The 13th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena

Conference Proceedings

Edited by

Monika Papież and Sławomir Śmiech

May 13–16, 2019
Zakopane, Poland

Conference is held under the patronage of
Department of Statistics, Cracow University of Economics, Poland
Committee of Statistics and Econometrics, Polish Academy of Sciences
Polish Statistical Association

Honorary Patronage of the Conference
Rector of the Cracow University of Economics
Prof. UEK dr hab. inż. Andrzej Chochół

Editors: Monika Papież, Sławomir Śmiech

All published papers have been reviewed by two referees before publishing using "double-blind" review process.

Suggested citation:
AUTHOR, A. (2019). Title of the paper. In M. Papież and S. Śmiech (Eds.), *The 13th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings*. Warszawa: Wydawnictwo C.H. Beck, pp. xx-xx. ISBN: 978-83-8158-734-1 (pdf)

The proceedings were not amended or proofread and editors are not responsible for the language used in papers

The publication is financially supported by the Polish Academy of Sciences

Publishing House:
Wydawnictwo C.H. Beck, 17 Bonifraterska St., 00-203 Warsaw, Poland

ISBN: 978-83-8158-734-1 (PDF)

Year of publishing: 2019

Number of pages: 258

Conference website: http://www.konferencjazakopianska.pl/

http://www.pliki.konferencjazakopianska.pl/proceedings_2019/

Organizer's website:

http://www.katstat.uek.krakow.pl/en/

List of Referees:

Paweł Baran, *University of Szczecin, Poland*
Beata Bieszk-Stolorz, *University of Szczecin, Poland*
Małgorzata Ćwiek, *Cracow University of Economics, Poland*
Grażyna Dehnel, *Poznan University of Economics and Business, Poland*
Krzysztof Dmytrów, *University of Szczecin, Poland*
Iwona Foryś, *University of Szczecin, Poland*
Magdalena Gostkowska-Drzewicka, *University of Gdansk, Poland*
Justyna Góral, *Institute of Agricultural and Food Economics – National Research Institute in Warsaw, Poland*
Mariusz Grabowski, *Cracow University of Economics, Poland*
Roman Huptas, *Cracow University of Economics, Poland*
Petr Hlaváček, *Jan Evangelista Purkyně University in Ústí nad Labem, Czech Republic*
Przemysław Jaśko, *Cracow University of Economics, Poland*
Maciej Jewczak, *University of Lodz, Poland*
Marek Kwas, *SGH Warsaw School of Economics, Poland*
Karol Korczak, *University of Lodz, Poland*
Daniel Kosiorowski, *Cracow University of Economics, Poland*
Paweł Lula, *Cracow University of Economics, Poland*
Ewa Majerowska, *University of Gdansk, Poland*
Justyna Majewska, *University of Economics in Katowice, Poland*
Iwona Markowicz, *University of Szczecin, Poland*
Małgorzata Markowska, *Wroclaw University of Economics, Poland*
Iwona Müller-Frączek, *Nicolaus Copernicus University in Torun, Poland*
Jacek Osiewalski, *Cracow University of Economics, Poland*
Viera Pacáková, *University of Pardubice, Czech Republic*
Barbara Pawełek, *Cracow University of Economics, Poland*
Włodzimierz Rembisz, *Institute of Agricultural and Food Economics – National Research Institute in Warsaw, Poland*
Michał Rubaszek, *SGH Warsaw School of Economics, Poland*
Aneta Rybicka, *Wroclaw University of Economics, Poland*
Victor Shevchuk, *Cracow University of Technology, Poland*
Elena Sibirskaya, *Plekhanov Russian University of Economics Moscow, Russia*
Agata Sielska, *SGH Warsaw School of Economics, Poland*
Elżbieta Sobczak, *Wroclaw University of Economics, Poland*
Małgorzata Stec, *University of Rzeszow, Poland*
Grażyna Trzpiot, *University of Economics in Katowice, Poland*
Marek Walesiak, *Wroclaw University of Economics, Poland*
Łukasz Wawrowski, *Poznan University of Economics and Business, Poland*
Sylwia Wiśniewska, *Cracow University of Economics, Poland*
Wojciech Zieliński, *Warsaw University of Life Sciences, Poland*

Table of Contents

PREFACE

A quantitative approach in the analyses of socio-economic processes is based on three viewpoints: that of mathematics, statistics and economic theory. All these methods use measurements in order to uncover relations or to verify relevant hypotheses.

The problems of contemporary socio-economics, in both global and regional scopes, especially the ones related to Central and Eastern European countries, are discussed during the 13th Professor Aleksander Zeliaś International Conference on Modelling and Forecasting of Socio-Economic Phenomena, which takes place in Zakopane on 13–16 May 2019. The conference gathers academics from Poland, the Czech Republic, Germany, Italy, Slovakia, and Ukraine.

The book presented here contains a selection of papers submitted to the conference and revised independently by two anonymous reviewers in the double-blinded review process. This book is a result of scientific and research work on issues related to modelling and forecasting of social and economic processes carried out recently both at Polish and foreign universities (in the Czech Republic, Germany, Italy, Slovakia).

The chapters use a quantitative approach to discuss valid economic problems or present extensions of statistical or econometrical methods. The book addresses various research areas which include, among others: classification problems – theory and applications, statistical methods in regional and international comparative analysis, econometric modelling, statistical modelling, social statistics, labour market analysis, financial econometrics, economic forecasting.

We hope that Readers of this volume will find in it original ideas, flashes of inspirations, and exciting and useful statistical and econometrical methods. We also believe that the book is a valuable voice in an interesting debate which sheds some light on the most critical problems of the modern world.

Monika Papież, Sławomir Śmiech

Application of the classification trees for separation of groups of persons threatened by the long-term unemployment

Beata Bieszk-Stolorz[1], Krzysztof Dmytrów[2]

Abstract

In order to conduct the effective labour market policy, identification of groups of persons threatened by the long-term unemployment is essential. The goal of the research is separation of these groups of persons and designation which long-term unemployed persons are more often de-registered for work and which ones resign from mediation of the labour office. Classification trees were used for separation of these groups. Characteristics that divided the unemployed persons were: gender, age, education, seniority and the number of registrations in the office. In the first stage of the research the groups of persons at greater risk of the long-term unemployment in comparison to other groups were separated. The de-registrations to work and removals for reasons attributable to the unemployed person were the most common. These two forms of de-registration among the long-term unemployed persons were analysed in the second stage. The individual data for persons de-registered from the Poviat Labour Office in Szczecin in years 2013–2017 were used. About 1/3 of all the analysed persons were sampled to the training dataset, and obtained results were applied for the test dataset. The quality of the classification was evaluated by means of accuracy, sensitivity and precision.

Keywords: *long-term unemployment, de-registration forms, classification trees, evaluation of classification*

JEL Classification: *C38, J64*

1. Introduction

In order to lead the efficient labour market policy, the identification of persons threatened by the long-term unemployment is essential. The share of these people in the total registered unemployment in 2017 in Poland and the Zachodniopomorskie voivodeship was almost 55% and in Szczecin – less than 47%. They are the persons looking for employment actively, but in vain. Their competences, professional education or age make it difficult for them to find a job offer and they are covered by various activation programmes. The effectiveness of these programmes is territorially diversified (Bieszk-Stolorz and Dmytrów, 2018b). The registered unemployment rate in Poland and the long-term unemployment rate has decreased since 2013 (fig. 1). Processes on the Polish labour market are similar to those on the Slovak and Hungarian ones (Hadaś-Dyduch et al., 2016). In the Visegrad Group countries during the financial crisis 2007–2009 especially young persons were threatened by the long-term unemployment and it had great impact on their future professional life (Pavelka, 2016). The unemployment is affected by the social policy of a country, carried out, inter alia, by the labour offices.

1 Corresponding author: University of Szczecin/Institute of Econometrics and Statistics, Department of Operations Research and Applied Mathematics in Economics, 64 Mickiewicza St., 71–101 Szczecin, beata.bieszk-stolorz@usz.edu.pl.

2 University of Szczecin/Institute of Econometrics and Statistics, Department of Operations Research and Applied Mathematics in Economics, 64 Mickiewicza St. 71–101 Szczecin, krzysztof.dmytrow@usz.edu.pl.

Activation programmes directed to the threatened groups of persons have a positive impact on the labour market. They should also be focused on the social integration and formation of the long-term unemployment persons' social skills (Fervers, 2018). However, extensive system of unemployment benefits may lead to extension of the unemployment duration (Bieszk-Stolorz and Markowicz, 2015).

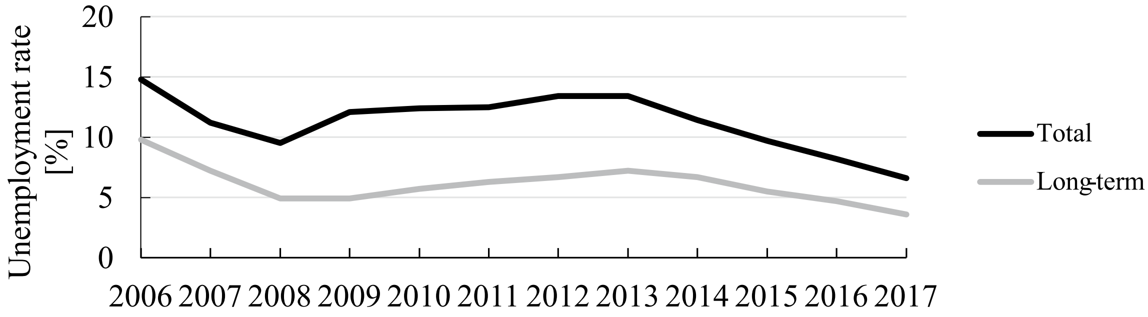


Fig. 1. Registered total and long-term unemployment rates in Poland in years 2006–2017

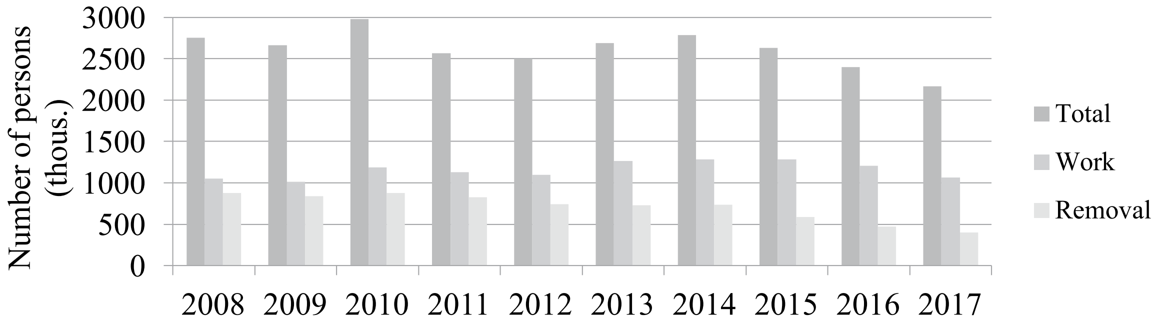


Fig. 2. Number of de-registrations from the labour offices in Poland in years 2008–2017

Not all unemployed are interested in co-operation with the office. Work is the most frequent cause of de-registration from the office. Removal due to lack of readiness for work is the second one. In the years 2006–2017 the lack of readiness for work constituted from 19% to 32% of all de-registrations in Poland (fig. 2, on the basis of data from Yearbooks of Labour Statistics). Some of the unemployed persons do not inform the office about finding employment. Hitherto performed analyses of the labour market in Szczecin indicated that gender did not significantly affect the probability of finding employment, but men were more intensively removed from the register. On the contrary, education and age of the unemployed persons strongly afffected both de-registrations to work or removals (Bieszk-Stolorz and Dmytrów, 2018a). The European Union data from the research on the income and living conditions (EU-SILC) for 24 European countries (2005–2012) indicate that the risk of the long-term unemployment increases particularly for persons with low qualifications and professions, single parents, immigrants and the disabled persons. Women, older persons and permanently employed are less affected by the short-term unemployment, but more by the long-term one (Heidenreich, 2015).

The goal of the research is separation (by means of the classification trees) of groups of persons threatened by the long-term unemployment and designation which long-term unemployed persons are more often de-registered to work and which ones – removed from the register.

2. Data used in the research

The anonymous data about the de-registered unemployed persons from the Poviat Labour Office in Szczecin in 2013–2017 was used in the research. It contains information about: gender, age, education, seniority, number of subsequent registrations, unemployment duration and cause of de-registration. The long-term unemployed persons were separated. They consisted of 22%, 22%, 20%, 17% and 14% of all registered unemployed in subsequent years, respectively (table 1). Data from the labour office contains dozens of causes of de-registrations. Among them, three groups were separated: work, removal and other causes. Other causes were far less numerous and, as earlier researches indicate, each of them separately had only marginal impact on the probability of de-registration (Bieszk-Stolorz, 2017).

Table 1. Number of unemployed persons de-registered from the Poviat Labour Office in Szczecin in 2013–2017

Year	Unemployed total (long-term)	Including	
		Work	Removal
2013	23971 (5336)	11058 (2160)	10692 (2604)
2014	24723 (5518)	11119 (2203)	11392 (2652)
2015	25881 (5266)	11201 (1960)	12672 (2695)
2016	23734 (4091)	10181 (1342)	11718 (2193)
2017	19931 (2706)	8031 (810)	10157 (1320)

In years 2013–2017 on the average 44% of persons registered in the Poviat Labour Office took job and 48% resigned from the co-operation with the labour office. On the contrary, among the long-term unemployed persons this proportion was far more disadvantageous: only 36% of them took job and even 50% of them were removed from the register.

3. Methodology of the research

The research was conducted in two stages. In the first stage homogeneous groups of the unemployed persons particularly threatened by the long-term unemployment (or those in which the ratio of the long-term unemployed was higher than for all population) were separated. In the second stage only the long-term unemployed persons were analysed. Only the ones who took job or were removed from the register were selected. These persons constituted the vast majority of the long-term unemployed persons. Their share decreased from 89% in 2013 to 79% in 2017 (table 1). Later on, the homogeneous groups of the long-term unemployed persons de-registered to work and removed more frequently than the whole population were separated.

The classification was done by means of the classification trees. Among available algorithms, the C&RT (Łapczyński, 2010) was used. Heterogeneity of obtained subsets was analysed by means of the χ^2 statistics (Gatnar, 2001). When constructing the tree, the two aspects should be considered (Capelli and Zhang, 2007): data division, or tree growing and pruning the tree. The more the data is divided, the more homogeneous groups will be obtained, which may result in *overfitting*, which decreases the predictive capacities of the model. The *tree pruning* methods are applied in order to prevent its extensive growth. The tree growing can be stopped if the decrease of heterogeneity of obtained subsets is smaller than certain, assumed value, if we assume certain, minimal group size, at which the division may occur or the maximum tree depth (Mudunuru, 2016). In the research it was assumed that the division will not occur if the size of the subset is less than 100.

In the first stage of the research the unemployment duration (t) until the moment of de-registration was the dependent variable. It was encoded as follows: if $t < 12$ months then the variable took the value 0 and 1 otherwise (in such case it was assumed that analysed person was long-time unemployed). Independent variables in both stages were encoded as follows:

- gender: woman – 1, man – 0;
- age (years): 18–24 – 1, 25–34 – 2; 35–44 – 3, 45–54 – 4, 55–59 – 5, 60–64 – 6;
- education: at most lower secondary – 1, basic vocational – 2, general secondary – 3, vocational secondary – 4, higher – 5;
- seniority: without seniority – 0, with seniority – 1;
- number of subsequent registrations: first– 0, subsequent – 1.

In the second stage the cause of de-registration was the dependent variable. It was encoded as follows: work – 1, removal – 0.

In both stages the whole population of de-registered unemployed persons was divided into the training (about 1/3 of the whole population) and test datasets. In every year the classification trees were constructed for the training set and on the basis of groups obtained for the training dataset, the groups of the unemployed persons particularly threatened by the long-term unemployment (in the first stage) and groups of persons de-registered to work and removed (in the second stage) were separated in the test dataset. The quality of classification was evaluated for the test dataset. It was based on the confusion matrix (table 2) (Fawcett, 2006).

Table 2. Confusion matrix

condition	actual negative	actual positive
predicted negative	true negative (*TN*)	false negative (*FN*)
predicted positive	false positive (*FP*)	true positive (*TP*)

On the basis of the confusion matrix the following measures were calculated:

1. accuracy:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

2. sensitivity:

$$TPR = \frac{TP}{TP + FN} \tag{2}$$

3. precision:

$$PPV = \frac{TP}{TP + FP} \tag{3}$$

4. the F_1 measure:

$$F_1 = \frac{ZTP}{ZTP + FP + FN} \tag{4}$$

Measure (1) is the basic measure of accuracy of classification. From the point of view of the research goal (selection of groups of persons particularly threatened by the long-term unemployment in the first stage and groups of long-term unemployed persons de-registered to work and removed in the second stage) measures (2) and (3) are more important. If we increase the sensitivity, the precision will be adversely affected and vice versa. The F_1 score, being the harmonic mean of sensitivity and precision, measures balance between them. Assessing the quality of classification, more attention was devoted to sensitivity – when classifying too many objects as positive, we will certainly be able to cover more of the real positives, however it will also bring more false positive objects. The reason for attaching more importance to the sensitivity is the fact that if the labour office wants to counteract the long-term unemployment, when the precision is relatively low, substantial part of funds will reach persons who do not need it, but, on the contrary, larger part of persons who actually need help will receive it.

4. Results of the analysis

Firstly, the groups of persons particularly threatened by long-term unemployment were separated (fig. 3). The quality of classification for the test dataset is presented in table 3.

Age, education and gender were the variables that divided the population of the de-registered unemployed persons with respect to the unemployment duration. In subsequent years persons threatened by long-term unemployment created various groups and it was difficult to notice a greater regularity. More frequently there were women, persons with at most secondary education and in older age groups. These unemployed persons should be particularly covered by the programmes of professional activation.

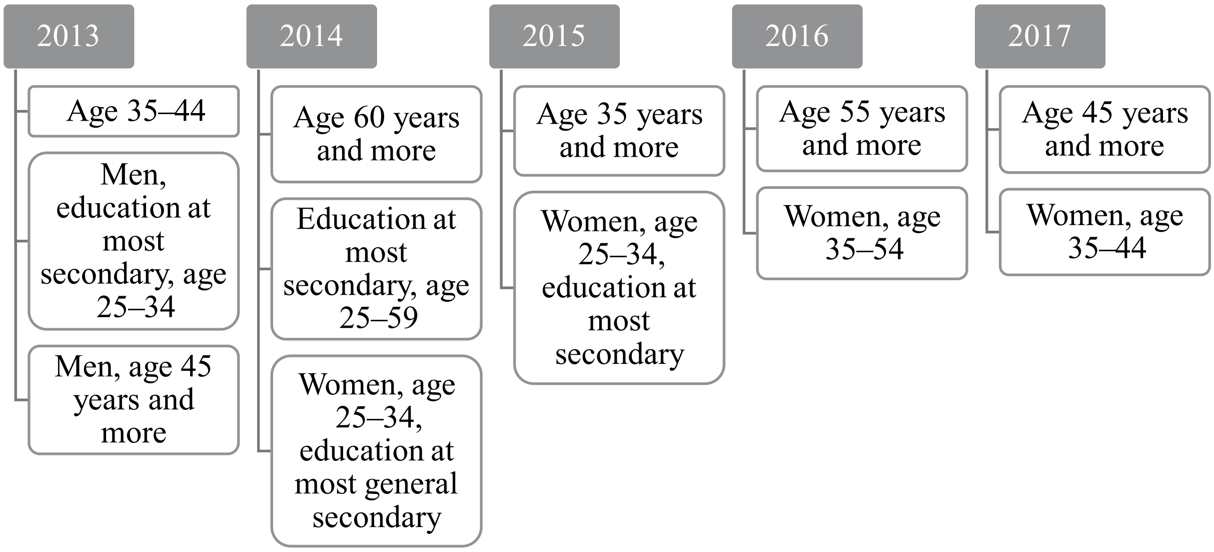


Fig. 3. Groups of persons particularly threatened by long-term unemployment in 2013–2017

Table 3. Quality of classification of persons particularly threatened by the long-term unemployment

Coefficient	2013	2014	2015	2016	2017
ACC [%]	59.88	63.23	61.55	67.20	55.36
TPR [%]	48.03	49.90	43.77	57.93	80.04
PPV [%]	27.66	62.51	54.46	27.38	20.66
F_1 [%]	35.11	55.50	48.53	37.18	32.85

The quality of classification was quite low – analysed unemployed persons constituted quite heterogeneous population. The accuracy was between 55.4% in 2017 and 67% in 2016. The most important from the point of view of social policy sensitivity was generally weak – only in two last years it was better (80% in 2017). Precision was generally at the very low level (except for the years 2014 and 2015). It means that most persons that need help will receive it. High sensitivity means that in 2017 80% of long-term unemployed persons who actually need help will receive it. The second stage of the research was classification of the long-term unemployed persons as taking job and removed due to reasons attributable to them. Selected groups of persons in subsequent years are presented on figures 4 and 5. Quality of classification for the test dataset is presented in table 4.

In case of the long-term unemployed persons de-registered to work and removed from the register, the following features: education, seniority, number of registrations (in 2013) and gender and age (both in 2015) were the ones that divided the population. Generally, long-term unemployed persons that were de-registered to work more frequently were the persons with at least secondary education and with seniority. In 2015 there were women. On the contrary, long-term unemployed persons removed from the register were the persons with lower level of education without seniority. In 2015, as in the case of de-registration to work, gender was also the significant feature that divided the population. On the contrary to the de-registration to work, in 2015 men were more frequently removed than women.

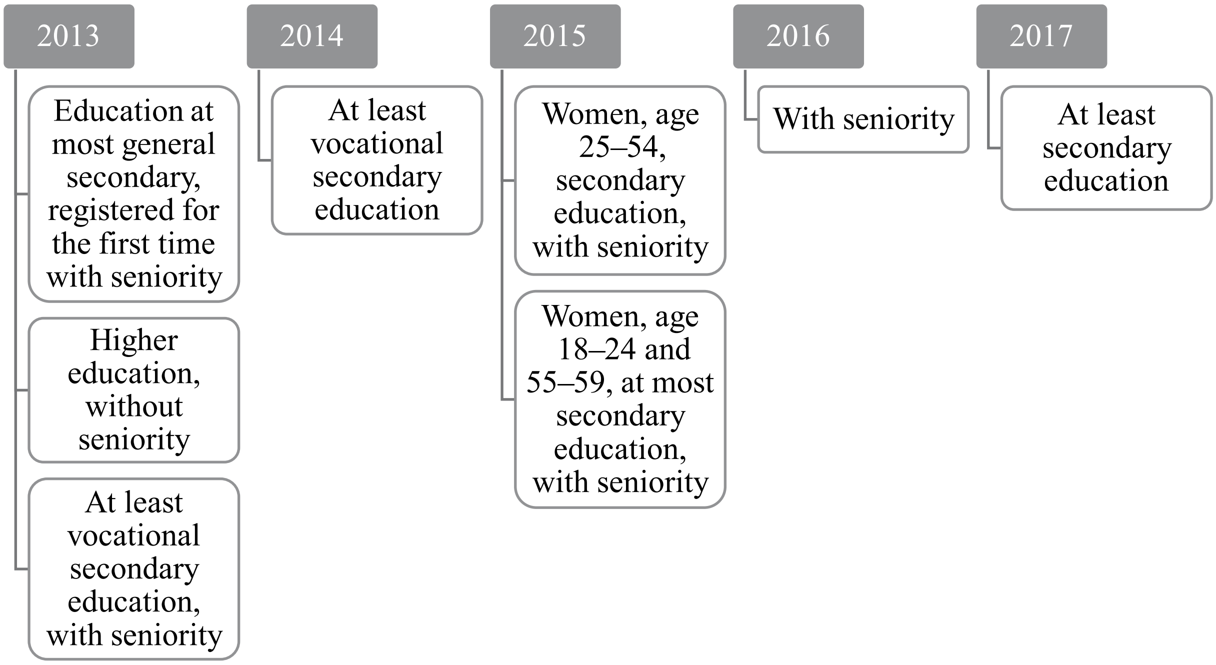


Fig. 4. Groups of long-term unemployed persons de-registered to work in 2013–2017

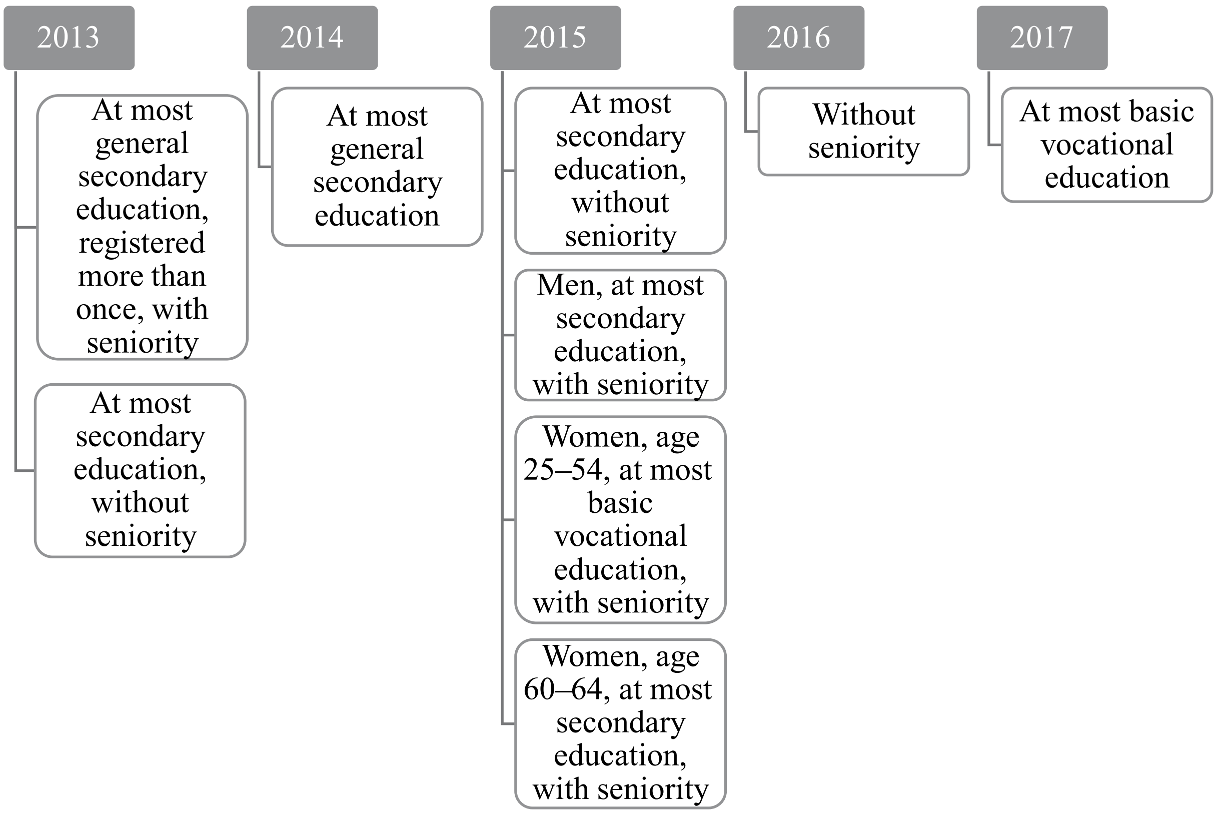


Fig. 5. Groups of long-term unemployed persons removed from the register in 2013–2017

Table 4. Quality of classification of long-term unemployed persons de-registered to work and removed from the register

Coefficient	2013	2014	2015	2016	2017
ACC [%]	63.42	63.23	61.55	52.33	56.95
TPR (work)/*PPV* (removal) [%]	49.58	49.90	43.77	79.69	69.50
PPV (work)/*TPR* (removal) [%]	61.90	62.51	54.46	42.93	45.80
F_1 [%]	55.06	55.50	48.53	55.80	55.21

Similarly, as in the first stage, also in case of the long-term unemployed persons de-registered to work or removed from the register, the population was highly heterogeneous. The accuracy of the classification was between 52% in 2016 and 63.4% in 2013. In years 2013–2015 the precision (for de-registration to work, in case of removal it was the sensitivity) was at the higher level and in the remaining years the sensitivity (for de-registration to work, in case of removal it was the precision) was higher. The F_1 measure, being the harmonic mean of those two, was between 48.5% in 2015 and 55.8% in 2016.

Conclusions

An attempt of classification of persons threatened by long-term unemployment and the long-term unemployed persons de-registered to work and removed from the register due to reasons attributable to them in the Poviat Labour Office in Szczecin in years 2013–2017 was made in the article. Conducted analyses indicated that the analysed population of the unemployed persons was highly heterogeneous, which made the quality of classification quite weak – only in last two years it improved, especially the sensitivity. In the first stage of the research, persons particularly threatened by long-term unemployment were classified with respect to age, education and gender. Persons with lower levels of education were more threatened by long-term unemployment. Also, more often there were women. Despite the insufficient quality of the classification, it was possible to separate the groups of persons to which the help within the counteracting the long-term unemployment could be directed. In the second stage of the research, also the values of the features characterising the long-term unemployed persons de-registered to work and removed from the register were specified (also despite the insufficient quality of the classification). In this case the persons were divided by the following variables: education, seniority and gender (in 2015). Generally persons with better education were de-registered to work, while persons with lower education were more frequently removed from the register. Identification of groups of persons threatened by long-term unemployment on the local labour market is important due to creation of effective labour market policy.

References

Bieszk-Stolorz, B., & Markowicz, I. (2015). Influence of Unemployment Benefit on Duration of Registered Unemployment Spells. *Equilibrium. Quarterly Journal of Economics and Economic Policy,* 10(3), 167–183.

Bieszk-Stolorz, B. (2017). Cumulative Incidence Function in Studies on the Duration of the Unemployment Exit Process, *Folia Oeconomica Stetinensia* 17(1), 138–150.

Bieszk-Stolorz, B., & Dmytrów, K. (2018a). Application of the Survival Trees for Estimation of the Propensity to Accepting a Job and Resignation from the Labour Office Mediation by the Long-Term Unemployed People. In: Nermend, K., Łatuszyńska, M. (eds.), *Problems, Methods and Tools in Experimental and Behavioral Economics. CMEE 2017*. Springer Proceedings in Business and Economics. Springer, Cham, 141–154.

Bieszk-Stolorz, B., & Dmytrów, K. (2018b). Efektywność form aktywizacji zawodowej w przekroju wojewódzkim. *Wiadomości Statystyczne,* 12(691), 57–74.

Cappelli, C., & Zhang, H. (2007). Survival Trees. In: Härdle, W.,·Mori, Y., Vieu, P. (eds.), *Statistical Methods for Biostatistics and Related Fields*. Springer-Verlag, Berlin, 167–179.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters,* 27, 861–874.

Fervers, L. (2018). Can public employment schemes break the negative spiral of long-term unemployment, social exclusion and loss of skills? Evidence from Germany. *Journal of Economic Psychology,* 67, 18–33.

Gatnar, E. (2001). *Nieparametryczna metoda dyskryminacji regresji*. Wydawnictwo Naukowe PWN, Warszawa.

Hadaś-Dyduch, M., Pietrzak, M.B., & Balcerzak, A.P. (2016). Wavelet Analysis of Unemployment Rate in Visegrad Countries. *Globalization and Its Socio-Economic Consequences, 16th International Scientific Conference, Conference Proceedings, University of Zilina, The Faculty of Operation and Economics of Transport and Communication, Department of Economics, 5th–6th October 2016*, 595–602.

Heidenreich, M. (2015). The end of the honeymoon: The increasing differentiation of (long-term) unemployment risks in Europe. *Journal of European Social Policy,* 25(4), 393–413.

Łapczyński, M. (2011). *Drzewa klasyfikacyjne i regresyjne w badaniach marketingowych*. UEK Kraków.

Mudunuru, V.R. (2016). Modeling and Survival Analysis of Breast Cancer: A Statistical, Artificial Neural Network, and Decision Tree Approach. *Graduate Theses and Dissertations*. http://scholarcommons.usf.edu/etd/6120 (19.01.2019).

Pavelka, T. (2016). Long-term unemployment in Visegrad countries. In: Loster, T., Pavelka, T. (eds.), *Proceedings of the 10th International Days of Statistics and Economics*, 1408–1415.

Quality of life and the labour market in Poland

Magdalena Brudz[1], Maciej Jewczak[2]

Abstract

Work is an essential activity in human's life, i.a. due to the amount of time devoted to it, but also because it is a means of meeting both economic and social needs. Having a job or losing it (or lack of employment) is one of the determinants of the level of human welfare, and in recent years dynamic changes on both the supply and demand side have forced multi-faceted research, e.g. related to the quality of employment, cultural diversity or the impact of local factors on the labour market.

The aim of the article is to assess and compare regional (NUTS2) labour markets in Poland in terms of quality of life with the possible spatial effects of this process/relationship. To achieve the research objective, a set of synthetic indicators describing the quality of life in Polish regions was constructed for the needs of the study. The analysis considered processes in the labour market in years 2012–2017 and the tendencies were described with the dynamic Shift-Share Method with spatial modification (SSSA). Within this framework, it was possible to include spatial interactions of regional labour markets between "neighboring" geographic objects (according to the accepted matrix of spatial weights **W**). The approach allowed for an in-depth analysis of labour market conditions in terms of quality of life, as well as verification of hypotheses regarding the regional determinants and spatial differentiation of labour demand and supply, as well as the significant impact of quality of life on labour assets in voivodships.

Keywords: *labour market, regional diversity, social indicator, quality of life*

JEL Classification: *C38, E24, I31, J23, R12*

1. Introduction

It is common for modern societies that people change their place of residence more often in the pursuit of better work and better life quality. However, from the individual's point of view a better life (or job) is defined and interpreted differently, with full regard to the subjective evaluation. It is undeniable that nowadays it is not enough to earn well and have a good job to live happily, have a better life; a wider perspective is crucial and more aspects should be taken into account (Tomkiewicz, 2018).

In 1994 the World Health Organization (WHO) created its Quality of Life section (named WHOQOL), which defined the quality of life term (QOL) as individual perception of one's life, which takes into account cultural conditions and values connected with personal goals, expectations, norms and problems (WHO, 1997). Complexly, quality of life affects physical and mental health, relationships with others and environmental characteristics important for an individual. This definition was the reference point in the proposed study both in theoretical and empirical considerations. It should be also noted that while constructing an indicator of the quality of life, the results should always be interpreted from a subjective perspective. Each person is the

[1] Corresponding author: University of Lodz, Department of Spatial Econometrics, 3/5 POW St., 90–255 Lodz e-mail: magdalena.brudz@uni.lodz.pl.

[2] University of Lodz, Department of Operational Research, 37 Rewolucji 1905 St., 90–214 Lodz, e-mail: maciej.jewczak@uni.lodz.pl.

best judge in his/her case and makes a comprehensive assessment of the quality of life, taking into account personal views and beliefs, which most often result from the place of residence or the hierarchy in the social structures (Hawthorne and Osborne, 2005). Therefore, the presented provinces ratings compiled in the research that compare the quality of life in particular voivodships with the labour market should be perceived as a multidimensional, but a subjective study.

The article analyses the regional results for Polish voivodships in terms of the activity of employees in productive age and the estimated quality of life levels in these regions. As a consequence of the research, in the final part of the paper, spatial units were divided into 4 subgroups. Data on the population's economic activity were sourced from the Labor Force Survey (LFS), whereas the information on the quality of life came from the Central Statistical Office (CSO). The degree of activity was investigated on the basis of collected data for years 2012–2017, with the use of Spatial Shift-Share Analysis (SSSA). The synthetic indicator of the Quality of Life Index (LQI) was estimated to obtain information on the quality of life for every province. The combined and multi-dimensional approach allowed to verify the research objective concerning changes in the professional activity of the population in relation to the population quality of life in given regions.

2. Criteria for building a synthetic index

In order to accomplish the research objective, it was necessary to construct a suitable tool. For this purpose, an LQI measure was constructed that allowed to estimate the quality of life in Polish voivodships. Furthermore, the changes in quality of life indicators were compared with the labour market performance, i.e.: with the tendencies and rates acquired from the SSSA analysis calculated for the total number of professionally active people in the analysed regions.

The most challenging part of obtaining the essential results for LQI was the selection of the most important criteria. Here, the LQI measure was constructed according to the "Better Life Index" described in The Organization for Economic Co-operation and Development documentation (OECD). Therefore, the adopted methodology is convergent with the current recommendations (the 10-step scheme for determining indicators) of the OECD and Joint Research Centre (JRC) of the European Commission (OECD, 2008). However, not all characteristics listed by the OECD were included in presented research,; some of them were omitted intentionally and replaced by other determinants. The research included only objective and available data sources; subjective data, such as results of representative surveys were not included in the study.

Due to the fact that the quality of life is a multidimensional phenomenon, many obvious factors, such as housing conditions and wages, were taken into account for the construction of the indicators. However, an essential element of the analyses was the inclusion of intangible factors, i.e. education, the state of the environment or overall security (Agénor and Lim, 2018). The determinants were generally selected on the basis of recommendations published in reports and studies of the OECD and JRC of the European Commission mentioned already above. The selection of variables was also limited by the availability of reliable Polish regional statistical data. The analysis was based on 58 quantitative characteristics, grouped into 13 thematic categories, according to the CSO classification illustrating various aspects of the quality of life. Each of the category consisted of 1 to 5 determinants.

The LQI estimating procedure included categories listed below and assumed determination of the impact of groups (with weights values in brackets for indicating the impact of each category):

- Public finances (w_1 = 0.1103): revenues of the budgets of the voivodships: revenues of the voivodship budgets in PLN, income per inhabitant; revenues and expenditures of Local Government Units: EU funds for financing programs and projects per 1 inhabitant in PLN, share of investment expenditures of communes and poviats (NUTS4) in total expenditure; expenditure on the budgets of the voivodships: total budget expenditure in PLN, capital expenditure on investment in PLN, expenditure per capita in PLN.
- Security (w_2 = 0.0357): state organization and administration of justice: offences identified by the police in completed preparatory proceedings.
- Housing and communal economy (w_3 = 0.1362): housing allowances: the number of paid housing allowances, the number of housing allowances paid in PLN; housing resources: housing resources, residential buildings, average usable floor space per apartment, average usable floor area of a flat per 1 person, flat for 10,000 population.
- Trade and gastronomy (w_4 = 0.1038): gastronomic establishments: restaurants, bars, canteens and food outlets; shops and petrol stations: shops, petrol stations, the number of inhabitants per store.
- Culture and art (w_5 = 0.1833): centers, houses and cultural centers, clubs and common rooms: a total number; stage and exhibition activities: theaters and music institutions; cinemas: permanent cinema, Multiplexes; museums: museums including departments.
- Health protection, social welfare and benefits for the family (w_6 = 0.0769): outpatient health care: general clinics (public, non-public), clinics for 10 000 population; pharmacies and pharmacy points: pharmacies, population on a publicly accessible pharmacy; medical staff: doctors with the right to practice a medical profession for 10 000 population, nurses and midwives per 10 000 population; hospitals: population per bed in general hospitals; nurseries: children in wages and children's clubs for 1 000 children aged up to 3 years.
- Industry and construction (w_7 = 0.0307): housing construction: apartments put into use per 10 000 population.
- Labour market (w_8 = 0.0786): jobs: number of newly created jobs, number of job vacancies.
- Condition and protection of the environment (w_9 = 0.0304): emission of air pollutants from particularly disruptive plants: emission of gas pollution, emission of dust pollution – in total, total dust emission (Poland = 100), total dust emission per 1 km sq. area; green areas: parks, green areas of the estate, green areas – the total area, the participation of parks, greens and green areas of the estate in the total area.
- Education (w_{10} = 0.0677): primary education; junior high education; secondary education, including general secondary schools, vocational and general vocational schools, basic vo-

cational schools, first-cycle schools and vocational schools special together, post-secondary schools, including colleges: gross scholarization rate for each school type separately; higher education: students of higher schools per 10 000 population.

- Transport and communication (w_{11} = 0.0948): public roads: public roads in total, roads with a hard surface per 100 km sq., roads with a hard surface of 10 000 population; public transport: rolling stock inventory (buses, trams, trolleybuses), length of public transport lines per 1 000 population; bicycle paths: total bike paths, bicycle paths for 10 000 km sq., bicycle paths for 10 000 population.
- Pre-school education (w_{12} = 0.0258): kindergarten, pre-school points: children in kindergartens and pre-school points for 1 000 children aged 3 to 6 years.
- Earning (w_{13} = 0.0258): remuneration: average monthly gross salary in PLN.

3. Survey method

The factors listed in 2nd Section, have been initially standardized and used to estimate the *LQI* indicator for each province and each period separately. Before calculating the cumulative *LQI*, for each of the 13 categories, individual *iLQI*s were firstly calculated. The procedure was based on the non-pattern method (Łuniewska and Tarczyński, 2006), as follows:

$$iLQI_{it} = \frac{\sum_{j=1}^{k} z_{jt}^{i}}{\sum_{j=1}^{k} \max_{i} \left\{ z_{jt}^{i} \right\}} \tag{1}$$

where: $z_{jt}^{i} = v_{jt}^{j} + \max_{i} \left\{ v_{jt}^{i} \right\}$, $i = 1, 2, \ldots, n$; $t = 1, 2, \ldots, T$, v_{jt}^{i} – variables normalised by formula: $v_{jt}^{i} = \frac{x_{jt}^{i}}{\bar{x}_{jt}}$ if x_{jt}^{i} is a stimulant or $v_{jt}^{i} = \frac{\bar{x}_{jt}}{x_{jt}^{i}}$ if x_{jt}^{i} is a destimulant.

The obtained individual *iLQI* indicators were further weighed according to the non-parametric Spearman's correlation coefficients which allowed to determine the strength and direction of the relationship between selected factors (Gaca and Sawiłow, 2014). In order to obtain information about the group's total impact (the weight of category), individual variables have been standardised as follows:

$$w_j = \frac{|\rho_{m+1j}|}{\sum_{j=1}^{m} |\rho_{m+1j}|}\,. \tag{2}$$

Is allowed to omit insignificant correlations which in turn allowed for a reduction in the size of the database. The *iLQI* indicator reaches values between [0; 1] and if it is closer to unity, it indicates a higher weighted quality of life. An attempt was made to estimate the weights and the *LQI* indicator was calculated according to the following formula (3):

$$LQI_{it} = w_j iLQI_{it}\,, \tag{3}$$

where: w_j – weights for categories.

After calculating the weights and indicators of quality of life, voivodships were additionally grouped into four classes due to the *LQI* values, as follows (Nowak, 1990; Kompa and Witkowska, 2010):

- I – very high quality of life when $LQI_{it} > \overline{LQI} + S_{LQI}$;
- II – a high level of quality of life $LQI + S_{LQI} > LQI_{it} > \overline{LQI}$;
- III – average level of quality of life $\overline{LQI} > LQI_{it} > \overline{LQI} - S_{LQI}$;
- IV – low quality of life $LQI_{it} < \overline{LQI} - S_{LQI}$,

where: LQI_{it} – a synthetic measure according to which classification is made, $\overline{LQI}$ – average value of the indicator, S_{LQI} – standard deviation.

The analysis of the labour market, i.e. professionally active people, included the Spatial Shift-Share Analysis (SSSA). The method was proposed because of the spatial units' inseparability, while a single voivodship does not constitute for a separate area. A single region has an impact on neighboring ones, i.e. their economic development, as well as structural changes or competitive position, due to the *First Law of Geography* (Tobler, 1970) indicating that "*everything is related to everything else, but near things are more related than distant things*". This feature is especially applicable to employment mobility and a search for better life quality of families/households. In the study of the labour market using SSSA, a binary matrix of spatial weights **W** was used for voivodships, according to the direct bordering scheme (Suchecki, 2010). This allowed to construct the equation of decomposition of the growth rate has the following form as follows (Nazara and Hewings, 2004) (4):

$$tx_{ri} = tx.. + (\mathbf{W}tx_i - tx..) + (tx_{ri} - \mathbf{W}tx_i)\ , \tag{4}$$

where: **W** is a standardized matrix of spatial weights, $i = 1, \dots, S$; $(tx..)$ is the global effect indicator for a given region; $(\mathbf{W}tx_i - tx..)$ is the values of the rates of structural changes, $(tx_{ri} - \mathbf{W}tx_i)$ is the local component, also known as local competitiveness indicators.

The *Shift-Share* methods are widely used to analyse selected markets, including the labour market. As an example, even the European Commission (Denis et al., 2004; Murtin and Robin, 2018) used a *Shift Share* tool to compare the growth rate of labour productivity in the EU and US economies (labour productivity was determined by the growth rate of productivity of particular industries and changes in the employment structure).

4. Results from the empirical analysis

Looking for links between the quality of life of inhabitants of voivodships in Poland and the situation on the labour market, the quality of life indexes *LQI* were assessed and the pace of tendencies on the labour market with *Shift Share Analysis* was determined. Tab. 1 illustrates the level of quality of life with the assumption of weighing component groups. On this basis, a ranking and classification of spatial objects were created to compare it with the unweighted approach (*uwLQI*).

