme Value Theory Approach. In: *Financial Markets Principles of Modelling Forecasting and Decision-Making, FindEcon*. Łódź: Uniwersytet Łódzki, pp. 73-86.

[12] Fałdziński, M. (2017). Backtesting value-at-risk for multiple risk levels: a lexicographic ordering approach. In: *Mathematical methods in economics 2017: proceedings of the 35th International Conference.* Hradec Králové: University of Hradec Králové, pp. 155-160.

[13] Fałdziński, M. and Pietrzak, M.B. (2015). The Multivariate DCC-GARCH Model with Interdependence Among Markets in Conditional Variances' Equations. *Przeglad Statystyczny,* 62(1), pp. 397-413.

[14] Fałdziński, M., Balcerzak, A.P., Meluzín, T., Pietrzak, M. B. and Zinecker, M. (2016). Cointegration of Interdependencies among Capital Markets of Chosen Visegrad Countries and Germany. In: *34th International Conference Mathematical Methods in Economics MME 2016 Conference Proceedings.* Liberec: Technical University of Liberec, pp. 189-194.

[15] Heryán, T. and Ziegelbauer, J. (2016). Volatility of Yields of Government Bonds Among GIIPS Countries During the Sovereign Debt Crisis in the Euro Area. *Equilibrium. Quarterly Journal of Economics and Economic Policy,* 11(1), pp. 61-74

[16] Kruk, H. and Waśniewska, A. (2017). Application of the Perkal method for assessing competitiveness of the countries of Central and Eastern Europe. *Oeconomia Copernicana*, 8(3), pp. 337–352.

[17] Kubiszewska, K. (2017). Banking concentration in the Baltic and Western Balkan states - selected issues. *Oeconomia Copernicana*, 8(1), pp. 65-82.

[18] Lisowska, R. (2016). The Potential of Business Environment Institutions and the Support for the Development of Small and Medium-sized Enterprises. *Entrepreneurial Business and Economics Review*, 4(3), pp. 85-101.

[19] Meluzín, T., Pietrzak, M. B., Balcerzak, A. P., Zinecker, M., Doubravský, K. and Dohnal, M. (2017). Rumours Related to Political Instability and their Impact on IPOs: The Use of Qualitative Modeling with Incomplete Knowledge. *Polish Journal of Management Studies*, 16 (2), pp. 171-187.

[20] Meluzín, T., Zinecker, M., Balcerzak, A.P., Doubravský, K., Pietrzak, M. B. and Dohnal, M. (2018). The timing of initial public offerings – non-numerical model based on qualitative trends. *Journal of Business Economics and Management,* 19(1), 109-125.

[21] Meluzin, T. and Zinecker, M. (2016). Trends in IPOS: the evidence from cee capital markets. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 11(2), pp. 327-341.

[22] Meluzin, T. and Zinecker, M. (2014). Research into Determinants Influencing 100 Decisions on the Polish Capital Market. *Betriebswirtschaftliche Forschung und Praxis*, 66(6), pp. 652-671.

[23] Navid, B.J. and Shabantaheri, P. (2017). The development of capital market and bank risk: The Case of Iran. *International Review*, 1-2, pp. 99-107.

[24] Osińska M., Fałdziński M. and Zdanowicz T. (2009). Econometric Analysis of the Risk Transfer in Capital Markets. The Case of China. *Argumenta Oeconomica*, 2(29), pp. 139-164.

[25] Pietrzak, M. B., Balcerzak, A. P., Gajdos, A. and Arendt, Ł (2017a). Entrepreneurial environment at regional level: the case of Polish path towards sustainable socio-economic development. *Entrepreneurship and Sustainability Issues, 5*(2), pp. 190-203.

[26] Pietrzak, M. B., Fałdziński, M., Balcerzak, A. P., Meluzín, T. and Zinecker, M. (2017b). Short-term Shocks and Long-term Relationships of Interdependencies among Central European Capital Markets. *Economics & Sociology*, 10(1), pp. 61-77.

[27] Skalická Dušátková, M., Zinecker, M. and Meluzin, T. (2017). Institutional Determinants of Private Equity Market in Czech Republic. *Economics & Sociology,* 10(4), pp. 83-98.

[28] Śliwiński, P. and Łobza, M. (2017). The impact of global risk on the performance of socially responsible and conventional stock indices. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(4), pp. 657-674.

[29] Smolović, J.C., Lipovina-Božović, M., and Vujošević, S. (2017). GARCH models in value at risk estimation: empirical evidence from the Montenegrin stock exchange. *Economic Research-Ekonomska Istraživanja*, 30(1), pp. 477-498.

[30] Sosnowski, T. (2017). Earnings management and the floatation structure: empirical evidence from Polish IPOs. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(4), pp. 693-709.

[31] Strielkowski, W., Tcukanova, O. and Zarubina, Z. (2017). Globalization and economic Integration: the role of modern management. *Polish Journal of Management Studies*, 15(1), pp. 255-261.

[32] Szumilo, N., Bienert, S., Łaszkiewicz, E., Pietrzak, M. B. and Balcerzak, A. P. (2018). The real alternative? A comparison of German real estate returns with bonds and stocks, *Journal of Property Investment & Finance,* 36(1), pp. 19-31.

[33] Vukovic, D., Hanic, E. and Hanic, H. (2017). Financial integration in the European Union - the impact of the crisis on the bond market. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(2), pp. 195-210.

[34] Vychytilová, J. (2015). Linkages among U.S. Treasury Bond Yields, Commodity Futures and Stock Market Implied Volatility: New Nonparametric Evidence. *Journal of Competitiveness*, 7(3), pp. 143-158.

[35] Zinecker, M., Balcerzak, A. P., Fałdziński, M., Meluzín, T. and Pietrzak, M. B. (2016). Application of DCC-GARCH Model for Analysis of Interrelations among Capital Markets of Poland, Czech Republic and Germany. In: *Proceedings of the International Scientific Conference Quantitative Methods in Economics Multiple Criteria Decision Making XVIII.* Vratna: Letra Interactive, pp. 416-421.

[36] Zygmunt, A. (2017). Innovation activities of Polish firms. Multivariate analysis of the moderate innovator countries. *Oeconomia Copernicana*, 8(4), pp. 505-521.

Author's address

doc. Ing. Marek Zinecker, Ph.D.
Brno University of Technology, Faculty of Business and Management, Kolejní 2906/4, 61200 Brno, Czech Republic
emai: zinecker@fbm.vutbr.cz

doc. Ing. Tomáš Meluzín, Ph.D.
Brno University of Technology, Faculty of Business and Management, Kolejní 2906/4, 61200 Brno, Czech Republic
emai: meluzint@fbm.vutbr.cz

Marcin Fałdziński, Ph.D.
Nicolaus Copernicus University in Toruń, Faculty of Economic Sciences and Management, Ul. Gagarina 13a, 87-100 Toruń, Poland
emai: marf@umk.pl

Adam P. Balcerzak, Ph.D.
Nicolaus Copernicus University in Toruń, Faculty of Economic Sciences and Management,
Ul. Gagarina 13a, 87-100 Toruń, Poland
email: adam.balcerzak@umk.pl

Michał Bernard Pietrzak, Ph.D.
Nicolaus Copernicus University in Toruń, Faculty of Economic Sciences and Management,
Ul. Gagarina 13a, 87-100 Toruń, Poland
email: michal.pietrzak@umk.pl

INTERACTION TERMS IN REGRESSION MODELS: SOME PRACTICAL ISSUES

Jan Zouhar, Martina Zouharová,
University of Economics, Prague

Abstract
Interaction terms have become a standard part of an empiricist's regression toolkit. Nevertheless, there are numerous practical issues regarding the use of interaction terms that still seem to confuse empiric researchers. In particular, the paper focuses on questions and problems related to (i) interpretation and hypothesis testing, (ii) multicollinearity, (iii) marginal effects in nonlinear models, and (iv) treatment of missing values. The aim of this paper is to provide practical guidance that will help empiricists avoid potential pitfalls related to these issues.

Keywords: *moderation, interaction effects, multicollinearity, marginal effects, multiple imputation*

JEL Classification: C51
AMS Classification: 62P20

1 INTRODUCTION

With the massive body of existing empiric literature, novel research questions are not easy to come by. In economics and social sciences at large, empiric research has covered such a vast range of cause-and-effect relationships that scholars have difficulty finding new variables to play the role of x and y in the $x \rightarrow y$ scheme.

There is a way to avoid the lengthy search, though. Instead of studying a yet unexplored $x \rightarrow y$ relationship, many authors have opted to delve into a more detailed description of a relationship that has been studied previously, perhaps even in a similar environment and context. The development of the research hypothesis typically goes from a straightforward "*H1: an increase in x has a positive effect on y*" to a more detailed "*H2: an increase in x has a positive effect on y, but the size of the effect depends on z; namely, for individuals who score high in z, the positive effect of x on y is much stronger than for the rest.*"

In social sciences, scholars often use the term *moderation*. In H2 above, z is said to have a *moderating effect* on the relationship between x and y. The prevalence of moderation-oriented studies can be verified by a quick Google Scholar search of the phrase "moderating effect"; ours returned a whopping 18,300 works published since 2017 only.

Quantification of the moderating effects is almost exclusively confined to regression analyses. Technically, there is a small step from a regression that allows to test H1 to one that tests H2: all one needs to do is add the explanatory variables z (if it was not already in the model) and xz, a product of x and z. In the regression jargon, this product is typically referred to as an *interaction term* (e.g. Greene, 2003, Ch. 7; Wooldridge, 2013, Ch. 6).

Even though the addition of an interaction term to a regression model is straightforward, practitioners seem to struggle with the subsequent econometric routine, as our peer-review work and a review of empiric literature revealed. The aim of this short note is to point out some practical issues that seem to confuse empiric researchers.

The rest of the paper is structured as follows. Section 2 focuses on elementary problems related to the interpretation of coefficients on the interaction term, and stresses some issues regarding hypothesis testing and multicollinearity detection. Section 3 identifies possible reasons why empiric studies of moderating effects avoid advanced econometric tools, such as the calculation of marginal effects in nonlinear models or imputation techniques to handle missing data. Section 4 concludes.

2 ELEMENTARY ISSUES WITH INTERACTION TERMS

Throughout the text, key points will be presented using cross-sectional models of participation in *nascent entrepreneurship*, an early stage of starting up a new venture; see (Lukeš, et al., 2012) for a more detailed definition. In this section, we consider the linear probability model

$$startup = \beta_0 + \beta_1 female + \beta_2\, age + \beta_3\, (female \times age) + \mathbf{x}\boldsymbol{\gamma} + u, \qquad (1)$$

where *startup* is a dummy variable indicating that the respondent is actively starting up a new (nascent) business; *female* is a gender dummy; *age* is expressed in years; $\mathbf{x}$ is a vector of controls that include categorical variables *education*, *region* and *year of survey*; u is the random error; and Greek letters denote model parameters (scalars or vectors).

2.1 Correct interpretation, hypothesis testing, and omitted constituent terms

In (1), the coefficient on *female*, β_1, is not interesting in its own right; it captures the difference between the startup probability of men and women who are 0 years old. The coefficient on *age*, β_2, measures the effect of an additional year of age on the startup probability of a male person. The coefficient on the interaction term, β_3, explains how the gender gap in startup probabilities evolves with age, or how the effect of age on startup probability varies with gender.

Firstly, it should be obvious from here that p-values and significance stars obtained for the *female* variable are of little use: why would we want to run a test about the gender gap for newborn babies? To see whether there exists a gender gap in (1), we need to run a test of *joint significance* of *female* and *female* × *age.* Secondly, the p-value in the *age* row of a regression table clearly reveals nothing about the "effect of age on a typical respondent"; this would imply that the typical respondent is necessarily a male person. We have encountered both these misconceptions repeatedly in peer reviews.