Table 1. Results of the LQI in 2012–2017 and comparison of classifications

No.	2012	2013	2014	2015	2016	2017	uwLQI	Class
1	MZ	MZ	MZ	MZ	MZ	MZ	MZ	I
2	SL	SL	SL	SL	SL	SL	SL	I
3	DL	DL	DL	DL	DL	DL	DL	II
4	MP	MP	MP	MP	MP	MP	MP	II
5	PO	WP	WP	WP	WP	PO	PODL	II
6	WP	KP	KP	KP	KP	WP	WM	II
7	KP	LBE	LBE	LBE	LBE	KP	WP	II
8	LBE	LBS	LBS	LBS	LBS	LBE	KP	III
9	LBS	LD	LD	LD	LD	LBS	LBE	III
10	LD	PODK	PODK	PODK	OP	LD	LBS	III
11	OP	PODL	PODL	PODL	PODK	OP	LD	III
12	PODK	PO	PO	PO	PODL	PODK	PODK	III
13	PODL	WM	WM	WM	PO	PODL	PO	III
14	WM	ZP	ZP	ZP	WM	WM	ZP	III
15	ZP	OP	OP	OP	ZP	ZP	OP	IV
16	SW	SW	SW	SW	SW	SW	SW	IV

Note: DL – Dolnoslaskie, KP – Kujawsko-Pomorskie, LBE – Lubelskie, LBS – Lubuskie, LD – Lodzkie, MP – Malopolskie, MZ – Mazowieckie, OP – Opolskie, PODK – Podkarpackie, PODL – Podlaskie, PO – Pomorskie, SL – Slaskie, SW – Swietokrzyskie, WM – Warminsko-Mazurskie, WP – Wielkopolskie, ZP – Zachodnio-pomorskie

For comparison, in the table above, the *uwLQI* column contains average *LQI* values in 2012–2017 without weighting processes. The first 4 provinces and the last 3 did not differ in terms of their position in the ranking. For most voivodships, the scaling resulted in a higher position. The exceptions are two provinces, Warminsko-Mazurskie (13th position with weights, 6th without weights) and Podlaskie (11th position with weights, 5th without weights). The results indicate that adjustment by weight variables changes the classification for spatial units in terms of quality of life.

The *Shift Share Analysis* summarizes the processes that took place on the labour market in years 2012–2017. It was assumed that positive changes in the regional market affect the quality of life of citizens. The comparison was presented in the Table 2 below.

Research confirmed that the high values of the quality of life indicator are not necessarily accompanied by dynamic changes in the labour market. The highest positive changes in the situation on regional labour markets characterized spatial objects classified previously into the 2nd or 3rd group in terms of the *LQI* indicator. Moreover, voivodships ranked as the highest were characterised by negative changes on the labour market, which would confirm the assumptions concerning the stronger impact of other factors than having an employment on the quality of life of residents. The changes in the labour market in Swietokrzyskie, Lodzkie and Lubuskie voivodeships should also be assessed negatively. For the most part, negative changes in regional labour markets should be identified with local factors and unfavourable neighbouring scheme with other regions, e.g.: Lodzkie and Mazowieckie.

Table 2. Results of the LQI in 2012–2017 and comparison of classifications

Provinces	Class of LQI	Summary of SSSA effects results		
		net	structural	local
Mazowieckie	I	–2.36%	–1.45%	–0.91%
Slaskie	I	–6.05%	0.29%	–6.33%
Dolnoslaskie	II	8.41%	–2.71%	11.12%
Malopolskie	II	5.06%	1.61%	3.45%
Podlaskie	II	0.01%	–1.02%	1.03%
Warminsko-Mazurskie	II	4.58%	4.71%	–0.13%
Wielkopolskie	II	12.31%	–0.24%	12.55%
Kujawsko-Pomorskie	III	3.96%	–3.64%	7.60%
Lodzkie	III	–12.60%	1.35%	–13.95%
Lubelskie	III	0.23%	0.08%	0.15%
Lubuskie	III	–10.69%	4.79%	–15.48%
Podkarpackie	III	–0.03%	–1.02%	0.99%
Pomorskie	III	11.95%	4.20%	7.75%
Zachodniopomorskie	III	18.01%	2.08%	15.94%
Opolskie	IV	6.29%	–3.81%	10.10%
Swietokrzyskie	IV	–15.88%	–4.86%	–11.02%

Conclusions

The analysis gave information on the overall quality of life and the labour market performance in Polish provinces. Because of the construction, it should be noted that the proposed indicator of the quality of life LQI should be treated as information point due to limitations related to its nature and subjectively included components in the design phase. Although the indicator is based on reliable information, it cannot perfectly reflect the quality of life in voivodships, because of the multidimensional characteristic of the phenomenon.

However, it should be also emphasised that it is possible to make the results more reliable by, for instance, adding certain weighting of structure components. In this scenario, it is more likely to reduce or even avoid the compensativeness effect which means that a relatively high score for some categories can compensate for low scores for most other categories. As an example, one can imagine a city with a highly developed economy, where it is easy to get an attractive job, but at the same time the location is also characterized by a very poor environmental conditioning, which makes it troublesome to live there. It was also observed in the research sample when proper scaling (weighting) boost more regions, which are usually classified as bad conditions (especially economical). Fig. 1 comperes the ratios of weighted and unweighted estimates of LQI indicators, in a way that when the ratio exceeds unity than the weighted measure gives better, accurate results.

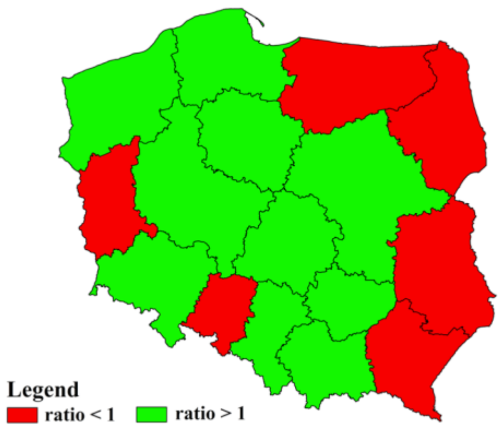


Fig. 1. Weighted or unweighted – spatial comparison

Positively, one should interpret the connections of quality of life indicators and the situation (the tendencies) on the labour market. Although the results may raise some doubts, the lack of clear connections of spatial units with higher levels of quality of life of the inhabitants with the situation on the labour market, indicates that it is not only the professional situation, such as having a good job, that affects the situation of the residents (compare in Table 2 *Class of LQI* and *net effect*).

References

Agénor, P.R., & Lim, K.Y. (2018). Unemployment, growth and welfare effects of labor market reforms. *Journal of Macroeconomics*, 58, 19–38.

European Commission. (2019). *Constructing a composite indicator*. https://ec.europa.eu/jrc/en/coin/10-step-guide/overview (accessed on: 10.01.2019).

Denis, C., McMorrow, K., & Röger, W. (2004). An analysis of EU and US productivity developments (a total economy and industry level perspective). *European Commission, Economic Papers* 208, Brussels, Belgium.

Gaca, R., & Sawiłow, E. (2014). Zastosowanie współczynników korelacji rang Spearmana do ustalenia wag cech rynkowych nieruchomości. *Rzeczoznawca Majątkowy*, 82, 24–30.

Hawthorne, G., & Osborne, R. (2005). Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. *Australian and New Zealand journal of public health*, 29(2), 136–142.

Kompa, K., & Witkowska, D. (2010). Zastosowanie wybranych mierników syntetycznych do porównań poziomu rozwoju społeczno-gospodarczego w krajach Unii Europejskiej, in: Jajuga, K., & Walesiak, M. (eds.), *Taksonomia 17. Klasyfikacja i analiza danych – teoria i zastosowania*, *Nr 107*, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław, 71–79.

Łuniewska, M., & Tarczyński, W. (2006). *Metody wielowymiarowej analizy porównawczej na rynku kapitałowym*. Wydawnictwo Naukowe PWN.

Murtin, F., & Robin, J.M. (2018). Labor market reforms and unemployment dynamics. *Labour Economics*, 50, 3–19.

Nazara, S., & Hewings, G.J. (2004). Spatial structure and taxonomy of decomposition in shift-share analysis. *Growth and change*, 35(4), 476–490.

Nowak, E. (1990). *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych.* PWE, Warszawa.

OECD, JRC European Commission (2008). *Handbook on constructing composite indicators: methodology and user guide*. OECD Publishing.

Suchecki, B. (ed.) (2010). *Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych*. Wydawnictwo C.H. Beck, Warszawa.

Tobler, W.R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46(sup1), 234–240.

Tomkiewicz, J. (2018). The labour market and income distribution in post-socialist economies – Non-obvious regularities. *Communist and Post-Communist Studies*, 51(4), 315–324.

Students' quality of life in the context of competencies

Małgorzata Ćwiek[1], Mariusz Grabowski[2], Renata Oczkowska[3], Robert Szydło[4]

Abstract

The matter of factors that determine the effectiveness of the employee and his professional success has been examined since the very beginning of the existence of management science. Despite the lack of universal consent for a full list of these factors, most researchers agreed to include knowledge, skills and attitudes among them. These are the core elements of competencies. They determine not only professional success itself but also other areas such as self-esteem, student's standard of living, interpersonal relations and coping with life, which together form the quality of life.

The aim of the article is to examine the impact of various competencies on the quality of life of students. The study was based on primary data of quantitative and qualitative nature. They were analysed by the usage of single, double and multidimensional statistics. As a result of conducted analyses, a significant, positive influence of professional, personal and communication competencies was revealed.

Keywords: *competencies, quality of life, factor analysis, analysis of variance*
JEL Classification: *C12, I31, J53, M12, M54*

1. The approach towards competencies

The matter of factors that determine the effectiveness of the employee and his professional success has been examined since the very beginning of the existence of management science. In 1960s and 1970s the concept of competencies was introduced to the literature by R. White and D.C. McClelland (Oczkowska et al., 2017), and later on developed by Boyatzis (1982). Klieme and Hartig (2007) claim that the use of the term competencies has three origins: Weber's sociology, Chomsky's linguistic theory and functional-pragmatic tradition (Klieme et al., 2008). At the same time, Glaesser (2018) mentions Chomsky's understanding of competencies (1968) is slightly different from the rest of educational researches. Some of the authors prefer general and some specific approach towards competencies. Selected definitions are presented in Table 1 and Table 2.

The general approach defines certain personal capabilities that are demonstrated in overall performance or effectiveness. It does not seem to point out to any specific observable skills that can be measured or developed.

1 Cracow University of Economics, Department of Statistics, 27 Rakowicka St. 31-510 Cracow, Poland, malgorzata.cwiek@uek.krakow.pl.

2 Cracow University of Economics, Department of Computational Systems, 27 Rakowicka St. 31-510 Cracow, Poland, mariusz.grabowski@uek.krakow.pl.

3 Corresponding author: Cracow University of Economics, Department of Labour Resources Management, 27 Rakowicka St. 31-510 Cracow, Poland, renata.oczkowska@uek.krakow.pl.

4 Cracow University of Economics, Department of Labour Resources Management, 27 Rakowicka St. 31-510 Cracow, Poland, robert.szydlo@uek.krakow.pl.

Table 1. General approach towards competencies

Author	Definition
Boyatzis (1982)	Underlying characteristic of the person that leads to or causes effective or superior performance.
Thierry, Sauret, Monod (1994)	Inner human potential (competencies which consist of knowledge, skills, values, motives and attitudes) disclosed when translated into the external effects thereof, for example while performing one's professional duties.
Levy-Leboyer (1997)	Set of behaviours which enable some workers to do the job better than the others.
Pocztowski (2001)	Stable personal characteristic that creates the cause-reaction relations with excellent work results.

Table 2. Specific approach towards competencies

Author	Definition
Pawlak (2003)	Competence means both the ability to work as confirmed by documents, as well as those abilities and skills that the candidate can confirm by performing the assigned tasks properly.
Filipowicz (2004)	All knowledge, skills and attitudes that allow completing tasks at the appropriate level.
Penc (2005)	Competence which consists of knowledge (formal qualifications) and practical skills acquired in the course of work at a given position (professional experience and formal entitlements to a specific activity).
Tyrańska (2015)	Main attributes of the person, its personality traits, education, professional experience, knowledge, skills and abilities as well as practical aspects of their use in the implementation of tasks at the workplace.

On the contrary to the general approach, the specific approach towards competencies (Table 2) includes specific components that can be measured or built upon.

It is important to state that the term of competencies is not coherent. For some of the researchers competencies are only the combination of knowledge, skills and attitudes. Others understand the term in a much wider sense by adding a formal aspect of competencies (Pawlak, 2003), values and motives (Thierry et al., 1994) or education, professional experience and personality traits (Tyrańska, 2015). Based on the presented approaches, the definition of competencies as the result of knowledge, skills, attitudes, including personal characteristics of a person, which allow completing tasks at the expected quality, has been accepted. Knowledge, skills and attitudes are common to each definition – they are incontestable parts of competencies. Competencies have an impact on various aspects of our lives. They not only determine professional success itself but also contribute to self-esteem, student's standard of living, interpersonal relations and coping with life, which together form the quality of life.

Quality of life may be considered as a multidimensional concept that consists of health, marriage, family life, friends, home, employment, level of life, free time, education and generally perceived standard of living. Naderifar et al. (2019) claim that quality of life is an all-inclusive concept involving various aspects of life such as financial status, occupation, love, religion, and also the physical, mental, and social health. Eurostat approach indicates that Quality of Life consists of only two dimensions, objective living conditions and subjective well-being. Comprehensive data coming from the medicine studies equate quality of life with health (Solati et al., 2019), but it is only one of the dimensions in most of the approaches. As already mentioned, similar to the concept of competencies, the definition of quality of life and its measurements is not coherent. In spite of all, most approaches are focusing on including various aspects. In the following research the concept of combining various aspects has been implemented.

2. Data and research method

Data used in the research originate from the data collection obtained by the researchers themselves. The questionnaire research was conducted with the use of CSAQ (Computerized Self-Administered Questionnaire), i.e. by means of a computer based questionnaire where respondents could give their answers directly. There were 252 students who participated in the research from among undergraduate and graduate studies (respectively 62% and 21%) as well as the students from the combined bachelor's degree/master's degree programs (13%), doctoral studies (3%) and post-graduate studies (1%). Women represented 82% of the respondents' population. The age range was between 18 and 42 years, while the average age was at 22.9 years.

In order to identify and interpret the latent factors which determine the variability of the observed diagnostic variables, the factor analysis was utilized. The starting point of the analysis is the construction of entry data matrices. Each of the entry variables after its standardization is represented as a linear combination of the unobservable variables, known as the major factors, which carry the information common for the entry variables, and the unique factor, which carries the exclusive information for the entry variable, not present in any other entry variable. It is assumed that the common and unique factors are not correlated. In consequence, the variance of each entry variable can be represented by the variance explained by the common factors as well as by the unique factor (Panek and Zwierzchowski, 2013):

$$S^2(z_j) = h_j^2 + d_j^2 = \sum_{l=1}^{s} w_{jl}^2 + d_j^2 = 1, \quad j = 1, 2, \ldots, m,$$

where: h_j^2 – resources of common variability for j-variable, d_j^2 – resources of unique variability. The factor analysis eliminates the influence of the unique factor in favour of the common factors, and, at the same time, minimizes the influence on the values of the entry variables other than the common factors. The influence is eliminated by replacing the R correlation matrix of the diagonal coefficients of the correlation with the common variability resources. In the result, the below reduced correlation matrix is obtained (Panek and Zwierzchowski, 2013):

$$\tilde{R} = \begin{cases} \tilde{r}_{jj'} = r_{jj'}, & \text{for } j \neq j' \\ \tilde{r}_{jj} = h_j^2, & \text{for } j = j' \end{cases}, \quad j, j' = 1, 2, \ldots, m.$$

The reduced correlation matrix serves as the basis to determine the loadings in the subsequent model equations. The loadings were estimated by means of the principal axis method. The obtained results were subject to the rotation using Varimax rotation with Kaiser normalisation.

3. Students' competencies in the light of empirical studies

The respondents evaluated their competencies by answering 22 questions with regards to various skills and abilities. They were rating their responses with the Likert scale – 1 (very poor) and 5 (very good). The factor analysis was used in order to identify and interpret the latent factors determining the variability of the observed diagnostic variables.

Table 3. Matrix of rotated elements

Variable	Element					
	1	2	3	4	5	6
Showing initiative	0.731	0.151	0.126	0.240	–0.075	0.167
Coping with stress	0.692	0.126	0.222	–0.017	0.088	0.023
Collaboration	0.640	0.005	0.364	0.189	0.05	0.184
Innovativeness	0.502	0.303	–0.137	0.467	0.17	0.069
Client-oriented	0.473	0.360	0.098	0.198	–0.035	–0.004
Analytical skills	0.190	0.699	0.100	0.003	0.111	0.102
Goal orientation	0.053	0.603	0.290	0.458	0.092	0.111
Subject matter expertise related to future job	0.247	0.589	0.177	0.038	–0.023	0.152
Quality-driven	–0.016	0.546	0.378	0.343	0.168	0.155
Organising one's own work	0.192	0.220	0.802	0.179	–0.014	0.087
Time management	0.224	0.121	0.757	0.11	–0.009	0.014
Adaptability	0.399	0.308	0.526	0.057	0.070	0.200
Engagement	0.190	0.080	0.322	0.674	–0.146	0.245
Mobility	0.278	0.009	0.050	0.673	0.156	–0.097
Learning	–0.002	0.357	0.223	0.554	0.060	0.363
Knowing more than one foreign language	–0.043	0.130	0.014	–0.007	0.755	0.097

Variable	Element					
	1	2	3	4	5	6
Knowing English	–0.069	–0.095	0.138	–0.094	0.741	0.126
Ability to work with MS Office, OpenOffice, Google Docs	0.154	0.082	–0.052	0.258	0.610	0.123
Programming skills	0.270	0.405	–0.266	0.146	0.540	–0.067
Writing skills	0.038	0.006	0.200	0.238	0.206	0.766
Speaking skills	0.464	0.108	0.093	0.079	0.096	0.674
Awareness of cultural diversity	0.058	0.367	–0.06	–0.07	0.113	0.625

The first factor is highly correlated with such variables as showing initiative, coping with stress, innovativeness, collaboration and client-orientation. This factor can be tagged as Personal Competencies. The second factor is linked with the Professional Competencies correlated with such variables as: analytical skill, goal-orientation, expert knowledge related to the future employment and being quality-driven. The third factor interpreted as the Competencies of Self-Organization comprises the ability to organize one's own work, time management and adaptability. Engagement and mobility as well as learning compose another factor that can be labelled as the Self-Development Competencies. The fifth factor describes Language Competencies that are correlated with knowing more than one foreign language, knowing English as the foreign language, having a working knowledge of MS Office, OpenOffice, Google Docs and computer programming. The last factor being Communication Competencies is correlated with the following variables: writing skills, speaking skills and cultural diversity awareness.

The internal consistency of the determined factors was tested with a reliability analysis, the results of which are summarized in Table 4.

Table 4. Reliability analysis of the determined factors

Factor No	Factor Name	Number of positions	Cronbach's Alpha
1	Personal Competencies	5	0.766
2	Professional Competencies	4	0.724
3	Self-Organization Competencies	3	0.760
4	Self-Development Competencies	3	0.618
5	Language Competencies	4	0.822
6	Communication Competencies	3	0.834

The reliability of factors presented in the above table is at the adequate level except for factor 4 (Self-Development Competencies). Further analysis proves that Mobility variable should be removed from this factor. Once removed, the value of Cronbach's Alpha for factor 4 increased up to 0.682 which is at a satisfactory level. Descriptive statistics of particular factors have been collated in Table 5.

Self-evaluation of specific competencies among the students is diverse (compare Table 5). The highest value was assigned to factor 4 – the average value is at 4.27. This factor has the lowest value of the coefficient of variation. Factor 5 was evaluated with the lowest average (3.12) and at the same time the highest value of the coefficient of variation (24.8).

Table 5. Descriptive statistics of students' competencies evaluation

Factor No.	Respondents Number	Average	Standard Deviation	Coefficient of Variation
1	252	3.62	0.71	19.59
2	252	3.65	0.62	16.88
3	252	4.03	0.72	17.98
4	252	4.27	0.66	15.53
5	252	3.12	0.77	24.80
6	252	4.01	0.69	17.09

4. Competencies and the quality of students' lives

The quality of students' lives, similar to the competencies, was rated with reference to the 5-level scale. Students were asked to evaluate the quality of their lives in comparison to their peers. The average level of competencies among the population of the students who evaluated the quality of their lives at various levels is depicted in Figure 1.

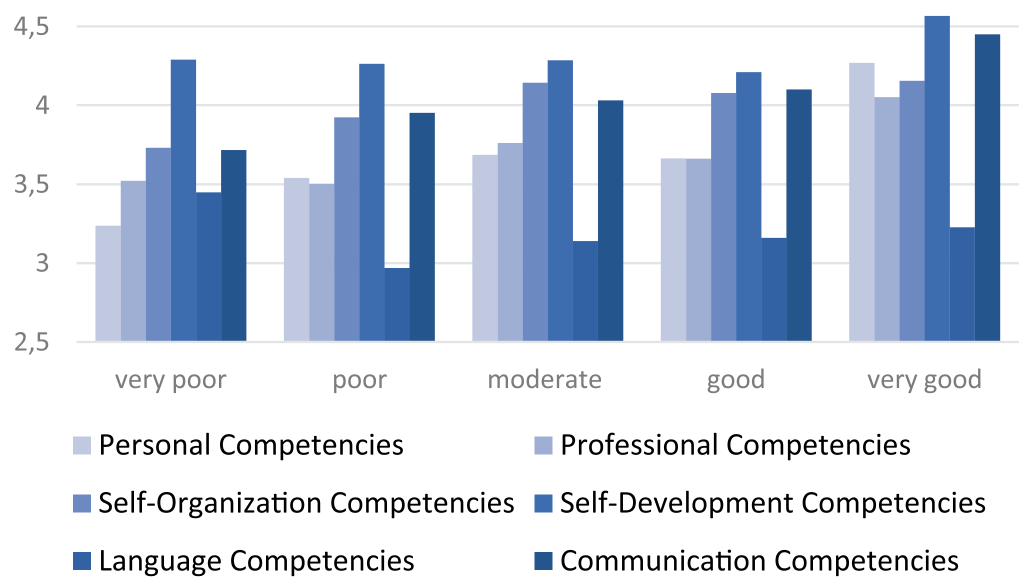


Fig. 1. Evaluation of the competencies by the students who declare various levels of life quality

There is an overall tendency among the evaluators which proves that the high rating of the competencies corresponds with the high evaluation of the quality of life. One exception is observed with regards to factor 5 – the highest average for this factor is observed among the respondents who rated their quality of life at the lowest. This exception can be linked to the fact that foreign language skills and computer programming engage a lot of time and there is not much time left for the students to engage in other life activities, which leads to the lower evaluation of the life quality.

In order to check whether there is a significant difference between the levels of competencies among the participants who declared various levels of the quality of life, a one-way analysis (ANOVA) was applied (Aczel and Sounderpandian, 2009). Factors 3 and 6 were excluded from the analysis as they did not meet the criteria of variance homogeneity. The parameters of the variance are presented in Table 6.

Table 6. Parameters of variance analysis

Variable	Between Group Effect	Within Group Effect	Value of Statistics F	p-value
Factor 1	1.921	0.481	3.994	0,004
Factor 2	1.125	0.369	3.053	0.018
Factor 4	0.446	0.433	1.030	0.392
Factor 5	1.011	0.593	1.705	0.150

The hypothesis of the equality of the average values of Personal Competencies and Professional Competencies should be rejected (where p-value <0.05). Accordingly, it is correct to assume that the quality of life is an important differentiator in the self-evaluation of students' competencies. Taking into account both of the above factors, the participants declaring high levels of life quality also declare high levels of competencies. However, no relation was proved between factors 4 and 5 and the students' quality of life.

Conclusions

Revenue and the resulting consumption are recognized as key determinants of quality of life in the literature (Medcalfe 2018, Jorgenson and Schreyer, 2017, Reeskens and Vandecasteele, 2016). The study showed a significant relation between employee and personal competences as well as the quality of life of students. However, the influence of self-organization, self-development, language and communication competencies on the quality of life has not been demonstrated. This lack of statistical significance may result from inadequate competence assessment or be associated with low experience of respondents on the labour market. The authors plan to extend the students' study, which will result in the analysis of the impact of competences on the quality of life of students in individual years of study, as well as to examine the dependence of competences and quality of life among working and non-working students.

Acknowledgements

The Project was financed by the Ministry of Science and Higher Education within "Regional Initiative of Excellence" Programme for 2019–2022. Project no.: 021/RID/2018/19. Total financing: 11 897 131.40 PLN.

References

Aczel, A.D., & Sounderpandian J. (2009). *Complete Business Statistics*. 7 ed., Boston: McGraw-Hill/Irwin.

Boyatzis, R.E. (1982). *The Competent Manager: A Model for Effective Performance*. NY: John Wiley and Sons.

Chomsky, N. (1968). *Language and Mind.* NY: Brace & World.

Filipowicz, G. (2004). *Zarządzanie kompetencjami zawodowymi.* Warszawa: PWE.

Glaesser, J. (2018). Competence in educational theory and practice a critical discussion. *Oxford Review of Education,* 45*,* 70–85.

Jorgenson, D.W., & Schreyer, P. (2017). Measuring Individual Economic Well-Being and Social Welfare within the Framework of the System of National Accounts. *Review of Income and Wealth,* 63, 460–477.

Klieme, E., & Hartig, J. (2007). The concept of competence in social and educational sciences. *Zeitschrift für Erziehungswissenschaft,* 10, 11–29.

Klieme, E., Hartig, J., & Rauch, D. (2008). *The concept of competence in educational contexts*. In: J. Hartig, E. Klieme & D. Leutner (Eds.), *Assesment of competencies in educational contexts,* 2–22. Gottingen: Hogrefe.

Levy-Leboyer, C. (1997). *Kierowanie kompetencjami. Bilans doświadczeń zawodowych.* Warszawa: Poltext.

Medcalfe, S. (2018). Economic Well-Being in US Metropolitan Statistical Areas. *Social Indicators Research,* 139(3), 1146–1167.

Naderifar, M., Tafreshi, M.,Z., Ilkhani, M., Akbarizadeh, M.,R., & Ghaljaei, F. (2019). Correlation between quality of life and adherence to treatment in hemodialysis patients. *Journal of Renal Injury Prevention,* 8, 22–27.

Oczkowska, R., Wiśniewska, S., & Lula, P. (2017). Analysis of the competence gap among vocational school graduates in the area of smart specialization in Poland. *International Journal for Quality Research,* 11, 945–966.

Panek, T., & Zwierzchowski, J. (2013). *Statystyczne metody wielowymiarowej analizy porównawczej. Teoria i zastosowania*. Warszawa: Oficyna Wydawnicza Szkoła Główna Handlowa.

Pawlak, Z. (2003). *Personalna funkcja firmy – procesy i procedury kadrowe*. Warszawa: Poltext.

Penc, J. (2005). *Role i umiejętności menedżerskie, sekrety sukcesu i kariery.* Warszawa: Difin.

Pocztowski, A. (2001). *Wokół pojęcia kompetencji i ich znaczenia w zarządzaniu zasobami ludzkimi*. In: B. Urbaniak (Ed.), *Gospodarowanie pracą*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Reesknes, T., & Vandecasteele, L. (2016). Hard times and European youth. The effect of economic insecurity on human values, social attitudes and well-being. *International Journal of Psychology,* 52, 19–27.

Solati, K., Mardani, S., Shmadi, A., & Danaei, S. (2019). Effect of mindfulness-based cognitive therapy on quality of life and self-efficacy in dialysis patient. *Journal of Renal Injury Prevention,* 8, 28–33.

Thierry, D., Sauret, C., & Monod, N. (1994). *Zatrudnianie i kompetencje* w *przedsiębiorstwach w procesie zmian.* Warszawa: Poltext.

Tyrańska, M. (2015). *Koncepcja systemu oceny kompetencji kadry menedżerskiej* w *przedsiębiorstwie.* Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.

Unit level models in the assessment of monthly wages of small enterprises employees

Grażyna Dehnel[1], Łukasz Wawrowski[2]

Abstract

Sample surveys are considered to be the most important source of information about monthly wages of employees in small enterprises. The sample size is usually sufficient for precise estimation of parameters at the level of provinces at the most. However, information about local economic conditions at lower levels of territorial aggregation is required to support the development of entrepreneurship. Therefore, an attempt was made to estimate monthly wages of employees at the district level. The study described in the article involves the method of composite estimation as part of the approach based on unit-level models. The aim of the study was to estimate the average monthly wage in small trade enterprises at the level of districts. The study was based on data from a monthly business survey conducted by the Statistical Office in Poznan. Data from administrative registers were used as auxiliary variables. The adoption of the new solutions in the area of business statistics is expected to increase the scope of statistical outputs and improve the efficiency of estimates.

Keywords: *small area estimation, indirect estimation, unit-level model, administrative registers, business statistics*

JEL Classification: *C13, C51, M20*

1. Introduction

The average wage is one of the basic macro-economic indicators. The analysis of average wage levels is necessary for the assessment of the economic and social situation of the country. There is an evident relation between the level of average monthly wages and the labour market as well as economic and social growth, given that wages are the main source of income for most households. They are a key factor enabling individuals to meet their needs. Wage growth boosts consumption and constitutes an important stimulus to economic development.

Analysis of average wage levels and their dynamics would not be complete without the inclusion of the spatial dimension. At the international level, the level of wages, as an important component of labour costs, affects the position of a given country in the international labour market. At the national and regional level, information about the variation in wages can be used in the economy to plan actions aimed at increasing work productivity and improving the competitive position of enterprises and for a more rational management of human resources (Karaszewska, 2003). Taking into account the fact that the human capital is increasingly becoming crucial for competitiveness, information about wages is valuable and helpful at the local level and can also benefit individual enterprises, as wages vary depending on company size and type of activity (Baran and Markowicz, 2018; Dehnel, 2017). In Poland, information about monthly financial results of enterprises by NACE section is published only at the country and

[1] Corresponding author: Poznan University of Economics and Business, Department of Statistics, Niepodległości 10, 61-875 Poznan, Poland, grazyna.dehnel@ue.poznan.pl. ORCiD 0000-0002-0072-9681.

[2] Poznan University of Economics and Business, Department of Statistics, Niepodległości 10, 61-875 Poznan, Poland, lukasz.wawrowski@ue.poznan.pl. ORCiD 0000-0002-1201-5344.

province level. This study is expected to provide information about average wages of business units at the local level cross-classified by NACE section. The aim of the study was to estimate the average monthly wage in small enterprises at the level of districts. The analysis was limited to trade enterprises employing from 10 to 49 employees. Owing to data availability, the analysis was conducted only for the year 2011.

The article consists of five main parts. It starts with an introduction to the subject, which is followed by part describing data sources used for the estimation. The third part includes methodological considerations of the analysis. The fourth part contains a summary of the results and their interpretation. The article ends with conclusions and suggestions for further work. The article is a continuation of the study presented in (Dehnel and Wawrowski, 2018, Dehnel and Wawrowski, 2019).

2. DG1 as Business Survey

The study is based on data from the monthly DG1 survey, which is the main source of information about Polish entrepreneurs. In the survey, a 10% sample of enterprises employing between 10 and 49 is asked to complete a questionnaire about basic company characteristics (Dehnel, 2016). The sampling design of the DG1 survey enables direct estimation using the HT estimator to obtain precise estimates at province level or for NACE sections.

3. Empirical Bayes method for wage estimation

The classic approach to the estimation of total or mean values from survey data relies on the direct estimator proposed by Horvitz and Thompson (1952). It is design-unbiased and design-consistent if the sample size in domain (n_d) is sufficiently large. However, in the case of very small n_d, this estimator is very inefficient and cannot be used for non-sampled domains, i.e. when $n_d = 0$ (Wawrowski, 2016).

Disadvantages of the direct estimator can be overcome by applying small area estimation methods. One of them is the Empirical Bayes method, based on a nested error model, which was introduced by Molina and Rao (2010) for estimating poverty indicators. However, it can be applied for any indicator based on a continuous variable. In this section, we introduce the theoretical background of the Empirical Bayes approach for estimating the mean wage using a nested error linear regression model.

Consider a random vector $y = (Y_1,\ldots,Y_N)'$ which contains values of a random variable associated with N units of a finite population. Let y_s be defined as a sub-vector of with sampled elements and y_r as a sub-vector with out-of-sample elements. After sorting the units, the vector can be written as $y = (y_s', y_r')'$. The aim is to predict the real value of function $\delta = h(y)$ using only sample data y_s.

For the predicted value $\hat{\delta}$ the mean squared error is defined as:

$$MSE(\hat{\delta}) = E_y[(\hat{\delta} - \delta)^2] \tag{1}$$

where E_y denotes the expectation with respect to the joint distribution of the population vector y. The best predictor (BP) of δ is a function of y_s that minimises (1) and it is given by the conditional expectation

$$\delta^B = E_{y_r}(\delta \mid y_s) \tag{2}$$

where the expectation is taken with respect to the conditional distribution of y_r.

For purposes of wage estimation, we can assume that there is a one-to-one transformation $Y_{dj} = T(E_{dj})$ of the wage variables E_{dj}, for j-th unit (company) in d-th domain (district), such that vector y which contains values of the transformed variables Y_{dj} for all population units satisfies $y \sim N(\mu, V)$.

Let $\overline{W}_{dj}$ denote a random variable representing the mean wage calculated based on Y_{dj}. Then $\delta = \overline{W}_d$ and it follows that the BP of $\overline{W}_d$ is given by:

$$\overline{W}_d^B = E_{y_r}(\overline{W}_d \mid y_s) \tag{3}$$

The mean wage can be decomposed in terms of sample and out-of-sample elements:

$$\overline{W}_d = \frac{1}{N_d}\left\{\sum_{j \in s_d} \overline{W}_{dj} + \sum_{j \in r_d} \overline{W}_{dj}\right\} \tag{4}$$

where r_d denotes the set of out-of-sample elements belonging to area d. Now, introducing the conditional expectation inside the sum, the BP becomes:

$$\hat{\overline{W}}_d^B = \frac{1}{N_d}\left\{\sum_{j \in s_d} \overline{W}_{dj} + \sum_{j \in r_d} \hat{\overline{W}}_{dj}^B\right\} \tag{5}$$

As values of vector $\hat{\overline{W}}_d^B$ are unknown, they must be estimated. It can be done because $y = (y_s', y_r')'$ is normally distributed with the mean vector $\mu = (\mu_s', \mu_r')'$ and covariance matrix partitioned conformably as

$$V = \begin{pmatrix} V_s & V_{sr} \\ V_{rs} & V_r \end{pmatrix} \tag{6}$$

the conditional distribution of y_r given y_s is

$$y_r | y_s \sim N(\mu_{r|s}, V_{r|s}), \tag{7}$$

where $\mu_{r|s} = \mu_r + V_{rs}V_s^{-1}(y_s - \mu_s)$ and $V_{r|s} = V_r - V_{rs}V_s^{-1}V_{sr}$.

This estimation can be done using a Monte Carlo simulation involving a large number L of vectors y_r generated from the conditional distribution of y_r given y_s (Rao and Molina, 2015). Let $Y_{dj}^{(l)}$ be the value of out-of-sample observation Y_{dj}, $j \in r_d$, obtained in the l-th simulation, $l = 1, \ldots, L$. A Monte Carlo approximation of the best predictor of for is then given by:

$$\hat{\overline{W}}_{dj}^B = E_{y_r}[h(Y_{dj}) | y_s] \approx \frac{1}{L}\sum_{l=1}^{L} h\left(Y_{dj}^{(l)}\right), \quad j \in r_d. \tag{8}$$

The resulting predictor, denoted $\hat{\overline{W}}_d^{EB}$ is called empirical best predictor (EBP) of $\overline{W}_{dj}$. The EBP of the mean wage $\overline{W}_d$ is given by:

$$\hat{\overline{W}}_d^{EB} = \frac{1}{N_d}\left\{\sum_{j\in s_d}\overline{W}_{dj} + \sum_{j\in r_d}\hat{\overline{W}}_{dj}^{EB}\right\} \tag{9}$$

A nested error linear regression model (Battese et al., 1988) is a model for all areas, which describes the relation between transformed variable Y_{dj} and vectors x_{dj} with p auxiliary variables. Moreover, it includes a random area-specific effect u_d and residual errors e_{dj}:

$$Y_{dj} = x'_{dj}\beta + u_d + e_{dj}, \quad j = 1,\ldots,N_d,\ d = 1,\ldots,D, \tag{10}$$

where $u_d \overset{iid}{\sim} N(0,\sigma_u^2)$ and $e_{dj} \overset{iid}{\sim} N(0,\sigma_e^2)$, u_d and e_{dj} are independent.

This model is used for estimating the conditional distribution of y_r given y_s (7) in the empirical best predictor.

The model MSE of $\hat{\overline{W}}_d^{EB}$ is given by:

$$\mathrm{MSE}\left(\hat{\overline{W}}_d^{EB}\right) = E\left[\left(\hat{\overline{W}}_d^{EB} - \overline{W}_d\right)^2\right] \tag{11}$$

which can be decomposed as the sum of the model variance and the model bias. MSE is estimated using the parametric bootstrap method for finite populations proposed by Gonzales-Manteiga et al. (2008).

The above approach is based on unit-level data which are richer than area-level data, and the model is fitted with a much larger sample. Moreover, this method can be applied to any indicator defined as a function of the variable Y_{dj}. On the other hand, it depends on model assumptions and can be affected by unit-level outliers. The use of Monte Carlo simulations and the parametric bootstrap to obtain estimates is computationally intensive (Guadarrama et al., 2014).

4. Wage estimation at district level

The target level of estimation was the district level (LAU). In the dataset, out of a total of 379 districts, 366 were represented in the sample. 3568 enterprises from those districts were sampled in the DG1 survey. The number of sampled companies in these domains varies from 1 (in 28 districts) to 288 (the capital city of Warsaw) companies with the median of 6 enterprises. The corresponding statistics for the population are a little bit larger: the minimum is 4, the median is 29 and the maximum – 1468 units. Because only one enterprise was sampled in 28 districts, it was impossible to obtain direct estimates of the standard error for these districts, so, in fact, only 338 districts with complete direct estimates are included in the comparison of results.

The first stage of the EB method involves model specification. In our case, the dependent variable was the mean wage per employee in an enterprise. A set of variables from administrative registers were considered as potential auxiliary variables: i.e. gross revenue, the number of employees indicated at the moment of enterprise registration (DG1) and the number of employees from the Social Insurance Institution register. While gross revenue does not require further explanation, the two variables describing the number of employees can be confusing. Information from the first source can be outdated because this value is collected only once – at the time of registration and can change over time. The second one is more up-to-date but itcan be biased as well. For the target group of companies, with 10-49 employees, this variable ranges from 1 to 453. Because both sources are imperfect for modelling purposes, a new variable was calculated, which combines information from both. We assumed that values from the Social Insurance Institution within an interval 1-60 were plausible but in case the number of employees in this source was higher, we took the value from the DG1.

The next step was to estimate the linear nested error model. The dependent variable in that model was the log-transformed average wage, while the explanatory variables include the logarithm of gross revenue (X_1) and the number of employees calculated in the way described above (X_2). Variables used to estimate the monthly wage in small enterprises were chosen based on the assumption that firm-specific variables play an important role in wage determination (Currie and McConnell, 1992*)*. However, the number of variables was limited owing to data availability. The authors are aware that the results depend on the kind of variables taken into account but the main emphasis of the study was to show the possibility of applying a particular methodological approach. In addition, these variables were also tested for collinearity using the variance inflation factor and the results indicate a lack of collinearity. The parameters of the resulting model can be found in Table 1.

Table 1. Mixed model coefficients of the average wage at district level in Poland

Variable	Beta coefficient	Standard error	*t*-value
Intercept	6.0628	0.0530	114.426
X_1	0.1943	0.0064	30.370
X_2	–0.0034	0.0006	–5.309

All variables in the model are statistically significant. An increase in the company's gross revenue is associated with a slight increase in the average wage. However, for enterprises in the analysed NACE section and of this size category, an increase in the number of employees actually leads to a slight decline in the average wage. Random effects variance is equal $\sigma_u^2 = 0.01855$ and residual variance $\sigma_e^2 = 0.14300$. Figure 1 shows the distribution of random effects and residuals.

The random effects and residuals distribution is close to normal, but there are a few outliers. The biggest values of random effects are observed in the biggest cities (the maximum value of random effects is observed for the capital city of Poland – Warsaw) and their neighbouring districts.

After fitting and checking the model, the EB procedure was applied. Estimates of the average wage in districts were obtained after performing $L = 200$ Monte Carlo simulations; the same number of bootstrap replicates was used to compute MSE estimates. All computations were conducted using two R packages – *emdi* (Kreutzmann et al., 2018) and *nlme* (Pinheiro et al., 2018). A comparison of results is presented in Figure 2.

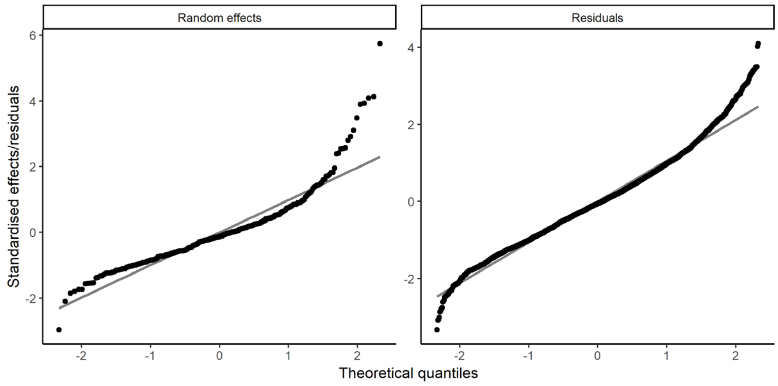


Fig. 1. Distribution of random effects and residuals

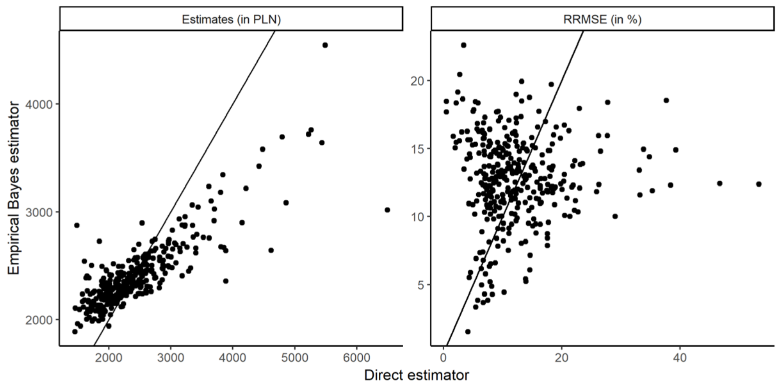


Fig. 2. A comparison of average wage estimates and RRMSE estimates

The EB estimates are highly correlated with the direct estimates (r = 0.84), but the ranges of the two sets of estimates are slightly different. The minimum of Horvitz-Thompson estimates equaled to 1228 PLN and the maximum 6494 PLN. The EB estimates range from 1890 PLN to 4545 PLN. It is worth emphasising that according to the EB estimates, the highest wages could be found in the city of Warsaw and its surrounding districts. However, the mean value was almost the same in both cases: 2411 PLN for direct estimates and 2406 PLN for empirical Bayes estimates.

In the case of relative root mean square error, a gain in estimation precision can be observed for 142 out of 338 districts. For the rest of domains, RRMSE values were slightly higher. Nevertheless, the maximum RRMSE of the direct estimator was 53%, while the maximum value for EB was 23%. More importantly, the EB method made it possible to estimate the average wage for those 41 districts which were either not present in the sample or contained only sampled one unit (making it impossible to calculate direct estimates). Figure 3 shows the spatial diversity of estimated wage values.

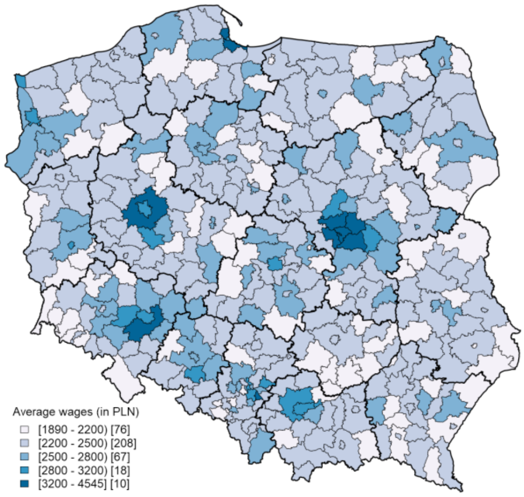


Fig. 3. Spatial diversity of empirical Bayes estimates of the average wage

Because of the right-skewed distribution of the average wage in districts, the range of class intervals presented in the map varies. Figure 3 shows that in most districts (208 out of 379) the average wage is in the interval (2200, 2500). The highest values of wage estimates are found in big cities and their neighbouring districts e.g. Warsaw, Poznan, Wroclaw and port cities of Gdansk and Gdynia.

5. Conclusion and further work

The study was the first attempt of applying the Empirical Bayes method to estimate average wages in small enterprises in the trade sector (NACE Rev. 2 section G). The crucial stage in this approach is model fitting, which, in this study, involved only two auxiliary variables. Unfortunately, access to unit-level enterprise covariates at this level of spatial aggregation is very limited. Nevertheless, the proposed approach improves the precision of average wage estimates and makes it possible to obtain estimates for territorial units not represented in the sample.

Further work will focus on estimating characteristics of enterprises representing other NACE sections, for which the sample size can be even smaller. Given the presence of outliers, another idea is to utilize a two-level robust M-quantile estimator (Chambers and Tzavidis, 2006) or its three-level variant (Beręsewicz et al., 2018) to estimate average and median wages in Poland.

Acknowledgements

The project is financed by the Polish National Science Centre DEC-2015/17/B/HS4/00905.

References

Baran, P., Markowicz, I. (2018). Analysis of intra-Community supply of goods shipped from Poland. *The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings – The Socio-Economic Modelling and Forecasting*, 1, 12–21.