Another elementary mistake is an omission of one of the constituent parts of the interaction. For instance, changing (1) to

$$startup = \beta_0 + \beta_1\, age + \beta_2\, (female \times age) + \mathbf{x}\boldsymbol{\gamma} + u, \qquad (2)$$

yields a model with a built-in assumption that the gender gap is zero at the age of zero. It is difficult to find a theory that backs such an assumption. In most cases, omitting an interaction's constituent term should be viewed as an unnecessary restriction that limits flexibility of the functional form.[1]

[1] There are exceptions, though. If a researcher needs to estimate an analogue to (1) in a longitudinal setting with unobserved heterogeneity, the usual fixed-effects estimator (e.g. Wooldridge, 2012, Ch. 14) precludes the use of time-invariant regressors (such as gender). Yet, it is perfectly possible to estimate a coefficient on an interaction

2.2 Interaction terms and multicollinearity

In our recent peer review, an author insisted on omitting one of the constituent terms because of multicollinearity. As the interaction term is in a functional relationship with its components, a certain degree of correlation will typically be induced when the interaction term is added. Practitioners can thus expect to see large variance inflation factors, or VIFs (e.g. Wooldridge, 2013, Ch. 3), for the three interaction-related terms. For example, estimating (1) using the dataset from (Lukeš, et al., 2012) yields VIF factors of 10.76, 2.42, and 12.80 for *female*, *age*, and *female* × *age*, respectively. These VIFs would normally be considered as alarming, surpassing the usual rule-of-thumb thresholds that range from 2.5 (Allison, 2012) over 5 (Rogerson, 2001) to 10 (Hair, et al., 1995).

As Allison (2012) points out, this is perfectly natural and should not be a cause for concern, let alone a reason to omit any of the constituent terms. If we are distressed about large VIFs of the interaction variables, we can actually manipulate them by centering the interaction variables (i.e., subtracting the sample mean from the given variable). In the example above, using a centered *age* variable decreases the VIFs to 1.07, 2.42, and 2.41, respectively; with both variables centered, the VIFs are 1.03, 1.07, and 1.08.

Does this imply we should automatically center all variables that enter interactions? Not at all. A closer look reveals it would be a superfluous exercise. Nothing of interest is gained by centering. In a model with intercept, centering will affect neither the coefficient on the interaction term, nor its standard error (and *p*-value). Coefficients on other explanatory variables – those contained in **x** in (1) – will not be affected, either. A change occurs only for the coefficients and standards errors of the constituent terms, which should be obvious: after all, the coefficients on the centered variables have a different interpretation than in the uncentered case. The confidence intervals for, say, the gender gap at 45 years of age, will nevertheless be the same, with and without centering. Moreover, coefficient interpretation is often less straightforward with centered variables. The bottom line is that practitioners should not feel worried only because large VIFs crop up after including interactions in the regression.

3 AVOIDANCE OF ADVANCED ECONOMETRIC TECHNIQUES IN REGRESSIONS WITH INTERACTION TERMS

The previous section dealt with elementary issues that seem to trouble practitioners in the phase of manuscript preparation of revision; few of the resulting mistakes make it to eventual publications. Evidence about these issues comes mostly from peer reviews and internet forums.

Issues presented in the present section are of a different kind. We identified them as a part of a broader attempt to map methodological pitfalls in current empiric research of entrepreneurship. Concretely, we reviewed the methodology in 21 empiric papers that appeared in two 2017 issues of each of the following leading entrepreneurship-oriented journals: *Journal of Business Venturing*, *Entrepreneurship Theory and Practice*, *International Small Business Journal*, and *Journal of Small Business Management*.

We identified several methodological problems, most of which are not related to the present discussion and will be dealt with elsewhere. What caught or attention in connection with

---

of a time-varying and a time-invariant variable, i.e. a model analogous to (2) can still be used, and the criticism given above would not apply. See (Wooldridge, 2010, Ch. 10) for more details.

interactions, though, was the inverse relationship between (i) the use of interaction terms in regressions and (ii) the use advanced econometric techniques. It is true that the presence of interactions will sometimes raise methodological questions, especially if coupled with advanced techniques. It seems that many practitioners simply give up on these techniques and use a safe elementary approach, rather than try and find a satisfactory answer to the questions at hand. We find this lack of effort disappointing. In the sequel, we will try to demonstrate this phenomenon using two concrete examples: the calculation of marginal effects in nonlinear models and the use of imputation techniques for missing data.

3.1 Interactions versus marginal effects in nonlinear models

Consider a general probit model in the standard form,

$$\Pr(y = 1 \mid \mathbf{x}) = \Phi(\mathbf{x}\boldsymbol{\beta}), \tag{3}$$

where y is a binary dependent variable (with the value of 1 being referred to as a "success"), $\Phi(\cdot)$ is the cdf of $N(0, 1)$, and $\mathbf{x}$ and $\boldsymbol{\beta}$ are the vectors of explanatory variables and coefficients, respectively. In today's regression packages, parameter estimation in the probit model is obtained just as easily as in the linear probability model. Interpretation of estimated effect sizes in probit, however, is not trivial. If one wants to avoid the malpractice denounced by Ziliak and McCloskey (1996) as "sign econometrics" (interpretation of only the direction of an effect, rather than its size), it is often advised that *marginal* (or *partial*) *effects* for all explanatory variables should be obtained (Cameron and Trivedi, 2005, Ch. 5; Wooldridge, 2010, Ch. 2).

For a binary variable x_j, the *marginal effect* (ME_j) is defined as the difference in success probability if x_j switches from 0 to 1, i.e. $\mathrm{ME}_j = \Phi(\mathbf{x}_{-j}\boldsymbol{\beta}_{-j} + \beta_j) - \Phi(\mathbf{x}_{-j}\boldsymbol{\beta}_{-j})$, where $\mathbf{x}_{-j}$ and $\boldsymbol{\beta}_{-j}$ denote the vectors $\mathbf{x}$ and $\boldsymbol{\beta}$ without their j-th elements. For a count variable x_j, ME_j is defined analogously as the effect of a unit change in x_j. For a continuous x_j, ME_j is defined as $\mathrm{ME}_j = \partial\Pr(y = 1 \mid \mathbf{x})/\partial x_j = \phi(\mathbf{x}\boldsymbol{\beta})\beta_j$, where $\phi(\cdot)$ is the pdf of $N(0, 1)$; for a small change in x_j, the resulting change in success probability will thus approximately be $\mathrm{ME}_j\Delta x_j$.

Marginal effects depend on the values of explanatory variables for the respondent at hand. Two different statistics are being applied to aggregate marginal effects across the whole sample: *average marginal effects* (AME) and *marginal effects at mean* (MEM). There seems to be a growing consensus that the former is a more appropriate measure. Firstly, unlike MEM, it does not rely on the inelegant concept of a fictitious "mean respondent" – e.g. somebody who is 45 % male and 55 % female. Secondly, AME can be viewed as a sample counterpart to the expectation of the effect across the whole population, which resonates with standard inferential thinking (see Wooldridge, 2010, Ch. 2, for a more detailed argument). For a continuous x_j, the average marginal effect is defined as the sample mean of estimated marginal effects, i.e.

$$\mathrm{AME}_j = n^{-1}\sum\nolimits_{i=1}^{n}\phi(\mathbf{x}\hat{\boldsymbol{\beta}})\hat{\beta}_j\,, \tag{4}$$

where n is the sample size and hats denote coefficient estimates. For count and binary variables, average marginal effects are defined analogously.

In the last decade, most regression packages (Stata, R, SAS, SPSS to name a few) have implemented user-friendly functions that obtain marginal effects in nonlinear models. Even though practitioners have appreciated (and utilized) the practical merits of marginal effects in empiric research, they seem to be reluctant to use the concept in models with interactions. The following example explains why.

Consider the following probit counterpart to the linear probability model (LPM) from (1):

$$\Pr(startup = 1 \mid female, age, \mathbf{x}) = \Phi(\beta_0 + \beta_1 female + \beta_2 age + \beta_3 (female \times age) + \mathbf{x}\boldsymbol{\gamma}). \quad (5)$$

Let us assume that the goal is to test the hypothesis about a moderating effect of gender on the relationship between age and startup probability. The usual procedure involves carrying out a *t*-test to see whether the interaction term is significant in the equation.[2]

First two columns in Table 1 show estimates of the β coefficients for both (1) and (5). The LPM coefficients interpret nicely, but the usual criticism of LMPs applies. The probit model is usually strongly preferred, but the magnitudes of the coefficients do not reveal much about effect sizes. Last two columns show the AMEs. (The calculation of marginal effects in the LPM was similar to that in the probit model.) In the row with the interaction term, *female* × *age*, a blank space appears. That makes perfect sense, of course: the AME of *age* averages the predicted marginal effects for all respondents. For different individuals, the magnitude of predicted marginal effects varies with gender, but these differences are averaged out in the AME.

Table 1: Regression results for models (1) and (5). Data: (Lukeš, et al., 2012). Own calculation.

	Coefficient estimates (β)		AME (percentage points)	
	LPM	probit	LPM	probit
age	−0.00104**	−0.00802**	−0.0814***	−0.0942***
	(0.000462)	(0.00373)	(0.0254)	(0.0250)
female	−0.0572**	−0.309	−4.073***	−4.057***
	(0.0241)	(0.216)	(0.810)	(0.786)
female × *age*	0.000393	−0.00130		
	(0.000545)	(0.00531)		
n	3655	3655	3655	3655

Notes: (i) LPM = linear probability model, AME = average marginal effects. (ii) Control variables in all regressions included year of survey and the respondents' education and region (coefficients not reported). (iii) Standard errors, in parentheses, are based on 100 replications of non-parametric bootstrap. (iv) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

For many practitioners, this is a disappointing feature of the functions that report marginal effects. The interaction term appears dropped in the tables, and there are no significance stars to relate to.

It is vital to understand that the culprit here is not the marginal-effects approach, but our attempt to produce nice-looking and concise tables, whereby we aggregate the marginal effects to the point where the interaction effects cannot be seen. In fact, a big advantage of marginal effects is that they can be averaged for different subpopulations, thus enabling us to produce informative plots like Figure 1. Several authors (e.g. Greene, 2010) suggest that a visual inspection of these plots is the best way to study the interaction effects. But, of course, the plots do not produce the significance stars that make our results easily marketable.

[2] This approach in not undisputed, though. There has been a large discussion about interaction terms in the logit and probit model among political scientist, see the summary of Berry, DeMeritt and Esarey (2010). Economists largely agree that the interaction term is needed to model moderation, see (Ai and Norton, 2003; Greene, 2010).