Battese, G.E., Harter, R.M., & Fuller W.A. (1988). An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data. *Journal of the American Statistical Association*, 83, 28–36.

Beręsewicz, M., Marchetti, S., Salvati, N., Szymkowiak, M., & Wawrowski, Ł. (2018). The use of a three level M-quantile model to map poverty at LAU 1 in Poland. *Journal of the Royal Statistical Society Series A*. 4, 1077–1104.

Chambers, R., & Tzavidis, N. (2006). M-quantile models for small area estimation. *Biometrika*, 93(2), 255–268.

Currie, J., McConnell, J. (1992). Firm-Specific Determinants of the Real Wage. *The Review of Economics and Statistics*, 74(2), 297–304.

Dehnel, G. (2016). M-estimators in Business Statistics. *Statistics in Transition–New Series*, 17(4), 749–762.

Dehnel, G. (2017). GREG estimation with reciprocal transformation for a Polish business survey w: M. Papież, S. Śmiech (red.). *The 11th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena: Conference Proceedings*. May 9–12, 2017. Zakopane. Poland. Foundation of the Cracow University of Economics. Kraków. 67–75.

Dehnel, G., Wawrowski, Ł. (2018). Robust estimation of revenues of Polish small companies by NACE section and province. *The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings – The Socio-Economic Modelling and Forecasting*, 1, 110–119.

Dehnel, G., Wawrowski, Ł. (2019). Robust estimation of wages of small enterprises. *Statistics in Transition – New Series* (in press).

González-Manteiga, W., Lombardia, M.J., Molina, I., Morales, D., & Santamaria L. (2008). Bootstrap mean squared error of a small-area EBLUP. *Journal of Statistical Computation and Simulation*, 78, 443–462.

Guadarrama, M., Molina, I., & Rao, J.N.K. (2016). A comparison of small area estimation methods for poverty mapping. *STATISTICS IN TRANSITION new series and SURVEY METHODOLOGY. Joint Issue: Small Area Estimation 2014*, 17(1), 41–66.

Horvitz, D.G., & Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260), 663–685.

Karaszewska, H. (2003). *Ewolucja wynagrodzeń w Polsce w okresie zmian systemu ekonomicznego*. Wydawnictwo Uniwersytetu Mikołaja Kopernika.

Kreutzmann, A., Pannier, S., Rojas-Perilla, N., Schmid, T., Templ, M., & Tzavidis, N. (2018). *emdi: Estimating and Mapping Disaggregated Indicators*. R package version 1.1.4, https://cran.r-project.org/package=emdi.

Molina, I., & Rao, J.N.K. (2010). Small area estimation of poverty indicators. *The Canadian Journal of Statistics*, 38, 369–385.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., R Core Team (2018). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1–137, https://CRAN.R-project.org/package=nlme.

Rao, J.N.K., & Molina, I. (2015). Small area estimation. John Wiley & Sons.

Wawrowski, Ł. (2016). The Spatial Fay-Herriot Model in Poverty Estimation. *Folia Oeconomica Stetinensia*, 16(2), 191–202.

Analysis of the structure and spatial diversity of co-operatives in Poland active after 1990

Iwona Foryś[1]

Abstract
The aim of the paper is comparison activity of cooperatives in Poland after the political changes. The research is aimed at answering the question whether there have been significant changes in the number of active cooperatives, their industry structure and the domination of particular industries in the spatial layout. The study used multidimensional analysis methods as well as methods for assessing the stability of structures. Results of studies: As a result of the study on the number of active housing co-operatives in the subsequent years 1975–2017, it was shown that in 1990, when political transformation took place in Poland, it cannot be said that the popularity of housing co-operatives decreased. The number of active co-operatives followed a growing trend throughout the period under review. Only in 1986 and 2002 can a lower dynamics be observed in this trend. It has been shown that the dominant sector of activity of co-operatives is housing, food trade and agricultural processing. In addition, the largest number of co-operatives is located in the Mazowieckie and Wielkopolskie Voivodeships, while in the latter case the tendency is growing. The research is based on information concerning all 12,754 national economy entities were registered in the Polish business register REGON kept by the Central Statistical Office (as on 30 April 2017) with the legal form of co-operatives.

Keywords: *co-operatives, legal form of economic activity, structure stability*
JEL Classification: *R31, C38*

1. Introduction

A co-operative is a form of joint action of individual entities, which has been popular since the 18th century. It promotes the cooperation of dispersed entities in a joint project in order to achieve benefits for all participants. In Polish law, a co-operative is defined as a voluntary association of an unlimited number of persons, with variable staff and a variable share fund which, in the interests of the members, manages a joint economic activity and may also conduct social and educational activities.

Foryś (2011, 2017) demonstrated that the main idea of co-operatives was distorted in Poland in the years of the planned economy. In this context, it was only political changes and economic changes after 1990 that could change the attitude of society towards co-operatives. The legal form of running - business, badly associated in the previous political system, should expire in a new system (see, e.g. Batóg and Foryś, 2018). However, even superficial analyses indicate that this is still a popular form of activity in Poland. Hence, the research question concerning the activity of co-operatives after 1990. Have the political changes limited the activity of co-operatives in particular groups of economic activity, or are there areas in which co-operatives are no longer popular? Additionally, the problem of spatial location of active co-operatives in Poland can be considered.

[1] University of Szczecin, Faculty of Economics and Management, Institute of Econometrics and Statistics, 64 Mickiewicza St.,71-101 Szczecin, iwona.forys@usz.edu.pl.

The main research objective is to compare the activity of co-operatives in Poland after political changes. The survey aims to answer the question whether there have been significant changes in the number of active co-operatives in the spatial layout, their industry structure. Basic statistics analysis methods were used in the study, as well methods for assessing the stability of structures and dynamic models.

Data from the public statistics resource of the REGON system were used in the study. The information for 12,754 active co-operatives includes (as of 30 April 2017): date of establishment and commencement of activity, section of the classification of economic activity, number of employees and the gmina where the head office of the co-operative is located.

2. Literature review

Cooperation in economic, scientific and social initiatives (see, e.g. Aiken, Cornforth and Spear, 2009) aims at achieving higher benefits than individual action (see, e.g. Kaiser, 2002). This form of joint action is co-operative activity. Its place in the economy became the subject of lively discussion in post-communist countries after 1990.

In his research, Hunt (1992) emphasized that these are not profit-oriented organizations, but a joint action for co-operative members. They are organizations open to people, to local activities, not to the concentration of capital. Co-operatives primarily use local sources of supply, human resources, in line with the idea of sustainable development. The co-operative principles underline the openness of these organizations, without any discrimination, voluntary affiliation and co-decision-making.

In the world there are organizations that cultivate co-operative principles, such as the International co-operative Alliance (ICA), the largest and oldest non-governmental organization founded in 1895, representing more than one billion co-operative members, or co-operatives Europe – the main European co-operative organization since 2006. The latter in 2013 took steps to support the potential of co-operatives to create smart growth and to promote the idea of co-operative activity. Financial support for the development of this type of activity is particularly important, especially in developing countries with a high level of income inequality among households. More than 300 of the world's largest co-operatives have an annual turnover of more than $2.5 trillion and more than 250 million people earn their livelihoods through co-operatives (ICA COOP, 2013).

Since its inception, the co-operative movement has focused on agricultural production, crafts, food industry, housing, savings and credit unions, and over time also social and educational activities. In the past centuries, cooperativeness evolved and adapted to the realities of a modern market economy, preserving the principles inherent in this legal and organizational form to the present day.

In Poland co-operatives flourished the most after regaining independence in 1918 when all forms of co-operatives were established, including housing co-operatives offering flats for the poorer population, which could not afford to buy privately owned apartments.

In the years 1918–1939, registry courts in Poland recorded about 12,000–16,000 active co-operatives, on the basis of the Act on co-operatives passed on 29 October 1920 which was considered one of the best and most modern in Europe in those years. It is estimated that before the outbreak of World War II, every fifth citizen of the Republic of Poland was a member of a co-operative, in retail trade the share of co-operatives reached up to 5%, in the purchase of agricultural products up to 12%, and 20% of savings deposits were placed in co-operative banks and credit unions (according to the Central Statistical Office).

After the end of the Second World War, the co-operative movement was incorporated into the idea of the new system and, as a result, it was deprived of its fundamental principles: self-government and independence. The co-operative movement and co-operatives were dependent on political decisions and central authorities. It was only after 1980 that the co-operative movement in Poland began to regain independence, but with simultaneous limitation of financial support from the state. The problem continued to worsen in the 1990s (Brzozowska-Wabik, 2013), in the period of investment regression, liquidation of State Farms (PGR) and growing competition from other market forms of economic activity. The area of co-operative activity which did not find its place in the new reality, were, inter alia, housing co-operatives, which from investment became entities managing the possessed resources. In addition, these resources are increasingly less frequently owned by the co-operatives themselves (Foryś, 2017).

Today, the global co-operative movement has its supporters, especially in the case of agricultural production and craft co-operatives (Becker and Dietz, 2004), as well as in underdeveloped countries (King and Ortmann, 2007). Attention is paid to innovative organizational solutions, improved profitability of production, cheaper direct marketing. This form of action guarantees co-operatives direct access to the customer and avoids monopolistic lowering of prices below production costs by large corporate distributors. On the other hand, it guarantees the supply of competitive, high-quality products (Bijman and Hendrikse, 2008). It also allows the implementation of innovative sales techniques directly by the producer, while maintaining the benefits of proximity to the markets. Similarly, in many countries co-operative construction (for rent) remains popular due to the competitive quality of housing and lower maintenance costs (e.g. Germany, Scandinavia).

Borzoga and Spear observed that the problem of continued low popularity of this form of economic and social activity remains in post-communist countries where co-operative ideas have been distorted. Hence, the research topic undertaken in the study is important in the context of co-operative perspectives in Poland.

3. Data

The data used in the study refer to 12,754 active co-operatives in Poland as on April 30th 2017, registered in the REGON system in the years 1975–2017. In the first step, an analysis of the surveyed entities was made with regard to two key parameters: the section of the classification of economic activity (PKD) and the location of the seat of the co-operative. Since in almost all cases the year (date) of formation of the co-operative coincides with the date of commencement of the activity, the first one was accepted for the study.

Analysing the years 1975–2017, one can observe a growing trend in the number of active co-operatives, while the dynamics of this growth in individual years is variable (Fig. 1).

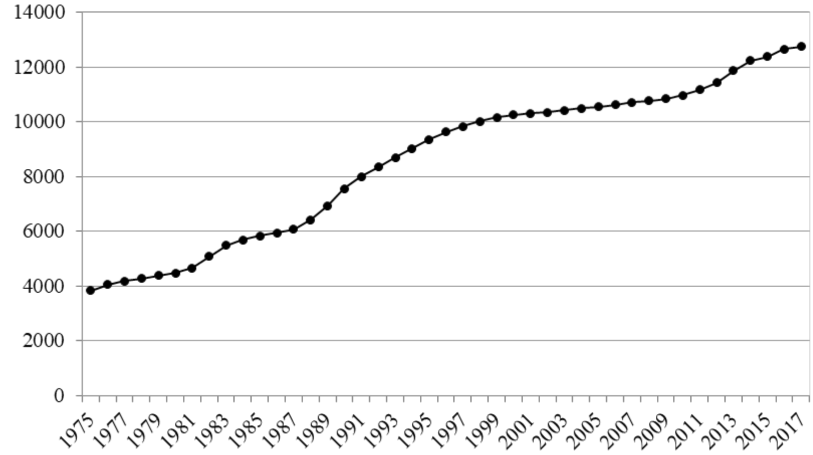


Fig. 1. Dynamics of the number of active co-operatives in Poland in the years 1975–2017

In particular groups of co-operatives, due to the number of employees, the discussed trend was diversified (Tab. 1).

Table 1. Dynamics of the number of active co-operatives in Poland including the number of employees in 1975–2017

Years/Stratum Number of employers	S1 not specified	S2 0–9	S3 10–49	S4 50–249	S5 250 and more	Total
1975–1986	0	2 475	2 135	1 203	128	5 941
1987–2002	0	648	3 594	158	7	4 407
2003–2017	434	166	1 792	13	1	2 406
1975–2017	434	3 289	7 521	1 374	136	12 754

In Poland there are most active medium-sized and small co-operatives employing up to 49 people. These two groups account for 84.76% of the total number of co-operatives between 1975 and 2017. The largest number of co-operatives was established and operated in 1987–2001, in the group employing 10–49 employees, i.e. in the period of political transformations in Poland.

In the next step, an attempt was made to assess the spatial diversity of the dynamics of co-operative activity in Poland, broken down by voivodeships (Tab. 2).

Table 2. Spatial structure of active co-operatives in Poland in 1975–2017* (%)

Voivodeship	1975–1986	1987–2002	2003–2017	1975–2017
Dolnośląskie	6.5	8.7	5.0	7.0
Kujawsko-Pomorskie	5.0	6.0	5.2	5.4
Lubelskie	7.3	7.3	5.2	6.9
Lubuskie	2.5	2.6	3.3	2.7
Łódzkie	6.5	3.9	6.9	5.7
Małopolskie	7.8	5.4	5.4	6.5
Mazowieckie	16.9	22.6	13.4	18.2
Opolskie	3.5	2.1	3.9	3.1
Podkarpackie	5.5	3.7	5.2	4.8
Podlaskie	3.6	1.8	2.9	2.9
Pomorskie	4.9	5.6	4.3	5.0
Śląskie	7.7	5.6	7.5	6.9
Świętokrzyskie	2.9	1.9	2.4	2.4
Warmińsko-Mazurskie	3.3	4.8	6.1	4.3
Wielkopolskie	11.2	10.1	20.3	12.6
Zachodniopomorskie	4.8	8.1	3.2	5.7
Polska	100.0	100.0	100.0	100.0

* – for 4 months

In the years 1975–2017, the largest number of the surveyed co-operatives operated in the Mazowieckie (18.2%) and Wielkopolskie (12.6%) Voivodeships with the oldest tradition in the history of co-operative activity). However, in the Wielkopolskie Voivodeship in the years 2003–2017 this share was twice as high as in the previous period, which proves the strengthening of this form of economic activity, while in the Mazowieckie Voivodeship the tendency is opposite. The lowest co-operative activity is reported by entities in the Lubuskie and Świętokrzyskie Voivodeships.

Then the measure of structure differentiation *(V_3)* was used for the data contained in Table 2, proposed by M. Walesiak (Walesiak, 1983). For the three sub-periods analysed, the indicator values are very close to zero (1.6E-01 and 2.1E-01, respectively), which indicates that the spatial structure of co-operatives is not diversified in the analysed sub-periods.

The analysis is complemented by co-operative activity in particular industries. The NACE sectional breakdown (Tab.3) has been used for this purpose. Definitely, most of the surveyed co-operatives conduct business activity in the area of real estate market (36.6%) – these are mainly housing co-operatives. The second significant area is wholesale and retail trade; repair of motor vehicles and motorcycles (18.1%) – these are mainly Społem consumers' co-operatives. Other industries are represented on a much smaller scale.

Table 3. Structure of active co-operatives according to NACE code in Poland in the years 1975–2017* (%)

NACE codes	Type of business	1975–1990	1991–2002	2003–2017	1975–2017
A	Agriculture, forestry and fishing	15.6	4.1	9.2	11.9
C	Manufacturing	8.2	1.9	8.1	6.7
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	20.2	5.1	26.6	18.1
K	Financial and insurance activities	7.6	1.5	1.4	5.1
L	Real estate activities	33.1	69.7	8.5	36.6
	Other	15.3	17.7	46.2	21.6
	Total	100	100	100	100

However, the Spearman's rank correlation does not confirm the strong relationships between the surveyed variables: voivodeship, industry, number of employees and year of establishment. The determined coefficients have different characters but all of them are close to zero.

4. Models of the dynamics of the number of active housing association in spatial arrangement

Preliminary results (Fig. 1) indicate an increase in the number of housing association in the analysed period and division into sub-periods, however, the data contained in Table 2 (and Appendix) suggest variable dynamics in individual voivodships and division into three sub-periods: 1976–1990, 1991–2002 and 2003–2016. The first year was omitted, in which the number of active housing association in all voivodships lags several times from the other years and incomplete 2017 (4 months). For these sub-periods, the parameters of the trend function were estimated, the table below presents the grouping of voivodships due to similar trends in the number of active housing association (Table 4).

In all voivodships, a decrease in the number of cooperatives after 1990 and in the years 2003–2016 (except for the Mazowieckie Voivodship) can be noticed. On the other hand, different trends in the two groups are in the years 1976–1990, which is the result of historical events. Further research, difficult due to long ranks, should focus on seeking cause and effect relationships of observed trends.

Table 4. Spatial dynamics of active co-operatives in Poland in the years 1976–2016

Group	Voivodeship	1976–1990	1991–2002	2003–2016
I	Lubuskie, Łódzkie, Opolskie, Świętokrzyskie, Podkarpackie	constant	decrease	increase
II	Mazowieckie	increase	decrease	decrease
III	Dolnośląskie, Kujawsko-Pomorskie, Lubelskie Małopolskie, Podlaskie, Pomorskie, Śląskie, Warmińsko-Mazurskie,Wielkopolskie, Zachodniopomorskie	increase	decrease	increase

Conclusions

The analysis of empirical data does not confirm that in 1990 there was a clear shift away from the co-operative form of running a business despite the fact that it was the year of significant political and economic changes in Poland. According to the authors referred to in the article, the image of co-operatives, which has been distorted over the years of the central economy, should discourage entities to conduct business activity in this form. The conducted empirical study does not confirm this assumption in the next periods. The values of differentiation of structures are very close to zero, which indicates that the spatial structure of co-operatives is not diversified in the analysed sub-periods.

The obtained results are noteworthy for the dynamic growth of the number of housing co-operatives in the whole period under review, with two moments of slowdown in the growth rate in 1990 and 2002. They are related to the general economic situation in the country rather than to political changes. The Spearman's rank correlation does not confirm the strong relationships between voivodeship, industry, number of employees and year of establishment.

Housing co-operatives (section L) have a very strong share in co-operative ventures. Although they build new flats to a limited extent in the growing competition of developers, their role is still significant in the area of managing these resources. An equally interesting conclusion is the concentration of co-operative undertakings in two voivodeships: Mazowieckie and Wielkopolskie, which may be related to the historical attachment to this form of management. In addition, the concentration of co-operative industries in each of these voivodeships indicates this. To sum up, we can state that despite numerous other forms of conducting business activity, there are still areas in Poland in which co-operatives find their place. However, there are also economic sectors in which the share of co-operatives is negligible, e.g. sector B or O (Tab. 3).

The important role and merits of co-operatives have been recognized by including (at the request of Germany) co-operatives in the UNESCO List of the Intangible Cultural Heritage of Humanity. By contributing to a fairer development of globalization processes, co-operatives contribute to reducing inequalities and discrimination with regard to economy, gender, race, religion or political orientation. The research on co-operatives should therefore be continued, especially for future law regulation and for government policy in local economics area and activities based on local social initiatives. The results of the similar research are also the subject of continuous interest of co-operative international organizations.

References

Aiken, M., Cornforth, Ch. & Spear, R. (2009). *The Governance challenges of Social Enterprises: Evidence from a UK Empirical Study*, Wiley Online Library.

Borzaga, C. & Spear, R. (2004). *Trends and challenges for co-operatives and social enterprises in developed and transition countries*, Publisher: Edizioni31, Trento – Italy.

Becker, W. & Dietz, J. (2004). R&D cooperation and innovation activities of firms—evidence for the German manufacturing industry, *Research Policy*, 33(2), 209–223.

Bijman, J. & Hendrikse, G. (2008). Co-operatives in chains: institutional restructuring in the Dutch fruit and vegetable industry, *Journal on Chain and Network Science*, 3(2), 95–107.

Bijman, J. & Bitzer, V. (2016), *Quality and innovation in food chains*, Wageningen Academic Publishers.

Brzozowska-Wabik, J. (2013). Spółdzielczość to nie przeżytek, *Indos 2013*, 21(135), Biuro Analiz Sejmowych, Warszawa.

Hendrikse, G. (2008). Co-operatives in chains: institutional restructuring in the Dutch fruit and vegetable industry, *Journal on Chain and Network Science*, 3(2), 95–107.

Foryś, I. (2011). Apartment Price Indices on the Example of Cooperative Apartments Sale Transactions, *Folia Oeconomica Stetinensia*, 1(9), 72–82.

Batóg, B. & Foryś, I. (2018). Analysis of Transaction Prices on one of Housing Estates in Szczecin: Do the Buyers Differentiate Prices on Local Markets with Respect to the Kind of the Right to Own? *Proceedings* Edited by: M. Papież and S. Śmiech, Publisher: Foundation of the Cracow University of Economics.

Foryś, I. (2017). Spółdzielczość mieszkaniowa w Polsce po transformacji gospodarczej, Studia i Prace WNEiZ US, 50(1) *Metody ilościowe w ekonomii*, 31–41, Wydawnictwo Uniwersytetu Szczecińskiego, Szczecin.

Hunt, G.C. (1992). Division of Labour, Life Cycle and Democracy in Worker Co-operatives, *Economic and Industrial Democracy*, 13(1), 9–43.

ICA COOP (2013). *Exploring The Co-operative Economy. Report 2013*, COOP International Cooperative Alliance, European Research Institute on Cooperative and Social Enterprises (EURICSE). www.monitor.coop.

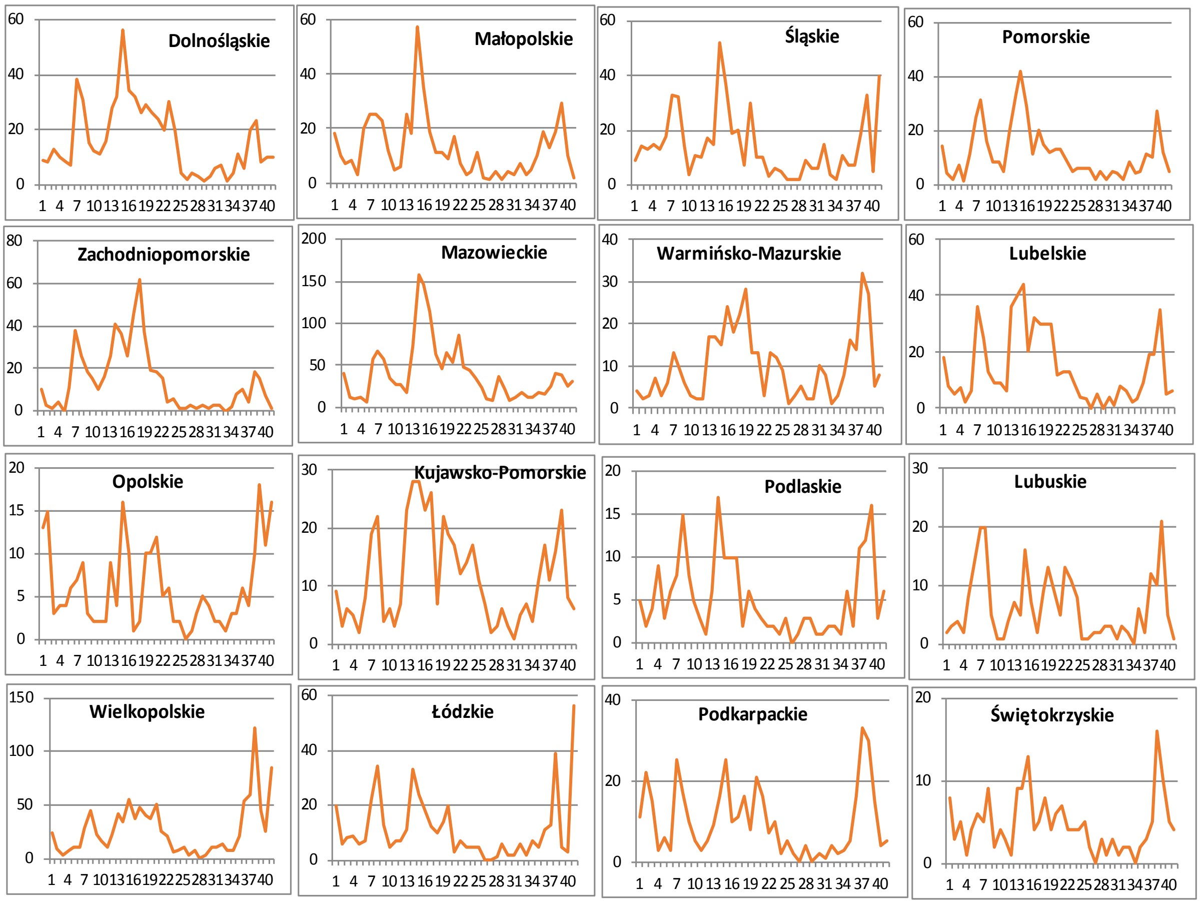
Kaiser, U. (2002). An empirical test of models explaining research expenditures and research cooperation: evidence for the German service sector, *International Journal of Industrial Organization*, 20(6), 747–774.

King, R.P., & Ortmann, G.F. (2007). Agricultural Cooperatives: History, Theory and Problems, *Agrekon*, 46(1). 18–46.

Walesiak, M. (1983). Propozycja rodziny miar odległości struktur udziałowych, *Wiadomości Statystyczne*, 10(29), 23–24.

APPENDIX

Fig. 1. Spatial dynamics of active co-operatives in Poland in the years 1975–2017*



* – for 4 months

Effect of solar energy policy on solar power efficiency in the European Union countries

Katarzyna Frodyma[1], Monika Papież[2], Sławomir Śmiech[3]

Abstract

The aim of this study is to assess the impact of solar energy policy on the efficiency of solar power in the EU countries. We apply two-stage Simar and Wilson, 2007 approach for cross-sectional data representing 24 EU countries. In the first stage, an input-oriented bias-corrected DEA model is applied to assess the efficiency of solar power potential regarding different aspects of solar power generation. In the baseline model, the global horizontal irradiance (GHI), which represents solar potential, is considered as the input variable. Solar-generated electricity is the output variable. Other specifications take into account environmental and economic aspects in addition to solar-generated electricity. In the second stage, the bias-corrected solar power efficiency scores are used in a truncated regression (Simar and Wilson, 2007) to explain the impact of solar energy policy on the relative solar power efficiency scores in the EU countries in 2016. Two main conclusions can be drawn. Firstly, additional benefits of solar power are limited to the group of countries with a relatively high share of solar power. Secondly, economic incentives positively affect the efficiency of solar power potential.

Keywords: *bias-corrected Data Envelopment Analysis (DEA) model, solar power generation*

JEL Classification: *C59, C61, Q2, Q01*

1. Introduction

Over the last two decades the EU member states have been implementing common energy policy regarding the development of renewable energy sources. The national renewable targets for each EU member state are specified by the European climate and energy package. Each EU country has created its own national renewable energy action plan to meet this target.

Two aspects of solar power development should be considered. Firstly, the EU countries are very diverse in terms of their solar potential, which has a significant impact on the supply of a given energy source. As a result, the high viability of investment is expected. Secondly, each country provides its own energy policy and supports the development of e.g. solar power by offering specific incentives, such as feed-in tariffs, green certificates, promotional loans, investment grants, and tax exemptions. However, a broader perspective reveals that the development of renewable energy sources leads to positive changes in the natural environment and decreases the costs of importing energy sources. The assessment of the viability of investment in solar power should take all these factors into account.

[1] Corresponding author: Cracow University of Economics, Department of Statistics, 27 Rakowicka St., 31-510 Cracow, Poland, frodymak@uek.krakow.pl.

[2] Cracow University of Economics, Department of Statistics, 27 Rakowicka St., 31-510 Cracow, Poland, papiezm@uek.krakow.pl.

[3] Cracow University of Economics, Department of Statistics, 27 Rakowicka St., 31-510 Cracow, Poland, smiechs@uek.krakow.pl.

A number of studies try to explain the dynamics of the development of renewable energy. They usually examine the role of different aspects of energy and climate policy. Best and Burke (2018) investigate the role of carbon pricing in the adoption of solar or wind power and Nicolini and Tavoni (2017), Li et al. (2017), Dijkgraaf et al. (2018), Best and Burke (2018) and García-Álvarez, et al. (2018) analyse the effect of the feed-in tariff (FIT) efficacy on the development of solar power. Romano et al. (2017) examine the influence of informal institutions on renewable strategy development, fiscal incentives, and public investments. All these studies focus on the capacity (generation) of renewable energy achieved by countries or the share of renewable electricity generation in total electricity generation. However, only several of them admit that energy policy can result in inefficiency, redundancy or overlapping of renewable energy (see Del Río and Mir-Artigues, 2014). The efficiency of renewable energy is a frequently studied issue. Most papers use Data Envelopment Analysis (DEA) as the empirical framework. There are only few studies which analyse solar power efficiency (Sueyoshi and Wang, 2017 and Frodyma et al., 2018). However, to the best of our knowledge, the effect solar energy policy experts on efficiency has not been studied so far.

The aim of this study is to assess the impact of solar energy policy on the efficiency of solar power in the EU countries. We apply two-stage Simar and Wilson, 2007 approach to cross-sectional data representing 24 EU countries.

In the first stage, an input-oriented bias-corrected DEA model is applied to assess the efficiency of solar potential regarding different aspects of solar power generation. In the baseline model, the global horizontal irradiance (GHI), which represents solar potential, is considered as the input variable. Solar-generated electricity is the output variable. Other specifications take into account environmental and economic benefits in addition to solar-generated electricity.

In the second stage, the bias-corrected solar power efficiency scores are used in a truncated regression (Simar and Wilson, 2007) to explain the effect of solar energy policy on the relative solar power efficiency scores in the EU countries. Three bundles of policies regarding economic instruments, policy support and regulatory instruments related to solar energy policy target are considered.

Our contribution to the literature is manifold. Firstly, the approach adopted in the study is based on a comprehensive view of the solar power efficiency on the country level. The study assesses not only the technical efficiency, which transforms solar power investment into electricity generation, but additionally considers economic and environmental benefits, which seem to play a crucial role and have motivated the EU legislation in this area. Secondly, the study assesses the effect of three areas of solar energy policy on the efficiency of solar power in the analysed countries. The assessment of their influence could prove beneficial for policymakers, as it points at possible constraints to be taken into consideration while deciding on renewable energy policy.

2. The two stage Simar and Wilson procedure

Data envelopment analysis (DEA), proposed by Charnes at al. (1978), is a mathematical programming approach to the construction of frontiers and the measurement of the efficiency of

DMU concerning the constructed frontiers. Technical efficiency measures the ability of a DMU to achieve the possible maximum level of output Y conditional on an input level X. There are two approaches in basic DEA: input-oriented and output-oriented. The first one maximises the proportional reduction of inputs X to maintain the outputs constant Y. The second one maximises the proportional increase of outputs Y for constant inputs X. A large number of papers explain the DEA estimates of efficiency with an additional set of exogenous factors. Regression models are used in this context.

Statistical properties of the efficiency scores and regression applied were extensively studied by Simar and Wilson (2002). Since the DEA efficiency scores are serially correlated, the standard inference procedure is invalid. They proposed a two-stage procedure. In the first stage, a consistent bootstrap procedure to correct the bias and to provide a statistical inference of the technical efficiency measures is applied. In the second stage, a truncated bootstrapped (Tobit) regression is used, which, according to Simar and Wilson (2007), provides the only feasible inference tool in the second stage.

3. Data

The empirical part of this study incorporates two above mentioned stages based on different datasets.

The first one, used in the bias-corrected DEA method, contains an input variable and a subset of three output variables. The selection of input and output variables is based on literature (Sueyoshi and Wang, 2017). Global horizontal irradiance (GHI), which is the total amount of shortwave radiation received from above by a surface horizontal to the ground, is considered as the input variable (Best and Burke, 2018). This value is of particular interest to photovoltaic installations and includes both direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI). The data for Global Horizontal Irradiance are obtained from the Global Solar Atlas[4].

Solar-generated electricity (TWh per capita) in 2016 (GEN) and two variables which describe the environmental and economic benefits resulting from the replacement of conventional energy with solar power are considered as the output variables. The economic aspect of solar power[5] (ECON) measures the economic benefit of avoided costs of generating electricity from fossil fuels after replacing conventional energy with solar power (per capita), while the environmental aspect of solar power[6] (ENV) measures the environmental benefit of avoided carbon dioxide (CO2) emissions by replacing conventional energy with solar power (per capita). All data describe 24 European Union countries (excluding Latvia and Estonia, which do not gener-

[4] https://globalsolaratlas.info/ accessed on 15.03.2019.

[5] $ECON = \frac{CFF}{\sum_{i=1}^{3} GEN_i} \cdot GEN$, where: $CFF = \sum_{i=1}^{3} GEN_i \cdot price_i$; where i indicates: coal, oil, natural gas, GEN_i – total primary energy supply generated from i energy sources in 2016, GEN – gross electricity generation from solar per capita in 2016.

[6] $ENV = \frac{CO_2}{TGEN} \cdot GEN$, where: CO_2 – values in fossil CO_2 is expressed in Mt CO_2/year, TGEN is total gross electricity generation, and GEN is gross electricity generation from solar power per capita in 2016.

ate any solar power in the analysed period, and Finland and Sweden for which there is no global horizontal irradiance data) in 2016 and are obtained from the European Commission webpage[7].

The second dataset used in the truncated bootstrapped regression is directed at assessing solar energy policies affecting the efficiency of solar power in 24 EU countries. The efficiency scores of the analysed countries obtained from the bias-corrected DEA method are used as the dependent variable. Three bundles of policies regarding economic instruments (EI), policy support (PS) and regulatory instruments (RI) which are linked with the solar energy policy targets are considered as independent variables. Following Marques and Fuinhas (2012), Aguirre and Ibikunle, 2014; Polzin et al., 2015, and Liu et al., 2019, the three bundles of policy instruments are measured by the number of active policies in a country per year regarding the solar energy policy target (in previous studies they were called "accumulated number of renewable energy policies and measures (ANPM)"), and they are obtained from the IEA's Global Renewable Energy Policies and Measures[8].

Economic Instruments (EI) include the following types of policy instruments: direct investment, fiscal and financial incentives support (feed-in tariffs, grants and subsidies, loans, tax relief, taxes, user charges) and market-based instruments (GHG emissions allowances, green and white certificates). These types of policy instruments are designed to reduce the investors' risk and provide a tool for trading and meeting renewable energy obligation among producers. They promote direct investment aimed at reducing the capital cost of investments in renewable energy.

Policy Support (PS) includes two types of policy instruments which define strategies and outline specific programs aimed at promoting renewable capacity inside a country, i.e. institutional creation and strategic planning.

Regulatory Instruments (RI) include the following types of policy instruments: auditing, codes and standards, monitoring, obligation schemes and other mandatory requirements. These policy instruments impose requirements on the minimum amount of electricity supplied mainly from renewable sources.

4. Results

To analyse the efficiency of the EU countries regarding their solar power investments, the study considers four models with global horizontal irradiance as the input variable, and selected combinations of the output variables. The models are presented in Table 1. The X indicates the variables which are used in particular models.

[7] Energy datasheets: EU-28 countries (https://ec.europa.eu/energy/en/data-analysis/country) accessed on 12.03.2019.

[8] International Energy Agency: https://www.iea.org/policiesandmeasures/renewableenergy/ accessed on 15.02.2019.

Table 1. Input-output variables of four models

	M1	*M2*	*M3*	*M4*
	GEN	*GEN_ECON*	*GEN_ENV*	*ALL*
GHI	X	X	X	X
GEN	X	X	X	X
ECON		X		X
ENV			X	X

The results of the bias-corrected DEA analysis, i.e. the bias-corrected efficiency scores (θ) of the EU countries in 2016 in all five models, are reported in Fig. 1 a-c. Fig. 1a-c present the efficiency scores obtained by model M1-GEN and by M2-GEN_ECON, M3-GEN_ENV, and M4-ALL models. The countries are ranked in a descending order according to the value of the bias-corrected efficiency scores obtained from the first (baseline) model with electricity generation (M1-GEN) as the only output variable. These figures allow for comparing the results of the first model (M1-GEN) with the remaining models which include additional aspects.

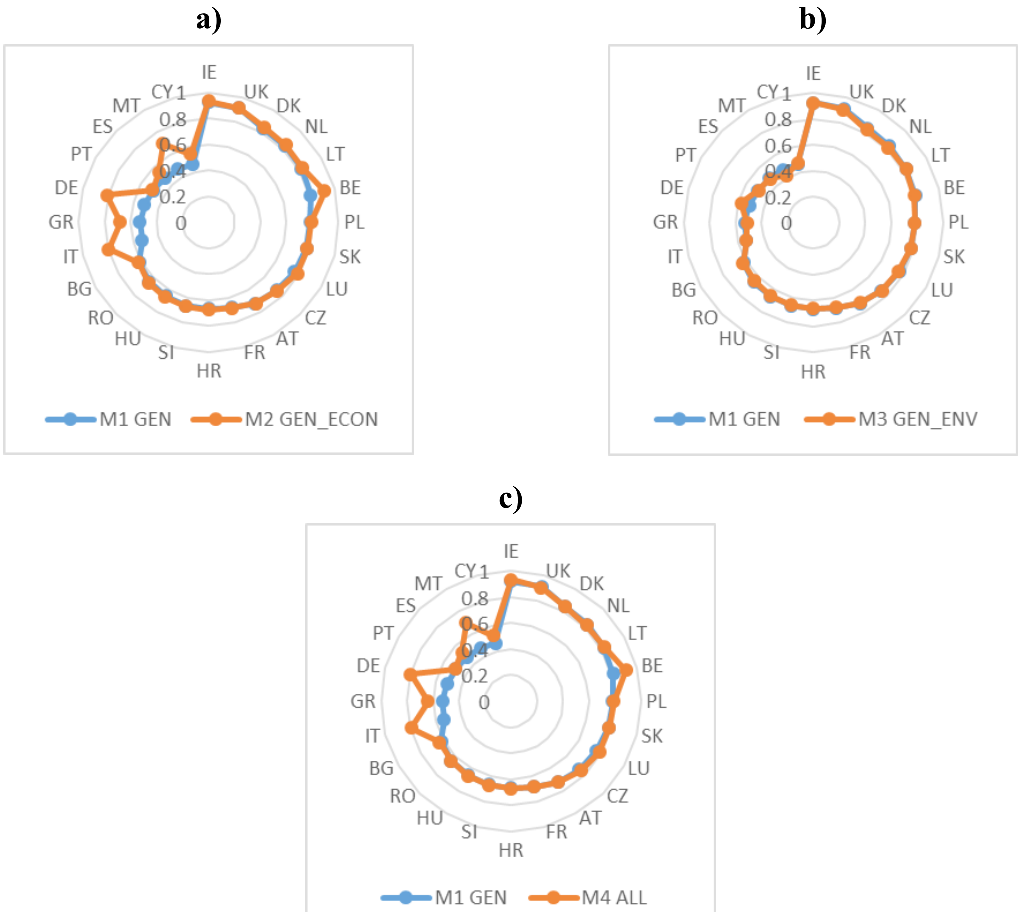


Fig. 1. The bias-corrected efficiency scores of the EU countries in 2016 obtained by M2-GEN_ECON, M3-GEN_ENV, M4-ALL models

Fig.1 demonstrates that in the baseline model, (model M1-GEN), Ireland and the United Kingdom are the most efficient countries in terms of their solar power potential. Their efficiency scores are the highest in all analysed countries and range between 0.902 and 0.915. Cyprus and Malta are the least effective countries.

Fig. 1a indicates that Ireland, Belgium, and the United Kingdom are the most efficient countries in terms of solar power potential when solar-generated electricity and the economic output (M2-GEN_ECON) are taken into account. Moreover, Germany, Italy, and Malta gain the most, and their relative efficiency scores increase by 56%, 50%, and 49%, respectively. An increase in the relative efficiency scores and a higher position in the rank are observed in countries using oil or gas as the main energy source (i.e. Ireland, in which in 2016 as much as 50% of total electricity generation comes from heating gas, in Italy – 44%, in Malta in which in 2016 as much as 84% of total electricity generation comes from heating oil power plants). Fig. 1b shows that Ireland, the United Kingdom, Denmark, Lithuania, Netherlands, and Belgium are the most efficient countries when solar-generated electricity and the environmental output are investigated (model M3-GEN_ENV). In addition, Germany and Cyprus gain the most in terms of environmental benefits. Fig. 1c points out that Denmark, Lithuania, and the Netherlands join the group of most efficient countries when all benefits are considered (model M5-ALL). In other words, Germany (where the efficiency score increases by 57%), Italy (493%) and Malta (48%) gain the most. In this model, Portugal turns out to be the least efficient country.

Table 2. The results of the truncated bootstrapped regression models

	M1	*M2*	*M3*	*M4*
	GEN	*GEN_ECON*	*GEN_ENV*	*ALL*
EI	0.0086	0.0205**	0.0102	0.0204**
SP	–0.0296	0.016	–0.0323	0.0148
RI	–0.0529**	–0.0455***	–0.0529***	–0.0478***
const	0.7201***	0.6962***	0.7117***	0.6944***

Note: Calculations obtained using STATA program; ***, **, * indicate statistical significance at 1, 5 and 10 per cent level of significance, respectively. Bootstrapped regression models are performed using the Stata "simarwilson" command.

The next stage of the study investigates the impact of renewable energy policy on the solar power efficiency in the analysed countries. Table 2 reports the results of the truncated bootstrapped regression models.

The results (Table 2) reveal that the effect of solar energy policy on solar power efficiency in the EU countries is significant, yet ambiguous. In two models (M2-GEN_ECON and M4-GEN_ALL) the bundles of policies regarding economic instruments (EI) (i.e., direct investment, fiscal and financial incentives supports, feed-in tariffs and market-based instruments, etc.) have a significant and positive effect on the solar power efficiency in the EU countries. It implies that the impact of fiscal and financial incentives, feed-in tariffs, grants and subsidies on solar power efficiency is notably strong. These types of policy instruments have large positive effects on solar energy promotion and play a substantial role in the assessment of the efficiency of solar potential when economic and environmental benefits are considered in addition to solar-generated electricity.

In contrast with economic instruments, the results obtained in the study demonstrate that the relative efficiency of solar potential in particular countries is negatively correlated with the number of active policies regarding regulatory instruments. So, the bundles of regulatory instrument policies (RI), which include auditing, codes and standards, monitoring, obligation schemes and other mandatory requirements, exert a negative and significant effect on the solar power efficiency in the EU countries in all models. It means that packages of policy instruments that impose requirements on a minimum amount of electricity, mainly from renewable sources, do not stimulate the EU countries to use their own solar potential.

However, the bundles of policy support (PS), which covers institutional creation and strategic planning, have a positive but insignificant impact on the efficiency of solar potential when economic and environmental benefits are considered in addition to solar-generated electricity. When the other two models are considered, this impact of policy support on solar power efficiency is negative, but also insignificant.

To sum up, a positive impact on the solar power efficiency of each country is exerted mostly by economic instruments (EI), mainly fiscal and financial incentives, such as feed-in tariffs (FIT). These feed-in tariffs spur deployment and technological diversity and lower risks associated with renewable energy technologies faced by private sectors.

Conclusions

Two general conclusions can be drawn from the first step of the study. Firstly, the countries which have favourable conditions for solar energy development are rather inefficient. It means, however, that they do not utilize their natural potential, thus they should provide much more capacity in solar energy. Secondly, the environmental and economic benefits of solar power are significant in countries with relatively large shares of solar power in total electricity production (Italy and Malta) or in countries witch use coal (environment aspect) or oil (economic aspect) as important energy sources. In the remaining group of countries additional outcomes are rather meaningless.

In the second step, the three bundles of policies regarding economic instruments, policy support and regulatory instruments (RI) linked with solar energy policy target are used to explain the impact of solar energy policy on the efficiency of solar power in the European Union countries. The role of these political factors is significant, albeit rather ambiguous.

Solar power efficiency is positively related to the package of policies regarding economic instruments. This result is in line with other studies, e.g. the ones conducted by Nicolini and Tavoni

(2017), Li et al. (2017), Dijkgraaf et al. (2018), Best and Burke (2018) and García-Álvarez, et al. (2018) who confirm a positive effect of the presence of feed-in tariff policies on the development of solar energy in the European countries. The bundle of regulatory instrument policies has a negative effect on the development of solar energy. Similarly, Li et al. (2017) find that the regulatory instruments have a negative yet insignificant impact on the renewable energy development. Policy support including institutional creation and strategic planning has a positive but insignificant impact on the solar power efficiency when – in addition to solar-generated electricity – economic and environmental benefits are taken into account. When solar-generated electricity is considered as the output variable, the impact of policy support on the efficiency of solar potential is negative but insignificant. This result is rather unexpected, as it might be logically assumed that a clear long-term energy strategy is conducive to investment, as investors appreciate a long-term framework with a clear vision. On the other hand, Polzin et al. (2015) report a negative impact of policy support and regulatory instruments on renewable energy sources development.

Acknowledgments

Publication was financed from the funds granted to the Faculty of Management at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

Aguirre, M., & Ibikunle, G. (2014). Determinants of renewable energy growth: A global sample analysis. *Energy Policy*, 69, 374–384.

Best, R., & Burke, P.J. (2018). Adoption of solar and wind energy: The roles of carbon pricing and aggregate policy support. *Energy Policy*, 118, 404–417.

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

Del Río, P., & Mir-Artigues, P. (2014). Combinations of support instruments for renewable electricity in Europe: A review. *Renewable and Sustainable Energy Reviews,* 40, 287–295.

Dijkgraaf, E., van Dorp, T.P., & Maasland, E. (2018). On the Effectiveness of Feed-in Tariffs in the Development of Solar Photovoltaics. *Energy Journal*, 39(1).

Energy datasheets: EU-28 countries. European Commission webpage: https://ec.europa.eu/energy/en/data-analysis/country (accessed on 10.01.2018).