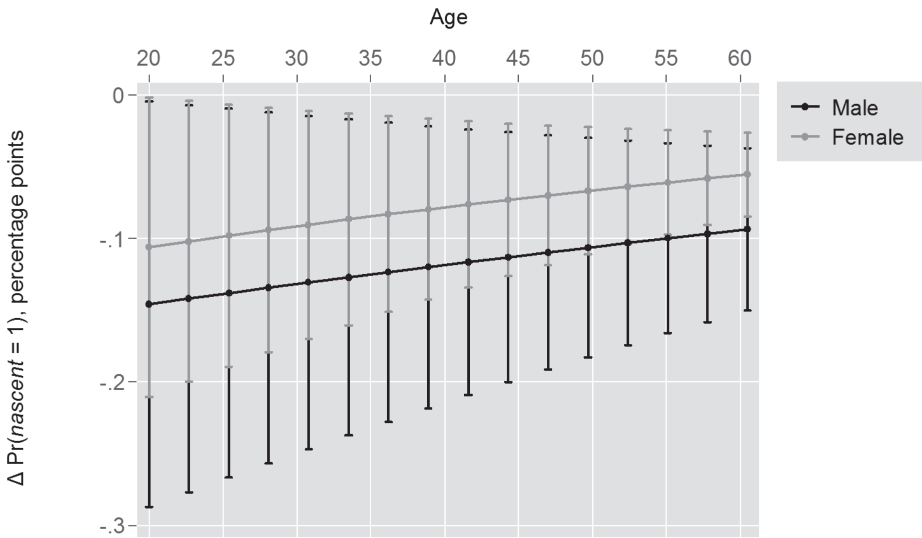


Figure 1: Average marginal effect of *age* for males and females, at different values of age.

3.2 Interactions and multiple imputation of missing values

Scholars have repeatedly pointed out that empiric researchers seem reluctant to adopt state-of-the-art imputation techniques to handle missing data (Davidsson and Gordon, 2012). Presumably, the main obstacle in using these methods is the sheer fact that in many regression packages, the functions that perform these techniques are either missing or not easy to operate.

The presence of interaction terms makes matters even worse for a typical practitioner. The practical obstacles are strengthened by methodological issues. The leading question regarding interactions is how to impute the interaction term in case that the component variables contain missing values. Do we first impute *female* and *age* and then simply calculate the value of *female* × *age*? This is the first solution that springs to one's mind, and one that most software packages implement by default.

It turns out that this procedure typically produces biased estimates. As White, Royston and Wood (2010) argue, in most cases the preferred approach is one that imputes the missing values of the interaction term as if it was a separate variable on its own, not one functionally related to its constituent parts (the so-called *just-another-variable* approach). Unsure whether they can justify their procedure, many researchers just give up on imputation techniques altogether. Sadly, in many cases, this is the worst solution at hand.

4 CONCLUSIONS

Studies of moderating effects are very unlikely to lose their current prevalence in empiric analyses. It is fair to say that the current level of econometric tools used in conjunction with interaction terms does not do justice to the popularity of moderation-driven research hypotheses. We hope we have done our small bit to remedy this situation by pointing out typical obstacles where practitioners seem to stumble.

Acknowledgements: This work was supported by the Faculty of Informatics and Statistics, University of Economics, Prague, grant no. IP400040.

References

[1] Ai, C., Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, 80(1), 123-129.

[2] Allison, P. (2012). When can you safely ignore multicollinearity? *Statistical Horizons*, [blog] 10 September. Available at: <https://statisticalhorizons.com/multicollinearity> [Accessed 20 February 2018].

[3] Berry, W. D., DeMeritt, J. H. and Esarey, J. (2010). Testing for interaction in binary logit and probit models: Is a product term essential? *American Journal of Political Science*, 54(1), 248-266.

[4] Davidsson, P. and Gordon S. R. (2012). Panel studies of new venture creation: A methods-focused review and suggestions for future research. *Small Business Economics*, 39(4), 853-876.

[5] Greene, W. H. (2003). *Econometric Analysis*. 5th ed. Upper Saddle River: Prentice Hall.

[6] Greene, W. (2010). Testing hypotheses about interaction terms in nonlinear models. *Economics Letters*, 107(2), 291-296.

[7] Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. and Black, W. C. (1995). *Multivariate Data Analysis*. 3rd ed. New York: Macmillan.

[8] Lukeš, M., Zouhar, J., Jakl, M. and Očko, P. (2013). Faktory ovlivňující vstup do podnikání: začínající podnikatelé v České republice. *Politická ekonomie*, 61(2), pp. 229-247.

[9] Rogerson, P. A. (2001). *Statistical methods for geography*. London: Sage.

[10] White, I. R., Royston, P. and Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, 30(4), 377-399.

[11] Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. 2nd ed. Cambridge: MIT press.

[12] Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach*. 5th ed. Boston: Cengage Learning.

[13] Ziliak, S. T., McCloskey, D. N. (2004). Size matters: The standard error of regressions in the American Economic Review. *The Journal of Socio-Economics*, 33(5), pp. 527-546.

Authors' addresses

Jan Zouhar, Ph.D.
University of Economics, Prague, Department of Econometrics
W. Churchill Sq. 1938/4, 130 67 Prague 3 – Žižkov
Czech Republic
email: zouharj@vse.cz

Martina Zouharová
University of Economics, Prague, Department of Econometrics
W. Churchill Sq. 1938/4, 130 67 Prague 3 – Žižkov
Czech Republic
email: em.zouharova@gmail.com

COMPARISON OF RADIAL AND SBM DEA MODELS: EFFICIENCY EVALUATION OF COMPANIES SELLING SPORTS EQUIPMENT IN THE CZECH REPUBLIC

Petra Zýková, University of Economics, Prague

Abstract
This paper deals with analysis and comparison between two groups of data envelopment analysis (DEA) models – traditional radial model with the assumption of constant returns to scale (CCR model) and slack-based measures of efficiency and super-efficiency (SBM model). These two models split decision making units into two subsets – efficient and inefficient. The second group can be easily ranked according to efficiency score of the units. The efficient units have identical efficiency scores and cannot be ranked directly. That is why this paper also aims to compare two important super-efficiency models derived from conventional CCR and SBM models – they are Andersen and Petersen model and SBM super-efficiency model. The results of all models are illustrated on a real data set – companies selling sports equipment and accessories in the Czech Republic. All calculations are realized using original procedures written in LINGO modelling language.