Frodyma, K., Papież, M., Śmiech, S. (2018). Efficiency of investments in solar power in the EU countries. *The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena: Conference Proceedings*, *The Socio-Economic Modelling and Forecasting*, 1, 140–149. Foundation of the Cracow University of Economics, Cracow.

García-Álvarez, M.T., Cabeza-García, L., & Soares, I. (2018). Assessment of energy policies to promote photovoltaic generation in the European Union. *Energy*, 151, 864–874.

Global Wind Atlas: https://globalwindatlas.info/ (accessed on 31.01.2019).

International Energy Agency: https://www.iea.org/policiesandmeasures/renewableenergy/ accessed on 15.02.2019.

Li, S.J., Chang, T.H., & Chang, S.L. (2017). The policy effectiveness of economic instruments for the photovoltaic and wind power development in the European Union. *Renewable Energy*, 101, 660–666.

Liu, W., Zhang, X., & Feng, S. (2019). Does renewable energy policy work? Evidence from a panel data analysis. *Renewable Energy,* 135, 635–642.

Marques, A.C., & Fuinhas, J.A. (2012). Are public policies towards renewables successful? Evidence from European countries. *Renewable Energy,* 44, 109–118.

Nicolini, M., & Tavoni, M. (2017). Are renewable energy subsidies effective? Evidence from Europe. *Renewable and Sustainable Energy Reviews*, 74, 412–423.

Polzin, F., Migendt, M., Täube, F.A., & von Flotow, P. (2015). Public policy influence on renewable energy investments – A panel data study across OECD countries. *Energy Policy,* 80, 98–111.

Romano, A.A., Scandurra, G., Carfora, A., & Fodor, M. (2017). Renewable investments: The impact of green policies in developing and developed countries. *Renewable and Sustainable Energy Reviews,* 68, 738–747.

Simar, L., & Wilson, P.W. (2002). Non-parametric tests of returns to scale. *European Journal of Operational Research*, 139(1), 115–132.

Simar, L., & Wilson, P.W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics,* 136(1), 31–64.

Sueyoshi, T., & Wang, D. (2017). Measuring scale efficiency and returns to scale on large commercial rooftop photovoltaic systems in California. *Energy Economics*, 65, 389–398.

The factors shaping the capital structure of the stock-exchange-listed companies in the Visegrad Group countries

Magdalena Gostkowska-Drzewicka[1], Ewa Majerowska[2]

Abstract

The purpose of this paper is an analysis of the impact of selected internal factors on the capital structure of the companies listed on the stock exchanges in the Visegrad Group countries. Using a dataset of the stock-exchange-listed entities in 1998–2016, we argue that the shaping of the capital structure in these companies is explained by the pecking order theory. The results of panel data modelling highlight that, despite the fact that the examined entities operate in one geographical region, in developing economies with the EU, impact of internal factors affecting the capital structure of the stock-exchange-listed companies in the Visegrad Group countries is not the same.

Keywords: *capital structure, pecking order theory, trade-off theory, transition economies, panel data estimation*
JEL Classification: *G32, C23*

1. Introduction

The issue of shaping the capital structure has been the subject of numerous studies for nearly 60 years. Despite the fact that so many years have passed, it remains a valid subject and is considered one of the key research trends in the area of corporate finance. Research results on the selection of financing sources are ambiguous and often mutually exclusive. Many theories explaining the issues related to shaping the structure of capital are known presently,, however, science has not developed any unambiguous solutions that could be used in economic practice. Most studies have recognized that companies act differently depending on their size, growth opportunities, debt cost, liquidity, industry effect, etc., which leads them to display different financial behaviors (Koralun-Bereźnicka, 2018; Acedo-Ramirez et al., 2017). New factors affecting the decisions regarding the shape of the capital structure are still being identified. It leads to modification of the existing concepts and development of the new ones, based on separate assumptions. One of the reasons behind the diversity of the factors shaping the capital structure are the different economic conditions and the diversified degree of development of the capital markets in individual countries or groups of countries, as exemplified by the Visegrad Group. During the period of economic transformation, these countries represented a similar level of socio-economic development. However, despite all the similarities, later on, these countries did not develop at the same pace. These differences concerned, among others, the conditions and the level of financial market development. Therefore, the aim of the article is as follows:

1. Analysis of the impact of selected factors, i.e. the growth rate, liquidity, the non-debt tax shield, profitability, company size and the structure of assets, on the capital structure of the companies listed on the stock exchanges in the Visegrad Group countries.

---

[1] Corresponding author: University of Gdansk, Department of Corporate Finance, 101 Armii Krajowej, 81-824 Sopot, magdalena.gostkowska-drzewicka@ug.edu.pl.

[2] University of Gdansk, Department of Econometrics, 101 Armii Krajowej, 81-824 Sopot, ewa.majerowska@ug.edu.pl.

2. To obtain an answer to the question whether the capital structure of stock-exchange-listed companies in the Visegrad Group countries is shaped in accordance with the assumptions of the trade-off theory or the pecking order theory.

Implementation of the aim formulated as such required verification of the following hypotheses:

1. The growth rate is negatively related to the amount of debt.
2. The relation between financial liquidity and the amount of debt is negative.
3. The non-debt tax shield is negatively related to the amount of debt.
4. There is a negative relation between profitability and the amount of debt.
5. The size of an enterprise is positively related to the amount of debt.
6. The structure of assets and the amount of debt are positively related.

The above hypotheses were verified using appropriate statistical tests. Calculations were carried out using the Gretl package. The analysis covered 328 companies listed on the stock exchanges in the countries of the Visegrad Group, during the years 1998–2016.

The article consists of an introduction, three parts and an ending. The first part deals with the shaping of the capital structure, in the context of the trade-off theory and the pecking order theory. The second part is an overview of the studies based on these concepts. Further, the methodology and the results of the research are presented, summarized in the conclusion part.

2. The shaping of the capital structure in the light of the trade-off theory and the pecking order theory

The static trade-off theory of capital structure is a development of the work by Modigliani and Miller (1958, 1963). In the following years it was modified multiple times, inter alia by Jensen and Meckling (1976) and Myers (1977, 1984). The static trade-off theory assumes that the optimal structure of capital in an enterprise results from balancing the interest-rate tax benefits with the so-called costs of financial difficulties (costs of bankruptcy), along with the agency costs related to equity and foreign capital. An optimal capital structure ensures maximisation of the market value of an enterprise, and at the same time – a minimal weighted average cost of capital. In the trade-off theory, an optimal ratio of interest-bearing liabilities to equity, i.e. the financial leverage, is determined. In practice, however, a fixed, optimal structure of capital is not possible, because a company's financial situation and the market environment are subject to constant changes.

A different approach to shaping of the capital structure is presented in the pecking order theory. This concept was first introduced by Donaldson (1961) and later on confirmed by Myers and Majluf (1984). In the same year, this theory was further refined by Myers (1984). The pecking order theory refers to the preferences of enterprises regarding the use of particular sources of financing. The order is as follows: first, internal sources are used, i.e. retained earnings as well as surpluses of cash and short-term financial assets. After exhaustion of these, enterprises reach for external sources. They issue debt securities in the first place,, followed by hybrid securities, and only at the end – shares. In this theory, the target capital structure is not defined. The choice of individual sources of financing has decisive impact on the capital structure (Duliniec, 2015).

3. Overview of the research based on the trade-off theory and the pecking order theory

Mazur (2007) and Jaworski and Czerwonka (2017) confirmed that under the Polish conditions the pecking order theory serves as best explanation of the capital structure shaping in the entities examined. Hartwell and Malinowska (2018) argued that neither the trade-off nor the pecking order theories fully explain corporate capital structure in Poland. The authors indicate that the strength of property rights and stock market capitalization are driving forces behind corporate financing decisions. Majerowska and Gostkowska-Drzewicka (2017) found that the corporate performance of the companies is in line with pecking order theory.

Czech enterprises operating in the agricultural, industrial and construction sectors shaped the capital structure in accordance with the presumptions of both the pecking order theory and the trade-off theory. The size of these companies' debt was affected by the sectoral affiliation and the type of the ownership structure (Poulová, 2017). Bauer (2004) arrived at more definite conclusions, analysing Czech stock-exchange-listed companies from the non-financial sector. According to the author's research, the size of debt in the examined entities was negatively related to profitability, which is in line with the pecking order theory.

The capital structure of Slovakian enterprises from the non-financial sector is best explained by the trade-off theory (Režňáková et al., 2010). Research has shown a positive relation between the amount of debt and the structure of assets, the profitability and the company size. In turn, debt and growth opportunities, liquidity, and non-debt tax shield were related negatively.

In the 1990s, Hungarian public companies shaped the capital structure in accordance with the pecking order theory (Nivorozhkin, 2002; Dević and Krstić, 2001). Similar conclusions follow from a later research (De Jong et al., 2008). They showed a negative relation between the amount of debt and profitability, liquidity and the income tax rate. In turn, debt and company size were positively related.

4. Methodology and research results

On December 15th, 2017 there were 542 companies listed on the Stock Exchanges of the Visegrad Group countries. The sample excluded entities that did not submit complete financial statements during the period considered. Companies in bankruptcy or under restructuring were also rejected. In addition, only the entities that were continuously listed for a period of at least 5 years were admitted to the study. Ultimately, 328 enterprises were qualified for the research, i.e. almost 61% of the pre-selected entities (Table 1).

Based on the above-mentioned literature, to verify the hypotheses listed above, estimation of a linear model of panel data was proposed, in the following form:

$$Y_{it} = f(X_{1it}, \ldots, X_{6it}, Z_{1it}, \ldots, Z_{4it}, \xi_{it}) \quad (1)$$

where the endogenous variable Y is the amount of debt (D) represented by the ratio of total debt to total assets. The exogenous variables of the model are: X_1 – growth rate (GR), percentage change in sales revenues, in relation to the previous year; X_2 – financial liquidity (LIQ), the ratio of current assets to current liabilities; X_3 – non-debt tax shield (NDTS), the depreciation to total assets ratio; X_4 – profitability (ROA), the ratio of total assets to total equity; X_5 – size (SIZE), a natural logarithm from total assets; X_6 – asset structure (TANG), the ratio of tangible assets to the to total assets. The variables $Z_1, \ldots, Z_4$ are the dummy variables that take the value of 1 when the company comes from the country selected, zero in other cases. The subscript i denotes the number of the company under consideration, t the number of the period, while is the random component. The proposed model (1) was estimated using a pooled OLS, taking into account all the selected factors shaping the capital structure. The results of the estimation are included in tab. 2.

Table 1. Numerical amounts of the companies listed on the Stock Exchanges of the Visegrad Group countries during the years 1998–2016 (as of December 15th, 2017)

Specification	Total number of companies	Number of the companies qualified for the analysis	Number of the companies that were rejected
stock-exchange-listed companies (Poland)	477	280	197
stock-exchange-listed companies (Hungary)	40	30	10
stock-exchange-listed companies (Czech Republic)	21	15	6
stock-exchange-listed companies (Slovakia)	4	3	1
Total	542	328	214

Statistical tests indicated existence of some statistically non-significant variables at the level of 0.05. Also the diagnostic tests suggested introducing a fixed effect model. Ultimately, by including all companies from all the countries under examination, it can be concluded that profitability has significant negative impact on the level of indebtedness. It means that the capital structure shaping in the companies listed on the Stock Exchanges in the Visegrad Group countries is best illustrated by the pecking order theory. Firstly, according to this concept, entities achieving high profitability usually show lower level of indebtedness, because they have large possibilities of financing their activity using internal sources. Secondly, the pecking order theory does not clearly indicate the direction of the relation between the size of an enterprise and its debt. This dependence can be both positive and negative. In the case of the examined entities, it is positive, which is consistent with the assumptions of this concept. Thirdly, the relation between the debt ratio and the non-debt tax shield is positive, which is consistent both with the trade-off theory and the pecking order theory, but it is not statistically significant. Furthermore, dummy variables which identify the company's affiliation to a given country were introduced to the model. The results (presented in the table 1) point to statistically significant differences in the level of indebtedness of the companies from the Czech Republic and Hungary, in relation to Polish companies (lower level of indebtedness). Taking into account only the model's statistically significant variables, it can be concluded that, in general, the level of indebtedness of Hungarian companies was higher than that of the companies from other countries. The negative relation between two of the selected factors of capital structure and the level of indebtedness was also confirmed, which is consistent with the pecking order theory. In turn, the positive relationship between the level of indebtedness and the structure of assets is consistent with the trade-off theory.

Table 2. Estimates of the panel regression for the entire sample

	pooled	**pooled**	**pooled**	**fixed effects**	**fixed effects**
const	–2.2592***		–2.2862***	–10.1414***	–11.1380***
GR	0.0000	0.0000		0.0000	
LIQ	0.0004	–0.0004		0.0010	
NDTS	0.0549	0.0549		–4.3949	
ROA	–6.9919***	–6.9719***	–6.9918***	–7.0555***	–7.0607***
SIZE	0.2254***	0.2254***	–0.2335***	0.8546***	0.9173***
TANG	0.2650	0.2650		–0.3466	
CZECH	0.0231	–2.2362***			
SLOVAKIA	0.0681	–2.1911**			
HUNGARY	–1.7838***	–4.0431***	–1.7888***		
POLAND		–2.2592***			
Joint significance test	1.8736#	na	1.8662#	1.9869#	1.9820#
Breusch-Pagan test	8.3574#	na	8.6841#	12.6463#	12.8939#
Hausman test	39.4238#	na	29.6675#	47.0778#	40.2492#

*)**)***) statistically significant at the significance level of 0.1; 0,05 and 0.01
#) at 0.05 significance level, the null hypothesis should be rejected

In the next step, the proposed model for the companies coming from each country was estimated separately. The rationale for this was the significant differences in the number of the companies in each country under analysis. Estimates of the best models are included in tab. 3.

Table 3. Panel regression estimates for each country separately

	CZECH REPUBLIC	**SLOVAKIA**	**HUNGARY**	**POLAND**
	fixed effects	**pooled**	**fixed effects**	**fixed effects**
const	–1.7331***	3.4403***	0.6077***	0.1369***
GR		0.0721*	–0.0000***	
LIQ		0.2593***	–0.004**	
NDTS			–2.9339***	–8.1535***
ROA	–0.2129***	–1.8262***	–1.9654***	–7.6549***
SIZE	0.1395***	–0.2297***		
TANG				
Joint significance test	49.1380#	0.0498	2.7012#	1.8663#
Breusch-Pagan test	547.180#	1.5726	26.5406#	79.6269#
Hausman test	24.0261#	na	19.2485#	38.7805#

*)**)***) statistically significant at the significance level of 0.1; 0.05 and 0.01
#) at 0.05 significance level, the null hypothesis should be rejected

The above estimates, only in the case of the data concerning the companies from Slovakia, indicated correctness of the pooled OLS estimation method. With regard to other countries, the use of a fixed effects estimator turned out to be most proper. In the case of the companies from the Czech Republic, the level of indebtedness is negatively affected by the level of ROA, while the company size is affected positively. It means that these companies shape their capital structure in accordance with the pecking order theory.

The debt of Slovakian companies is positively influenced by their liquidity and growth rate, and negatively by the ROA and the company size. It should be noted that the positive direction of the relation between liquidity and indebtedness is consistent with the trade-off theory. Other dependencies, however, are characteristic of the pecking order theory. It can be therefore assumed that this concept best describes the shaping of the capital structure in Slovakian companies.

Companies from Hungary show a decline in debt due to the increase in such factors as the growth rate, liquidity, the non-debt tax shield and the ROA. In the case of these entities, the negative direction of the relation between the growth rate and indebtedness suggests that they shape their capital structure in accordance with the assumptions of the trade-off theory. However, the other dependencies are specific of the pecking order theory. It means that this concept quite better describes the shaping of the capital structure in Hungarian companies.

With regard to Polish companies an increase in the value of such factors as the non-debt tax shield and the ROA cause a decline in corporate debt. The direction of these dependencies clearly confirms that for these entities the pecking order theory is applicable. This conclusion is consistent with the study on the capital structure of Polish enterprises carried out by other authors.

Conclusions

The empirical analysis of the capital structure factors in the companies listed on the Stock Exchanges of the Visegrad Group countries indicates that the shaping of the capital structure in these entities is explained by the pecking order theory. Thus, with regard to the Czech companies, it was proved that the level of indebtedness is negatively influenced by the level of ROA, while the company size is positively affected by it. In Polish companies, the level of indebtedness was negatively related to the non-debt tax shield and profitability. It can be therefore assumed that in both countries, the stock-exchange-listed companies shape their capital structure in accordance with the pecking order theory.

In the case of Hungarian and Slovakian companies, inconclusive results were obtained. In Slovakian companies, the level of indebtedness is positively affected by the growth rate, while profitability and the company size affect it negatively, which is consistent with the pecking order theory. Conversely, the positive relation between financial liquidity and indebtedness is characteristic of the trade-off theory. In contrast, in Hungarian companies, the negative relations between indebtedness and such factors as financial liquidity and profitability are consistent with the pecking order theory. The negative relation between the non-debt tax shield and debt is consistent with both concepts. In turn, the negative relation between the growth rate and indebtedness points to substitution theories.

It can be concluded that out of the six hypotheses set out in the introduction, only in the case of Hungarian companies there are no reasons for rejecting the first and the second hypotheses, since the growth rate and the financial liquidity are negatively related to the amount of debt. There are no reasons to reject the third and the fourth hypotheses, i.e. the tax shield and profitability are negatively related to the level of indebtedness, in terms of Hungarian and Polish companies. The fifth hypothesis, assuming a positive relation between the size of the enterprise and its debt, was not rejected in the case of the companies from the Czech Republic. The last hypothesis, assuming a positive relation between the structure of assets and indebtedness, with regard to the companies from all countries, analysed separately, was rejected.

Analysing the models estimated for all the companies and taking into account the "country effect", there are no reasons for rejecting the fourth and the fifth hypotheses, while the first, the second, the third and the sixth one should be rejected.

Ultimately, it follows that, despite the fact that the examined entities operate in one geographical region, in developing economies with the EU, impact of individual factors affecting the capital structure of the stock-exchange-listed companies in the Visegrad Group countries is not the same.

References

Acedo-Ramirez, M.A., Ayala-Calvo, J.C., & Navarrete-Martinez, E. (2017). Determinants of Capital Structure: Family Businesses versus Non-Family Firms. *Finance a Uver: Czech Journal of Economics & Finance*, 67(2), 81–103.

Bauer, P. (2004). Determinants of capital structure: empirical evidence from the Czech Republic. *Czech Journal of Economics and Finance (Finance a uver)*, *54*(1–2), 2–21.

De Jong, A., Kabir, R., & Nguyen, T.T. (2008). Capital structure around the world: The roles of firm-and country-specific determinants. *Journal of Banking & Finance*, 32(9), 1954–1969.

Devic, A., & Krstic, M. (2001). Comparable analysis of the capital structure determinants in Polish and Hungarian enterprises–empirical study. *Economics and Organization*, 1(9), 85–100.

Donaldson, G. (1961). *Corporate debt capacity: A study of corporate debt policy and the determination of corporate debt capacity*. Division of Research, Harvard Graduate School of Business Administration, Boston.

Duliniec, A. (2015). Wybór źródeł finansowania a optymalna struktura kapitału w przedsiębiorstwie. *Finanse, Rynki Finansowe, Ubezpieczenia*, 74(855), 73–82.

Hartwell, C.A., & Malinowska, A. (2018). Firm-level and institutional determinants of corporate capital structure in Poland: New evidence from the Warsaw stock exchange. *Finance a Uver-Czech Journal of Economics and Finance*, 68(2), 120–143.

Jaworski, J., & Czerwonka, L. (2017). Determinanty struktury kapitału przedsiębiorstw notowanych na GPW w Warszawie. Sektor usług. *Annales Universitatis Mariae Curie-Skłodowska, sectio H–Oeconomia*, *51*(4), 133–142.

Jensen, M.C., & Meckling W.H. (1976). The Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics,* 3(4), 305–360.

Koralun-Bereźnicka, J. (2018). Determinants of capital structure across European countries. In: *Contemporary Trends and Challenges in Finance: Proceedings from the 3rd Wroclaw International Conference in Finance*, 199–209. Springer International Publishing.

Majerowska, E., & Gostkowska-Drzewicka, M. (2017). Determinants of corporate performance: modelling approach. *Dynamic Econometric Models*, 17(1), 115–127.

Mazur, K. (2007). The determinants of capital structure choice: evidence from Polish companies. *International Advances in Economic Research*, 13(4), 495–514.

Modigliani, F., & Miller, M.H. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review,* 48(3), 261–297.

Modigliani, F., & Miller, M.H. (1963). Corporate Income Taxes and the Cost of Capital: A Correction. *American Economic Review,* 53(3), 433–443.

Myers, S.C. (1977). Determinants of corporate borrowing. *Journal of financial economics*, *5*(2), 147–175.

Myers, S.C. (1984). The capital structure puzzle. *The journal of finance*, 39(3), 574–592.

Myers, S.C., & Majluf, N.S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of financial economics*, 13(2), 187–221.

Nivorozhkin, E. (2002). Capital structures in emerging stock markets: the case of Hungary. *The Developing Economies*, 40(2), 166–187.

Poulová, L. (2017). Capital Structure of Czech Joint Stock Companies. *Český finanční a účetní časopis*, 2017(4), 25–39.

Režňáková, M., Svoboda, P., & Polednáková, A. (2010). Determinants of capital structure: Empirical evidence from Slovakia. *Ekonomický časopis (Journal of Economics)*, 3(58), 237–250.

Verification of usability of the price cobweb model in current market conditions

Justyna Góral[1], Włodzimierz Rembisz[2]

Abstract

Specific relations between the price and supply in the market of vegetable products are determined by many causes of environmental and biological nature as well as by the market situation. It results in a specified alternation of these values and affects income implemented by producers in this market. The basis of this analysis is the law of supply (the price rise results in the rise in the supply volume while the price fall – in the fall of the supply) and its flexibility in relation to the price. The objective of these studies is to verify the assumptions of the cobweb model (used originally in closed economies) as a relevant theorem describing the mechanism of the agricultural market. A hypothesis has been adopted that the usability of this theorem is still high. The word "still" refers to repealing an assumption regarding the closed economy. Currently, we have the wheat market which is integrated and subjected to globalisation processes. It changes assumptions regarding the conditions of the price cobweb theorem.

Keywords: *market equilibrium, cobweb theorem, wheat price, supply*

JEL Classification: *C02, C62, C65*

1. Introduction

One of the most relevant issues is the mechanism of price shaping and its impact on choices made by these entities. The point here is particularly the impact of the price mechanisms on the producers' decisions with given choices made by consumers. It is important in the case where producers are price takers. Such conditions are most often characteristic of agricultural producers. Moreover, also an assumption on the relative homogeneity of the product (homogeneous basic products) refers to them. It, in turn, implies similar effect for the market of choices made by consumers and producers, and consequently – almost typical establishment of relations between the demand and supply curves. The specificity of activity of agricultural producers is also the fact that their production is of seasonal and reproducible nature. It is accompanied by a typical postponement of the production effect i.e. the time which lapsed between making a decision and obtaining an effect. This phenomenon takes place for reasons independent from the producer himself, but is of strictly biological nature and results from the specific nature of the course of the agricultural production process in the given geographical latitude. In this period of postponement, economic conditions may change. In the conditions of the market price determination it means that there can be a difference between market prices of products at the time of making a decision and their products at the time of completing the production process (moment of obtaining the production effect, final supply volume). It is determined by many factors related to the environmental-biological-climatic

[1] Corresponding author: Institute of Agricultural and Food Economics – National Research Institute, Mathematics Application in Agricultural Economics Department, 20 Świętokrzyska St., 00-002 Warsaw, Poland, justyna.goral@ierigz.waw.pl.

[2] University of Economics and Human Sciences in Warsaw, Management and Finance Department, 55 Pawia St., 01-030 Warsaw, Poland, wrembisz@gmail.com.

determination of nature of production processes in agriculture. It is also related to the above-mentioned economic conditions and seasonal cycles. Moreover, we should also mention the impact of political and regulatory factors which, to a greater or smaller extent distort the natural market mechanism. Finally, we should also include individual psychological predispositions of producers (behavioral factors[3] – part of economic psychology or behavioral economics[4]), including so-called herd behaviour (group thinking syndrome).

In the economics of agriculture is a long tradition to study and specify principles regarding price variability in relation to the aforementioned conditions or to isolate the impact of individual price change determinants. Seasonal and economic principles of price changes enjoy the greatest popularity. In fact, it is more important to get an answer to a basic question – how do agricultural producers respond to these price changes or to forecasts on potential price changes and their levels? One of such possibilities is the "price cobweb theorem" (model). Originally, it was formulated in the conditions of domination of national economies as relatively closed systems with the additional share of external turnover. Currently, market entities in Poland function in the system of the integrated EU market which is characterised by being opened to the global market. Thus, the objective of the studies was to verify the assumptions of the cobweb model in the current market conditions. A hypothesis has been adopted that the usability of the cobweb theorem is still high and its assumptions are useful in interpreting relations between prices and the supply of offered commodities (supply of wheat in the Polish market).

2. Market conditions and the cobweb theorem

In the economics, a lot of attention is devoted to price fluctuations in markets. Their level is a result of the game of demand and supply (markets' striving for achieving the level of equilibrium). These adaptation processes in markets are illustrated by the cobweb model. The price cobweb theorem relates to an analysis of a mutual relation between the price and supply over time. First and foremost, the time refers to the response of supply. Here, we can identify three situations:

1) In the very short period (several months) – the price is the function of demand;
2) In the short period (e.g. 1 year of the given production and commercial cycle) – there are adaptation processes of supply under the existing involvement of production factors (manufacturing potential, the supply may increase thanks to its better use);
3) In the long period (several years) – the supply may increase in relation to the increased involvement of factors (increased production capacity) as a result of investments.

[3] Classical economics (A. Smith, D. Ricardo, J.S. Mill) referred to psychologism while Neo-Clas-sical economics consciously departs from it for the benefit of the logic of choice. However, human behaviour and decision-making processes should not be analysed only on the basis of the theory of economics, particularly Neo-Classical. People are often guided by emotions or are not careful enough in striving for obtaining full information at the moment of making decisions.

[4] The *homo oeconomicus* model was subject to modifications extending between two research approaches: psychologism and antipsychologism (economic psychology, behavioural economics).

The cobweb theorem refers mainly to the first two conditions. This model has been formulated at the same time by: Hanau (1928), Ricci (1930), Schultz (1930) and Tinbergen (1930). The term "cobweb" was used by Kaldor (1934) for the first time, while in the theory of economics this model was more broadly explained by Ezekiel (1938). Ezekiel described a mechanism adjusting the supply and demand using present or delayed prices in the market. The proposed approach as a way of explaining the reality enjoyed great interest in the economics of agriculture and in the economics (AL-Daami and AL-Hiyali, 2017). In Poland, the cobweb model in agriculture was applied for the first time by Schmidt and Mandecki in 1933.

The cobweb model allows to interpret price effects related to shifting the production effect in time, in fact – the impact of the supply level on the price in relation to the price, at which a production decision was made (Finkenstadt, 1995). It is important in the case of such agricultural markets as the cereal market (homogeneous and standardised) with the production and supply effect which is clearly postponed in time. After the accession of Poland to the European Union, the Polish cereal market as well as other agricultural markets is a part of the integrated European market. In the past, these were the conditions of the domestic market with the given supply volumes and on the other hand – with the given demand volume and also with the specific information effectiveness of the market. In the cobweb model, it was implicitly assumed that these decisions have herd conditions (i.e. all response in the same way to a given price signal).

The supply volume is a flexible variable. It is determined not only by the national production but also by the easily available import. As far as the demand is concerned, the situation is similar – the export as part of the EU market significantly increases the scale and flexibility of this variable. This flexibility is also determined by the fact that there are currently much better possibilities of storing agricultural products. Currently applied technologies are also more advanced and transport of products is so efficient and effective that the flexibility of supply and demand significantly increases. Moreover, we should also take into account the hypothesis of rational expectations and the fact that in fact agricultural producers may correct errors made in the past. They have quicker access to market information and latest market forecasts (also those on prices). The aid of agricultural advice is much more developed. Moreover, widely used are supply forward contracts or the increasingly developing form of production organisation for contracts concluded. The above-mentioned reasons and beneficial changes may reduce the usability of the cobweb model in agricultural markets. We should also add the formal aspect of the original approach to this theorem which applied to linear forms of the demand and supply functions. The linearity of these two variables is currently replaced by non-linear relations (Brock and Hommes, 1997). Currently, in addition to the more static approach, we also have a dynamic approach.

3. Essence of the price cobweb model

It is assumed that the cobweb model illustrates the mechanism of achieving market equilibrium. In a traditional form, it was one of the simplest dynamic economic models describing changes in prices and supply of products for the given demand in the individual market. The name or the term "cobweb" is derived from the shape of the lines. The objective of the model is to determine such a price path for a given commodity so that for each period the demand is balanced by the supply. Various oscillations describing price changes were identified. We deal with periodic oscillations when the demand flexibility is equal to the supply flexibility. On the other hand, we talk about explosive oscillations when the demand flexibility is lower than the supply flexibility. In the last two cases, the market is unstable. The dynamic cobweb model presents price changes and changes in the demand and supply volume in a period longer than one production and commercial cycle. These changes are characterised by the cyclic movement, regular fluctuations with alternating periods of rise and fall. The cobweb model points to the dynamics of the market mechanism which at the time of inequilibrium strives for eliminating both the demand and supply surplus (Fig. 1) and for returning to the state of relative equilibrium. Then, there are e.g. damped fluctuations of prices and quantity of products in the market which means that the market returns to equilibrium when the demand flexibility is higher than the supply flexibility (Lichtenberg and Ujihara, 1989; Onozaki, Sieg and Yokoo, 2000; Włodarczyk, 2006; Lenart and Mazur, 2017; Mazur, 2018).

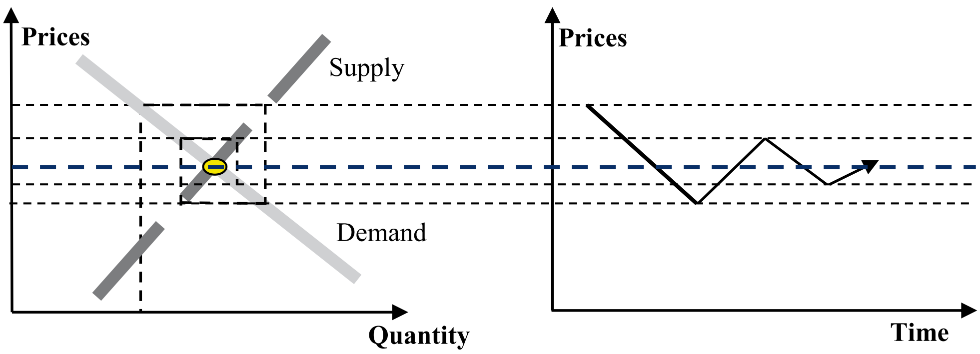


Fig. 1. Dynamic cobweb model (damped oscillations, stable market)

4. Analytical approach to the price cobweb model

Originally, in the cobweb model the functions of demand and supply in relation to the price were linear functions and had the following analytical form:

$$D_t = a - b * P_t, \tag{1}$$

$$S_t = c + d * P_{t-1}, \tag{2}$$

where: D_t – demand volume, S_t – supply volume, P_t – price, a, b, c, d – parameters, t – time.

The above relations illustrate one of the basic assumptions of the cobweb model. It is assumed that the demand is shaped by current prices (in the t period) and the supply – by previous prices (delayed by one period ($t - 1$)). This assumption was adopted based on observations of decisions of customers (demand side) who purchase products, raw materials or services on a basis of current prices in the market. In turn, producers (supply side) make decisions on the production volume in the previous period and thus based on prices delayed by one period. The prerequisite to achieve equilibrium is presented by the following relation:

$$D_t = S_t \tag{3}$$

After making a conversion including the previous relations, we have:

$$a - b * P_t = c + d * P_{t-1}. \tag{4}$$

It is assumed that the producer making a decision assumes that the price in the period of implementing the production (supply) is the same as in the period of making a decision ($P_t = P_{t-1}$). This price is called the $\overline{P}$ equilibrium price (this can be defined as a projected price) and is used to solve the equation:

$$a - b * \overline{P} = c + d * \overline{P} \Rightarrow \overline{P} = \frac{a-c}{b+d}. \tag{5}$$

In practice, such an equilibrium price may not occur. The type of fluctuations is dependent on the value adopted by the relation:

$$\frac{-d}{b} \quad \text{(where: } b \text{ and } d \text{ are parameters).} \tag{6}$$

The fluctuations can be: (1) damped – if $d < b$; (2) explosive – if: $d > b$; (3) regular – if: $d = b$. Damped fluctuations take place where the market, with every subsequent period, approaches the state of achieving equilibrium. Regular fluctuations occur when the market oscillates around the equilibrium price (where once the price is higher than the equilibrium price by the value and then the price is lower than the equilibrium price by the value). Explosive fluctuations take place where the market, with every subsequent period, goes away from the state of equilibrium.

In the classical cobweb model, producers make decisions on the volume of their future production based on current prices. However, producers may, as it is assumed in the modification of the approach, estimate their production based on the price (hereinafter designated as P_t^e) projected in the future period (Finkenstädt 1995). Then the model assumes a new form:

$$S_t = c + d * P_t^e, \tag{7}$$

where: P_t^e is illustrated by the formula:

$$P_t^e = P_{t-1}^e + w(P_{t-1} - P_{t-1}^e). \tag{8}$$

Here, the parameter is a coefficient of expectations included within the range of $(0; 1\rangle$ (Jakimowicz 2010). Where $w - 1$ then the projected price in the t period is equal to the price in the $t - 1$ period. Looking at the above formula, we can convert it as follows:

$$P_t^e = (1 - w)P_{t-1}^e + wP_{t-1}. \tag{9}$$

In turn, when: $w = 0$ then $P_t^e = P_{t-1}^e$.

Linear functions of the demand and supply do not always reflect the reality properly. For example, in a situation where producers have to incur some fixed costs. Therefore, non-linear functions have been proposed, both for the demand and for the supply:

$$D_t = \frac{a}{P_t}. \tag{10}$$

According to this formula, the demand is a reciprocal of the price, corrected by the a parameter. In turn, the non-linear function of the supply is as follows:

$$S_t = \text{arctg}(\lambda * P_t^e), \tag{11}$$

where: λ is a non-negative parameter regulating the supply function (Hommes 1991).

5. Verification of the model using empirical data

Empirical data came from the IAFE-NRI database prepared based on data of the Statistics Poland. It has been assumed that the demand volume is determined by the total of domestic consumption and export. The supply volume is a total of production, import and initial stocks. Thus, the assumption regarding the closed economy has been repealed in empirical studies. This analysis refers to partial equilibrium (one commodity market) and is also a static analysis (Table 1). Data have been collected in form of time series (2004–2016). The functions of trends for the demand, supply and wheat price have been identified. As it results from Fig. 2, the better adjustment was obtained using the exponential function (demand) and power function (price). Only for the supply, the linear function best explained relations and illustrated the trend over time.

Table 1. Correlation between the wheat supply volume in the "*t*" and its price in the "*t* − 1" period

Specification	Years	Supply in the "*t*" period	Price in the "*t* − 1" period
Years	1.00	0.85	0.73
Supply in the "*t*" period	0.85	1.00	**0.61**
Price in the "*t* − 1" period	0.73	**0.61**	1.00

The formula No. 1 was verified based on the empirical data.. However, it was not confirmed in the case of the data analysed here (Fig. 3). The price rise did not reduce the demand and even stimulated it to some extent. One of potential reasons for this phenomenon could be the growing export of wheat in the analysed period. It should be noted that the positive effect of the accession of Poland to the EU (May 1st, 2004) was a gradual improvement in the trade balance in the wheat market. This process was best visible at the turn of 2014 and 2015, when the domestic consumption was 8,443 thousand tones while the export was 4,921 thousand tones cereals.

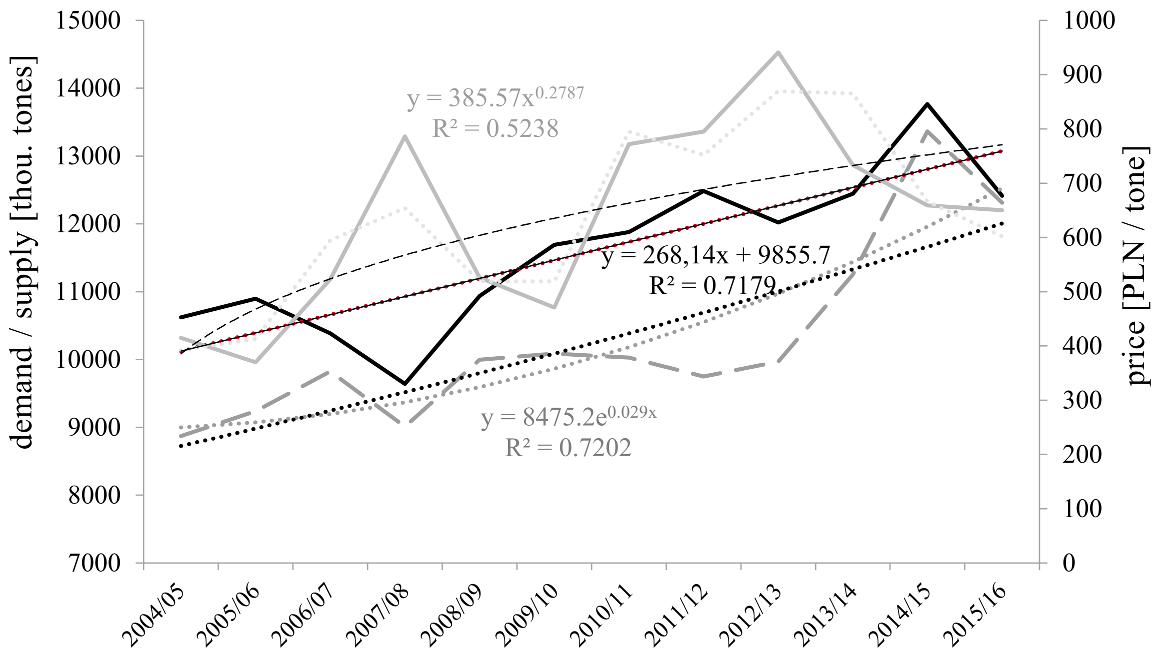
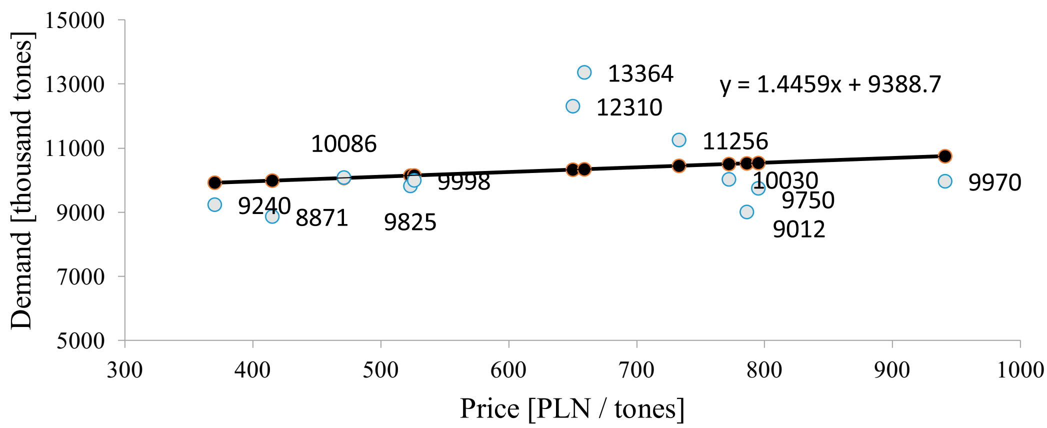


Fig. 2. Functions of trends for the demand, supply and wheat price in Poland (2004–2016)

The supply was analysed using the above-mentioned formula No. 2 (referring to the classical cobweb based on the linear function). The results of estimation are shown in Table 2. The λ and w values in Table 2 optimise the result, i.e. they maximise R^2 and also minimise the standard error. If we are to use methods using the expected price P_t^e, we should determine its initial value, as in the formula No 6 (and 7) referred to P_{t-1}^e. In this paper, the following was adopted:

$$P_{t_0}^e = \frac{S_{t_0-1} * P_{t_0-1}}{S_{t_0}}.. \tag{12}$$



* grey colour marks real data, black colour marks data estimated based on the formula No. 1

Fig. 3. Volume of demand for wheat in relation to the wheat price in the years 2004–2016

It was also assumed that at the beginning of the analysed period there was market equilibrium i.e. $S_{t_0-1} * P_{t_0-1} = S_{t_0} * P_{t_0}$ which after converting leads to the formula No. 12.

Table 2. Results of estimating the wheat supply function in Poland in the years 2006–2016

Specification	Approach_1	Approach_2	Approach_3	Approach_4
Coefficient of determination ($\boldsymbol{R}^2$)	32.79%	99.44%	74.87%	99.79%
Standard error	995.91	924.91	608.99	561.54
Parameter ($\boldsymbol{\lambda}$)	–	0.0064	–	0.0028
Coefficient of expectations ($\boldsymbol{w}$)	–	–	0.1994	0.1602

The approach No. 4, where during statistical verification the best parameters were obtained (satisfactory level of the coefficient of determination and of the standard error), turned out to be the most appropriate one. The coefficient of determination (R^2 showed the high level of explaining the variability of supply by the selected statistical model (Approach_4). The estimated model took the following form: $S_t = 12186.12 * \text{arctg}(0.0028 * P_t^e)$.

Therefore, it can be concluded that in making decisions on the production volume for the next year agricultural producers were most often guided by the level of the price forecast for the next year ($t + 1$) rather than by the current market price (t). This phenomenon was in contradiction with the idea and assumptions of the price cobweb model.

Conclusions

In the economics, much attention is devoted to price fluctuations in markets. Their level is a result of the game of demand and supply and in particular of markets' striving for achieving the level of equilibrium. These adaptation processes in markets are illustrated by the cobweb model where the dynamising factor is the delay of supply in response to the current price. Producers usually respond with a shift in time by one period ($t + 1$). Its objective is to determine a proper price path for this good so that the supply and demand are fully implemented in each year (i.e. market clearing price)[5]. This path can be written as follows: $\{P(t)\}_{t=0}^{\infty}$.

The agricultural sector was an original example showing the mechanism of functioning of the price cobweb, although now, due to a possibility of storing most agricultural products, this principle refers only to products which are perishable or unfit for storage. Therefore, the scope of use and suitability of the model in current economic (market) conditions is reduced.

Agricultural producers, due to the duration of the production process, are somehow forced to make decisions in conditions of uncertainty as to the future price. Therefore, they build on the level of the price from the previous period. Moreover, they usually, do not take into account their previous experiences (the occurrence of short market memory and continuing to make the same errors).

[5] Non-linear deterministic models appeared in economics in the mid-20th century. They are based on an assumption that a source of fluctuations and instabilities, observed in the economic reality, can be not external distortions but endogenous factors. The possibility of using the theory of chaos in economics was mentioned for the first time in 1975. Since that moment, chaos has been found in many existing models of economic phenomena (Orzeszko, 2005; Prokhorov, 2008).

References

AL-Daami, H.H.M. & AL-Hiyali, A.D.K. (2017). Stability analysis in equilibrium of the demand and supply for wheat crop in Iraq during the period 1971–2013 using cobweb model. *Iraqi Journal of Agricultural Sciences,* 48(5), 1326–1337.

Brock, W.A., & Hommes, C. H. (1997). A rational route to randomness. *Econometrica,* 65(5), 1059–1095. *Communications in nonlinear science and numerical simulation,* 57, 402–414.

Ezekiel, M. (1938). The Cobweb Theorem. *Quarterly Journal of Economics,* 52(2), 255–280.

Finkenstadt, B. (1995). *Nonlinear Dynamics in Economics: A Theoretical and Statistical Approach to Agricultural Markets*. Lecture Notes in Economics and Mathematical Systems, Springer, 426.

Górka, J. (2009). *Ekonomia matematyczna: materiały do ćwiczeń*. Warszawa: Wydawnictwo C.H. Beck, 29–47.

Hommes, C.H. (1991). Adaptive learning and roads to chaos: the case of the cobweb. *Econo-mics Letters,* 36(2), 127–132.

Jakimowicz, A. (2010). Źródła *niestabilności struktur rynkowych*. Warszawa: Wydawnictwo Naukowe PWN.

Lenart, Ł. & Mazur, B. (2017). Business cycle analysis with short time series: a stochastic versus a nonstochastic approach. *The 11th Prof. A. Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena – Conference Proceedings,* 212–222.

Lichtenberg, A.J. & Ujihara, A. (1989). Application of nonlinear mapping – theory to commodity price fluctuations. *Journal of Economic Dynamics & Control,* 13(2), 225–246.

Mazur, B. (2018). Cyclical fluctuations of global food prices: a predictive analysis. *The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena – Conference Proceedings,* 286–297.

Onozaki, T., Sieg, G. & Yokoo, G. (2000). Complex dynamics in a cobweb model with adaptive production adjustment, *Journal of Economic Behavior and Organization,* 4, 101–115.

Orzeszko, W. (2005). *Identyfikacja i prognozowanie chaosu deterministycznego w ekonomicznych szeregach czasowych*. Warszawa: Polskie Towarzystwo Ekonomiczne, 29–36.

Prokhorov, A.B. (2008). Nonlinear Dynamics and Chaos Theory in Economics: a Historical Perspective, *Quantile,* 4, 79–92.

Stępień S. (2015). Wahania cykliczne na rynku mięsa wieprzowego w skali globalnej. *Roczniki Naukowe SERiA, 17*(3), 362–366.