Keywords: *Data envelopment analysis, SBM model, efficiency, super-efficiency*

JEL Classification: C44
AMS Classification: 90C05, 90C90

1 INTRODUCTION

The efficiency evaluation of the set of decision making units (DMUs) is an assignment which can be solved in several ways. An application of data envelopment analysis (DEA) is one of them. DEA models have been first developed by Charnes et al. (1979) based on the concept introduced by (Farrell, 1957). The main aim of this paper is to compare two important groups of DEA models - traditional radial model with the assumption of constant return to scale (CCR model) on one hand and slack-based measure model (SBM model) on the other hand. These two models split DMUs into two subsets – efficient and inefficient. Inefficient units can be easily ranked according to efficiency scores but efficient ones cannot be ranked due to their identical maximum efficiency scores. That is why many attempts how to deal with the problem of ranking of efficient units have been published. One of the most important group of models for ranking the efficient units are super-efficiency models. In our study we will compare two super-efficiency models – they are traditional Andersen and Petersen (1993) super-efficiency model derived from CCR model and SBM super-efficiency model derived from SBM model.

The paper is organized as follows: The next section presents definitions of all models used in the study, i.e. traditional radial and SBM models and their super-efficiency modifications. The following section contains numerical illustration modelled on the real data set – evaluation of efficiency of 22 companies selling sports equipment in the Czech Republic. The final section of the paper summarizes the results of all models.

2 DEA MODELS

DEA models are a general tool for efficiency and performance evaluation of the set of homogenous DMUs that spend multiple (w) inputs and transform them into multiple (t) outputs. Measure of efficiency (efficiency score) of this transformation is one of the main results of application of DEA models. Let us denote $\mathbf{Y} = (y_{rj}, r = 1,\ldots,t, j = 1,\ldots,n)$ a positive matrix of outputs and $\mathbf{X} = (x_{kj}, k = 1,\ldots,w, j = 1,\ldots,n)$ a positive matrix of inputs. The efficiency score of the unit under evaluation DMU_{j_0} is as follows:

$$\theta_{j_0} = \frac{\sum_{r=1}^{t} u_r y_{rj_0}}{\sum_{k=1}^{w} v_k x_{kj_0}} \tag{1}$$

where u_r, $r = 1,\ldots,t$ and v_k, $k = 1,\ldots,w$ are positive weights of the outputs and inputs respectively. Traditional DEA models maximize the efficiency score of the j_0-th unit (1) under the assumption that efficiency scores of all other units of the set do not exceed 1.

2.1 CCR–I model

The optimization model derived from (1) is often denoted as CCR (Charnes, Cooper, Rhodes) input-oriented envelopment model and its linearized version is the following:

$$\begin{aligned}
&\text{Maximize} && U_{j_0} = \sum_{r=1}^{t} u_r y_{rj_0} \\
&\text{subject to} && \sum_{r=1}^{t} u_r y_{rj} \le \sum_{k=1}^{w} v_k x_{kj}, \quad j = 1,\ldots,n \\
& && \sum_{k=1}^{w} v_k x_{kj_0} = 1, \\
& && u_r \ge \varepsilon, \; r = 1,\ldots,t \\
& && v_k \ge \varepsilon, \; k = 1,\ldots,w.
\end{aligned} \tag{2}$$

Model (2) will be used further as the first model of our study. It is the traditional radial model with the assumption of constant returns to scale technology. Its multiplier version is computationally more efficient. It is the dual to model (2) and can be written as follows:

$$\begin{aligned}
&\text{Minimize} && U_{j_0} = \theta_{j_0} - \varepsilon\left(\sum_{k=1}^{w} s_k^- + \sum_{r=1}^{t} s_r^+\right) \\
&\text{subject to} && \sum_{j=1}^{n} x_{kj}\lambda_j + s_k^- = \theta_{j_0} x_{kj_0} \quad k = 1,\ldots,w \\
& && \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{rj_0}, \; r = 1,\ldots,t \\
& && \lambda_j \ge 0, \; j = 1,\ldots,n, \\
& && s_k^- \ge 0, \; k = 1,\ldots,w, \\
& && s_r^+ \ge 0, \; r = 1,\ldots,t.
\end{aligned} \tag{3}$$

In this model $\lambda = (\lambda_1, \ldots, \lambda_n)$, $\lambda \geq 0$, is a vector of weights assigned to particular DMUs, $\mathbf{s}^- = (s_1^-, \ldots, s_w^-)$ and $\mathbf{s}^+ = (s_1^+, \ldots, s_t^+)$ are vectors of slack/surplus variables, θ_{j_0} is the efficiency score of the unit under evaluation j_0, and ε is an infinitesimal constant. Efficient units identified by this model have efficiency scores $\theta_{j_0} = 1$ and all slack/surplus variables are equal to 0. Inefficient units have efficiency scores lower than 1. The model is not able to rank efficient units because of their identical efficiency scores. For this purpose super-efficiency has been proposed. They will be discussed further in this section.

2.2 SBM model

The second DEA model which is used in our analysis is slack-based measure model developed by Tone (2001). This model measures efficiency using slack and surplus variables whose positive values express the distance of the unit under evaluation from the efficient frontier constructed by the model. The input-oriented form of Tone's SBM model is as follows.

$$
\begin{aligned}
&\text{Minimize} && \tau_{j_0} = 1 - \frac{1}{w}\sum_{k=1}^{w}\frac{s_k^-}{x_{kj_0}} \\
&\text{subject to} && \frac{1}{t}\sum_{r=1}^{t}\frac{s_r^+}{y_{rj_0}} = 0 \\
& && \sum_{j=1}^{n} x_{kj}\lambda_j + s_k^- = x_{kj_0},\ k = 1,\ldots,w \\
& && \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ = y_{rj_0},\ r = 1,\ldots,t, \\
& && \lambda_j \geq 0,\ j = 1,\ldots,n, \\
& && s_k^- \geq 0,\ k = 1,\ldots,w, \\
& && s_r^+ \geq 0,\ r = 1,\ldots,t,
\end{aligned}
\tag{4}
$$

where τ_{j_0} is the efficiency score of the unit under evaluation j_0. Efficient units by this model have efficiency scores $\tau_{j_0} = 1$, i.e. all slack/surplus variables equal to 0. Inefficient units have efficiency scores τ_{j_0} lower than 1.

2.3 Andersen and Petersen super-efficiency model

In order to rank efficient units in DEA models many models based on various methodological approaches have been proposed in the past. An important group of these models are so-called super-efficiency models. The very first super-efficiency model is the model introduced in (Andersen and Petersen, 1993). This model is derived from conventional CCR model which means that it assumes constant return to scale. Its main idea is to remove the unit under evaluation from the set of units and measure its distance from the new efficient frontier. Model is as follows:

$$
\begin{aligned}
&\text{Minimize} && U_{j_0} = \theta_{j_0} \\
&\text{subject to} && \sum_{j=1,\neq j_0}^{n} x_{kj}\lambda_j + s_k^- = \theta_{j_0} x_{kj_0},\ k = 1,\ldots,w
\end{aligned}
\tag{5}
$$

$$\sum_{j=1,\neq j_0}^{n} y_{rj}\lambda_j - s_r^+ = y_{rj_0},\ r = 1,\ldots,t,$$
$$\lambda_j \geq 0,\ j = 1,\ldots,n,$$
$$s_k^- \geq 0,\ k = 1,\ldots,w,$$
$$s_r^+ \geq 0,\ r = 1,\ldots,t.$$

This model assigns efficiency scores θ_{j_0} greater than 1 to originally efficient units, which is here called super-efficiency score. Inefficient units have efficiency scores θ_{j_0} lower than 1 and are identical to efficiency scores obtained by CCR-I model (3). We can directly rank the whole set of units according to θ_{j_0} values. This relative simplicity is the main reason why Andersen and Petersen super-efficiency model is being applied very often.

2.4 SBM super-efficiency model

Andersen and Petersen model has several significant drawbacks – e.g. the models may not be feasible under the assumption of non-constant returns to scale technology or the model returns often results that are difficult to interpret. As an alternative to Andersen and Petersen model a slack-based measure of super-efficiency was proposed in (Tone, 2002). This model is, in its original formulation, a non-linear model. Due to computational problems with non-linear models either its input or output orientation is used. The input oriented version of SBM super-efficiency model is as follows:

Minimize $$\delta_{j_0} = \frac{1}{w}\sum_{k=1}^{w}\frac{x_k^*}{x_{kj_0}}$$

subject to $$\sum_{j=1,\neq j_0}^{n} x_{kj}\lambda_j + s_k^- = x_k^*,\ k = 1,\ldots,w, \qquad (6)$$
$$\sum_{j=1,\neq j_0}^{n} y_{rj}\lambda_j - s_r^+ = y_r^*,\ r = 1,\ldots,t,$$
$$x_k^* \geq x_{kj_0},\ k = 1,\ldots,w,$$
$$y_r^* = y_{rj_0},\ r = 1,\ldots,t,$$
$$\lambda_j \geq 0,\ j = 1,\ldots,n,$$
$$s_k^- \geq 0,\ k = 1,\ldots,w,$$
$$s_r^+ \geq 0,\ r = 1,\ldots,t.$$

where $\mathbf{x}^* = (x_1^*,\ldots,x_w^*)$ is the vector of optimal inputs, $\mathbf{y}^* = (y_1^*,\ldots,y_t^*)$ is the vector of optimal outputs and δ_{j_0} is the super-efficiency score. SBM super-efficient units by this model have super-efficiency scores δ_{j_0} greater than 1. Originally SBM inefficient units have efficiency scores δ_{j_0} = 1. This is the difference between Andersen and Petersen super-efficiency model, where inefficient units have the same efficiency scores as in CCR-I model. That is why in order to rank all DMUs SBM models must be applied in two stages. The first stage is an application of the SBM model (4). This stage allows ranking of inefficient units. The SBM efficient units by model (4) undertake evaluation by SBM super-efficiency model (6) which is the second stage of the evaluation process. In this way it is possible to obtain a ranking of the whole set of units.

3 A NUMERICAL ILLUSTRATION

The use of above mentioned models is illustrated on a data set of 22 companies selling sports equipment and accessories in the Czech Republic. In the model we use three inputs (nominal capital in thousands of CZK, borrowings in thousands of CZK, and the number of employees) and one output (sales in thousands of CZK). The data set was obtained from publicly available annual accounts 2014. The complete list of companies is given in Table 1.

	Company	CCR-I	SBM
1	SPORTISIMO, s.r.o.	0.3614	0.3207
2	HERVIS Sport a móda, s.r.o.	0.2889	0.1873
3	Intersport ČR s.r.o	1.0000	1.0000
4	HUDYsport a.s.	1.0000	1.0000
5	NIKE Czech s.r.o.	1.0000	1.0000
6	Adidas ČR s.r.o.	1.0000	1.0000
7	JB SPORT, spol. s r.o.	1.0000	1.0000
8	DRAPS s.r.o.	0.1075	0.0631
9	Rock Point a.s.	0.2726	0.1671
10	Columbia Sportswear Czech s.r.o.	0.6261	0.4259
11	Hannah Czech a.s.	0.1314	0.0798
12	Puma Czech Republic s.r.o.	0.2467	0.2383
13	ALPINE PRO, a.s.	0.4651	0.2351
14	DECASPORT s.r.o. (Decathlon)	0.5834	0.3887
15	SEVEN SPORT s r.o. (Insportline)	1.0000	1.0000
16	NABOSO s.r.o.	0.3038	0.2389
17	MONTANA SWISS s.r.o.	1.0000	1.0000
18	HUSKY CZ s.r.o.	0.3473	0.2214
19	Gerberos, spol. s.r.o. (Harfa sport)	0.4515	0.2659
20	A3 SPORT s.r.o.	0.5652	0.3601
21	PONATURE s.r.o. (kilpi)	1.0000	1.0000
22	Piccollo spol. s r.o. (LOAP)	0.5428	0.3589

Table 1: Efficiency scores of companies.