Włodarczyk, J. (2006). Wahania cen w liniowych i nieliniowych modelach pajęczyny. *Studia Ekonomiczne Akademii Ekonomicznej w Katowicach – Teoretyczne i pragmatyczne zagadnienia ekonomii,* 43, 71–93.

On the impact of intraday trading volume on return's volatility – a case of the Warsaw Stock Exchange

Roman Huptas [1]

Abstract
The aim of this paper is to present results of preliminary research in which the influence of intraday trading volume on return's volatility on the Warsaw Stock Exchange is empirically examined. This study investigates whether the effect of intraday trading volume on return's volatility is homogenous by dividing trading volume into its expected and unexpected components. We use 10-minute intraday data and measure return's volatility by the exponential generalized autoregressive conditional heteroscedasticity (EGARCH) structure using expected and unexpected components of trading volume as explanatory variables. We found that both higher expected and unexpected trading volumes are connected with a higher conditional return's volatility. We also observed that unexpected volume shocks have a significantly larger effect on return's volatility than changes in expected volume. Moreover, when volume is split into its expected and unexpected components and then incorporated into the conditional variance specification, GARCH effects are definitely reduced.

Keywords: *volume-volatility relationship, trading volume, return volatility, ACV model, EGARCH*
JEL Classification: *C58, C11, C22*

1. Introduction

The information role of trading volume in explaining price volatility and returns has generated a lot of interest for a long time. Hence, the dynamic relation between asset returns, return's volatility and trading volume has been a subject of a considerable amount of research. It must be stressed that understanding of the volume-volatility relation might eventually lead to better volatility forecasting and a new and better way for modelling returns distributions. Moreover, analysis of the role of trading volume in explaining return's volatility is of importance to policy makers and investors to better understand the response of the market to information shocks and dissemination of new information among market participants.

There are at least two theories in market microstructure literature that explain the volume-volatility relationship. The first one refers to the mixture of distributions hypothesis (MDH) developed by Clark (1973) and Epps and Epps (1976). According to this theory, a joint distribution of volume and volatility is conditional upon the flow of information into the market. Thus, both trading volume and volatility react and change contemporaneously in response to the arrival of new information.

On the other hand, the sequential information arrival hypothesis (SIAH) was advanced by Copeland (1976) and Jennings et al. (1981). This theory assumes that new market information is disseminated sequentially to traders. Therefore, the process in which new information is impounded into the price can spread out over time. It explains the lead-lag relation between volatility and trading volume.

[1] Corresponding author: Cracow University of Economics, Department of Statistics, 27 Rakowicka St., 31-510 Cracow, Poland, e-mail: huptasr@uek.krakow.pl.

Beyond theoretical considerations, there is an extensive literature on the empirical aspects of the volatility-volume and return-volume relations. Empirical evidence on the volatility-volume relationship was found in numerous papers. The main results of theoretical frameworks were confirmed primarily by early studies of Clark (1973), Epps and Epps (1976), Jennings et al. (1981) and Karpoff (1987). However, the volatility-volume relation was also found by many other studies, among which are the following: Lamoureux and Lastrapes (1990), Gallant et al. (1992), Bessembinder and Seguin (1993), Jones et al. (1994), Gallo and Pacini (2000), Chen et al. (2001), Girard and Biswas (2007), Chevallier and Sevi (2012), Slim and Dahmene (2016), Ftiti et al. (2017) among others.

Many research papers usually document the positive correlation between trading volume and price volatility. However, in many cases there are not identical conclusions of the empirical research. Results range from lack of, to weak and strong relationships between return's volatility and volume. Research even points out the possibility of a negative relation between volatility and trading volume. Furthermore, studies on this topic were conducted mostly for daily, weekly or monthly data. The application of intraday data is rather limited (Chevallier and Sevi, 2012; Slim and Dahmene, 2016; Ftiti et al., 2017). Moreover, the majority of empirical studies concentrate on well developed markets, especially on the U.S. market. Research on the volume-volatility relation for developing countries of Middle and Eastern Europe, such as Poland is still relatively sparse. The papers on the Polish financial market include Bohl and Henke (2003), Gurgul et al. (2005), Doman (2011), Bień-Barkowska (2012). Therefore, further insight should be obtained through different econometric methods as well as for the high frequency data and less developed markets, including Poland.

This study provides additional empirical evidence on the relations between price volatility and trading activity measured by the trading volume. The aim of this research is to present results of a pilot study in which the influence of intraday trading volume on return's volatility on the Warsaw Stock Exchange is empirically examined. Following Bessembinder and Seguin (1993), this study investigates whether the effect of intraday trading volume on return's volatility is homogenous by dividing trading volume into its expected (anticipated) and unexpected (unanticipated) components and allowing each component to have a separable effect on return volatility. In this study in order to define expected and unexpected trading volume variables, the autoregressive conditional volume (ACV) model of Manganelli (2005) is applied to describe the observed trading volume. To model return's volatility, the exponential generalised autoregressive conditional heteroscedasticity (EGARCH) structure is employed using expected and unexpected components of trading volume as explanatory variables. An empirical study of the intraday volume-volatility relationship is performed for 10-minute intraday volume and return data of the main index of the Warsaw Stock Exchange. To estimate considered model, Bayesian approach is adopted. The Markov chain Monte Carlo (MCMC) methods including Metropolis-Hastings (MH) algorithm are suitably used to obtain posterior densities of interest.

2. Research methodology

In order to examine the influence of trading volume on return's volatility, we must first distinguish between the expected and unexpected trading volume. It was well documented that the trading volume is highly autocorrelated, indicating that it is also highly forecastable. Thus, the expected volume is the result of more persistent fluctuations in liquidity whereas unexpected volume should approximate a new information arrival to the market. Trading volume decomposition into its expected and unexpected components is typically performed by means of ARMA models (Bessembinder and Seguin, 1993; Bjonnes et al., 2003; among others). In this study in order to define expected and unexpected trading volume variables, the ACV model of Manganelli (2005) is applied to describe the observed trading volume. In particular we use the linear ACV(1,1) model with the generalised gamma distribution for the error term. This model for the volume v_i, $i = 0, 1, 2,...,N$ (with N standing for the total number of observations) can be written as follows:

$$v_i = \Phi_i \cdot \varepsilon_i, \tag{1}$$

$$\Phi_i = E(v_i | \Im_{i-1}, \theta), \tag{2}$$

where $\Im_{i-1}$ denotes the set of information available at time t_{i-1},θ is the vector of unknown parameters, Φ_i represents the conditional expected trading volume, ε_i denotes an error term and $\{\varepsilon_i\} \sim i.\ i.\ d.\ GG\,(\lambda, \gamma, v)$ with parameter $\lambda = \left(\Gamma\left(\frac{v}{\gamma}\right)/\Gamma\left(\frac{1+v}{\gamma}\right)\right)^{\gamma}$ such that expected value $E\,(\varepsilon_i) = 1$. The conditional expectation of trading volume has the following representation:

$$\Phi_i = \omega + \alpha \cdot v_{i-1} + \beta \cdot \Phi_{i-1}, \tag{3}$$

where $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta < 1$. Accordingly, the expected volume $v_{exp,i}$ is defined as an estimate of the conditional expectation of volume $\widehat{\Phi}_i$, whereas the unexpected volume $v_{unexp,i}$ is defined as ratio of observed volume to expected volume $v_{unexp,i} = v_i / \widehat{\Phi}_i$.

Now we proceed to dynamic specification for the 10-minute financial logarithmic returns. We assume a simple AR (1) structure for the intraday returns:

$$r_i - \delta = \rho \cdot (r_{i-1} - \delta) + u_i, \tag{4}$$

$$u_i = \sigma_i \cdot \xi_i, \tag{5}$$

where $-1 > \rho < 1$, ξ_i is the innovation term, $\{\xi_i\} \sim$ i. i. d. t (0; 1; v), $v > 2$ and σ_i^2 is the conditional variance of the returns. By t (0; 1; v) we denote Student's *t* distribution with zero mean, unit precision and an unknown number of degrees of freedom $v > 2$. To model return's volatility, we specify the EGARCH structure with expected and unexpected components of trading volume as explanatory variables. In fact, we propose an EGARCH (1,1)-type specification of the conditional variance, the dynamics of which evolves according to the following equation:

$$\ln \sigma_i^2 = \omega_G + \alpha_{1G} \cdot \xi_{i-1} + \alpha_{2G} \cdot \left(|\xi_{i-1}| - \frac{2\Gamma\left(\frac{v+1}{2}\right)\sqrt{v-2}}{\Gamma\left(\frac{v}{2}\right)(v-1)\sqrt{\pi}} \right)$$

$$+ \beta_G \cdot \ln \sigma_{i-1}^2 + \eta_1 \cdot v_{exp,i} + \eta_2 \cdot v_{unexp,i}\,. \tag{6}$$

This specification allows for an asymmetric response of σ_i^2 to volatility shocks in the innovation term ξ_{i-1} when parameter α_{1G} differs from zero. The use of an EGARCH-type model is also justified by the advantage of keeping the volatility component positive regardless of the sign of the right-hand side components in the volatility equation. The absence of non-negativity constraints on the parameters also facilitates numerical estimation.

3. Bayesian estimation of the AR-EGARCH-X and ACV models

In order to estimate parameters of the proposed models, the Bayesian approach is applied. Bayesian estimation of the AR-EGARCH-X and ACV models outlined above requires certain prior assumptions. We assume that all parameters – whenever possible – are *a priori* independent. Moreover, in order to express the lack of prior knowledge, fairly diffuse prior distributions are assumed, so that the data dominates the inference about the parameters through the likelihood function. Specifically, for all parameters of Equation (3) we propose the normal distributions with zero mean and standard deviation of five, adequately truncated, due to relevant restrictions imposed on the parameters in the model. For the ACV model with the generalised gamma innovations, the prior density for parameter γ is also specified as density of the normal distribution with zero mean and standard deviation of five adequately truncated whereas the prior density for parameter v is specified as density of the normal distribution with zero mean and standard deviation of ten adequately truncated. The prior for the parameter ρ is set to be uniform over the $(-1;1)$ interval. Prior for degrees of freedom v is practically uniform and assumes that the parameter is restricted to the range (2; 102) since it ensures that the conditional variance exists. The upper bound is used for numerical convenience only and the restriction has very limited empirical consequences. For remaining parameters of Equations (4) and (6) we propose the normal distributions with zero mean and standard deviation of five, adequately truncated if necessary.

The inference was conducted using MCMC techniques. The MH algorithm with a multivariate Student's *t* candidate generating distribution with three degrees of freedom and the expected value equal to the previous state of the Markov chain was used to generate a pseudo-random sample from the posterior distribution. The covariance matrix was obtained from initial cycles, which were performed to calibrate the sampling mechanism. Convergence of chain was carefully examined by starting the MCMC scheme from different initial points and checking trace plots of iterates for convergence to the same posterior. Acceptance rates were sufficiently high and always exceeded 50%, indicating good mixing properties of the posterior sampler. The final results and conclusions were based on 100,000 draws, preceded by 50,000 burn-in cycles. All codes were implemented by the author and ran using the GAUSS software, version 13.0.

4. Data description

The empirical analysis is based on 10-minute intraday data of the WIG20 Index of the Warsaw Stock Exchange (WSE). Our sample consists of transaction prices and trading volumes matched for each time interval. The analysis covers the period from January 12th, 2018 to April 11th, 2018 (62 trading days), and is based on transactions carried out in the continuous trading phase which in the case of the WSE in 2018 starts at 09:00 and ends at 17:20 (GMT+1). The data are obtained from the Thomson Reuters Eikon Database. From the data, we generate the 10-minute index return series by taking the log of the ratio of last transaction prices in successive intervals.

It is well documented in the financial literature the existence of intraday periodicities in intraday returns and trading volumes. The intraday periodic patterns for intraday returns and volumes were estimated using the Nadaraya-Watson kernel estimator of regression of the variable on the time of the day. Finally, intraday returns and trading volumes were deseasonalised by dividing plain data by diurnal factor to obtain diurnally adjusted data.

5. Empirical results

Bayesian estimation results of the proposed AR-EGARCH-X model, including marginal posterior means and standard deviations (in parentheses), are reported in Table 1.

Table 1. Posterior means and standard deviations (in parentheses) of parameters in AR-EGARCH-X model

Parameter	Pure AR-EGARCH model	AR-EGARCH-X model with expected and unexpected volume variables
		The return equation
δ	−0.0311 (0.0197)	− 0.0206 (0.0185)
ρ	−0.0507 (0.0155)	− 0.0688 (0.0161)
		The volatility equation
ω_G	0.0286 (0.0156)	− 1.3159 (0.1424)
α_{1G}	−0.0383 (0.0120)	− 0.0649 (0.0329)
α_{2G}	0.0971 (0.0233)	0.1783 (0.0498)
β_G	0.9658 (0.0176)	0.1714 (0.0780)
η_1	----------	0.7417 (0.1349)
η_2	----------	1.0537 (0.0723)
υ	3.2935 (0.1761)	4.5526 (0.3995)

Looking at the results for the return equation parameters, it can be noted that the posterior mean of the autoregressive coefficient ($\boldsymbol{\rho}$) is negative and equal to – 0.0688. Moreover, the posterior distribution of $\boldsymbol{\rho}$ is well-separated from zero and features relatively little dispersion indicated by the standard deviation of about 0.0161. Moreover, the conditional normality of the intraday returns is strongly overridden by the data. The posterior mean and standard deviation of the degrees of freedom equal about 4.5526 and 0.3995, respectively. Therefore, our results confirm that allowing for fat tails of the conditional distribution may be crucial for empirically adequate statistical modelling with the use of GARCH-type processes.

As far as the variance equation is concerned, it can be seen that the posterior results for parameter β_G imply fairly weak volatility persistence. The posterior mean of β_G stands at about 0.1714. The asymmetry effect is slightly negative, but it seems to be statistically insignificant. The posterior mean of α_{1G} is equal to – 0.0649 and is accompanied by a relatively large posterior standard deviation of about 0.0329.

Finally, we analyse the estimation results for parameters η_1 and η_2 pertaining to the effects of trading volume on conditional variance. The posterior distributions of both parameters are well-separated from zero, indicating statistical significance of trading volume variables. The posterior mean of η_1 (related to the expected trading volume, $v_{exp,i}$) is positive and equal to 0.7417, suggesting that positive impact of the expected volume on return's volatility. Thus, higher expected trading volume is connected with a higher volatility. The posterior mean of η_2 (related to the unexpected trading volume, $v_{unexp,i}$) is also positive, amounting to about 1.0537. The posterior distribution of the parameter under consideration reveals rather a relatively small dispersion which is implied by the standard deviation of about 0.0723. It follows that higher unexpected trading volume is connected with a higher volatility. It can also be noted that unexpected shocks have larger effect on return's volatility than changes in expected volume. Thus, surprises in trading activity measured by trading volume have a larger effect on return's volatility than forecastable trading activity. Moreover, when volume is split into its expected and unexpected components and then incorporated into the conditional variance specification, GARCH effects are definitely reduced compared to pure EGARCH specification (see Table 1).

Conclusions

The main objective of this paper is to present results of a pilot study in which the influence of intraday trading volume on return's volatility on the WSE is empirically examined. We used 10-minute intraday data and measured return's volatility by the exponential generalized autoregressive conditional heteroscedasticity structure using expected and unexpected components of trading volume as explanatory variables.

Our main findings can be summarised as follows. The results suggest that trading volume has a significant impact on the return volatility of the main index of the WSE. We found that both higher expected and unexpected trading volume are connected with a higher conditional return's volatility. We also showed that unexpected volume shocks have significantly stronger effect on return's volatility than changes in expected volume. Moreover, when volume is divided into its expected and unexpected components and then incorporated into the conditional variance specification, GARCH effects are definitely reduced. These findings demonstrate the importance of splitting the total trading volume into expected and unexpected components. Finally, it must be stressed that our analysis was limited only to the main index of the WSE. So it would be of great interest to examine more indices from other well-developed and emerging markets in further research. Moreover, the issue that needs further examination is to analyse particular stocks included in the WIG20 index, which display substantial cross-sectional differences resulting from different capitalization. It will allow us to better understand the differences across various stocks and market structures.

Acknowledgements
This work was financed from the funds granted to the Faculty of Management at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

Bessembinder, H., & Seguin, P.J. (1993). Price volatility, trading volume and market depth: Evidence from futures markets. *Journal of Financial and Quantitative Analysis*, 28, 21–39.

Bień-Barkowska, K. (2012). Does It Take Volume to Move the EUR/PLN FX Rates? Evidence from Quantile Regressions. *Dynamic Econometric Models,* 12, 35–52.

Bjonnes, G.H., Rime, D., & Solheim, H.O. (2003). Volume and volatility in the FX market: Does it matter who you are? *CESifo Working Paper*, No. 783.

Bohl, M.T., & Henke, H. (2003). Trading volume and stock market activity: the Polish case. *International Review of Financial Analysis*, 12, 513–525.

Chen, G., Firth, M., & Rui, O.M. (2001). The dynamic relation between stock returns, trading volume, and volatility. *The Financial Review*, 38, 153–174.

Chevallier, J., & Sevi, B. (2012). On the volatility-volume relationship in energy futures markets using intraday data. *Energy Economics,* 34(6), 1896–1909.

Copeland, T.E. (1976). A model of asset trading under assumption of sequential information arrival. *Journal of Finance*, 31, 1149–1168.

Clark, P. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41, 135–155.

Doman, M. (2011). *Mikrostruktura giełd papierów wartościowych*. Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.

Epps, T.W., & Epps, M.L. (1976). The stochastic dependence of security price changes and transaction volumes: Implications for the mixture of distribution hypothesis. *Econometrica*, 44, 305–321.

Ftiti, Z., Jawadi, F., & Louhichi, W. (2017). Modelling the relationship between future energy intraday volatility and trading volume with wavelet. *Applied Economics,* 49(20), 1981–1993.

Gallant, A.R., Rossi, P.E., & Tauchen, G. (1992). Stock prices and volume. *The Review of Financial Studies*, 5, 199–242.

Gallo, G., & Pacini, B. (2000). The effects of trading activity on market volatility. *European Journal of Finance*, 6, 163–175.

Girard, E., & Biswas, R. (2007). Trading volume and market volatility: Developed versus emerging markets. *Financial Review*, *42*, 429–459.

Gurgul, H., Majdosz, P., & Mestel, R. (2005). Joint dynamics of prices and trading volume on the Polish stock market. *Managing Global Transitions*, 3(2), 139–156.

Jennings, R.H., Starks, L.T., & Fellingham, J.C. (1981). An equilibrium model of asset trading with sequential information arrival. *Journal of Finance*, 36, 143–161.

Jones, C.M., Kaul, G., & Lipson, M.L. (1994). Transactions, volume, and volatility. *The Review of Financial Studies*, 7(4), 631–651.

Karpoff, J. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, 22, 109–126.

Lamoureux, C., & Lastrapes, W. (1990). Heteroskedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance*, 45, 220–229.

Manganelli, S. (2005). Duration, volume and volatility impact of trades. *Journal of Financial Markets*, 8, 377–399.

Slim, S., & Dahmene, M. (2016). Asymmetric information, volatility components and the volume-volatility relationship for the CAC40 stocks. *Global Finance Journal,* 29, 70–84.

A Study of Usefulness of Selected Robust Model Based Clustering Techniques in a Digital Development Modelling

Daniel Kosiorowski[1], Przemysław Jaśko[2], Ewa Szlachtowska

Abstract

This paper critically discusses three robust model based clustering techniques in a context of their applicative usefulness in a process of specifying a two dimensional model generating spatio-temporal data related to a digital economy. We among others study TCLUST, OTRIMLE algorithms and certain algorithms, which are available in the mclust R package. Theoretical considerations are illustrated by means of empirical issues related to a preliminary analysis of spatial phenomena of a digital economy. Additionally, we present results of simulation studies involving spatial processes departing from regularity.

Keywords: *Robust Clustering, Spatio-Temporal Models, Mixture Models*

JEL Classification: *C14, C53, R10*

1. Introduction and problem formulation

Sustainable socio-economic development of modern societies belongs to hot topics in a current public and scientific debate. A motivation of this paper relates to a problem of specifying a spatial model describing "digital development" in a certain region of space, for example in a geographic or administrative region of a country. The paper concentrates on a choice of an appropriate robust model-based clustering technique, which shall provide functions, which are further used in an estimation of a functional regression model describing a digital development of the region. It should be pointed out that in robust modelling we concentrate on an influential majority of cases, having more or less formalised idea in mind, which does and which does not belong to a data generating mechanism (Kosiorowski and Zawadzki 2019). Despite of the fact that one can find several promising clustering procedures in the literature, which are described by authors as robust, an issue of robustness of a clustering procedure is still an open problem (Hennig, 2004). One cannot find a comprehensive study which compares these procedures in a context of modelling of economic spatial-temporal phenomena. In order to fill that gap we among others conducted a simulation research, in which clusters were generated by a mixture of skewed Student T distributions, a noise has "irregular" spiral shape support and additionally samples contained outliers. We aimed at modelling a situation of an existence of "asymmetric centres of gravity", e.g., cities, a net of roads and an object playing a role of a capital.

A sequence of mappings $E = \{E_n\}_{n\in\mathbb{N}}$ is called a general clustering method, if E_n maps a collection of entities $X^n = \{x_1, ..., x_n\}$ to a collection of subsets $\{C_1, ..., C_G\}$ of X^n. It is as-

[1] Daniel Kosiorowski: Cracow University of Economics, Department of Statistics, 27 Rakowicka St., 31-510 Cracow, Poland, daniel.kosiorowski@uek.krakow.pl.

[2] Przemysław Jaśko: Cracow University of Economics, Department of Computational Systems, 27 Rakowicka St., 31-510 Cracow, Poland, jaskop@uek.krakow.pl.

sumed that entities with different indexes can be distinguished. For a disjoint clustering method (DCM) $C_i \cap C_j = \emptyset$ for $i \neq j \leqslant G$. Most popular DCM yield partitions $\bigcup_{j=1}^{G} C_j = X^n$.

We consider a following general spatio-temporal model of a digital development

$$m\colon \mathbb{R}^d \times T \ni (z,t) \underset{m}{\rightarrow} P(z,t) \in \mathcal{F} \tag{1}$$

where $z \in \mathbb{R}^d$ denotes spatial coordinates, $T = [0, t_*)$ denotes time, $\mathcal{F}$ denotes a certain family of probability density functions defined on $\mathbb{R}^d$ and $P(z, t)$ denotes an element of the family indexed by a space point and time point t (for a detailed presentation of general issues related to a specification of (1) see chapter nine of Kokoszka and Reimherr, 2017).

Assume, we have a collection of samples from the above model $X^n_{z1,t1}, X^n_{z2,t2}, \ldots, X^n_{zG,tG}$, indexed by points in the space and time. For each sample we perform robust model based clustering, $E^m_{z1,t1}, E^m_{z2,t2}, \ldots, E^m_{zG,tG}$. Taking into account, that we consider model based clusterings, the sequence of clusterings denotes a sample of estimates of mixture densities (well-defined functions) indexed by the spatial-temporal parameters. The sequence enables us for an estimation of a specific form of the model (1) using known methods of estimation (e.g. using functional principal component method of estimation of a linear model of a type "vector-function" or functional kernel regression, see Ramsay and Silveman, 2005).

2. A comparison of robust model based clustering methods

The term "model-based cluster analysis" was coined by Banfield and Raftery (1993) for clustering based on finite mixtures of Gaussian distributions and related methods. In multidimensional case the standard Gaussian mixture model is to assume that data $X^n = \{x_1, ..., x_n\}$ are modelled as drawn i. i. d. from a distribution with density

$$f(x;\theta) = \sum_{j=1}^{G} \pi_j \phi(x; \mu_j, \Sigma_j) \tag{2}$$

where $\phi(\cdot, \mu_j, \Sigma_j)$ is the density of a Gaussian distribution with mean vector μ_j and covariance matrix Σ_j, π_j is the proportion of the j-th mixture component, $\sum_{j=1}^{G} \pi_j = 1$. The parameter vector θ contains all proportions, means and covariances.

The most popular estimator of θ is the maximum likelihood estimator (ML). It leads to a natural clustering rule: *classify an observation to the mixture component maximizing its posterior probability for a class membership*.

3. Selected robust model based clustering methods

MCLUST. The MCLUST (Scrucca et al. 2016), uses a Gaussian mixture model which has a single term representing a noise, and is given by

$$\prod_{i=1}^{n} \left[\pi_0 \frac{\mathbf{I}\{x_i \in S\}}{V} + \sum_{j=1}^{G} \pi_j \phi(x_i; \mu_j, \Sigma_j) \right], \tag{3}$$

where V is a volume of a convex hull S of the sample X^n and $\mathbf{I}\{\cdot\}$ denotes an indicator function, x_i represents an observation, G denotes number of components, π_j is the probability that an observation belongs to the j-th component ($\pi_j \geqslant 0, \sum_{j=1}^{G} \pi_j = 1$), $\phi(\cdot; \mu_j, \Sigma_j)$ denotes density of the p-dimensional normal distribution with parameters (μ_j, Σ_j).

Model selection is obtained via maximization of Bayesian Information Criterion (BIC)

$$BIC \equiv \ell_{\mathcal{M}}(X^n, \theta^*) p_{\mathcal{M}} \log(n), \tag{4}$$

where $\ell_{\mathcal{M}}(X^n, \theta^*)$ is the maximized loglikelihood for the model and the data X^n, $p_{\mathcal{M}}$ is the number of parameters to be estimated in the model $\mathcal{M}$, and n is number of observations.

TCLUST. The second algorithm we consider is the TCLUST (Fritz et al. 2012). Various aspects of its empirical usefulness were studied in Szlachtowska et al. (2016).

OTRIMLE. The third algorithm we consider is the RIMLE (Coretto, Hennig 2013), which maximizes a pseudo-likelihood, based on the improper pseudo-density of the form

$$\psi_\delta(x; \theta) = \pi_0 \delta + \sum_{j=1}^{G} \pi_j \phi(x; \mu_j, \Sigma_j), \tag{5}$$

where $\phi(\cdot; \mu_j, \Sigma_j)$ is the p-dimensional Gaussian density with mean $\mu_j \in \mathbb{R}^p$ and covariance matrix Σ_j, π_0, $\pi_j \in [0,1]$ for $j = 1, 2, \ldots, G$, $\pi_0 + \sum_{j=1}^{G} \pi_j = 1$, and $\delta > 0$ is the improper uniform density representing outliers.

This improper uniform density, which is not spanned on the predefined support set (unlike in the MCLUST, where support set S is selected to cover the data sample X^n, is not aimed to model the noise component, but it's rather treated as a technical tool to account for the points, which are in low density areas for Gaussian components. In contrast to the MCLUST model, extreme points in the data sample, won't have impact on the uniform density, which in the RIMLE model takes the improper form.

The parameter vector θ contains all Gaussian components parameters and each of the proportion parameters, including π_0. In the RIMLE model the δ and the number of Gaussian components G are treated as fixed.

Given the sample X^n the improper pseudo-log-likelihood function takes the form

$$\ell_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} \log \psi_\delta(x_i; \theta), \tag{6}$$

and for prespecified value of δ the RIMLE estimator is defined as

$$\hat{\theta}_n^{RIMLE}(\delta) = \arg\max_{\theta \in \Theta_n} \ell_n(\theta), \tag{7}$$

where Θ_n is a constrained parameter space defined as

$$\Theta_n = \left\{ \theta : \pi_j \geqslant 0 \forall j \geqslant 1, \pi_0 + \sum_{j=1}^{G} \pi_j = 1;\ \frac{\lambda_{\max}(\Sigma^\theta)}{\lambda_{\min}(\Sigma^\theta)} \leqslant \gamma;\ \frac{1}{n} \sum_{i=1}^{n} \tau_0(x_i; \theta) \leqslant \pi_{\max} \right\}. \tag{8}$$

The parameter space Θ_n is defined by (Correto and Hennig, 2013) to ensure existence of the RIMLE estimator $\hat{\theta}_n^{RIMLE}(\delta)$. To obtain boundedness of the pseudo-log-likelihood criterion function, along with constraints on the proportion parameters and the eigenratio constraint, additional constraint called the "noise proportion" is needed in the RIMLE case.

To establish the eigenratio constraint, we need to introduce some notation. Given $\lambda_{j,k}$ is the k-th eigenvalue of the j-th component covariance matrix Σ_j, we define the set of eigenvalues $\Lambda(\Sigma^\theta) = \{\lambda_{j,k}: j = 1,2,\ldots, G; k = 1,2,\ldots, p\}$, from which we select respectively minimal and maximal element: $\lambda_{\min}(\Sigma^\theta) = \min_{j,k} \Lambda(\Sigma^\theta)$, $\lambda_{\max}(\Sigma^\theta) = \max_{j,k} \Lambda(\Sigma^\theta)$. Eigenratio constraint is aimed at preventing from the degeneracy of model "regular" components distributions, in other words, from having some components with Gaussian distributions, concentrated in the vector subspace of $\mathbb{R}^p$, resulting in the singular covariance matrix parameters for them, which in turn translates into the infinite value of the pseudo-likelihood function. In the RIMLE case, eigenratio constraint alone, cannot preclude forming "regular" components with distributions concentrated on single points, and all other points fitted by the improper uniform component. In order to mitigate such kind of situations, the "noise proportion" constraint is added to the definition of RIMLE parameter space. Considered constraint is defined using pseudo posterior probabilities of the noise component for consecutive sample observations.

Under the RIMLE model with parameter θ, noise component pseudo posterior probability $\tau_0(x_i; \theta)$, conditional on the observation $x_i \in X^n$, is given by

$$\tau_0(x_i;\theta) = \frac{\pi_0\delta}{\psi_\delta(x_i;\theta)} = \frac{\pi_0\delta}{\pi_0\delta + \sum_{j=1}^{G}\pi_j\phi(x_i;\mu_j,\Sigma_j)}. \tag{9}$$

Using ergodic arguments, X^n sample mean of the pseudo posterior probabilities $\frac{1}{n}\sum_{i=1}^{n}\tau_0(x_i;\theta)$, can be treated as an approximation to the expected proportion of noise points, under model parameter θ. So, constraint $\frac{1}{n}\sum_{i=1}^{n}\tau_0(x_i;\theta) \leqslant \pi_{\max}$, imposes that expected noise proportion should be no higher than $\pi_{\max} \in (0,1)$.

As can be seen the "noise proportion" constraint is sample dependent, which in turn translates into the dependency of the parameter space Θ_n on the sample X^n. Pseudo posterior probabilities $\tau_j(x_i; \theta)$, $j = 1,2,\ldots, G$ for the Gaussian components are defined analogously. The observation x_i is assigned to the cluster whose index corresponds to the highest pseudo posterior $J(x_i; \theta) = \arg\max_{j\in\{1,2,\ldots, G\}}\tau_j(x_i; \theta)$. To fix the value of δ for the noise component improper uniform density, on which estimator $\hat{\theta}_n^{RIMLE}(\delta)$ depends, (Correto, Hennig, 2013) proposed the OTRIMLE (Optimally Tuned RIMLE) procedure. Under the OTRIMLE approach, improper density level δ is selected according to the formal criterion, which allow to make trade-off between conformity of components' empirical distributions with Gaussian distribution and a proportion of outliers. Optimal level for the improper density is given by

$$\delta_n = \arg\min_{\delta \in [0,\delta_{\max}]} D(\delta) + \beta\pi_{0,n}, \tag{10}$$

where $\delta_{\max} > 0$ is a prespecified value, and $\beta \geqslant 0$ is the penalty parameter for increasing the proportion of outliers component, while $D(\delta)$ measures departures from Gaussianity of "regular component" empirical data clusters, resulting under the model improper constant density level of δ. With the observation assignment rule for J the RIMLE estimator $\hat{\theta}_n^{RIMLE}(\delta)$, value of the $D(\delta)$ is based on the Kolmogorov distances between the components' empirical distributions of square Mahalanobis distances and the chi-square distribution with degrees of freedom, which is expected for them under Gaussianity of components. Under the OTRIMLE procedure, to find the optimal value of δ_n, the RIMLE estimator $\hat{\theta}_n^{RIMLE}(\delta)$ value is computed over the candidate set, within certain "golden section search algorithm". To compute values of the RIMLE estimators the ECM-algorithm (Expectation Conditional Maximization algorithm) is used, for which pseudocode is presented in (Correto and Hennig, 2013). This kind of algorithm imposes fulfillment of the RIMLE parameter space constraints in each of its iteration.

4. A comparison of computational algorithms for considered robust estimators

Table 1. A summary comparison of the clustering algorithms

Aspect \ Algorithm	MCLUST	TCLUST	OTRIMLE
An approach to take noise into account	Uniform distribution with a predefined support dependent on a convex hull of a sample X^n.	A noise is jointly modelled with outliers.	Using an improper uniform distribution.
Conditions for the identification of the estimator	Different kinds of parametrisations for cluster dispersion matrices (imposing constraints of cluster distributions).	Conditions on ratios of cluster dispersion matrices eigenvalues.	Conditions on ratios of cluster dispersion matrices eigenvalues, and on proportion of noise.
Consistency of the estimator	By default ML estimator enhanced with BIC is used.	By default ML estimator based on the MCD is used.	
Affine equivariance of the estimator	Maximised objective function is affine equivariant, but initialization steps are not.	Depending on the predefined constraints on scatter matrices.	Not, for the basic form of the estimator. Yes, for the modified RIMLE.
Algorithm for finding an optimal value of the objective function	Expectation-Maximization (EM)	Classification EM	Expectation-Conditional Maximization (ECM)

Aspect \ Algorithm	**MCLUST**	**TCLUST**	**OTRIMLE**
Breakdown point	Problems with outliers at extreme positions.	Depends on a configuration of data in a sample.	Depends on a configuration of data in a sample.
Implementation in `R`	`mclust`	`tclust`	`otrimle`

OTRIMLE criterion for the δ constant, allows for a trade-off between a departure of components of empirical distributions from normality and proportion of outliers in a sample. Thanks to it, OTRIMLE is less prone to existence of the extreme points in the sample data than the MCLUST, where support for the noise component uniform density is a set including all the sample points (most often convex hull or hyperrectangle).

Simulation studies. In order to investigate small and moderate sample properties of the algorithms in a context of exploring spatial phenomena of digital economy we conducted extensive simulation studies. We among others generated samples from 2D mixtures of three component skewed Student T distributions "noised" by distribution with two-spiral support. The samples consisted of 3330 points, from which 3000 were "clean" points, 300 noise points and 30 outlying observations.

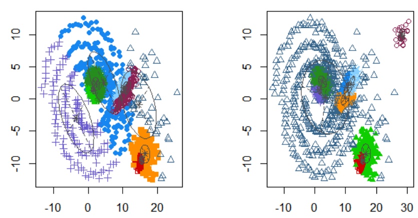


Fig. 1. Results of clustering using MCLUST

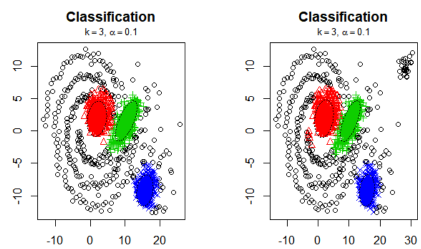


Fig. 2. Results of clustering using TCLUST

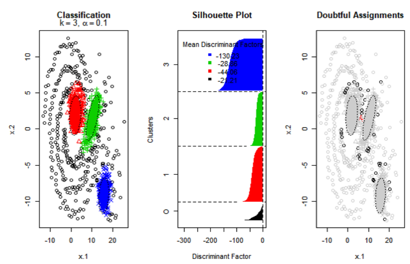


Fig. 3. Properties of TCLUST clustering for dataset with 10% spiral noise

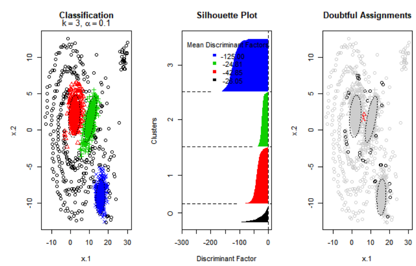


Fig. 4. Properties of TCLUST clustering for dataset with 10% spiral noise with 1% outliers

Figure 1 presents results of the clustering using MCLUST for "clean data with noise" (left panel) and "noised data with outliers" (right panel). Figures 2–4 present analogous situation for the TCLUST algorithm and figures 5–6 for the OTRIMLE algorithm. The figures may be treated as an illustration of a one experiment from a collection of 1000 experiments.

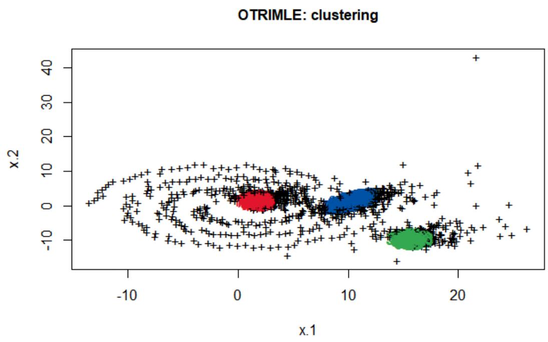


Fig. 5. Results of clustering using OTRIMLE

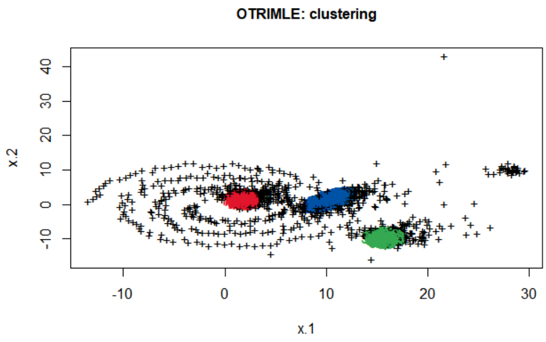


Fig. 6. Results of clustering using OTRIMLE

Empirical example. We used considered algorithms in the image segmentation task for the 24-bit RGB digital raster image, with the resolution of 200 × 200 pixels, representing night lights satellite image of Poland. We aimed at establishing spatial clusters, with similar level of activity, measured by the night light intensity. We assumed that the light intensity correspond to a degree of development. To extract data from the image we used R package magick. Data were downloaded from the NOAA Database (https://www.ngdc.noaa.gov/eog/download.html). Figures 7 and 8 present results of the clustering obtained via MCLUST and TCLUST algorithms correspondingly. For the algorithms we assumed data sample partition into six clusters (indicated by the BIC for MCLUST model) and a set of outliers.

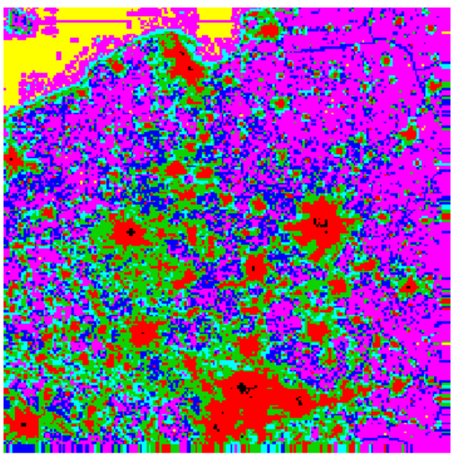


Fig. 7. Night light intensity–results of MCLUST

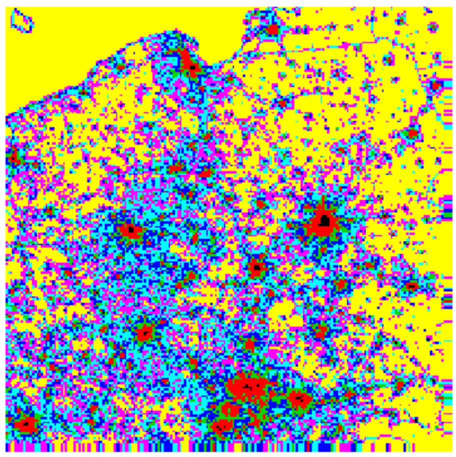


Fig. 8. Night light intensity–results of TCLUST

For both methods results, cluster numbers are sorted according to the decreasing order of its mean nightlight brightness. Set grouping outlying observations, associated with the brightest pixels is numbered as zero. First we describe results of the MCLUST algorithm.

Table 2. Pixels (area) partition between clusters by MCLUST and TCLUST

Cluster\ Fraction of pixels	0 (black)	1 (red)	2 (green)	3 (blue)	4 (cyan)	5 (magenta)	6 (yellow)
MCLUST	0.18	13.96	22.46	16.86	9.87	32.76	3.91
TCLUST	0.20	4.13	6.23	10.95	19.03	19.65	39.81

The brightest pixels associated with the main parts of Polish biggest cities (Upper Silesian conurbation, Warsaw, Poznan, Cracow, Gdansk) were assigned as outlying observations (these pixels are marked by the black colour on the maps). The first cluster is formed by the brightest pixels (marked on the map by the red colour) not assigned as outliers, which are concentrated in the areas near the Polish biggest cities, especially area between Upper Silesia and Krakow and Warsaw suburbs are distinguished. Subsequent three clusters with decreasing mean brightness level, are continuously formed by pixels associated with places with increasing distance from the main cities (pixels from clusters 2 to 4 are marked on the map respectively by the colours: green blue and cyan). Cluster 5 groups pixels representing terrestrial areas, which are least illuminated, which are large parts of NE and NW Poland. The sixth cluster, groups pixels representing areas on the Baltic Sea, distant from the shores, which are not artificially enlightened. For the TCLUST (with assumed fraction of outliers equal 0.2%) clustering results seems to be much better than that for the previous method. Most of the clusters include spatially concentrated group of pixels. Pixels assigned as outliers (black), form areas which highly resemble administrative territories of the biggest Polish cities. First cluster contains pixels (in red) associated with strict metropolitan areas of the mentioned cities (the radius for this areas are much smaller than in MCLUST case). The cluster 2 contains comparatively to the MCLUST, tiny fraction of pixels (in green) concentrated mainly in the area of Upper Silesia and western part of Lesser Poland. Cluster 4 contains pixels (cyan) which form fairly vast, continuous areas outside the metropolitan areas of biggest cities. Cluster 5 embraces highly scattered pixels (in magenta), mainly in the western part of the country. Cluster 6 contains the largest fraction of pixels, representing the least illuminated terrestrial areas (north-eastern and eastern part of Poland and north-west) and the full analysed area of the Baltic Sea (MCLUST assigned to the sixth cluster only distant parts on the sea). To sum up, in the case of the MCLUST and TCLUST respectively, pixels associated with each of the six clusters (0 stands for “outlying” pixels group), constitute following fractions of the pixels, representing all of the analysed area (which is presented on the map).

Conclusions and further studies

We have critically studied three high quality model based clustering algorithms in the context of their applications in modelling of spatial phenomena. We cannot indicate a “total winner”, which uniformly maximizes all criteria of evaluation.

Acknowledgements

DK and PJ thanks for financial support from the Ministry of Science and Higher Education within "Regional Initiative of Excellence" Programme for 2019–2022. Project no.: 021/RID/2018/19. Total financing: 11 897 131,40 PLN. DK thanks for the support related to CUE grant for the research resources preservation 2019.

References

Banfield, J.D. and Raftery, A.E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics,* 49, 803–821.

Fritz H., García-Escudero L. A., Mayo-Iscar A., (2012), tclust: An R Package for a Trimming Approach to Cluster Analysis, *Journal of Statistical Software*, 47(12), 1–26.

Hennig, C. (2004). Breakdown points for maximum likelihood estimators of location – scale mixtures. *The Annals of Statistics*, 32(4), 1313–1340.

Coretto, P., & Hennig, C. (2013). Finding approximately Gaussian clusters via robust improper maximum likelihood. *arXiv preprint arXiv:1309.6895*.

Hennig, C. (2008). Dissolution point and isolation robustness: robustness criteria for general cluster analysis methods. *Journal of Multivariate Analysis*, 99(6), 1154–1176.

Kokoszka, Reimherr, M. (2017). Introduction to Functional Data Analysis, CRC, London.

Kosiorowski, D., & Zawadzki, Z. (2019). DepthProc An R Package for Robust Exploration of Multidimensional Economic Phenomena, *Journal of Statistical Software*, forthcoming.

Ramsay J., & Silverman, B. (2005). *Functional Data Analysis*. Springer.

Scrucca, L., Fop, M., Murphy, T.B., & Raftery, A.E. (2016). mclust 5: Clustering, classification and density estimation using gaussian finite mixture models. *The R journal*, 8(1), 289.

Szlachtowska E., Kosiorowski D., Mielczarek D., (2016). Ocena jakości aplikacyjnej odpornego algorytmu analizy skupień TCLUST na przykładzie zbioru danych dotyczących jakości powietrza w Krakowie, *Przegląd Statystyczny*, R. 63(1), 67–80.

The diversity and intensity of poverty in Poland

Maciej Jewczak[1], Karol Korczak[2]

Abstract

The aim of the paper was to investigate the phenomenon of poverty in Poland from the subregional perspective. As the source of data on households, the Social Diagnosis data for 2015 was used. The data sample consisted of 10 thousand of households with full information on their levels of equivalent income and expenses. The research focused on the structure of poverty among households differentiated by number of members and the geographical affiliation. The analysis was carried out in two stages. Firstly, households have been characterized in terms of three poverty categories: the statutory poverty, the extreme poverty and rarely estimated the social minimum rate. In this part, the purpose of the research was to determine the levels for indicated categories of poverty for subregions and then comparing them within the voivodships. An important part of the work allowed demonstrating intra-regional differentiation of poverty types. In the second stage of the research, an attempt to assess the intensity of poverty levels was made. Within this framework, the indicators of the intensity of the considered categories of poverty were estimated, based on the category of household income gap. The obtained results enabled us to examine the households' poverty depth due to their geographical location and the certain type of household.

Keywords: *absolute (extreme) poverty; statutory poverty, social minimum, poverty intensity*

JEL Classification: *C23, G15, I32, Q47, R19*

1. Introduction

Poverty is a relative term that can be difficult to define due to the large number of factors that affect this phenomenon. Being in poverty does not depend solely on the individual (or household) position. It is a concept affected by the social policy and the quality of social services implemented at local, regional, national, and international levels. As a result, poverty is a multifactorial and multidimensional term that not only affects the individual's situation, but also conditions the socio-economic of the society; therefore it is a problem that accompanies the development of humanity itself.

The poverty is a subject of research of various scientific disciplines, with the interest of social policy, economics and sociology. The social policy considers poverty as a problem to be solved by the state. Economics, on the other hand, focuses on ways of dealing with the limited access to or even the lack of important goods or services that are essential to living. In sociology being in poverty is related to the level of income needed to maintain a life or perform the physical labor. In this sense, this concept should be perceived as biological poverty, which in literature is also identified as absolute poverty and related strongly to the theory of basic needs and their fulfilment. However, due to the development of societies the range of the needs considered as essential is constantly expanding (Panek and Szulc, 2004).

[1] Corresponding author: University of Lodz, Department of Operational Research, 37 Rewolucji 1905 St., 90-214 Lodz, e-mail: maciej.jewczak@uni.lodz.pl.

[2] University of Lodz, Department of Computer Science in Economics, 37 Rewolucji 1905 St., 90-214 Lodz, e-mail: karol.korczak@uni.lodz.pl.

In order to be able to remedy and counteract poverty, it is necessary to recognize the causes of its occurrences. In literature, three groups of poverty causes are mentioned, such as: personal, subjective and structural (Kaźmierczak and Łuczyńska, 1996). Personal causes occur independently and include situations of disability (also mental), long-term illness and advanced age – situations in which people cannot survive on their own and are unable to provide for a minimum of subsistence or to participate in common social life. Also, the numerous numbers of children in a household is perceived as the cause; in such families, usually only one member is employed, or all adults are jobless (Kumor, 2011). Subjective causes of poverty include specific attitudes and characteristics of people, such as: laziness, lack of principles, different system of values, unwillingness to work or to educate, dishonesty and mismanagement (Kowalak, 2002). The last group of causes of poverty results from the state aid, while it often limits the motivation to work, individual's independence and can lead to moral depression (Skorowski et al., 2006).

2. Subject and Range of the Investigation

The poverty can be caused by the differences identified in welfare, ownership, income, level of intelligence, ambition, ingenuity, education and qualifications levels. There are also various factors affecting the socio-demographic situation of households, as well as the overall situation of the country and the region of residence (Skorowski et al., 2006).

The multidimensionality of poverty causes further restrictions in its analysis, for instance concerning its definition and interpretation. Poverty can be understood in absolute or relative terms. In Poland, the most common and documented are the three basic poverty categories, such as:

- statutory – it involves people whose level of expenditures is lower than the monetary limit that enables the recipients making use of social assisted benefits;
- absolute (extreme) – "*a condition characterized by severe deprivation of basic human needs, including food, safe drinking water, sanitation facilities, health, shelter, education and information. It depends not only on income, but also on access to social services*" (United Nations, 1996). This type of poverty in nature is usually long term and often inherited;
- social minimum – defined as the scope and the level of fulfilled needs; the social minimum standard should provide such conditions that at every stage of human's development enables him/her to reproduce life forces, to have and raise children and to maintain social bonds (Deniszczuk, 1977).

In literature, there is also a relative poverty distinguished that is perceived as meeting the average standard of living. If a family's income is not high enough, the household is considered to be in relative poverty. This type of poverty is defined differently in Poland than in EU; in Poland, the reference point is income lower than 50% of the average remuneration, whereas in some EU countries 60% of median. The differences in the definition and the lack of essential information on the relative poverty levels for households with the number of members from two to five have eliminated this type of poverty from further analyses.

Research on poverty focuses mostly on the situation where individual/household remains permanently in the low-income group (e.g.: associated with inheriting poverty). Commonly, the levels of poverty are estimated for higher level of data aggregation, for instance voivodship level (Krzysztofik et al., 2017). For that reason, an attempt was made to examine the phenomenon of poverty in Poland from the lower, subregional perspective – to Authors' best knowledge such approach is rarely considered in the analysis (Szymkowiak et al., 2017), however examples mentioned were based on the indirect estimation of poverty.

In proposed paper, not only the level of poverty was estimated. The research focused also on establishing the intensity[3] of poverty. Previous Authors' experiments to establish the poverty enclaves and the depth of poverty resulted in proposed *Multiplicative Indicator of Poverty Intensity* (*MIPI*) for selected 25 streets of Zgierz, Poland (Jewczak et. al., 2018). *MIPI* indicator allowed us to estimate the intensity of poverty from the social assistance perspective, taking into account the number of recipients (beneficiaries) and the amounts allocated to social benefits, which in turn depended on the conditions of household income. However, it was not possible to calculate *MIPI*, due to the lack of very detailed and, at the same time, hard-to-access data for the local level.

3. Data and research methodology

As the source of information on households, the Social Diagnosis data for 2015 was used (Integrated database. www.diagnoza.com). The research of the Council for Social Monitoring (the Social Diagnosis) is a widely used, representative source of data on the situation and condition of Polish households (e.g.: Czapiński and Panek, 2015; Baranowski et al., 2016). The data sample consisted of 10 thousand of Polish households differentiated by number of members and the geographical affiliation, with full information on their levels of equivalent income and expenses. However, only households with a maximum of 5 persons were selected for the study. It was due to the lack of appropriate conversion rates for larger units. Households with more than five members accounted for less than 4% of the research sample and were mainly the multi-family units with an internal structure difficult to classify.

The analysis was carried out in two main stages. Firstly, households were characterised in terms of three categories of poverty, accordingly to the assigned levels:

- statutory poverty – the limit of law poverty in 2015 ranged from 634 PLN for one person to 2420 PLN for five people household of monthly income (Ustawa z dnia 12 marca 2004 r. o pomocy społecznej (Dz.U. 2018 poz. 1508));

[3] In literature, some examples can be found that the terms such as the depth and the intensity of poverty should be clearly distinguished. However, these terms are most often used interchangeably and related only to the income category when estimating the poverty rates (e.g.: Atkinson et. al., 2002; Lar et al., 2012). For clarification, in the proposed study, the depth and the intensity of poverty should be considered interchangeably, without taking into account additional factors, i.e.: those not related to income.

- extreme poverty – the limit of extreme poverty in 2015 ranged from 545.76 PLN for one person to 2389.11 PLN for five people household of monthly income (Instytut Pracy i Spraw Socjalnych, 2016a);
- social minimum rate – the limit of extreme poverty in 2015 ranged from 1079.53 PLN for one person to 4198.12 PLN for five people household of monthly income (Instytut Pracy i Spraw Socjalnych, 2016b).

In this part, the purpose of the research work was to determine the levels of indicated categories of poverty for subregional levels and then comparing them within voivodships' boundaries, indicating intra-regional diversity. Households were evaluated accordingly to their poverty rates. Atkinson et al. (2002) indicated that poverty rate is the ratio of the number of households (or at the individual level – persons) whose income *per capita* falls below the given poverty line in relation to total number of households (in this research three types of poverty were analysed – in accordance with the above-cited classification).

In the second stage of the research, an attempt to assess the intensity of poverty was made. Within this framework, the indicators of the intensity of the considered categories of poverty were estimated and based on the household income gap. The poverty indicator (*PI*) measures its intensity and informs on the distance between the poor and the given type of poverty line (Szwarc, 2005) and is typically defined, as follows (Mussard and Alperin, 2011):

$$PI = \frac{1}{Q}\sum_{i=1}^{Q}\left(\frac{z - y_i}{z}\right), \tag{1}$$

where: z – is a level of poverty line, y_i – income level of an individual in the poverty zone, with the assumption that $y_i < z$.

The amount Σ_i $(z - y_i)$ is called the income gap with values ranging from 0 to 1 (0–100%) and its lower level indicates lower intensity of poverty. The depth of poverty is therefore directly related to household's income (Li, 2018). The lower the income a household achieves, the further it is below the given type of poverty line, which indicates major depth of poverty. Therefore, income growth reduces the depth of poverty (Iceland and Bauman, 2007). The households were classified as being under or above poverty line with the adjustment of an equivalence scale that weights the household according to the number of equivalent adults. If a couple is below the poverty line then both adults have the equivalent income of $\frac{2}{3}$ household income y, so the income gap amounts to $\left(z - \frac{2}{3}y_i\right)$.

The core of the research included three stages. Firstly, households were classified into three poverty subgroups: statutory, extreme and social minimum. Secondly, the average subregional poverty rates were estimated and the results were obtained in order to group the households according to both the spatial affiliation and the types of households. Thirdly, the intensity of each type of poverty was examined. In order to summarise the overall poverty level for subregions (in terms of the rates and its intensity) the Ward's agglomeration grouping method was applied, which uses the analysis of variance to estimate the distance between clusters (Ward, 1963).

4. Results of empirical analyses

The most challenging part of obtaining the essential results was the classification of households according to the appropriate poverty levels and the household type. In order to complete the research aim, it was necessary for each type of household, according to its type (e.g.: non-family, marriage type 2+0, etc.) to assign accurate levels of border income, which qualified the unit to be (or not) in a given type of poverty. The use of algorithms, described above, allowed us to determine sub-regional levels of rates for a given type of poverty.

4.1. Poverty rates and its intensity by households' structure

Analysing the structure of the households (Table 1), the worse situation of being in poverty was estimated for the largest households. One in two household with 5 members were below the social minimum poverty line, while for 4-person households this rate indicated nearly one in three. Households with more than 4 members registered the highest rates of being in extreme and statutory (that enables making use of social benefits) poverty.

Table 1. Poverty rates and its intensity by type and household number of members

Household size	Poverty rates			Poverty intensity		
	extreme	statutory	social min.	extreme	statutory	social min.
1	2.2	3.9	19.7	31.4	25.2	24.6
2	3.1	4.8	16.7	18.6	29.1	26.0
3	3.8	9.6	29.0	27.2	21.4	27.3
4	8.3	11.4	31.6	22.6	22.7	31.7
5	11.4	11.9	49.9	31.0	30.6	30.0

Research also indicated that the highest intensity of extreme poverty should be associated with individual, non-family households, which confirms indirectly the assumption about the poor situation of single-person households.

4.2. Subregional diversification of poverty rates and their intensity

In this section households were classified in two ways. Firstly, the attention was paid to the poverty rates (Fig. 1).

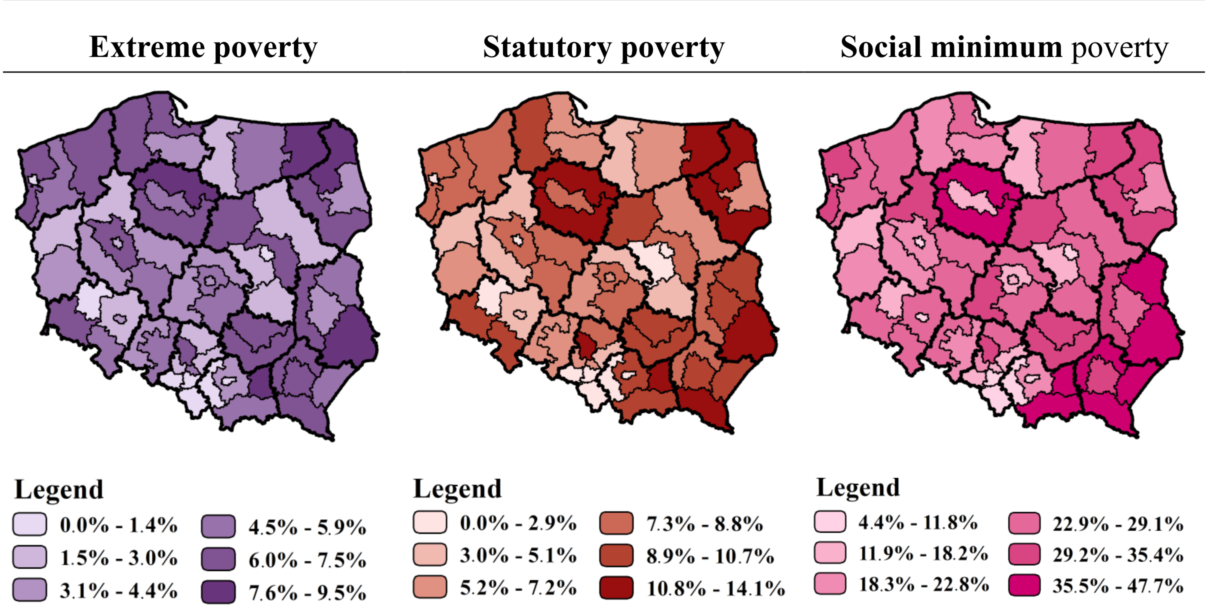


Fig. 1. Subregional poverty rates by type

However, the rates of being in a particular type of poverty might be high or low, the question arises on poverty intensity. For this purpose the poverty intensity indicator (*PI*) was calculated at subregional level (Fig. 2).

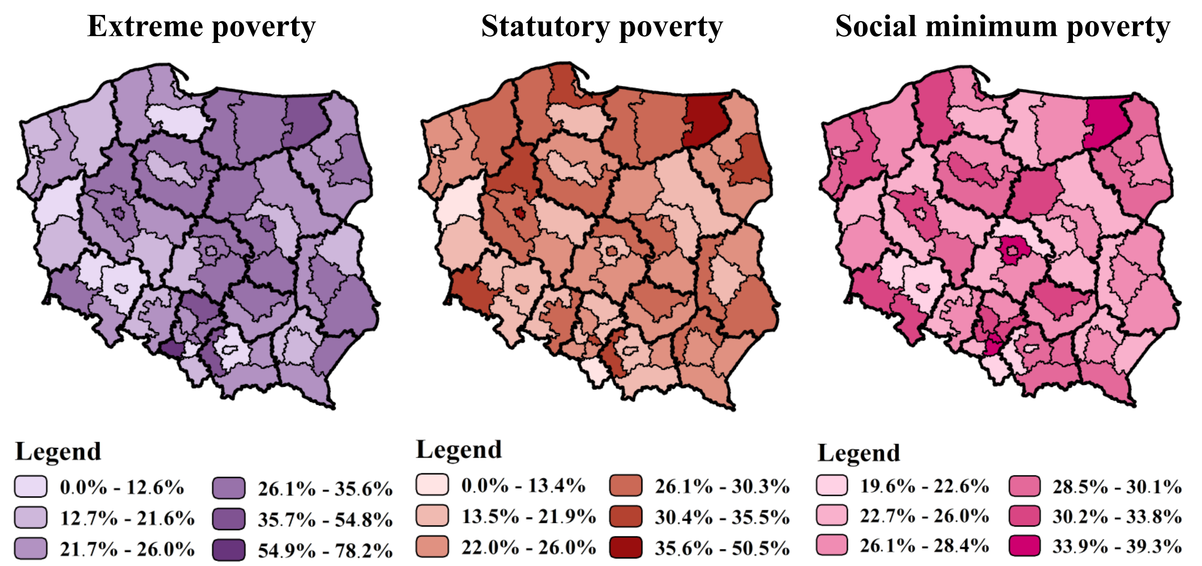


Fig. 2. Subregional poverty intensity by type

The results confirmed that despite high levels of subregional poverty, in most cases poverty turned out to be relatively low in intensity. On the other hand, those subregions that recorded in 2015 low levels of poverty rates, unfortunately, were characterised by a higher intensity of the phenomenon. For instance, for Rybnicki subregion, the extreme poverty rate was estimated on 0.74%, however the depth of this type of poverty amounted to the level of 78.16% (which was the highest for this type of poverty). Tyski subregion was characterized by one of the lowest social minimum rates (5.48%), but noted average *PI*=39.32% of its intensity. Such tendencies were observed for many objects, which supports the argument for thorough and two-dimensional analysis of the phenomenon of poverty: firstly, in absolute term and further supported by the intensity analysis (widely used *PI* indicator), which may indicate a different scale of the problem. The analyses on poverty allowed not only to estimate the subregional rates, but also to investigate the depth of the phenomenon. It results from the property that high-poverty locations are not necessarily high in the poverty intensity. This argument pointed previously Atkinson et al. (2002). Lack of any pattern or dependency between the level and the depth of poverty can be also support by low and insignificant values of non-parametric correlation coefficients, which were investigated by the Authors.

To summarise the distributions among subregions of poverty types and the intensity at the same time the cluster analysis with Ward's grouping was adopted and the results of classification was illustrated on Fig. 3.

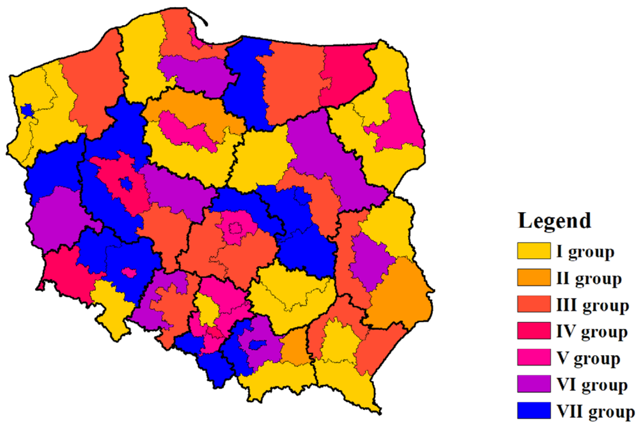


Fig. 3. Spatial similarity of poverty in Polish subregions
Note: colors indicate the belonging to a similar cluster

Taking all the information on poverty into account, it can be said that at the subregional level, the spatial diversity of poverty levels and intensity is rather high – with a bond length of about 16% of the maximum distance, 7 clusters were created, none of which was a single-item. The Ward's grouping allowed us to indicate the similarities of poverty rates and its intensity. Although there are no visible patterns (e.g.: in terms of spatial autocorrelation), the results of multivariate techniques provided information that may be beneficial from the point of view of social policy planning of services and financial support in certain locations.

Conclusions

Researchers most often focus on the economic aspects of the phenomenon of poverty, i.e.: deprivation of material living conditions (income, housing conditions and possessed goods). It is due to the easier rationalisation of poverty alleviation and the fact that problems, such as non-participating in society are also related to other factors. This research allowed us to demonstrate intra-regional differentiation of poverty types and, at the same time, to determine the households' poverty intensity due to their geographical location and the certain type of household. The results may be interpreted as important from the perspective, for instance, of the social policy or strictly in economic terms.

The data published by the Polish Central Statistical Office (*CSO*) present only the levels of poverty for provinces. However, proposed analysis indicated that the image of poverty at the higher level of data aggregation does not comply fully with the intra-look. Extremely valuable information resulting from the conducted research is also the poverty intensity (*PI*) of households differentiated in terms of the number of their members, based on the income gap. This finding supports the assumption of more detailed, local investigation on the phenomenon of poverty. It is also worth emphasising that the most common analyses on poverty, if they are carried out, concern typical households: one and four persons. Though, conducting such thorough analysis faces limitations such as lack of necessary data. Within these indications, it would be beneficial to move to the level of poviats or communes and carry out similar analyses for lowest possible levels of data aggregation.

The research confirmed the possibility of getting information on household or individual performance from different representative studies such as Social Diagnosis. The obtained results of the poverty levels are largely dependent on the quality of the collected statistical material, which has also been indicated by other Authors (e.g.: Atkinson et al., 2002).

Nonetheless, the poverty levels do not differ significantly from the results of research carried out by the *CSO*, except for the social minimum that is practically neglected or at best rarely estimated. In Authors' opinion, without observation of social minimum in view of its nature covering most aspects of life and functioning of the household, it is not possible to receive full information on poverty level, not mentioning its intensity.

References

Atkinson, T., Cantillon, B., Marlier, E. & Nolan, B. (2002). *Social Indicators: The EU and Social Inclusion*, Oxford Scholarship Online, Oxford University Press.

Baranowski, P., Gądek, A., Stelmasiak, D. & Wójcik, S. (2016). Wyjechać czy zostać? Determinanty zamiarów emigracji zarobkowej z Polski. *Gospodarka Narodowa*, 4, 69–89.

Czapiński, J. & Panek, T. (eds.) (2015). Social Diagnosis 2015: Objective and Subjective Quality of Life in Poland. *Contemporary Economics*, *9*(4).

Deniszczuk, L. (1977). *Wzorzec konsumpcji społecznie niezbędnej*, Studia i Materiały IPiSS, Zeszyt Nr 10, Warszawa.

Iceland, J. & Bauman, K.J. (2007). Income poverty and material hardship: how strong is the association? *The Journal of Socio-Economics*, 36(3), 376–396.

Instytut Pracy i Spraw Socjalnych (2016a). *Informacja o wysokości minimum egzystencji w 2015 r. (dane średnioroczne).* Warszawa: IPiSS.

Instytut Pracy i Spraw Socjalnych (2016b). *Informacja o wysokości minimum socjalnego w 2015 r. (dane średnioroczne).* Warszawa: IPiSS.

Jewczak, M., Suchecka, J., Korczak, K. & Melaniuk, M. (2018). Spatial analysis and assessment of effectiveness of selected social services. In: Papież, M. & Śmiech, S. (eds.), *Proceedings of the 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena*, 190–199.

Kaźmierczak, T. & Łuczyńska, M. (1996). *Wprowadzenie do pomocy społecznej*. Katowice: Wydawnictwo Śląsk.

Kowalak, T. (2002). *Polityka społeczna. Wybrane zagadnienia*. Białystok: Wyższa Szkoła Ekonomiczna.

Krzysztofik, R., Dymitrow, M., Grzelak-Kostulska, E. & Biegańska, E. (2017). Poverty and social exclusion: An alternative spatial explanation. *Bulletin of Geography. Socioeconomic Series*, *35*(35), 45–64.

Kumor, I. (2011). Ubóstwo – ujęcie teoretyczne. *Edukacja Humanistyczna*, 1(24), 101–110.

Lar, N., Calkins, P., Sriboonchitta, S., & Leeahtam, P. (2012). Policy-based analysis of the intensity, causes and effects of poverty: the case of Mawlamyine, Myanmar. *Canadian Journal of Development Studies*, 33(1), 58–76.

Li, L. (2018). Financial inclusion and poverty: The role of relative income. *China Economic Review*, 52, 165–191.

Mussard, S. & Alperin, M.N.P. (2011). Poverty growth in Scandinavian countries: A Sen multi-decomposition. *Economic Modelling*, 28(6), 2842–2853.

Panek, T. & Szulc, A. (2004). *Statystyka społeczna. Wybrane zagadnienia*. Warszawa: Szkoła Główna Handlowa.

Skorowski, H., Koral, J. & Gocko, J. (2006). *Świat u progu XXI wieku: wybrane zagadnienia z problematyki międzynarodowej: studium z katolickiej nauki społecznej*. Warszawa-Tyczyn: Wydawnictwo Wyższej Szkoły Społeczno-Gospodarczej.

Szymkowiak, M., Młodak, A. & Wawrowski, Ł. (2017). Mapping poverty at the level of subregions in Poland using indirect estimation. *Statistics in Transition*, 18(4), 609–635.

Szwarc, K. (2005). Próba zastosowania modeli ścieżkowych do analizy głębokości ubóstwa. In: Jajuga, K. & Walesiak, M. (eds.), *Taksonomia 12. Klasyfikacja i analiza danych – teoria i zastosowania*, *Nr 1076,* Prace Naukowe Akademii Ekonomicznej we Wrocławiu, Wrocław, 311–318.

United Nations (1996). *Report of the World Summit for Social Development*. New York.

Ustawa z dnia 12 marca 2004 r. o pomocy społecznej (Dz.U. 2018 poz. 1508).

Ward, J.H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244.

Energy poverty in the European Union. State of play

Lilia Karpinska[1]

Abstract

The article explores the concept of Pan-European composite index as a measure of energy poverty. Only indirect energy poverty indicators are selected. The study pursues two goals. The first one is to discover the possibilities of a composite energy poverty index as a simple tool delivered to policy- and decision-makers. The second one is to depict the most contributing energy poverty factors. To that end, the paper considers two dimensions of energy poverty, housing and income ones. The variables for the composite index and logistic regression are selected to cover both dimensions. Several assumptions are made with regards to the number and choice of indicators employed in the construction of the composite energy poverty index. The results show that energy poverty is strongly affected by general poverty prevalence in the EU26 countries. Hierarchical clustering reveals four groups of countries, including outliers. The EU-SILC micro-data is provided by the EU Commission within the Research Project Proposal 204/2018-EU-SILC.

Key words: *energy poverty, composite energy poverty index, indirect energy poverty metrics*

JEL Classification: *I31, I32, Q40, Q42*

1. Introduction

Energy poverty has recently received a lot of attention in the EU policy-makers primarily due to a wide range of political commitments starting from combatting social exclusion, health improvement, environmental protection, building stock renovation, vulnerable consumer protection, energy market integration, usage of renewables, prosumer role, households' energy efficiency and savings. The paper aims at describing the current situation of energy poverty prevalence in the EU. It is worth highlighting that energy poverty is a cross-cutting issue. There are a few EU policies mentioned in the founding EU treaties that cover at least one of the energy poverty dimension (Consolidated version of the Treaty on the Functioning of the European Union). Social policy, public health, consumer policy, environment, energy policy those are policy domains largely associated with the predicament. In most of the documents energy poverty is measured with indirect metrics from the EU-SILC database. The EU-SILC is a primary European survey on poverty and social exclusion. The key questions form the EU-SILC, which are related to insufficient energy consumption, are ability to keep home adequately warm; arrears on utility bills; leaks/damp/rot in the dwelling, and problems with the dwelling being too dark. The definition of energy poverty provided on the EU energy poverty observatory web-site, launched by the EU Commission in the early 2018, states that energy poor households experience inadequate level of essential energy services, which results in aggravation of diseases and poor wellbeing ("EU Energy Poverty Observatory", 2018). The EU Commission emphasises the distinct character of energy poverty and puts the phenomenon into

[1] Cracow University of Economics, 27 Rakowicka St., 31-510, Cracow, Poland, e-mail: lilea.lk@gmail.com

the broad context of energy consumer vulnerability and single energy market integration (Clean Energy for All Europeans).

2. Literature survey

Many studies on energy poverty are country-specific and based on national household energy consumption surveys. Not so much is done in relation to estimating energy poverty at the EU level not to mention the ongoing debate on the appropriate metrics. The literature on energy poverty in the EU could be roughly divided into two broad categories, policy-centered and metrics-centered. When it comes to energy poverty metrics in cross-countries comparisons, the vast majority of studies explore consensual metrics. Some authors deny possibility of single energy poverty measure as it is ill-suited to countries' specificity and could possible impede effective policy targeting (Deller, 2018).

Failures of direct energy poverty metrics are widely discussed in the literature (Moore, 2012; Heindl and Schuessler, 2015). Each indicator is based on several assumptions that are not always possible to maintain. Direct energy poverty metrics could be demanding in terms of data sophistication, i.e. detailed technical parameters of housing stock and data availability, which is the major problem for the EU analysis (Fuel poverty, 2018). Besides, direct energy poverty metrics provides inconsistent results ranging from too low to incredibly high rates of energy poverty depending on the thresholds and assumptions (Karpinska, 2018).

Cross-country comparisons are mainly constrained by lack of relevant micro data collected at the EU level. To that end EU-SILC is used. The survey is the only available micro-dataset suitable for Pan-European energy poverty assessment. The studies on the EU energy poverty prevalence go back to 2002. In the first Pan-European energy poverty assessment, the authors model the probability of energy poverty based on probit regression and build a composite measure of the predicament (Healy and Clinch, 2002). Various scenarios with different weights assigned to each variable are tested and different sets of variables are considered in the construction of a composite measure. The similar studies are conducted afterwards (Thomson, Bouzarovski and Snell, 2017). When exploring logistic regression and statistics from EU-SILC dataset, the authors discovered the propensity of households to energy poverty across the EU countries (Thomson and Snell, 2013). The research carried out under the auspices of the Commission, concludes that consensual-based metrics is the only easily accessible energy poverty metrics for the EU (Rademaekers et al., 2016).

Considering the limitations associated with single energy poverty metrics, the study discovers the possibility of a composite index as a practical and easily understandable tool for policy action. Besides, composite indices are proved to be the most influential and transparent measure, when it comes to energy poverty metrics. Composite or multidimensional energy poverty index (MEPI) was introduced within the context of energy access debate and achievement of the UN seventh millennium development goal, which is to "ensure access to affordable, reliable, sustainable, and modern energy for all" ("International Energy Agency", 2018). Initially the index was designed as a tool to estimate the depth and scale of modern energy services deprivation in the least developed and developing countries (Nussbaumer, Bazilian and Modi, 2012). The index is an aggregated measure that captures several dimensions of energy poverty. It resembles multiple human and sustainable development metrics used by the international organizations.

The study contributes to the literature on Pan-European energy poverty assessment by introducing hierarchical clustering to group the EU26 countries based on the results of a composite index and log regression estimates. Two classifications are conducted. The first one makes use of composite index results. The second one is based on the logistic regression estimates. The clusters are then compared. To conduct the first grouping, composite energy poverty index is computed at micro-level. The obtained results help to depict the profile of energy poor and to reveal the determinants of the predicament. This study supports the idea that energy poverty is complex and is related to both, energy efficiency of houses and income poverty.

The second grouping is performed based on log regression results. In line with previous researchers, log regression analysis is conducted for a set of EU-SILC variables. The set of variables used in a composite index and in log regression is almost the same. Ability to keep home warm variable is predicted with two variables related to energy efficiency of building stock, such as leaking roofs, and darkness in a house. The income component is covered by arrear on utility bills and poverty indicator.

3. Methodology and data

The study is based on the latest available cross-sectional EU-SILC variables from household data section. The EU-SILC survey is conducted by national statistical offices using harmonized procedures and is compiled by Eurostat. The latest data upgrade was released in November 2018. The number of observations depends on the country and varies between around 3,9 thousand for Malta and Luxembourg to 22 thousand for Italy. Household data for Poland-2017 contains around 13 thousand records. The survey is initially designed to cover income poverty and social inclusions with the purpose to monitor progress on 2020 goals. However, as discussed earlier the database contains questions that intuitively fit for energy poverty and vulnerability measurement. To build composite energy poverty index, three indirect energy poverty measures were selected accompanied with poverty indicator. Weighted sum model is adopted with equal weights being assigned to each indicator. The weights are arbitrary. The study makes usage of weighted additive models developed in the literature on the topic. The experts' cut off line is adopted at 0.5 points. Each indicator has weight of 0.25. The variables are transformed into binary, where 1 means energy poverty. Both types of "yes" answers to question HS021 are simplified to "yes" in a composite index analysis. Summary statistics of the chosen indicators is presented in Table 1.

Table 1. Descriptive statistics

Variable	Value	Dimension
HH040 Leaking roof, damp walls/ floors/foundation, or rot in window frames or floor	1-yes 2-no	Housing
HH050 Ability to keep home adequately warm	1-yes 2-no	Housing/ income poverty
HS021 Arrears on utility bills	1 yes, once 2 yes, twice or more 3 no	Income poverty
HS160 Problems with the dwelling: too dark	1 yes 2 no	Housing
HX080 Poverty indicator	0 when >= at risk of poverty threshold 1 when < at risk of poverty threshold	Income poverty

Each household is assigned composite energy poverty index. The chosen variables are HH040, HH050, HS021 and HX080. In addition, variables HH040, HS160, HS021 and HX080 are regressed against the probability of HH050. The log regression is computed for each country. The output variable has value 1 for energy poverty and 0 otherwise. In our case, all input variables except HS021 are binary. For clustering purposes, the study uses hierarchical (agglomerative) method. The results are presented on dendrograms depicting similarities within groups. Ward's minimum variance within the group method is deployed.

4. Results

In the first step of the analysis the composite energy poverty index is calculated. The index captures both the scale and the intensity of the phenomenon. As mentioned earlier four drivers of energy poverty are considered. All of them account for either energy efficiency of housing stock or income poverty of a household. Three indicators are objective in nature, because they impartially describe the real fact, whereas the ability to keep home warm is a subjective indicator that depends on various country-specific factors. This aspect of cultural and behavioral differences between countries should be considered in the first place when analysing the results of estimations. Energy poverty rate for EU26 is presented in Fig. 1. The rate ranges from 3,8% to 21,1%. Median is around 8%. The distribution is right-skewed, and the interquartile range is 7,97 pp. The worst performing countries are Bulgaria, Greece, Portugal, Cyprus, Lithuania, Latvia and Hungary. The lowest energy poverty rates are observed in Finland, Sweden, Czech Republic, Denmark, Netherlands, Slovakia, and Austria.

Energy poverty depth is the prevalence of the phenomenon across the respective equivalised income quartiles, i.e. the lower the quartile the deeper energy poverty is. The most affected countries are Bulgaria, Greece, Lithuania, Portugal and Croatia. The ratio of energy poor to non energy poor in each quartile is calculated. In the first income quartile median value of the ratio is 0,4017. In the second quartile the number of energy poor decreases, and the ratio varies between 0,0013 (Austria) to 0,4112 (Bulgaria). The distribution of the aforementioned ratio in the second quartile slightly differs, however the most vulnerable countries are practically the same. The upper quartiles in Austria and Belgium are not affected by energy poverty at all. The results of the discussed distributions are depicted in Fig. 2. Bulgaria is removed as an extreme outlier.

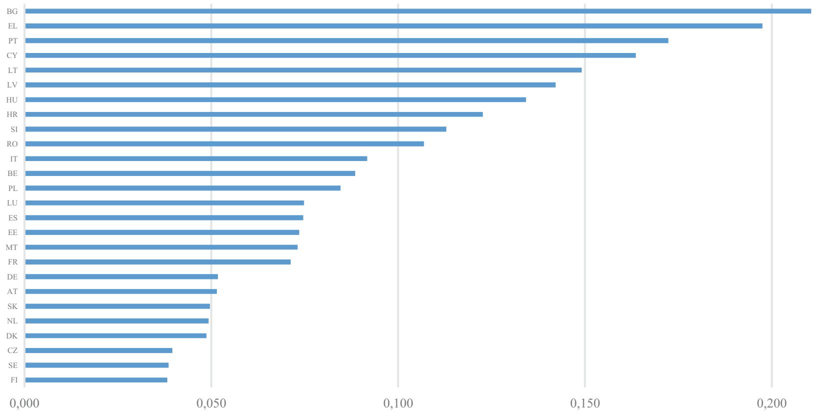


Fig. 1. Energy poverty rate

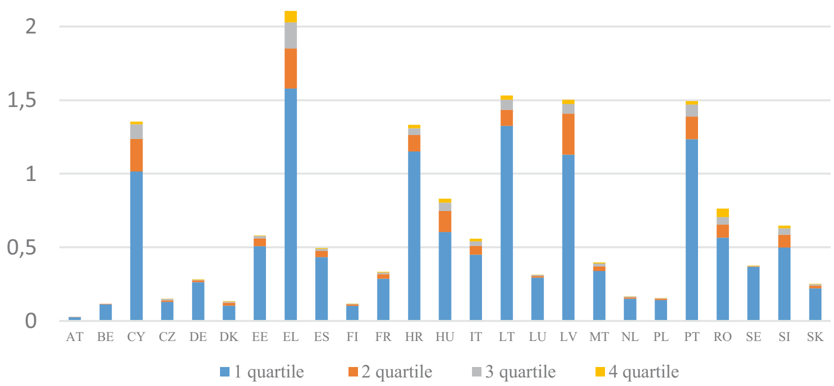


Fig. 2. Distribution of energy poverty by income quartile (without BG)

It is worth mentioning that some of the variables have greater impact on output than others. In Fig. 3 indicators have been sorted as being related mostly to energy efficiency dimension (HH040) or income poverty dimension (HS021, HX080).

Ability to keep home warm is equally divided between housing and poverty dimensions of energy poverty. Poverty domains are a bit of more importance in this analysis, meaning that countries experience energy poverty primarily due to general poverty. Croatia, Hungary, Poland and Slovenia appear to be in the forefront of energy poor countries where energy poverty is caused by poor dwellings.

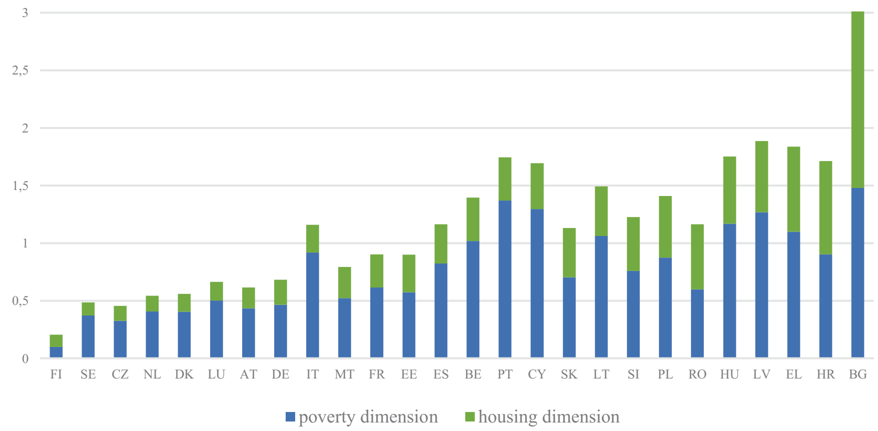


Fig. 3. Energy poverty determinants

In the second step of the analysis, logistic regression is computed[2]. In addition to index variables, HS160 is added. Ability to keep home warm is a predicted value. Given the limited number of determinants included into the model the predicting power of the model is around 53%. Variables have different significance levels across the countries. HX080 has high p-value for Lithuania and Luxembourg, slightly lower for Denmark and is not significant for Finland. It means that ability to keep home warm in those countries is impacted by other factors than income poverty. When it comes to HH040 and HS160 they are significant for most of the countries. It is worth highlighting that in the models HS021 has adverse impact on the odd ratio. In the logistic regression variables are not transformed to 0-1 binary variable. Hence, the interpretation depends on the initial values.

Table 2. Distribution of energy poverty determinants in groups of countries (min, max, median)

Countries/variables	Housing dimension		Income poverty dimension	
	HH040	HH050	HS021	HX080
AT, CZ, DE, DK, EE, ES, FI, FR, LU, MT, NL, RO, SE, SK	0.03/0.39/ 0.08	0.06/0.46/ 0.25	0.04/0.50/ 0.27	0.03/0.19/ 0.06
BE, EL, HR, HU, LV, PL, SI	**0.13/0.51/ 0.28**	**0.45/0.76/ 0.67**	0.21/0.72/ 0.44	0.17/0.55/ 0.22
BG	1.31	0.44	0.65	0.61
CY, IT, LT, PT	0.15/0.35/ 0.22	0.17/0.36/ 0.25	**0.68/0.89/ 0.71**	**0.14/0.43/ 0.29**
EU26	0.03/1.31/ 0.15	0.06/0.76/ 0.32	0.04/0.89/ 0.33	0.03/0.61/ 0.15

In the third step grouping of the EU26 countries is performed. Hierarchical clustering results are presented in Table 2 and 3. The analysis reveals four groups of countries, including two

[2] The results are available upon request.

groups of one outlier in each clustering. Distribution of variables across groups represents the most influential factors for each group of countries. Groups are not similar in each clustering. One representative group is identified in each clustering. As outlined previously, the study focuses on two dimensions of energy poverty, housing and income poverty-related. HH040 is an objective criterion used to describe poor houses. It has the highest values for the second group and the lowest for the first one. Another variable that is partially related to housing dimension is HH050, which has also the greatest impact in the second group, and a lower impact in the rest of the groups including Bulgaria. HS021 and HX080 are associated with income poverty. Income poverty is an issue in the fourth group, and the same group has highest arrears on utility bills. Bulgaria is an outlier and has extremely high values for all indicators, except HH050. It should be remembered that HH050 is the only indicator that has strong socio-behavioral specificity.

Grouping countries based on log regression estimates indicates that housing dimension plays an important role in the first group (HH040) and in the fourth group (HS160). Income poverty mostly affects the second group (HX080. There is strong evidence that income poverty is a key energy poverty driver. It is no surprise that energy poverty in the countries richer in terms of social welfare is driven mainly by energy inefficient housing stock.

Table 3. Distribution of logistic regression coefficients in groups of countries (min, max, median)

Countries/ coefficients	**Housing dimension**		**Income poverty dimension**		
	HH040	**HS160**	**HS021**	**HS021**	**HX080**
AT, BE, CZ, DE, ES, FR, HR, HU, LV, MT, NL, PL, RO, SI, SK	**0.32/ 1.08/ 0.79**	0.24/ 0.87/ 0.44	-1.28/ 0.08/ -0.40	0.23/ 1.24/ 0.76	-1.42/ -0.32/ -0.97
BG, CY, EL, IT, LT, LU, PT	0.18/ 0.51/ 0.37	0.15/ 0.70/ 0.24	-0.41/ 0.37/ -0.06	0.36/ 1.55/ 0.54	-0.74/ -0.15/- 0.53
DK	1.14	1.39	0.22	1.39	-0.71
EE, FI, SE	0.60/ 0.67/ 0.61	**0.62/ 0.88/ 0.84**	**-1.98/ -0.91/ -1.75**	-0.71/ 0.12/ 0.02	-1.27/ -0.26/ -0.81
EU26	0.18/1.14/ 0.69	0.15/ 1.39/ 0.45	-1.98/ 0.37/ -0.30	-0.71/ 1.55/ 0.73	-1.42/ -0.15/ -0.81

Conclusions

The paper relies on the EU-SILC database and makes use of self-reporting energy poverty indicators. Due to some shortcomings of direct metrics, the paper adopts indirect approach. Thus, concerns related to single energy poverty metrics, especially lack of cross-country comparability are mitigated. The paper addresses the problem of measuring energy poverty at the EU level exploring the concept of multidimensional energy poverty index and indirect metrics. The study begins by computing composite energy poverty index to obtain the distribution of the energy poverty across income groups and estimate the contribution of each variable to the final energy poverty score. The first aspect sheds light on the gap of energy poverty being the deepest in the first income quartile. The second aspect allows the author to characterise the underlying reasons of the predicament. In addition, log regression is conducted to predict the ability to keep home warm. The respective model for each EU country is built using three most popular indirect energy poverty metrics, and poverty indicator. The study concludes by comparing the results of hierarchical clustering of the EU countries based on log regression parameters on the one hand and composite index variables on the other.

To sum up, the main findings of the study are as follows. Firstly, housing and income poverty-related dimensions equally contribute to energy poverty prevalence. Secondly, the first income decile seems to be the most affected by energy poverty with the deepest gap notified for poor countries of the EU. Thirdly, energy poverty is also a concern in upper income quartile groups. In this regard, it is worth highlighting that energy poverty is distinct from income poverty. Fourthly, energy poverty in richer countries is determined more by poor housing conditions than income poverty. The analysis was constrained by a number of assumptions, including limitations of indirect energy poverty metrics, dependency of self-reported metrics on country-specific socio-behavioral patterns, finite number of energy poverty variables.

Acknowledgments

The access to the EU-SILC micro-data is granted by the EU Commission within the framework of the Research Project Proposal 204/2018-EU-SILC. The author gratefully acknowledges financial support from the National Science Centre in Poland (grant no. 2018/29/N/HS4/02813).

References

Deller, D. (2018). Energy affordability in the EU: the risks of metric driven policy, *Energy Policy, 119,* 168–182.

EU Energy Poverty Observatory, What is energy poverty? (2018, January 20), Retrieved from https://www.energypoverty.eu/about/what-energy-poverty

European Commission (2016). Clean Energy for All Europeans. COM/2016/0860 final. Consolidated version of the Treaty on the Functioning of the European Union, OJ C 326, 26.10.2012, 47–390.

Fuel poverty. Methodology handbook (2018). Department for Business, Energy & Industrial Strategy, London.

Healy, J., Clinch, J.P. (2002). Fuel poverty in Europe: a cross-country analysis using a new composite measure, *Environmental Studies Research Series*, Working Papers, ESRS 02/04.

Heindl, P., Schuessler, R. (2015). Dynamic properties of energy affordability measures, *Energy Policy,* 86, 123–132.

International Energy Agency, Sustainable Development Goal 7. (2018, January 20), Retrieved from https://www.iea.org/sdg/

Karpinska, L. (2018). Direct Energy Poverty Metrics: What are the Rates for Poland, *Folia Oeconomica Cracoviensia*, 59, 3–22.

Moore, R. (2012). Improving Hills approach to measuring fuel poverty, Consumer Focus, London.

Nussbaumer, P., Bazilian, M., Modi, V. (2012). Measuring energy poverty: Focusing on what matters, *Renewable and Sustainable Energy Reviews,* 16, 231–243.

Rademaekers, K., Yearwood, J., Ferreira, A., Pye, S., Hamilton, I., Agnolucci, P., Grover, D., Karasek, J., Anisimova, N. (2016). Selecting indicators to measure energy poverty, Final Report, Rotterdam.

Thomson, H., Bouzarovski, S., Snell, C. (2017). Rethinking the measurement of energy poverty in Europe, *Indoor and Print Environment*, 26(7), 879–901.

Thomson, H., Snell, C. (2013). Quantifying the prevalence of fuel poverty across the European Union, *Energy Policy,* 52, 563–572.

Dynamics of the relation between fine wine market and financial markets

Olga Kutera[1]

Abstract

We study the dynamic relation between the global and European stock markets, fine wine market and GBP/USD exchange rate by using VAR-DCC-GARCH framework and daily closing prices of LIVX50, S&P500 and FTSE100 indices from 2010 to 2018. Three versions of univariate GARCH models, namely standard, exponential (Nelson, 1991) and GJR (Glosten, Jagannathan, Runkle, 1993) were used in order to build best fitted multivariate dynamic conditional correlation model. Results of this study reveal the long-term time-varying links in volatility between the global and European markets. We found evidence of negative correlation between fine wines market and global stock markets in few periods. Most important results provide empirical evidence that fine wines can be hedge against declines of British currency and can help investors minimize risk to build optimal portfolios.