Efficiency scores of the companies obtained by CCR-I model (3) and SBM model (4) are presented in Table 1. The same group of companies was identified as efficient by both models but this conclusion does not hold true in general. They are Intersport ČR s.r.o, HUDY sport a.s., NIKE Czech s.r.o., Adidas ČR s.r.o, JB sport spol. s r. o., SEVEN SPORT s. r. o., MONTANA SWISS s. r.o., PONATURE s.r.o.

In order to discriminate among efficient units there are super-efficiency scores computed by Andersen and Petersen super-efficiency model (5) and SBM input oriented super-efficiency model (6). They are presented in Table 2. This table presents only the results of originally efficient units because for inefficient units Andersen and Petersen model returns the same efficiency score as CCR-I model and SBM super-efficiency model returns efficiency score equal to 1 for all originally inefficient units. The results obtained by both models show that among the most efficient units belong Adidas ČR, s r.o. and Intersport ČR, s r.o.

	Company	AP model	SBM super
3	Intersport ČR s.r.o	2.2348	1.7172
4	HUDYsport a.s.	1.2447	1.1261
5	NIKE Czech s.r.o.	1.0184	1.0078
6	Adidas ČR s.r.o.	2.4068	1.8739
7	JB SPORT. spol. s r.o.	1.6777	1.2871
15	SEVEN SPORT s.r.o. (Insportline)	1.5915	1.3547
17	MONTANA SWISS s.r.o.	1.2385	1.1447
21	PONATURE s.r.o. (kilpi)	1.4795	1.2342

Table 2: Super-efficiency scores of efficient companies.

Both applied models allow complete ranking of all companies. This ranking is given in Table 3. The data show that both rankings are very similar. The higher differences are in values of efficiency scores and super-efficiency scores. SBM super-efficiency scores are always lower than those obtained by Andersen and Petersen model. This is one of the properties of this model that was proved in (Tone, 2002). The ranking of companies by both models is almost identical in our case. This conclusion does not hold true in general.

4 CONCLUSIONS

Efficiency and performance assessment of the set of units is a very common task which is necessary to solve and as such it is of a high interest to researchers. This paper deals with application of two groups of DEA models for this purpose. Firstly, traditional radial CCR input-oriented model is applied and secondly, slack-based measure of efficiency model. These two models split set of the units into two disjoint sub-groups – efficient and inefficient, and give every unit its own efficiency score. These models are not able to rank efficient units, only inefficient ones. Therefore super-efficiency models have been proposed. Traditional Andersen and Petersen super-efficiency model is derived from CCR model and assigns super-efficiency scores to originally efficient units. Efficient units have super-efficiency score greater than 1, inefficient units have same efficiency score as obtained by CCR input-oriented model. SBM super-efficiency model returns super-efficiency score for all units. Efficient units have super-efficiency score greater than 1, inefficient units have super-efficiency score equal to 1. This is one of the main differences between these two super-efficiency models. One is able to get the ranking of all units according to super-efficiency scores obtained by Andersen and Petersen super-efficiency model but SBM model must be applied in two stages in order to obtain complete ranking.

Both DEA models – CCR-I and SBM – and their super-efficiency modification were applied to efficiency evaluation of 22 companies selling sports equipment in the Czech Republic. The study is a rather illustrative case and not serious study but the given results well demonstrate the main properties of both groups of DEA models. Even though the rankings of units obtained by both models are similar the concrete values of efficiency scores differ quite significantly. The numerical experiments were realized using original procedures written in LINGO modelling language.

	Company	AP model	SBM super
1	SPORTISIMO, s.r.o.	15	13
2	HERVIS Sport a móda, s.r.o.	18	19
3	Intersport ČR s.r.o	2	2
4	HUDYsport a.s.	6	7
5	NIKE Czech s.r.o.	8	8
6	Adidas ČR s.r.o.	1	1
7	JB SPORT, spol. s r.o.	3	4
8	DRAPS s.r.o.	22	22
9	Rock Point a.s.	19	20
10	Columbia Sportswear Czech s.r.o.	9	9
11	Hannah Czech a.s.	21	21
12	Puma Czech Republic s.r.o.	20	16
13	ALPINE PRO, a.s.	13	17
14	DECASPORT s.r.o. (Decathlon)	10	10
15	SEVEN SPORT s.r.o. (Insportline)	4	3
16	NABOSO s.r.o.	17	15
17	MONTANA SWISS s.r.o.	7	6
18	HUSKY CZ s.r.o.	16	18
19	Gerberos, spol. s.r.o. (Harfa sport)	14	14
20	A3 SPORT s.r.o.	11	11
21	PONATURE s.r.o. (kilpi)	5	5
22	Piccollo spol. s r.o. (LOAP)	12	12

Table 3: Complete ranking of companies.

Acknowledgements: This work was supported by the Internal Grant Agency of the Faculty of Informatics and Statistics, University of Economics, Prague, project F4/20/2018 (Data envelopment analysis models in economic decision making).

References

[1] Andersen, P. and Petersen, N. (1993). A Procedure for Ranking Efficient Units in Data Envelopment Analysis. *Management Science*, 39(10), pp. 1261-1264.

[2] Charnes, A., Cooper, W. and Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2(6), pp. 429-444.

[3] Farrell, M. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3), p. 253.

[4] Tone, K. (2001). A slacks-based measure of super-efficiency in data envelopment analysis. *European Journal of Operational Research*, 130(2), pp. 498-509.

[5] Tone, K. (2002). A slacks-based measure of super-efficiency in data envelopment analysis. *European Journal of Operational Research*, 143(1), pp. 32-41.

Author's address

Ing. Bc. Petra Zýková
University of Economics Prague, Faculty of Informatics and Statistics, Department of Econometrics
W. Churchill Sq. 4, 130 67 Prague 3
Czech Republic
email: petra.zykova@vse.cz