Keywords: *DCC-GARCH, wine investment, hedge, safe haven*

JEL Classification: *C58, G10, G15*

1. Introduction

For the last couple of years, alternative investment assets such as gold, diamonds, fine wines, art have been gaining more and more popularity among investors whose are trying to protect their portfolios against adverse market conditions. Fine wines and gold obtained specific meaning in the literature and become defined as hedge or safe haven assets.

According to Baur and Lucey (2010) safe haven asset is defined as uncorrelated or negatively correlated with a conventional portfolio in times of market turmoil, not on average. If an investor adds the safe haven asset to his equity portfolio, it reduces losses in extreme adverse market conditions. A strong (weak) hedge is defined as an asset that is negatively correlated (uncorrelated) with another asset or portfolio on average. In case of gold, the volatility is asymmetric but, unlike other solid ingredients of the portfolio, responds more to positive than a negative shocks on stock market. Therefore, Baur and Lucey (2010) conclude that gold can act as a safe haven against the fall in prices on the traditional financial markets.

In our study, we want to analyse dynamics of the relationship between fine wines and global and European stock markets. Empirical results will help us verify property of fine wines to be hedge or safe haven asset.

A number of studies highlighted the benefits of portfolio diversification using alternative forms of investments. Some of them referred to the relationship between fine wines and traditional financial assets.

Masset and Henderson (2010) analysed evolution of high-end wines prices from 1996 to 2007 and studied their properties from an investor's viewpoint. Using a polynomial goal programming model, they investigated how investors' preferences affect the portfolio allocation

[1] Cracow University of Economics, Department of Econometrics and Operational Research, 27 Rakowicka St., 31-510 Cracow, Poland, olga.sobolak@gmail.com.

and the return distribution. Results have shown that wine returns are only slightly correlated with other assets and could be used to reduce the risk of an equity portfolio. They summarised wine market as heterogeneous, fragmentary and limited in liquidity, which makes it difficult to establish a precise market price for any bottle of wine.

In other studies, Masset and Weisskopf (2010) used auction hammer prices over the period 1996–2009 and analysed risk, return and diversification benefits of fine wine. Including different investor types, they confirmed that wine in a portfolio produced higher returns and lower risk than the Russell 3000 equity index during that interval. The protective property of wine emerged during times of stress on the financial markets when the wine prices drops were smaller in comparison to other equities. The authors suggested wine returns are unrelated to economic market conditions and risk.

Baur (2012) examined the volatility of gold returns for a 30-year period for daily, weekly, monthly and quarterly and different currency denominations. He demonstrated that positive shocks increased the gold returns volatility more than negative shocks which can be characterised as abnormal compared to the findings reported for the volatility in equity markets. Baur (2012) argued that this effect is related to the safe haven property of gold, because the usual explanations for asymmetric volatility such as financial leverage and volatility feedback, cannot be applied to commodities in general and precious metals in particular.

Bouri (2014) applied TGARCH model on monthly data to measure reactions of wine returns on its volatility to previous positive and negative shocks. Empirical results showed that positive shocks increased conditional volatility more than negative innovations, leading to an inverted asymmetric volatility phenomenon. That is opposite to what could be found on the stock market. Bouri (2014) followed the methodology initiated by Baur (2012) and proposed the safe haven effect. Furthermore, the evidence provided show that adding wine to an equity portfolio is rewarded by an increase in risk reduction effectiveness.

Le Fur et al. (2016) examined time-varying risk premium associated with wine investments. They used monthly calculated Liv-ex Fine Wine 1000 index from 2003 to 2010 and conditional CAPM and DCC-GARCH model to calculate time-varying betas and time-varying risk premiums for the fine wine market. The authors concluded that Bordeaux fine wines are the most traded, less volatile during no-crisis periods and more volatile in times of financial crisis.

Bouri and Roubaud (2016) applied ARMA-DCC-GARCH model and analysed time-varying correlation between the returns of fine wine and stocks in the United Kingdom using monthly data from January 2001 to February 2014. They showed that during periods of market turmoil fine wine was a very weak safe haven against the UK stock market, even though there was a negative dynamic conditional correlation between fine wine and British stocks. It suggested the effectiveness of using fine wine as hedge against movements on UK stock market.

Cardebat and Jiao (2018) analysed the long-term relationship between the fine wine market and stock markets. They applied a cointegration approach on monthly database, including the Liv-ex Fine Wine Investables Index. Results suggested significant cointegration between emerging markets and fine wine markets. The slowdown of economic growth in emerging countries, especially Asia, could be a risk to the fine wine market. It is mainly because China appeared to be one of the main drivers of fine wine markets.

Bouri, Gupta, Wong, Zhu (2018) used the mean-variance and stochastic-dominance approaches to understand the role of wine investment within a portfolio of different assets. Their main findings suggested that wine is the best investment among all individual examined assets. To gain higher expected utility when short selling is not allowed, investors preferred to invest in low-risk with-wine portfolios than equal-weighted portfolios. That study revealed the possibility of earning abnormal returns when wine was included in the investment.

In our research, we want to analyse whether fine wine is correlated with stock markets and British currency and how the correlation varies in time. Based on Baur and Lucey (2010) definitions of hedge and safe haven assets we describe hedging property of fine wine.

2. Methodology

Engle (2002) proposed a new class of multivariate GARCH models, namely the DCC model. It can be viewed as a generalisation of the Bollerslev (1990) constant conditional correlation (CCC) model in which a vector r_t of k assets returns is assumed to be distributed with mean vector of 0 and a conditional variance–covariance matrix

$$H_t = D_t R_t D_t,$$
$$D_t = \mathrm{diag}\{\sigma_t\},$$
$$R_t = \rho_{ij},$$

where D_t is a diagonal matrix of time-varying standard deviation from the univariate GARCH model and R_t is a correlation matrix containing the conditional correlations, which do not vary over time. We denote $r_t = D_t\varepsilon_t$ where $\varepsilon_t \sim iidN\,(0,1)$ are residuals.

In contrast to constant conditional correlations, Engle (2002) allowed correlations to be dynamic and vary over time. Therefore, in (1) the correlation matrix R_t is not constant and the conditional variances of R_t must be equal to one. Other than this, requirements for the parameterisation of R_t, are the same as for H_t. The elements of D_t are modelled as univariate GARCH processes. DCC model offers a tractable way of modelling, simultaneously, both time-varying conditional volatilities and time-varying conditional correlations and the DCC-GARCH model for GARCH(M,N) order can be describe as

$$Q_t = \bar{Q}\left(1 - \sum_{m=1}^{M} a_m - \sum_{n=1}^{N} b_n\right) + \sum_{m=1}^{M} a_m\left(\varepsilon_{t-m}\varepsilon'_{t-m}\right) + \sum_{n=1}^{N} b_n Q_{t-n},$$
$$R_t = (Q_t^*)^{-1} Q_t (Q_t^*)^{-1},$$

where $Q_t = \{q_{ijt}\}$ is the conditional variance-covariance matrix of residuals and $\bar{Q}$ is its unconditional time-invariant variance-covariance matrix found in the first stage of the estimation process. The parameters a and b, respectively, determine the effects of shocks and dynamic correlations. (Q_t^*) is a diagonal matrix composed of the square root of the diagonal elements of Q_t.

In our examination, we use two stage estimation process (Engle, 2002). In the first step, residuals from VAR(1) are modelled by using three univariate GARCH models (basic, Bollerslev, 1986; exponential, Nelson, 1991; GJR, Glosten et al., 1993) with six different distributions (normal, t-Student, General Error Distribution, skewed normal, skewed t-Student, skewed General Error Distribution). We choose the best fitted models for each residuals series based on the information criteria. Later on, the standardised first-stage residuals are used to estimate the parameters of the dynamic conditional correlation equation (DCC-GARCH) by relying on chosen GARCH specifications. We use the R program for calculations.

3. Data and descriptive statistics

Analysis includes daily calculated rates of return of three indices and GBP/USD exchange rates over the period between March 1st, 2010 and December 28th, 2018. LVX50 index represents wine market, S&P500 and FTSE100 respectively represent changes on global stock market and European stock market.

The average rates of return are positive (FTSE100 = 0.0001, S&P500 = 0.0004, LVX50 = 0.0001). Due to several large drops in exchange rate related to United Kingdom European Union membership referendum, average return on GBP/USD is negative (-0.0001). Analysed data are skewed to the left and have asymmetric distributions with fat tail. Results of the Augmented Dickey-Fuller test show that all data series are stationary. Based on the Jarque and Bera test we can reject the null hypothesis of normality distribution at a significance level 0.01.

4. Estimation results

Kilian and Lütkepohl (2017) pointed that the appropriate choice of number of lags depends not only on the information criteria but also on the number of variables included, economic context and data frequency. Based on Shibata Information Criteria, daily frequency of our data and property to quickly respond for previous shock and we chose VAR(1) for modelling our data set.

At the first stage of the multivariate model building, we chose the optimal univariate GARCH models for every residuals data series based on the information criteria Accordingly, VAR(1)-EGARCH(1,1) model with skewed t-Student distribution was the best fit for S&P500 and exchange rates, VAR(1)-EGARCH(1,1) model with skewed GED distribution for FTSE100 and VAR(1)-GARCH(1,1) model with t-Student distribution for fine wines market.

We include these univariate GARCH model specifications in DCC-GARCH modelling. Based on a comparison of the log-likelihood values and information criteria across alternative lags and distribution specification we imply that the DCC(1,1) model with multivariate t-Student distribution as it was the best choice.

For the DCC-GARCH model, the estimated coefficients and are each positive and statistically significant respectively at level 1% (Table 1). Parameters sum to a value less than 1, indicating that the conditional correlations are dynamic and imply a persistence correlation. The volatility of recent returns has influence on the dynamic relationship between stock markets and fine wine market due to the value of the coefficient a_1. Value of the coefficient b_1 is slightly less than 1, meaning that the dynamic linkages between analysing markets can be long-term.

Table 1. Results of VAR(1)-DCC(1,1) estimation (rounded to at most four decimal places)

		Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
FTSE100	ω	–0.3800	0.0039	–97.246	<0.0001
	α_1	–0.1395	0.0157	–8.8591	<0.0001
	β_1	0.9608	0.0035	277.083	<0.0001
	γ_1	0.1787	0.0242	7.3863	<0.0001
	skewness	0.8963	0.0251	35.754	<0.0001
	shape	1.4998	0.0641	23.402	<0.0001
S&P500	ω	–0.4293	0.0043	–100.61	<0.0001
	α_1	–0.2182	0.0178	12.241	<0.0001
	β_1	0.9557	0.0007	1399.18	<0.0001
	γ_1	0.1747	0.0087	20.1778	<0.0001
	skewness	0.8713	0.0242	35.959	<0.0001
	shape	5.8312	0.6767	8.6170	<0.0001
LXV50	ω	<0.0001	<0.0001	0.2976	0.7666
	α_1	0.0468	0.0076	6.16485	<0.0001
	β_1	0.9487	0.0066	144.649	<0.0001
	df	6.0806	0.6829	8.9047	<0.0001
GBP/USD	ω	–0.0776	0.0012	–65.282	<0.0001
	α_1	–0.0262	0.0091	–2.8976	0.0004
	β_1	0.9926	<0.0001	23589.44	<0.0001
	γ_1	0.0639	0.0022	28.4774	<0.0001
	skewness	0.9491	0.0278	34.091	<0.0001
	shape	7.1033	1.2401	5.7280	<0.0001
DCC	a_1	0.0095	0.0020	4.8054	<0.0001
	b_1	0.9843	0.0040	247.56	<0.0001
	mshape	8.5622	0.6658	12.861	<0.0001

The dynamic conditional correlations between FTSE100 and S&P500 over the sample period are consistently positive (Fig. 1). We can assume that there is positive relationship between global and European stock markets. Dynamic relationships between British currency and FTSE100 index are different (Fig. 2). It can be divided into five periods, namely from 2010 to 2013, 2013-2H2015, 2H2015-2H2016, 2H2016-2017, 2017-end of 2018. In first and fourth periods, correlation is positive and react asymmetric on index movements. When FTSE100 index is declining, dynamic condition correlation raised and in opposite case – correlation fall, when index is growing. In second and fifth periods, correlation is negative or there is no correlation. The correlation pattern is similar to FTSE100 index movements but has larger amplitude of fluctuations. Third period covers one year gap before Brexit referendum. It is easy to see that correlation is negative, although in the beginning of 2016 correlation approaches 0 in response to decline of FTSE100 index. The average correlation coefficient is close to 0 and we might assume that British currency act as a neutral asset against FTSE100 index on average.

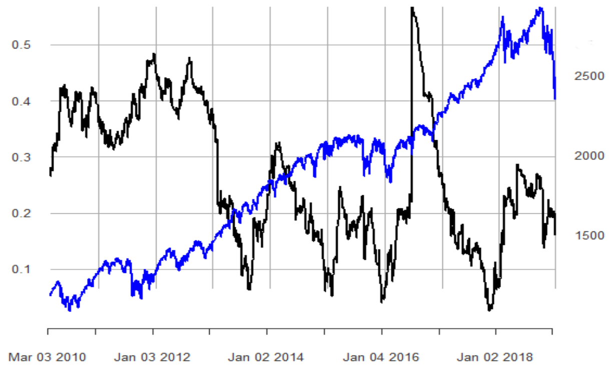


Fig. 1. DCC conditional correlations S&P500-GBP/USD (black); S&P500 index (blue)

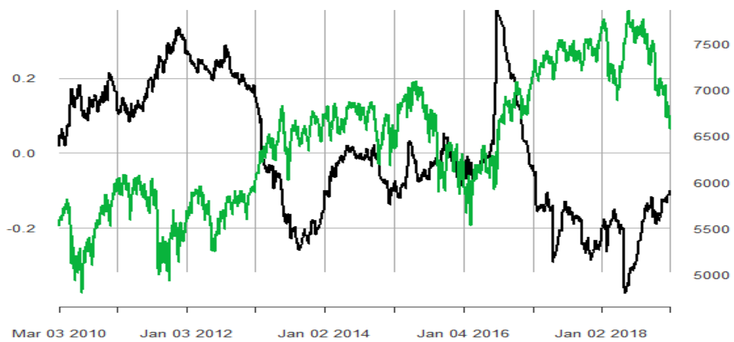


Fig. 2. DCC conditional correlations FTSE100-British currency (black); FTSE100 (green)

Correlations between fine wines and stock markets showed that there are periods when wines are negatively correlated or uncorrelated with FTSE100 and S&P500. From 2010 to 2013 fine wine market was mostly positively correlated with global stock market (Fig. 3), even though in period 2015-1H2018 correlation was negative. It is easy to capture that correlation movements are linked to S&P500 index. When stock index went down, the correlation changed from positive to negative or zero with a slightly delay. The average correlation coefficient is negative (-0.02), therefore we can assume that fine wine can be a hedge or act as safe haven asset against S&P500 index movements.

The correlation of fine wine and FTSE100 shows more periods when there is positive linkage between LVX50 index and European stock market (Fig. 4). It can be explained by the fact that both indices are quoted in pounds sterling and both listed on exchange located in London. The correlation movements are similar to FTSE100 index fluctuations as in the case of S&P500 index. When there is a drop on European stock market, the correlation changed from positive to negative or zero. Although, the average correlation coefficient is positive (0.02), fine wine cannot be a hedge or a safe haven asset. We might assume that it can act as a neutral asset against FTSE100 index movements.

Correlation of LVX50 and Pound sterling to American dollar exchange rate is mostly negative (average correlation coefficient is -0.09). Conditional correlations vary over time and are positive in periods of British currency appreciation and negative when GBP/USD exchange rate is declining (Fig. 5). We can assume that fine wine can be a hedge or act as safe haven against depreciation of the British currency.

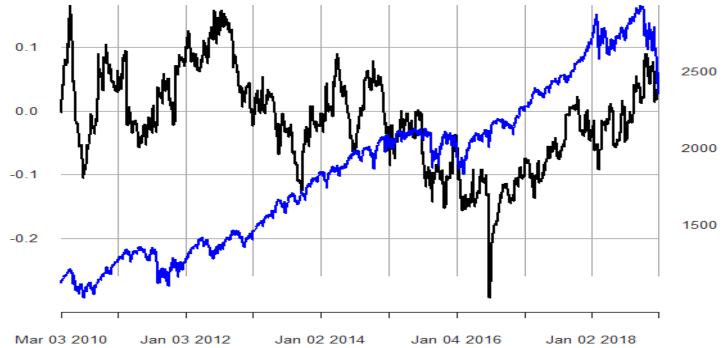


Fig. 3. DCC Conditional correlations S&P500-LVX50 (black); S&P500 index (blue)

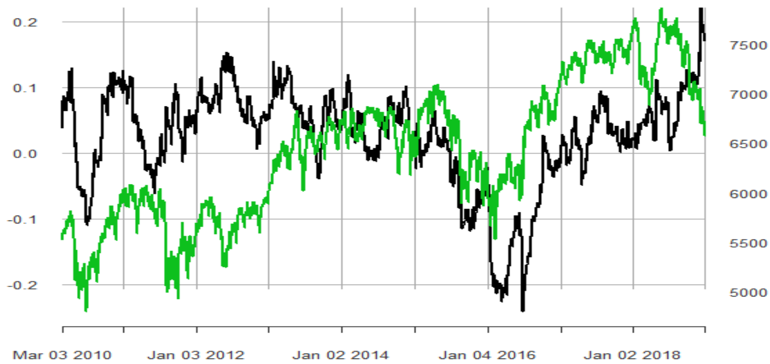


Fig. 4. DCC conditional correlations FTSE100-LVX50 (black); FTSE100 index (green)

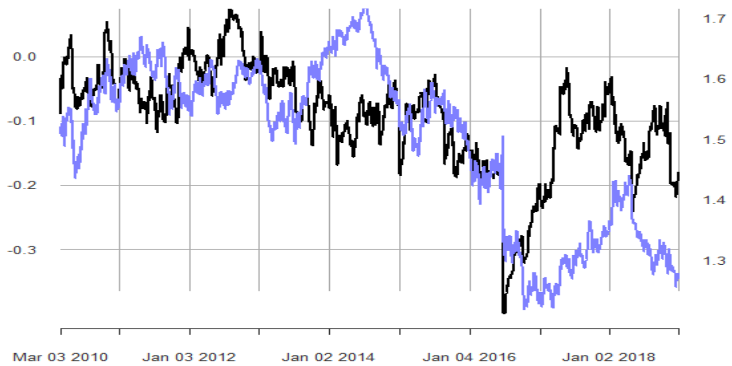


Fig. 5. DCC conditional correlations British currency-LVX50 (black); GBP/USD exchange rate (light blue)

5. Conclusions

This article examines the dynamics of the relation between daily fine wine prices, two stock indices and GBP/USD exchange rate using the dynamic conditional correlation approach.

First we used a VAR(1)-GARCH(1,1) to model the different error distribution margins. The adoption of this filtering method is motivated by the stylized facts of our data including serial dependence and volatility clustering. Estimated univariate GARCH models were used to build multivariate GARCH model, namely DCC model. Parameters of dynamic correlation were statistically significant in all cases indicating the importance of time varying co-movements. Results of our study reveal the long-term time-varying links in volatility between the global and European markets. We found evidence of property of fine wine to be hedge to global market and British currency. Moreover, fine wine can act as safe haven asset against S&P500 index and GBP/USD exchange rate declines. In the case of European stock market, both British currency and fine wine cannot be hedge and they can have only status of neutral assets which is opposite to Bouri and Roubaud (2016).

Our studies are an important contribution to the study of wine hedging properties. As investors tend to diversify their investment across different assets, results of our analysis would be crucial input for investors in portfolio diversification and hedging their stock positions in traditional financial assets by investing in fine wines.

References

Baur, D.G., & Lucey, B.M. (2010). Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold. *The Financial Review* 45(2010), 217–229.

Baur, D.G. (2012). Asymmetric volatility in the gold market. *The Journal of Alternative Investments* 14(4), 26–38.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31, 307–327.

Bollerslev, T. (1990). Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *Review of Economics and Statistics* 72, 498–505.

Bouri, E. (2014). Beyond the negative relation between return and conditional volatility in the wine market. Is fine wine particularly luscious for investors? *International Journal of Wine Business Research* 26(4), 279–294.

Bouri, E., Gupta, R., Wong, W.-K., & Zhu, Z. (2018). Is wine a good choice for investment? *Pacific-Basin Finance Journal* 51 (2018), 171–183.

Bouri, E., & Roubaud, D. (2016). Fine Wines and Stocks from the Perspective of UK Investors: Hedge or Safe Haven. *Journal of Wine Economics* 11(2), 233–248.

Cardebat, J.-M.,& Jiao L. (2017). The long-term financial drivers of fine wine prices: The role of emerging markets. *The Quarterly Review of Economics and Finance* 67, 347–361.

Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics* 20(3), 339–350.

Glosten, L., Jagannathan, R., & Runkle, D. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance* 48(5), 1779–1801.

Kilian, L., & Lütkepohl, H. (2017). Structural Vector Autoregressive Analysis. *Cambridge University Press*.

Le Fur, E., Ameur, H.B., & Faye, B. (2016). Time-Varying Risk Premiums in the Framework of Wine Investment. *Journal of Wine Economics* 11(3), 355–378.

Masset, P., & Henderson, C. (2010). Wine as an alternative asset class. *Journal of Wine Economics* 5(1), 87–118.

Masset, P., & Weisskopf, J.-P. (2010). Raise Your Glass: Wine Investment and the Financial Crisis. *American Association of Wine Economists. Working Paper* 57.

Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* 59(2), 347–370.

Analysis of the demand for competencies on the Polish labour market in the context of Industry 4.0

Paweł Lula[1], Sylwia Wiśniewska[2], Katarzyna Wójcik[3]

Abstract

The main purpose of the article is to analyse the demand for employee competencies on the labour market in Poland in the context of Industry 4.0. Due to the fact that employers search for candidates who possess a certain set of competencies, the aim of this type of research should be focused rather on the need for a certain sets of related competencies than on competencies considered independently. Taking thesis formulated above as a starting point, authors uses the term "competency schema" understood as a set of interrelated competencies together with information defining the significance of each competency and the nature and strength of links between them. The results of the research allowed us to identify the main competency schemes on the Polish labour market, to analyse existing competency schemas in spatial terms (including the division of Poland into voivodships), to analyse existing competency schemas due to the most popular job positions and to analyse identified competency schemas in the light of the assumptions of the Industry 4.0 concept. The research was based on the analysis of job offers regarding the Polish labour market. Identification of employers' expectations towards candidates' competencies was carried out by means of exploratory text analysis, while network analysis was used to study competency schemas. All calculations were carried out using programmes written in the R language.

Keywords: *competencies, competency schemas, labour market*
JEL Classification: *J24, C81*

1. Introduction

Today, Industry 4.0 brings various challenges like technological, strategic, business and people-related ones. In the conditions of Industry 4.0, human resources become one of the main assets of a company and significantly determine its competitive advantage on the market. Moreover, the dynamics of changes on the labour market is increasingly faster, which also means a quick outdating of the existing competencies and a growing demand for new ones. Therefore, skilful determination of the expectations towards employees' competencies and the hiring competent employees appoints the direction of organisational development. In addition, it is particularly important for the enterprise competitiveness to monitor the demand for competencies, including identification of trends regarding changes of the demand for labour. Thus, it becomes important to create possibilities of identification of the employer demand for competencies and the supply of competencies.

[1] Corresponding author: Cracow University of Economics, Department of Computational Systems, 27 Rakowicka St., 31-510 Cracow, Poland, pawel.lula@uek.krakow.pl.

[2] Cracow University of Economics, Department of Human Resources Management, 27 Rakowicka St., 31-510 Cracow, Poland, sylwia.wisniewska@uek.krakow.pl.

[3] Cracow University of Economics, Department of Computational Systems, 27 Rakowicka St., 31-510 Cracow, Poland, katarzyna.wojcik@uek.krakow.pl.

Due to the fact that employers search for candidates who possess a certain set of competencies, the aim of this type of research should be focused rather on the need for certain sets of related competencies than on competencies considered independently. On the basis of this assumption authors use the term “competency schema” which can be defined as a set of interrelated competencies together with information indicating the significance of each competence and the nature and strength of relations between them. Competency schemas can be represented as weighted graphs, in which nodes represent competencies, edges describe the relations between them, weights assigned to nodes indicate the importance of individual competencies in the given schema, and weights assigned to edges inform about the strength of relationships between competencies. The main purpose of the article is to analyse the demand for employee competencies on the labour market in Poland in the context of Industry 4.0. The research was based on the analysis of job offers regarding the Polish labour market. Identification of employers’ expectations towards candidates’ competencies was carried out by means of exploratory text analysis, while network analysis was used to study competency schemas. All calculations were carried out using programmes written in the R language.

2. The characteristics of Industry 4.0

One of the most novel ideas that has been applied to companies in recent years is “Industry 4.0”. It is a wide term that implies a drastic change in the way companies operate. Industry 4.0 promotes, among other things, interoperability, autonomous decision-making, flexibility, agility, efficiency and cost reductions (Pérez, 2016). The term “Industry 4.0” was officially presented at the 2012 Hannover Fairs in Germany as one of ten “Future Projects” that form Germany’s High-Tech Strategy 2020. Originally it was meant to describe technological changes in manufacturing and to set out priorities of a consistent policy framework to preserve the global competitiveness of the German industry. The title 4.0 indicates that Industry 4.0 is considered as the fourth industrial revolution, a logical continuation of the previous three ones (Kusmin, 2018).

Hermann et al (2015) defines Industry 4.0 as a collective term for technologies of value chain organizations. He categorises its components into Internet of Things, Cyber Physical Systems, Internet of Services and Smart Factory. The main purpose of Industry 4.0 is to achieve improvement in terms of automation and operational efficiency, as well as effectiveness (Ślusarczyk, 2018). Industry 4.0 is a result not only of the development of new technologies but it is also a result of new entrepreneurial mindset. Basically, the Industry 4.0 rises from Industrial policies’ of the well-developed countries that have introduced innovations, entrepreneurship and human capital as the most important elements of contemporary business culture (Sterev, 2017).

The main characteristics of Industry 4.0 that are very important to contemporary organization are (Lee et al, 2014; Schuh et al., 2015; Pérez, 2016; Sterev, 2017):

- vertical integration of smart production systems,
- horizontal integration of global value chain networks,
- cross-disciplinary through-engineering value chain,

- acceleration through exponential technologies as artificial intelligence,
- virtualization which means that companies are able to monitor their physical processes,
- interoperability which means that all internal and external systems used in company are connected,
- autonomation which means that "Industry 4.0" technologies and concepts are enabling machines and algorithms of future companies to make decisions and perform learning-activities autonomously,
- real-time availability which means that for organisational tasks it is necessary that data is collected and analysed in a real time,
- flexibility which means that due to new and more complex demand requirements, processes such as products development, products production processes or decision making procedures need to be performed faster,
- service orientation which means that the companies' Web services can be utilised by other participants,
- energy efficiency which means that climate change and scarcity of resources are megatrends that must be taken into consideration by the future industry players.

3. Competencies and competency schemas in the theoretical approach

The subject literature does not agree as to what the professional competencies of employees are. The multiplicity of approaches results mainly from the fact that this issue is analysed by specialists representing different scientific disciplines: management, psychology and sociology. The issue of competencies appeared in the subject literature at the turn of the 1960s and 1970s with the publication of works of two management psychologists: White (1959) and McClelland (1973) (Oczkowska et al, 2017). White (1959) depicted the human trait he called competence. McClelland (1973) claimed that although intelligence has an undoubted impact on human behaviour, the characteristics of a person (such as motivation or self-perception) that can be observed during life and professional situations, determine whether the behaviour will be effective or ineffective better than intelligence itself. McClelland (1973) pointed out that neither psychological tests nor school grades and testimonies may predict whether a person will succeed at work. It prompted McClelland (1973) to look for other methods of predicting work results. The identification of ways of thinking and behaviours of people who achieve a high level of effects at work has become one of them.

The definitions of competencies presented in the subject literature are not homogeneous, they are interpreted differently. In the opinion of Boyatzis (1982) competencies are the capacity of a given person to display behaviours compliant with the requirements of the job position specified by the organisational environment parameters, which, in turn, yields the desired results. According to Levy-Leboyer (1996) competencies are related to practical actions in specific situations and have the ability to being developed in the work process. As a dynamic structure they evolve under the influence of changes in the economy and in human life. Therefore, in the opinion of Bengtsson (1996), competencies make people capable of changing professions or raise their level of understanding the technology. It is worth noting that the concept of competencies is a multi-dimensional in its nature and much wider than the concept of qualifications itself. This is emphasised by Pawlak (2003), according to whom, competencies mean both confirmed by documents ability to work, as well as the abilities and skills that the candidate can confirm by performing the assigned tasks correctly.

The authors formulate the thesis that the analysis of possessed or required competencies shouldn't be carried out separately for each competence, but should always involve a certain set of them taken together. In the given set of competencies, the importance of each of them can be diverse, and this aspect can be expressed in the model by different weights assigned to them. Also, the character and the strength of relations between consecutive competencies may be varied. Taking thesis formulated above as a starting point, the authors uses the term "competency schema" understood as a set of interrelated competencies together with information defining their significance and the nature and strength of links between them.

It is worth pointing out the difference between the concept of "competency schema" proposed here and the term "competency profile" often used in the subject literature. The competency profile is a list of competencies along with the level of their fulfilment (Whiddett et al., 2007; Juchnowicz et al., 2014). The meaning of the concepts of the competency schema and the competency profile is not equal. The competency schema defines not only competencies and the level of their fulfilment, but also defines the connections between competencies describing their character (direction) and their strength.

4. Graph-based approach to competency schema definition

We will consider a set of competencies:

$$\boldsymbol{C} = \{C_1, C_2, ..., C_M\}$$

and a set of job offers:

$$\boldsymbol{O} = \{O_1, O_2, ..., O_N\}$$

The analysis of offers enables to define the *offer-competency matrix* $\boldsymbol{M}$:

$$\boldsymbol{M} = [m_{ij}]$$

where $m_{ij} = 1$ if the j-th competency is mentioned in the i-th offer, and $m_{ij} = 0$ if the information about the j-th competency does not appear in the i-th offer.

We assume that two competencies C_i and C_j are connected if they appear in the same offer O_k. Connections between competencies are represented by the *competency co-occurrence graph* which has weighted and undirected character and is represented by the adjacency matrix $\boldsymbol{G}$:

$$\boldsymbol{G} = [g_{ij}]$$

where g_{ij} element informs how many times the competency is mentioned in the given offer set $\boldsymbol{O}$, and g_{ij} (where $i \neq j$) shows how many times competencies C_i and C_j appeared together in the same offer.

Some relations between competencies can be very infrequent and have insignificant character in a given job offer set. Assuming that irrelevant connections are represented by weights assigned to edges smaller than a given threshold (90th percentile calculated for weights assigned to edges), the *competency interrelationship graph* (*CIR graph*) can be defined by an adjacency matrix $\boldsymbol{R}$:

$$\boldsymbol{R} = [r_{ij}]$$

where $r_{ii} = g_{ii}$, $r_{ij} = g_{ij}$ if $g_{ij} \geq t$, and $r_{ij} = 0$ if $g_{ij} < t$.

The *CIR* graph represented by the adjacency matrix $\boldsymbol{R}$ shows the importance of every competency (diagonal elements of the adjacency matrix) and the importance of connections between them (off-diagonal elements of the adjacency matrix). In the general case, the *CIR* graph can be disconnected. It means that two competencies which are not connected by one path may exist simultaneously.

The *CIR* graph constitutes the basis for the *competency schema* definition. The competency schema can be defined as a connected component (connected sub-graph) of the competency interrelationship graph.

5. The analysis of the demand for competencies on the Polish labour market

The identification of competency schemas describing the demand for competencies on the Polish labour market is the main goal of the empirical part of the research.

To reach the goal stated above the authors used a set of job offers retrieved from the https://www.pracuj.pl/ web site. The set which contains 8000 job offers was prepared using web scraping technique. Job offers were retrieved in January 2019. Only offers active at the time of scrapping procedure were acquired. Later on, all offers prepared in languages different than Polish were omitted. The remaining set of offers was comprised of 6667 documents.

Using the exploratory text analysis technique presented in (Lula et al., 2018) the process of identification of competencies in every document was performed. The analysis was focused of identification of four groups of competencies: individual, social, managerial and professional. Cumulatively 61 various competencies were considered.

The above step allowed us to build the *offer-competency matrix* $\boldsymbol{M}$. Later on, the algorithm presented in the previous section was used in order to build *competency co-occurrence graph*. During consecutive step of the analysis nodes representing the most frequent competencies were found and used to build the *competency interrelationship graph* (*CIR graph*) presented in Fig. 1.

To facilitate the analysis of the visualization of the *CIR* graph, the importance of competencies was represented by the size of nodes, and the importance of connections was expressed by the width of edges. Competency schemas can be obtained by erasing edges representing insignificant connections (with weights smaller than a given threshold) and analysing the connectivity of the graph obtained as the result. For an exemplary threshold used for the calculation, insignificant edges were drawn in orange.

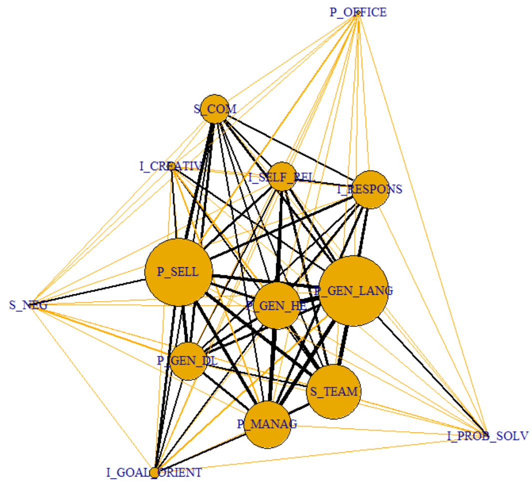


Fig. 1. General competency interrelation graph for the Polish labour market

The results show that the set of the most important competencies is composed of *P_SELL* (competencies related to selling and marketing), *P_GEN_LANG* (competencies relating to knowledge of foreign languages), *P_GEN_HE* (possession of higher education), *S_TEAM* (teamworking), *P_MANAG* (managerial competencies), *P_GEN_DL* (possession of a driving licence), *I_RESPONS* (responsibility), *S_COM* (communication skills), *I_SELF_REL* (self-reliance), *I_CREATIV* (creativity) and *I_GOAL_ORIENT* (goal orientation). All these competencies are connected very strongly. It is the most important competency schema describing the demand for competencies on the Polish labour market analysed as a whole. It is worth underlining that these competencies have rather general character and are not strongly connected with the Industry 4.0 concept.

Furthermore, the similar analysis of job offers related to *IT* sector was performed. This sector was chosen as a representation of an area characteristic for Industry 4.0 idea. In this study a set of 249 job offers was used. Using the same procedure as described above, the competency interrelation graph for *IT* sector was built. It was presented in Fig. 2.

The results show that for the *IT* sector the set of the most important competencies includes *P_GEN_LANG* (competencies relating to knowledge of foreign languages), *S_TEAM* (teamworking), *SEC_ICT* (competencies related to ICT sector), *P_SELL* (competencies related to selling and marketing), *I_RESPONS* (responsibility), *P_GEN_HE* (possession of higher education), *P_GEN_DL* (possession of a driving licence), *S_COM* (communication skills) and *I_SELF_REL* (self-reliance).

On the contrary to results obtained for the whole set of job offers, for *IT* sector professional competencies related to *ICT* area were considered as very important.

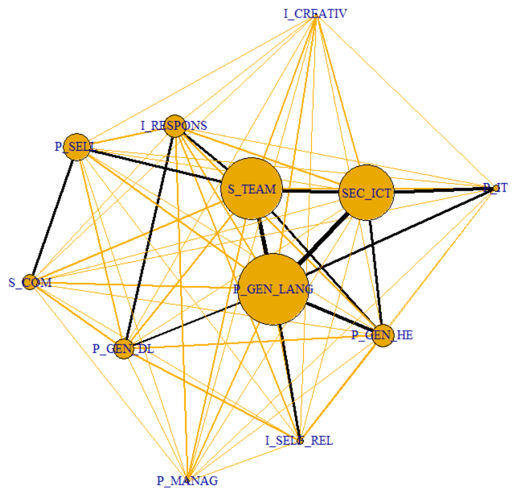


Fig. 2. Competency interrelation graph for *IT* sector of the Polish labour market

Conclusions

The results show that competency schemas constitute a useful tool for labour market analysis. Here they were used for description of the demand for competencies. But their applications are not limited to this task and the same concept may be used for the description of the competency supply. By the comparison of these two perspectives, the competency gap (and its representation in the form of competency schemas) can be estimated.

It is worth concluding that the scope of analysis has a huge impact on the results. Performing the analysis of the whole set of job offers, the results show that individual, social and managerial competencies play the crucial role on the labour market. Competency schemas identified during the study do not include professional competencies. But this fact should not be treated as a justification of the opinion that professional competencies play secondary role on the labour market and in the education process. The importance of professional competencies is clearly apparent when the scope of analysis is limited to one sector. Then professional competencies have crucial character.

The authors believe that the concept of competency schemas introduced here allows us to describe and analyse the demand and the supply for competencies better than the concept of competency profiles. Schemas do not only inform about all aspects of competencies which are described by profiles, but also show their co-occurence, which can play the crucial role in preparation of educational processes.

The authors are going to develop ideas presented here in their future works.

Acknowledgements

The Project has been financed by the Ministry of Science and Higher Education within "Regional Initiative of Excellence" Programme for 2019–2022. Project no.: 021/RID/2018/19. Total financing: 11 897 131,40 PLN.

References

Bengtsson, J. (1996). Rynki pracy przyszłości: wyzwania polityki edukacyjnej. *Nauka i Szkolnictwo Wyższe*, 7, 24–45.

Boyatzis, R.E. (1982). *The competent manager: a model for effective performance*. New York: John Wiley & Sons.

Hermann, M., Pentek, T., & Otto, B. (2015). Design principles for Industrie 4.0 scenarios: A literature review. *Working Paper*, 1, 1–16.

Juchnowicz, M. (ed.) (2014). *Zarządzanie kapitałem ludzkim. Procesy – narzędzia – aplikacje.* Warszawa: PWE.

Kusmin, K-L. (2018). Industry 4.0. http://www.tlu.ee/~pnormak/ISA/Analytical%20articles/2-Industry%204.0%20-%20Kusmin.pdf.

Lee, J., Kao, H-A., & Yang, S. (2014). Service innovation and smart analytics for Industry 4.0 and big data environment. *Procedia CIRP*, 16, 3–8.

Levy-Leboyer, C. (1996). *La Gestion des competences*. Paris: Les Editions d'Organisation.

Lula, P., Oczkowska, R., Wiśniewska, S., & Wójcik, K. (2018), Ontology-Based System for Automatic Analysis of Job Offers. *Information Technology for Practice*, 2018, 205–212

McClelland, D.C. (1973). Testing for competence rather than for "intelligence". *American Psychologist*, 28(1), 1–14.

Oczkowska, R., Wiśniewska, S., Lula, P. (2017). Analysis of the competence gap among vocational school graduates in the area of smart specialization in Poland. *International Journal for Quality Research*, 11(4), 945–966.

Pawlak, Z. (2003). *Personalna funkcja firmy – procesy i procedury kadrowe*. Warszawa: Poltext.

Pérez, D., Alarcón, F., & Boza, A. (2016). Industry 4.0: A classification scheme. In: *International Joint Conference – CIO-ICIEOM-IIE-AIM (IJC 2016)*, 1–9.

Schuh, G., Gartzen, T., Rodenhauser, T., & Marks, A. (2015). Promoting work-based learning through Industry 4.0. *Procedia CIRP*, 32, 82–87.

Sterev, N. (2017). Marketing leadership: The Industry 4.0 need of next generation marketing. *Trakia Journal of Sciences*, 15(1), 99–103.

Ślusarczyk, B. (2018). Industry 4.0: are we ready? *Polish Journal of Management Studies*, 17(1), 232–248.

Whiddett, S., & Hollyforde, S. (2007). Competencies. London: CIPD.

White, R. (1959). Motivation reconsidered: The Concept of Competence. *Psychological Review*, 66(5), 279–333.

Longevity risk factors: the perspectives of selected European countries

Justyna Majewska[1], Grażyna Trzpiot [2]

Abstract

A significant and ongoing debate about the economic, financial and social implications of longevity was initiated a few years ago. . Both individuals and governments are increasingly concerned about the effects of aging, however, their concerns differ. Individuals are more concerned about increased longevity, because it affects their own financial and labor market plan, whereas governments are more concerned about old-age dependency as an aspect of population aging. In this paper, in the first part set of economic, financial and demographic variables are discussed in a context of their impact on longevity. In the second part, using principal component analysis, we identify risk factors of longevity by means of macroeconomic and financial data collected from the last decade for selected countries from Europe.

Keywords: *longevity, risk, PCA*

JEL Classification: *C130, C180*

1. Introduction

The changing and ageing structure of European population is driven primarily by two factors. Firstly, improvements in life expectancy mean that people are living longer and reaching older ages. Along with this factor, there is a decrease in fertility, people are having fewer children and are having children later than ever. The observed improvements in longevity and changing structure of the population bring both opportunities and challenges for the economy, services, and society at national and local levels (ONS UK, 2018).

Institutions and individuals have to face some specific challenges related with longevity risk, i.e. the risk that actual survival rates and life expectancy will exceed expectation or pricing assumptions, resulting in greater-than-anticipated retirement cash flow needs (NAIC, 2019). For individuals longevity risk is the risk of outliving one's assets resulting in a lower standard of living, reduced care or just return to employment. For those institutions providing covered individuals with guaranteed retirement income, this is the risk of underestimating survival rates resulting in increased liabilities to cover promised payments.

For a few years discussions and research have been intensively carried out on how to deal with the economic and financial implications of longevity– in both academic circles and the business press (e.g., Olshansky et al., 2009; Bloom et al., 2010; Cocco and Gomes, 2011; IFM, 2012). It is a problem of shrinking populations, reduced economic growth, negative effect on

[1] Corresponding author: University of Economics in Katowice / Department of Demography and Economic Statistics, 50 1 Maja St., 40-287 Katowice, justyna.majewska@ue.katowice.pl.

[2] University of Economics in Katowice / Department of Demography and Economic Statistics, 50 1 Maja St., 40-287 Katowice, grazyna.trzpiot@ue.katowice.pl.

pension and health care system. Understanding the nature of longevity risk is crucial to model and manage this risk.

The goal of this paper is to examine a set of reasonable macroeconomic and financial variables for determining factors of longevity risk. As a research method we use principal component analysis. Sources of longevity risk are discussed on the example of three European countries with significant differences in level of economic growth, life expectancy, pensions and welfare systems.

2. Economic and financial implications of longevity

The impact of aging on the economy and financial markets has been analysed widely in the literature. There are different implications for individuals, for households, for insurers, for local and central governments. Concerns about the economic impacts of aging fall into three main areas: potential worker shortages, excessive expenditure on health services and old-age care and shortfalls in pension funding. The impact of aging on financial stability occurs largely through changes in investment demand, financial market asset prices and returns, and liabilities among individuals and institutions.

An aging population means that there will be fewer working-age people in the economy. It leads to a supply shortage of qualified workers, making it more difficult for businesses to fill in-demand roles. An economy that cannot fill in-demand occupations faces adverse consequences, e.g., declining productivity, higher labor costs, delayed business expansion and reduced international competitiveness. In some instances, a supply shortage may push up wages, thereby causing wage inflation and creating a vicious cycle of price or wage spiral. Given that demand for health care rises with age, countries with aging populations must allocate more money and resources to their health care systems. With health care spending as a share of GDP already high in most advanced economies, it is difficult to increase spending while ensuring care improvements and other social needs not to deteriorate in the case of publicly funded or government-administered health care systems. Countries with large elderly populations depend on smaller pools of workers in which to collect taxes to pay for higher health costs, pension benefits and other publicly funded programs. The combination of lower tax revenue and higher spending commitments on health care, pension and other benefits is a major concern for advanced industrialized nations.

Empirical studies uncovered evidence that population aging has an important impact on financial markets because of its expected impact on saving rates and the demand for investment funds (e.g. Poterba, 2001; Poterba , 2004). The rising demand for safe assets by the elderly may lead to safe asset shortages and an overpricing of safe assets. These effects may be counterbalanced by defined-benefit funds with funding gaps in the current low interest-rate environment, which may invest in risky assets to enhance expected returns. Population age structure can affect stock market prices and the real returns of different classes of financial assets. Relations between long-term government bond yields and demographics are well documented (see Andrews and Bonnar, 2018).

3. Selected variables associated with longevity risk

Empirical investigation of relations between longevity phenomenon and selected macroeconomic and financial variables is made for selected European countries with different level of economic growth and life expectancy, i.e. for Germany, Spain and Poland. From longevity perspective, life expectancy (at birth and at aged 65, for both sexes) in Poland is shorter than in Germany and Spain, while life expectancy is the highest in Spain. Spain is expected to become the world's second oldest country by 2050, behind Japan. According to HDI index[3] Germany – since 2010 – has been in the group of five the most developed countries, Spain – in the second ten, and Poland – in the third ten the most developed countries in the world (UNDP, 2018).

The selection of variables was preceded by an analysis of literature in the field of research on determinants of macroeconomic and financial implications of ageing. The analysis broadly focuses on the issues: economic welfare, current standard of living, current possible increase of welfare resulting from labour market, population longevity, and financial markets and investments opportunities.

In the process of identification of risk factors the following variables are taken into consideration:

1. Demographic old-age dependency ratio[4] – traditionally seen as an indication of the level of support available to older persons (those aged 65 or over, i.e. age when they are generally economically inactive) by the working age population (those aged between 15 and 64; expressed per 100 persons of working age (15–64)).
2. Life expectancy at birth[5] – the mean number of years that a new-born child can expect to live if subjected throughout his life to the current mortality conditions (age specific probabilities of dying; expressed in years).
3. Life expectancy at age 65[6] – the mean number of years still to be lived by a man or a woman who has reached the age 65, if subjected throughout the rest of his or her life to the current mortality conditions (age-specific probabilities of dying; expressed in years).
4. Consumer Price Index (CPI)[7] – the change over time in the prices of consumer goods and services acquired, used or paid for by households (measured in an index, 2015 base year).

[3] The Human Development Index (HDI) is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and have a decent standard of living. The HDI is the geometric mean of normalized indices for each of these three dimensions (http://hdr.undp.org/en/, access: 12.12.2018).

[4] In 2016 this ratio (per 100 person of working age 15–64) amounted to 32.4 (Germany), 28.7 (Spain), 24.2 (Poland). Projected ratio for 2050: 51.2 (Germany), 62.1 (Spain), 54.6 (Poland).

[5] In 2016 LE at birth (male and female): 81.0 (Germany), 83.5 (Spain), 78.0 (Poland).

[6] In 2016 LE at aged 65 (male and female): 19.8 (Germany), 21.6 (Spain), 18.5 (Poland).

[7] In 2016 CPI annual growth rate %: 0.48 (Germany), -0.20 (Spain), -0.58 (Poland). In general, CPI decreased over the last 7 years. CPI projected for 2020: 2.16 (Germany), 1.74 (Spain), 2.95 (Poland).

5. Real GDP per capita[8] – the ratio of real GDP to the average population of a specific year; a measure of economic activity used as a proxy for the development in a country's material living standards (a limited measure of economic welfare; per capita, in current prices).
6. Unemployment rate[9] – represents unemployed persons as a percentage of the labour force (the total number of people employed and unemployed) [% of active population].
7. Real effective exchange rates (REER)[10] – aims to assess a country's price or cost competitiveness relative to its principal competitors in international markets; changes in cost and price competitiveness depend not only on exchange rate movements but also on cost and price trends (indices).
8. Gross saving[11] – measures the portion of gross national disposable income that is not used for final consumption expenditure; gross national saving is the sum of the gross savings of the various institutional sectors (current prices).
9. Long-term government bond yields[12] – refer to central government bond yields on the secondary market, gross of tax, with residual maturity of around 10 years; the bond or the bonds of the basket have to be replaced regularly to avoid any maturity drift (%).
10. Long-term care (health) expenditures[13] – expenditures on a range of medical and personal care services that are consumed with the primary goal of alleviating pain and suffering and reducing or managing the deterioration in health status in patients with a degree of long-term dependency (share of current expenditures on health).
11. Currency exchange rates[14]: EUR/USD, EUR/PLN.
12. Stock market a main index[15]: DAX in Germany, IBEX35 in Spain, WIG20 in Poland.
13. Real Estate Funds and Equity/Dividend Funds: Unilmmo Deutchland and Allianz Vermögensbildung Deutschland (Germany), Seguffondo Inversion and Bankia Dividendo España FI (Spain), PZU UFK Investor Nieruchomości i Budownictwa and Investor FIO Subfundusz Akcji Spółek Dywidendowych (Poland).

Economic and demographic variables are derived from the Eurostat database (variables 1–9) and the OECD database (variable 10), stock quotes – from stock exchange (Frankfurt, Madrid, Warsaw) and other financial database (the Yahoo Finance) (variables 12 and 13). Time series were obtained for the time period 2010–2016. As the sample is short, some data were converted to monthly frequency (and then all variables were expressed as indices using a base

[8] In 2016 GDP (milions USD): 4.1M (Germany), 1.7M (Spain), 1.1M (Poland). Projected for 2030: 4.3M (Germany), 1.9M (Spain), 1.4M (Poland).

[9] In 2016 unemployment rate (% of labour force): 4.1 (Germany), 19.6 (Spain), 6.2 (Poland).

[10] In 2016 REER index (2010=100): 100.88 (Germany), 99.2 (Spain), 89.79 (Poland).

[11] In 2016 gross saving (2010=100): 116.97 (Germany), 101.64 (Spain), 113.85 (Poland).

[12] In 2016 long-term gov. bond yields (2010=100): 3.28 (Germany), 32.78 (Spain), 52.51 (Poland).

[13] In 2016 long-term health care exp. (2010=100): 113.69 (Germany), 97.18 (Spain), 95.62 (Poland).

[14] Rate of return 2010–2016: EUR/USD –26.34% , EUR/PLN +7.63%.

[15] Rate of return 2010–2016: WIG20 –18.45%; DAX +92.72%; IBEX35 -21.85%.

year of 2010), with maintaining the strength and direction of correlation between variables. The period does not cover years from the financial crisis of 2008–2009 to avoid unusual observations from financial market.

Relations between the above-mentioned variables and longevity are analysed in empirical studies. Some relations are clear, while others are still a subject of debate (in particular, the impact of longevity on inflation is unclear). Due to the complexity of these relations and their multidimensionality, a few confirmed consequences of longevity are worth mentioning (e.g. Bloom et al., 2010; Mandel et al., 2017; Rachel and Smith, 2015; Maestas at al., 2016; Acemoglu and Restrepo, 2017): reducing investment return, reducing public saving, reducing growth rates, reducing real interest rates, affecting labour supply and returns, reallocation of saving from riskier to safe assets may lead to potential mispricing of risk, running down assets may result in negative wealth effects.

Based on the results of Majewska and Trzpiot (2016), the above mentioned variables could be grouped into five clusters: standard of living risk, elderly needs risk, financial risk, longevity risk and long-term investment risk. However, it should be taken into account that the time period in their research covered years of financial crisis 2008–2009.

4. Principal component analysis of variables associated with longevity risk

The rotated Principal Component Analysis, an adaptive data analysis technique, was used to specify risk factors of longevity. Relations between some of considered variables suggest the existence of new variables – the principal components – that are the linear combinations of the original demographic, macroeconomic and financial variables. All calculations were made in R software environment.

As a result of PCA three components for each country separately were selected and treated as F_i risk factors. Risk factor loads are presented in Tables 1–3. Factors were simplified by orthogonal (varimax), which minimised the number of variables with high loadings on each factor (orthogonal rotation to transform the extracted factors into uncorrelated, independent factors to increase the interpretability of the factors). The correlations between the factors were explained by factor loadings, values greater than or equal to 0.4 were used to indicate significant correlations between the component and the variables. The components with eigenvalues (sum of the squared factor loadings) greater than or equal to 1 were retained for analysis (components with variances less than one produce negligible information than one of the original variables and, hence, are not worth retains).

For Germany (Table 1) the first principal component explains 59% of the variation, while all components – 83%. The first component is identified as the wealth risk because of the high positive factor loadings on GDP, gross savings, long-term care expenditures combined with a high negative weighting on long-term government bond yields and unemployment rate. All these variables are associated with standard of living risk and elderly needs risk. The second component was high loadings of variables that reflect longevity. Advancing age due to increased life expectancy itself is a risk factor. The last component explains 10% of total variance and was loaded only by REER and real estate fund and it would associate with financial market risk.

Table 1. Risk factor loads of principal components: Germany

Variable	*F*1	*F*2	*F*3
Old Age Dependence Ratio	0.93		
GDP	0.92		
Gross Saving	0.93		
CPI	0.80		
Long-term Care Expenditures	0.95		
Long-term Government Bond Yields	–0.89		
REER			0.92
Unemployment Rate	–0.90		
Dividend Fund	0.75		
Real Estate Fund			0.63
LE65		0.88	
LE birth		0.97	
DAX	0.75		
EUR/PLN	0.74		
EUR/USD	–0.85		
Cumulative Var	*0.59*	*0.73*	*0.83*

Table 2. Risk factor loads of principal components: Spain

Variable	*F*1	*F*2	*F*3
Old Age Dependence Ratio	0.93		
GDP		–0.94	
Gross Saving		–0.55	
CPI	0.83		
Long-term Care Expenditures		–0.66	
Long-term Government Bond Yields	–0.78		
REER	–0.61		
Unemployment Rate		0.93	
Dividend Fund	0.69		
Real Estate Fund	–0.91		

Variable	*F*1	*F*2	*F*3
LE65	0.85		
LE birth	0.94		
IBEX35			0.96
EUR/PLN			–0.65
EUR/USD	–	–	–
Cumulative Var	*0.44*	*0.68*	*0.82*

For Spain (Table 2) the first component was the highest loadings with variables that reflect elderly needs and longevity. Thus, this component was identified as long-term standard of living. It can be supported by noting that the factor loadings associated with long-term care expenditures, REER and real estate fund are negative. The second component proved to be a strong indicator of longevity risk related with GDP, gross savings, unemployment rate and long-tern expenditures. This component was clustered with standard of living and long-term investments risk factors. The last component explains 14% of total variance and has been positively loaded with IBEX35 returns and negatively – with EUR to PLN exchange rate. It would be associated with financial market risk.

Table 3. Risk factor loads of principal components: Poland

Variable	*F*1	*F*2	*F*3
Old Age Dependence Ratio	0.86		
GDP	0.89		
Gross Saving	0.77		
CPI	0.72		
Long-term Care Expenditures	–0.84		
Long-term Government Bond Yields	–0.91		
REER		0.80	
Unemployment Rate	–	–	–
Dividend Fund		0.92	
Real Estate Fund			–0.97
LE65	0.95		
LE birth	0.95		
WIG20		0.88	
EUR/PLN		–0.81	
EUR/USD	–	–	–
Cumulative Var	*0.54*	*0.77*	*0.87*

For Poland (Table 3) the first component was loaded with variables related with increasing life expectancy as well as economic well-being and explains 54% of total variance. Therefore, the component was identified as demo-economic risk. The second component was heavily loaded with dividend fund, WIG20 returns, REER and negatively with EUR to PLN exchange rate. It would be identified as financial market risk and explained 23% of total variance. The last component was loaded negatively only by a real estate fund and would be associated with individual wealth risk.

Conclusions

The current study focused on determining significant longevity risk factors through principal component analysis for three different European countries. The study performed PCA to reduce 14 inter-correlated demographic, economic and financial variables into groups of independent factors. For each country, within considered variables, 3 principal components with high explained variance (82%, 83% and 87% respectively for Spain, Germany and Poland) are sufficient to capture the nature of longevity risk.

The complex nature of longevity risk do not allow for constructing factors consisting of the same variables for each country. It is due to the fact that longevity increases at different rates across different countries. Therefore, impact of population aging differs across countries. This difference is linked to the state of health, age-structure profiles, lifestyle, progress in disease diagnosis and medical treatment, to mention a few. This is the reason why we need to have a real understanding of the causal factors underlying longevity, the ageing process, and the characteristics governing different populations. In our opinion, the risk factors will vary depending on the country selected for next the study.

Interestingly, some factors constitute a mix of demographic, financial and economic variables. It enables explanation of key differences between considered countries; however, some trends are noticeable. Countries are prepared for the challenges of population aging in varying degrees. Clustering of variables informs about diversification in both opportunities and challenges for the economy, services and society at national and local levels.

There is a need for further analysis made for various time periods, especially asset returns are sensitive to the start and end dates of the period covered by the analysis. Moreover, comparative analyses between countries with share similar demographic histories would be very important due to the growing popularity of coherent (multi-population) mortality (longevity) models, where a single population is modeled in reference to another coherent population. Little work has been done on the importance or selection of the reference group and there is no standard procedure on how to select populations for the reference population.

References

Acemoglu, D., & Restrepo, P. (2017). Secular Stagnation? The Effect of Aging on Economic Growth in the Age of Automation. *American Economic Review,* 107(5), 174–179.

Andrews, D., & Bonnar, S. (February, 2018). Population Aging, Implications for Asset Values, and Impact for Pension Plans: An International Study. Retrieved December 01, 2018, from https://www.cia-ica.ca/docs/default-source/2018/218014e.pdf.

Bloom, D.E., Canning, D., & Fink, G. (2010). Implications of population ageing for economic growth. *Oxford Review of Economic Policy,* 26(4), 583–612.

Bosworth, B., Bryant, R., & Burtless, G. (2004). The Impact of Aging on Financial Markets and the Economy: A Survey (July 1st, 2004). *The Brookings Institution*. Available at SSRN: https://ssrn.com/abstract=1147668.

Cocco, J.F., & Gomes, F. (2011). Longevity Risk, Retirement Savings and Financial Innovation (March 1st, 2011). *Netspar Discussion Paper 3*. Available at SSRN: https://ssrn.com/abstract=1888305

IFM The Financial Impact of Longevity Risk – 2012. (2012). Retrieved from https://www.aarp.org/livable-communities/learn/demographics/info-12–2012/financial-impact-of-longevity.html

Longevity risk. (2019). Retrieved January 2nd , 2019, from https://www.naic.org/cipr_topics/topic_longevity_risk.htm

Maestas, N., Mullen, K., & Powell, D. (2016). The Effect of Population Aging on Economic Growth, the Labor Force and Productivity. *Working Papers.* Rand Corporations.

Office for National Statistics UK. Living longer: How our population is changing and why it matters. (2018). Retrieved January 2, 2019, from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing/articles/livinglongerhowourpopulationischangingandwhyitmatters/2018-08-13

Olshansky, S.J., Carnes, B.A., & Mandell, M.S. (2009). Future trends in human longevity: Implications for investments, pensions and the global economy. *Pensions: An International Journal,* 14(3), 149–163.

Poterba, J.M. (2004). Impact of Population Aging on Financial Markets in Developed Countries. Paper presented to the Symposium on Global Demographic Change, Federal Reserve Bank of Kansas.

Poterba, J.M. (2001). Demographic Structure and Asset Returns, *Review of Economics and Statistics*, 83, 565–584.

Rachel, L., & Smith, T. (2015). Secular Drivers of the Global Real Interest Rate. *Bank of England Staff Working Paper 571*. J.P. Morgan Asset Management Multi-Asset Solutions.

Trzpiot, G., & Majewska, J. (2016). The Impact of Longevity on Long-term Investments Returns: Scenarios for Europe. Available at https://www.cass.city.ac.uk/__data/assets/pdf_file/0020/334082/L12-32-TRZPIOT-and-MAJEWSKA.pdf

UNDP (2018). Human Development Indices and Indicators 2018: Statistical update, UN, New York.

Use of mirror data in detecting irregularities in declared values of intra-Community trade by HS section

Iwona Markowicz[1], Paweł Baran[2]

Abstract

Official statistics on intra-Community trade in goods depend on data gathered within the Intrastat system. The database consists of declarations made by businesses dispatching goods to other EU member states or receiving goods from them. The database is collected by national statistical offices and then aggregated by Eurostat. Due to the specifics of data it is possible to assess its quality. Mirror data, i.e. data gathered by both sides of transaction, should be reported twice – by the country of the dispatcher and at the same time by the country of the receiver of traded goods. Thus, for every pair of countries the Intrastat system should contain the same values of trade in goods for the same period. Slight differences are permitted for various reasons, such as statistical thresholds or exchange rates. Generally, however, discrepancies in mirror data are due to errors in reporting or concealment of transactions. And it affects the quality of collected data.

The aim of the study was to analyse the quality of data on intra-Community trade by harmonised system (HS) sections in 2017. The data come from the Comext database provided by Eurostat. For each EU-28 country and each HS section, the quality of data was examined. For this purpose, author's own approach was used to determine aggregated indicators of data discrepancies with country-by-country aggregation. The research allowed us to identify those HS sections for which data quality is the lowest in all individual EU member states.

Keywords: *intra-Community trade, mirror data, HS section*
JEL Classification: *F14, C10, C82*

1. Introduction

Analysis of the activities of businesses must be based on data. Such data are often obtained through questionnaires prepared by national statistical offices (in the European Union, often following the model of Eurostat documents). This is the case with collecting data on the value and volume of intra-community trade of Polish companies. (Baran and Markowicz, 2018). The data is obtained on the basis of statistical declarations. Data collection for the Central Statistical Office is made by Intrastat Department of the Tax Administration Chamber in Szczecin, and the aggregate data are transmitted to Eurostat. It should be noted that these are statistical and not fiscal data. And the fact that there is virtually no sanction affects the timeliness and reliability of filling in forms. However, the Intrastat declaration has been the only source of information on intra-Community trade of Polish companies since 2004.

The problem of the quality of statistical data declared by stakeholders is not only a domain of foreign trade statistics. Similarly, sample surveys (DG1) conducted by the Central Statistical Office are currently the main source of information about revenues earned by small companies

[1] Corresponding author: University of Szczecin, Department of Econometrics and Statistics, 64 Mickiewicza St., 71-101 Szczecin, Poland, e-mail: Iwona.Markowicz@usz.edu.pl.

[2] University of Szczecin, Department of Econometrics and Statistics, 64 Mickiewicza St., 71-101 Szczecin, Poland, e-mail: Pawel.Baran@usz.edu.pl.

(Dehnel and Wawrowski, 2018). The questionnaire concerns the basic characteristics of the company. Likewise, in the study on the duration of companies, data of the Central Statistical Office from the REGON register is being used (Markowicz, 2014). These data are created on the basis of statistical declarations of entrepreneurs. Registers of applications (registration in the Labour Office) are also used as a basis for registered unemployment statistics (Bieszk-Stolorz and Markowicz, 2015).

As far as this type of research is concerned, an extremely important problem is to monitor the quality of the collected data. The way of collecting data in the Intrastat system in the form of mirror data makes it possible to compare the value (and also the quantity) of goods exported from one country and imported to another EU country. These values are declared in two countries by both parties to the transaction, so when referring to international trade data we mean data on particular trade flow as the value reported by the exporter and at the same time as the value reported by the importer – that is the very idea of mirror values (Ten Cate, 2014). Other bilateral data stored as mirror data include figures on direct foreign investment, foreign debt, and international migration. It is worth noting that the problem of discrepancies (often referred to as asymmetries) in mirror data is not new to economic researchers. Parniczky (1980) indicates that such research has been carried out at least since the 1920s, and Tsigas et al. (1992) date it back to the 1880s.

In (Ten Cate 2014) the discrepancies in reported bilateral statistical data are used to estimate the accuracy of the reporters. In (Ferrantino et al., 2011), the authors detect evading customs declarations with the use of discrepancies analysis. Ferrantino & Wang (2008) describe a measure of data asymmetry similar to the one being the basis of the aggregate index presented later in this paper.

Modeling discrepancies among reported data is a common problem in the compilation of macro-economic statistics (Wroe et al., 1999). In general there is some bookkeeping relation between the reported values which does not hold. According to Ten Cate, the solution consists of estimating the accuracy of the various reports and then finding the optimal adaption of the reported values.

The paper by Lejour et al. (2008) describes CPB's (a government agency in the Netherlands) contribution on bilateral services trade data to version 7 of the GTAP database (OECD Statistics of International Trade in Services). Among others, it uses reliability indices to determine the quality of the data reported by exporting and importing countries and briefly discusses alternative methods to make a choice between two available data being reported.

The aim of the study was to analyse the quality of data on intra-Community trade by harmonised system (HS) sections in 2017. The data come from the Comext database provided by Eurostat. For each EU-28 country and each section, the quality of data was examined. For this purpose, we propose an approach to determine aggregated indicators of data discrepancies with country-by-country aggregation. The research allowed us to identify those HS sections for which data quality is the lowest in all individual EU member states.

2. Methodology

Official statistics on intra-Community trade in goods depend on data gathered within the Intrastat system. The database consists of declarations made by businesses dispatching goods to other EU member states or receiving goods from them. Due to the specifics of data it is possible to assess its quality. Mirror data, i.e. data gathered by both sides of a transaction, should be reported twice – by the country of the dispatcher and at the same time by the country of the receiver of traded goods. Thus, for every pair of countries the Intrastat system should contain the same values of trade in goods for the same period. Slight differences are permitted for various reasons, such as statistical thresholds or exchange rates. Generally, however, discrepancies in mirror data are due to errors in reporting or concealment of transactions. And it affects the quality of collected data.

We used Comext data on intra-Community supplies for 2017 aggregated on country and section level. The goods listed in the database are grouped into 21 HS sections. The survey started with calculating data quality indices for each EU-28 country and each HS goods section.

The quality of data on ICS of a country A by HS section was calculated with the use of aggregated index of data quality:

$$ {}_{Z}W_{E^d}^{A,UE} = \frac{\sum_{i=1}^{n} \left| E_{A,B_i}^{d} - I_{B_i,A}^{d} \right|}{K} \tag{1} $$

where:

E_{A,B_i}^{d} – declared value of dispatches (supply) from country A to country B_i,

$I_{B_i,A}^{d}$ – declared value of acquisitions by country B_i delivered from country A (mirror data),

$K = \sum_{i=1}^{n} \frac{E_{A,B_i}^{d} - I_{B_i,A}^{d}}{2}$

$d = 1, \ldots, 21$ – HS section number.

The aggregate index takes values from the range from 0 to 2. The higher the value, the lower the quality of the analysed data is.

On the basis of data quality indices, a ranking of HS sections in each of the EU countries was established and compared. It enabled us to identify the HS sections with the highest and lowest data quality in the EU. The measurement of correlations between indicator vectors determined for individual countries allowed us to identify countries with a similar structure of data quality by section and to distinguish countries with a different structure.

3. Research results – the quality of data

The study used the intra-Community supplies (ICS) values in euros of each EU-28 country and the intra-Community acquisitions (ICA) values mirrored to them. In order to illustrate the scale, main directions and differences in the volume of trade between Member States, the volume of intra-Community trade is presented in Figure 1. (the graph was compiled using the circular library by Gu, (Gu et al., 2014)).

The largest stream of Community goods flows from the Netherlands to Germany. Germany also has the highest turnover (both ICS and ICA). Their main export destinations are France, Great Britain, Belgium and Italy. Spain is another important participant in intra-Community trade in goods. Poland is ranked 8th among 28 EU countries in terms of the volume of trade in goods. The Czech Republic and Austria complete the top ten Member States in terms of intra-EU trade volumes.

The survey was conducted for product groups divided by HS section (listed in Table 1).

For each country, the values of the ICS data quality indicators for each of the HS sections have been determined according to formula (1). The ICS sections in each country were then ranked according to the value of the index. Table 2 presents a summary of the sections with the highest and lowest data quality in intra-Community trade for each of the 28 Member States.

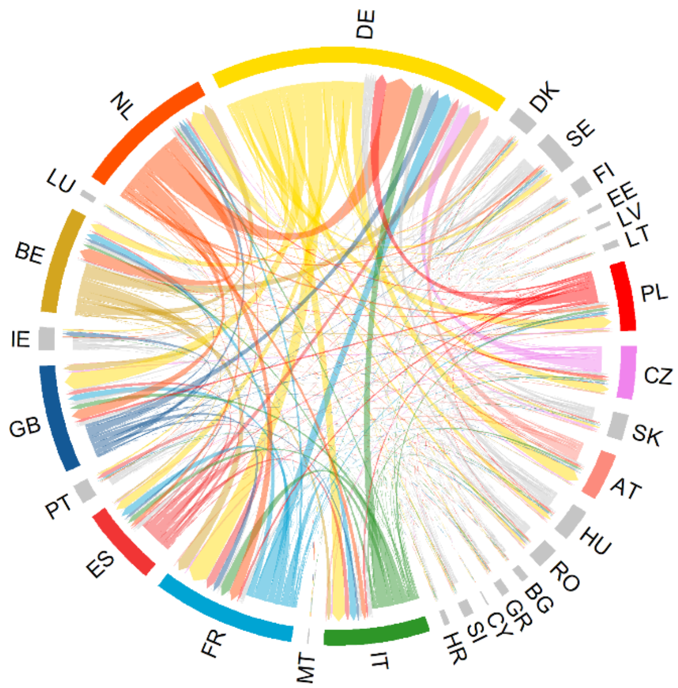


Fig. 1. Intra-Community trade – ICS and ICA by EU-28 countries

Some general schemes can be observed here. Among the sections with the lowest mirror data quality, the most common are sections XIX – weapons, XXI – works of art and special positions, and XIV – pearls, precious stones and metals. On the one hand, these are the sections with relatively small total turnover and therefore transactions within these product groups do not often exceed the statistical threshold. On the other hand, these sections are inherently susceptible to errors, misrepresentations and data gaps (e.g. arms trade can be included among special items, which affects total amounts in two sections).

Among the best-documented sections in terms of turnover reported by both sides are those in which there is a significant exchange of goods between large companies that take seriously the obligation to report intra-Community supplies. Sections I – animals and animal products, IV – prepared foodstuffs and beverages and VI – chemical industry products – are the most frequently observed in the first three sections according to data quality in different countries. High data quality is also recorded in sections II – vegetable products and VII – plastics, and to a lesser extent in sections XV – base metals and XVI – machinery, appliances, and equipment.

Table 1. HS sections and their descriptions

Section	Description
I	Live animals; animal products
II	Vegetable products
III	Animal or vegetable fats
IV	Prepared foodstuffs; beverages, spirits and vinegar; tobacco and manufactured tobacco substitutes
V	Mineral products
VI	Products of the chemical or allied industries
VII	Plastics and articles thereof; rubber and articles thereof
VIII	Raw hides and skins, leather, furskins and articles thereof; saddlery and harness; travel goods, handbags and similar containers; articles of animal gut (other than silkworm gut)
IX	Wood and articles of wood; wood charcoal; cork and articles of cork; manufactures of straw; basketware and wickerwork
X	Pulp of wood or of other fibrous cellulosic material; waste and scrap of paper or paperboard; paper and paperboard and articles thereof
XI	Textiles and textile articles
XII	Footwear, headgear, umbrellas, sun umbrellas, walking-sticks, seat-sticks, whips, riding-crops and parts thereof; prepared feathers and articles made therewith; artificial flowers
XIII	Articles of stone, plaster, cement, asbestos, mica or similar materials; ceramic products; glass and glassware
XIV	Natural or cultured pearls, precious or semi-precious stones, precious metals, metals clad with precious metal and articles thereof; imitation jewellery; coin
XV	Base metals and articles of base metal
XVI	Machinery and mechanical appliances; electrical equipment; parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles
XVII	Vehicles, aircraft, vessels and associated transport equipment

Section	Description
XVIII	Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus; clocks and watches; musical instruments; parts and accessories thereof
XIX	Arms and ammunition; parts and accessories thereof
XX	Miscellaneous manufactured articles
XXI	Works of art, collectors pieces and antiques

In order to indicate which countries have a similar data quality structure by section, the correlation between the vectors of the data quality indicators defined for all sections in the countries was measured. As many as 20 out of 28 countries were found to have a fairly clear correspondence of this structure, while the others differed from the latter and from each other.

Examples of pairs of countries with the highest correlation of data quality structure are Poland and Hungary (r_{xy} = 0.9670), Czech Rep. and Latvia (r_{xy} = 0.9663) or Estonia and Germany (r_{xy} = 0.9545). Figure 2 shows relations and connections between those pairs of countries for which the correlation coefficient between the vectors of the data quality indices in the sections was not less than 0.8. As can be seen, Cyprus (which differs most from the rest of EU-28 in terms of recorded data quality, as confirmed by other surveys), the United Kingdom, Luxembourg, Romania and Malta remain unrelated to other EU countries in this approach. In addition, the group of countries with an individual data quality structure includes Denmark, Ireland and Finland, each with 1, 2 and 3 links in the graph respectively.

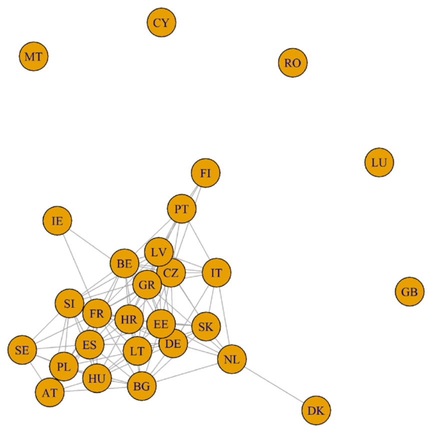


Fig. 2. Data quality correlations between EU-28 countries (only values exceeding 0.8 visible as connections)

Table 2. HS sections with the highest and lowest data quality – by country

Country	Sections with high data quality			Sections with low data quality		
	Rank 1	Rank 2	Rank 3	Rank 19	Rank 20	Rank 21
AT	I	IV	XV	V	XXI	XIX
BE	VI	IV	I	XII	XIX	XXI
BG	XII	XV	XIII	XIV	XXI	XIX
CY	I	VI	II	XII	IX	XXI
CZ	VI	I	XVII	XIX	XIV	XXI
DE	I	IV	X	XIX	XIV	XXI
DK	I	IV	XVIII	VI	XIX	XIV
EE	XX	III	XVI	XIV	XIX	XXI
ES	VIII	III	IV	V	XXI	XIX
FI	VII	IX	XII	III	XVII	XXI
FR	IV	VII	I	XIV	XIX	XXI
GB	XV	VI	IV	XXI	IX	XIX
GR	I	III	IV	XVII	XIX	XXI
HR	XIII	II	I	VIII	XIX	XXI
HU	XVI	IV	IX	XIV	XXI	XIX
IE	I	VI	III	XVII	XXI	XIX
IT	VI	III	VII	XX	XIV	XXI
LT	VI	I	IV	XIV	XIX	XXI
LU	XVIII	VII	XVI	V	XIX	III
LV	IX	IV	I	XIV	V	XXI
MT	XX	XVIII	VII	XXI	XIX*	III*
NL	VII	I	IV	XIX	XXI	XIV
PL	XVI	VI	II	III	XXI	XIX
PT	VI	V	XVI	XVII	XI	XXI
RO	I	II	VI	XIV	XIX*	XXI*
SE	III	IX	XV	XII	XXI	XIX
SI	VI	XIV	XV	VIII	XIX	XXI
SK	XII	XV	II	XIX	XIV	XXI

* both ranked on positions 20–21

Conclusions

HS sections with high data quality in each country and those sections where mirror data is reported asymmetrically were identified. Among the best reported commodity groups, those with large turnover and traded by large companies predominate. Low data quality is more often observed in the case of sections with lower turnover and those where counterparties on one side do not exceed the national statistical threshold. There are also those goods which are not classified unequivocally.

Both literature and EU documents underline the need to ensure the quality of statistical data on intra-Community trade in goods. The survey indicates, on the one hand, the need to improve the quality of data on specific product groups and, on the other hand, it has distinguished a group of countries for which the quality of data differs significantly from the others. There is no doubt that further work is needed on harmonising the procedure for collecting data and improving the quality of data across the EU.

References

Baran, P., & Markowicz, I. (2018a). Analysis of intra-Community supply of goods shipped from Poland. 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, 12–21.

Bieszk-Stolorz, B., Markowicz, I. (2015). Influence of Unemployment Benefit on the Duration of Registered Unemployment Spells. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 10(3), 167–183.

Cate ten A. (2014). The Identification of Reporting Accuracies from Mirror Data. *Jahrbücher für Nationalökonomie und Statistik*, 234(1).

Dehnel, G., & Wawrowski, Ł. (2018). Robust estimation of revenues of Polish small companies by NACE section and province. 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, 110–119.

Ferrantino, M.J., Liu, X., & Wang, Z., (2012). Evasion behaviors of exporters and importers: Evidence from the U.S.-China trade data discrepancy. *Journal of International Economics*, 86(1), 141–157.

Ferrantino, M.J., & Wang, Z. (2008). Accounting for discrepancies in bilateral trade: The case of China, Hong Kong, and the United States. *China Economic Review*, 19(3), 505–520.

Gu, Z., Gu, L., Eils, R., Schlesner, M., & Brors B, (2014). Circlize implements and enhances circular visualization in R. *Bioinformatics*, 30(19), 2811–2812.

Lejour, A., van Leeuwen, N., & ten Cate, A. (2008). *The quality of bilateral services trade data: contribution to GTAP7 database*. CPB Memorandum, CPB Netherlands Bureau for Economic Policy Analysis, 212. https://www.researchgate.net/publication/23968410_The_quality_of_bilateral_services_trade_data_contribution_to_GTAP7_database (accessed on 22.01.2019).

Markowicz, I. (2014). Business Demography – Statistical Analysis of Firm Duration. *Transformations in Business & Economics*, 13(2B), 801–817.

Parniczky, G. (1980). On the Inconsistency of World Trade Statistics. *International Statistical Review*, 48(1).

Tsigas, M.E., Hertel, T.W., & Binkley, J. K. (1992). Estimates of systematic reporting biases in trade statistics. *Economic Systems Research*, 4(4).

Wroe, D., Kenny, P., Rizki, U., & Weerakoddy, I. (1999). *Reliability and quality indicators for National Accounts aggregates*. http://epp.eurostat.ec.europa.eu, Office for National Statistics (UK).

Median classification of the European Union countries regarding the level of selected strategic goals' implementation – dynamic approach

Małgorzata Markowska[1], Danuta Strahl[2]

Abstract

Europe 2020 is the strategy covering the united Europe, in force for the current planning period (2010–2020). It is a strategy for smart, sustainable and inclusive growth (Europe 2010). This plan adopted three priorities: smart development, sustainable development and inclusive growth. The goals for each EU country, set in the National Reform Programme, were adopted to be implemented by these countries until 2020, based on socio-economic conditions and following negotiations with the European Commission.

The purpose of the study is to identify similar groups of countries regarding the level of the EU strategic goals' implementation in terms of smart development by applying the median classification and using the dynamic approach. The assessment period – taking into account the statistical data availability – covers the years 2010–2017, and the indicators used are: employment rate in the age group 20–64, the percentage of tertiary education graduates in the group.

Keywords: *Europe 2020, strategic goals, smart growth, the EU countries*
JEL Classification: *C19, F63, O52*

1. Introduction

Europe 2020. A European strategy for smart, sustainable and inclusive growth (Europe, 2010) is a strategy of the united Europe binding during the present decade, and includes three related priorities (Europe, 2010): smart growth: developing an economy based on knowledge and innovation; sustainable growth: promoting a more resource efficient, greener and more competitive economy; inclusive growth: fostering a high-employment economy delivering social and territorial cohesion.

For each of the priorities being implemented under the strategy, indicators were assigned for the purpose of assessment. It is expected that by 2020 the following levels of the indicators will be achieved (Europe, 2010):

– 75% of the population aged 20–64 should be employed;
– 3% of the EU's GDP should be invested in R&D;
– the "20/20/20" climate/energy targets should be met (including an increase to 30% of emissions reduction if the conditions are right);
– the share of early school leavers should be under 10% and at least 40% of the younger generation should have a tertiary degree;
– 20 million less people should be at risk of poverty.

[1] Wroclaw University of Economics, Department of Regional Economy, 118/120 Komandorska St., 53-345 Wroclaw, Poland, malgorzata.markowska@ue.wroc.pl.

[2] WSB University in Dabrowa Gornicza, Department of Management, 1c Cieplaka St., 41-300 Dabrowa Gornicza, dstrahl@wsb.edu.pl

The values given relate to the European Union as a whole. At the same time, each EU country committed itself in the National Reform Programme (NRP) to implement national goals, resulting from socio-economic conditions. Common "efforts" of the EU countries make up the ability to achieve the EU goals.

The aim of this study is to determine groups of countries similar to each other regarding the degree of implementation of the EU objectives in the scope of one of the priorities, i.e. smart development using the median classification – in the dynamic approach.

2. Method

Estimates of the level of full filling strategic goals are important issues for researchers. New measures are proposed (Pasimeni, 2013; Pasimeni and Pasimeni, 2016; Markowska, 2019), classifications performed (Stec and Grzebyk, 2018; Fura et al., 2017), and composite measures calculated (Hudrliková, 2013).

The method of classification described in the works by D. Strahl (Markowska and Strahl, 2003, Strahl, 2002) will be used to accomplish the principal purpose of the paper. The classification based on positional statistics, allows for value judgements. Multidimensional phenomena which are subjected to grouping can be described with a set of characteristics – variables. These variables are assigned to objects constituting sets. A data matrix is created, illustrating a selected phenomenon characterised by a set of m variables, marked with the symbols $X = \{X_1, ..., X_m\}$, observed on the objects of the study. For each variable X_j ($j = 1, 2, \ldots, m$) the median, i.e. the middle value will be calculated. Below and above the median there is 50% of the realisation of the variable.

The proposed classification procedure includes two variants: in the first of them a classification algorithm leads to the construction $(m + 1)$ of classes, in the second – to the construction of 2^m classes of possible combinations with m-variables, $m > 1$. The latter option may be applied when we want to consider several possible assessments to support the decision-making process (Markowska and Strahl, 2003).

In the first variant, the first class (S_1) contains these objects for which the values of all m variables X_j are equal to or higher than the median (maintaining the preference direction). The second class (S_2) includes objects whose values of $m - 1$ variables are equal to or higher than the median. The S_m class consists of objects for which the value of only one variable X_j from the set X is equal to or higher than the median. The S_{m+1} class includes these objects for which all variables are lower than the median.

In the second variant, S_1 class includes these objects from the set, for which the values of all m variables X_j are equal to or higher than the median. The S_2 class contains these objects from the set, for which the values of only $(m - 1)$ variables constituting one of the combinations $\binom{m}{m-1}$ are equal to or higher than the median. In the third class (S_3) there are objects for which the values of variables of the next $(m - 1)$-element combination are equal to or higher than the median. Having exhausted $(m - 1)$-element combinations we create classes for $(m - 2)$-element combinations and set a condition that they are equal to or higher than the median. We create S_g ($g = 2^m$) class, comprising objects for which the values x_{kj} of all variables X_j are lower than the median.

3. Variables

For the purpose of the work, it was assumed that (Europe 2020) smart development is based on obtaining positive effects in the field of education, research and innovation and the effective use of ICT techniques. With regard to both the given definition and the goals included in the Europe 2020 strategy, it was decided that the numerical illustration of the concept of smart development should include the following characteristics (EARLY – destimulant, the others are stimulants):

- EMPLO – employment rate in the 20–64 age group,
- TERTIARY – participation of people with higher education in the 30–34 age group,
- EARLY – early leaver from education and training, previously named early school leaver, refers to a person aged 18 to 24 who has completed at most lower secondary education and is not involved in further education or training; the indicator ‘early leavers from education and training’ is expressed as a percentage of the people aged 18 to 24 with such criteria out of the total population aged 18 to 24,
- BR_GDP – expenditures on R&D in relation to GDP.

Table 1 presents the target values of indicators set in the NRP, while table 2 shows the values of characteristics and basic statistics for the years 2010–2017.

Table 1. Target values of the indicators – EU and national goals

Country (acronym)	EMPLO	BR_GDP	EARLY	TERTIARY
the European Union (EU)	75.0	3.00	10.0	40.0
Austria (AT)	77.0	3.76	9.5	38.0
Belgium (BE)	73.2	3.00	9.5	47.0
Bulgaria (BG)	76.0	1.50	11.0	36.0
Cyprus (CY)	75.0	0.50	10.0	46.0
Czech (CZ)	75.0	1.00	5.5	32.0
Germany (DE)	77.0	3.00	10.0	42.0
Denmark (DK)	80.0	3.00	10.0	40.0
Estonia (EE)	76.0	3.00	9.5	40.0
Greece (EL)	70.0	1.20	10.0	32.0
Spain (ES)	74.0	2.00	15.0	44.0
Finland (FI)	78.0	4.00	8.0	42.0

Country (acronym)	EMPLO	BR_GDP	EARLY	TERTIARY
France (FR)	75.0	3.00	9.5	50.0
Croatia (HR)	62.9	1.40	4.0	35.0
Hungary (HU)	75.0	1.80	10.0	34.0
Ireland (IE)	69.0	2.00	8.0	60.0
Italy (IT)	67.0	1.53	16.0	26.0
Lithuania (LT)	72.8	1.90	9.0	48.7
Luxembourg (LU)	73.0	2.30	10.0	66.0
Latvia (LV)	73.0	1.50	10.0	34.0
Malta (MT)	70.0	2.00	10.0	33.0
Netherlands (NL)	80.0	2.50	8.0	40.0
Poland (PL)	71.0	1.70	4.5	45.0
Portugal (PT)	75.0	2.70	10.0	40.0
Romania (RO)	70.0	2.00	11.3	26.7
Sweden (SE)	80.0	4.00	7.0	45.0
Slovenia (SI)	75.0	3.00	5.0	40.0
Slovakia (SK)	72.0	1.20	6.0	40.0
United Kingdom (UK)	x	x	x	x
Maximum (country)	80 (SE, NL, DK)	4 (SE, FI)	16 (IT)	66 (LU)
Minimum (country)	62.9 (HR)	0.5 (CY)	4 (HR)	26 (IT)

S – standard deviation, V – coefficient of variation, x – a lack of goal in the National Reform Programme (NRP)
Source: own elaboration on the basis of data (https://ec.europa.eu/eurostat/web/europe-2020-indicators/europe-2020-strategy/overview)

When analysing national goals and their extreme values it is interesting to note that for example in Italy the high percentage of students leaving the education is accompanied by a low share of young people with higher education while in Croatia the lowest level of the goal related to the employment rate is accompanied by the lowest rate of early leavers from education.

The values of characteristics selected to assess the level of smart development in the EU countries were changing in the analysed period: mean values – of the analysed characteristics improved, as far as EU goals are concerned, minimum values in the EU countries decreased for EMPLO and increased for other characteristics, maximum values decreased for BR_GDP and increased for other characteristics, standard deviation was stable for two characteristics (EMPLO, BR_GDP) and decreased for others (EARLY, TERTIARY).

In order to achieve the aim of the work – grouping countries according to the degree of implementation of the smart development goals – it was necessary to determine for the EU

countries differences between the values of variables and the objectives: the EU objectives from the Europe 2020 strategy (2010) and national objectives from the NRP. Therefore, the following variables have been specified:
A_UE – difference between the value of the EMPLO and the values of the EU goal,
B_UE – difference between the value of the TERTIARY and the value of the EU goal,
C_UE – difference between the value of the EU goal and the value of the EARLY,
D_UE – difference between the value of the BR_GDP and the value of the EU goal.
and
A_ country – difference between the value of the EMPLO and the value of the NRP goal,
B_ country – difference between the value of the TERTIARY and the value of the NRP goal,
C_ country – difference between the value of the NRP goal and the value of the EARLY,
D_ country – difference between the value of the BR_GDP and the value of the NRP goal.

While carrying out the study, it was assumed that if the value of the variable in the country is more favourable than the goal (from the EU strategy in the first case and from the National Reform Programme in the second case), for the differences obtained, which are obviously positive, we will assume zero for the calculation of the median (cutting off in zero) (Markowska, 2019).

4. Results

As a result of a double use of positional classification with the median (Strahl 2002, Markowska, Strahl 2003) for the differences between the target level – firstly for the Europe 2020 goals (2010) and secondly for the goals from the NRP – and the values of variables (stimulants) and between the values of a variable and a goal (destimulants) five classes of countries were obtained. The first one (I) – comprising countries for which the differences in all characteristics of smart development are equal to or more favourable than their medians established for the EU countries. The second one (II) – countries for which the difference of one of the selected characteristics of smart development is lower than the median (for example class II ABC – is the class of countries for which the value of variable D_UE in comparison with goals from the Europe 2020 strategy or D_ country, when compared with goals from the NRP was lower than the median). Next classes are built in the same manner with the diminishing number of differences above median.

For most of medians of differences – aside from the median of difference between the value of the goal from the EU strategy and the value of the BR_GDP characteristic – the median was decreasing. Medians of differences (including “cut” values) are set in table 2.

Table 2. Medians of differences

Specification		2010	2011	2012	2013	2014	2015	2016	2017
The EU goals									
Median	A_UE	-7.80	-7.30	-6.80	-7.80	-7.00	-5.70	-4.10	-1.85
	B_UE	-5.20	-2.80	-0.85	0.00	0.00	0.00	0.00	0.00
	C_UE	-1.40	-1.15	-0.40	0.00	0.00	0.00	0.00	0.00
	D_UE	-1.58	-1.54	-1.67	-1.64	-1.66	-1.69	-1.74	-1.70
National goals									
Median	A_country	-5.10	-5.90	-4.85	-5.00	-4.40	-3.55	-2.45	-1.55
	B_country	-6.45	-5.70	-4.15	-3.65	-1.70	-0.90	-0.20	0.00
	C_country	-1.90	-1.35	-1.15	-0.45	0.00	-0.15	0.00	-0.05
	D_country	-0.80	-0.68	-0.72	-0.71	-0.70	-0.72	-0.74	-0.71

The results of the median classification for the goals of the Europe 2020 strategy (2010) are given in table 3 and for the purposes of the NRP in table 4. In these tables, for classes II, III and IV, there are variables for which the differences were equal to or more favourable than the median. It is significant considering the level of achievement of the EU objectives.

Table 3. Assignment of countries to groups – the EU goals – in the years 2010–2017

Class		2010	2011	2012	2013	2014	2015	2016	2017
I (ABCD)		FI, SE, DK, NL, LU, SI	FI, SE, DK, NL, EE, LU, SI	FI, SE, DK, NL, EE, SI	FI, SE, DK, NL, EE, FR, SI	FI, SE, DK, NL, IE, FR, AT	FI, SE, DK, NL, FR	FI, SE, DK, NL, AT	FI, SE, DK, NL, AT, SI
II	ABC			LT, LU	LV, LT, LU	LT, LU	LV, LT, LU, IE	LV, LT, IE	LV, LT
	ABD	UK, FR, BE	UK, FR	UK, FR	UK, BE	UK, EE	UK, EE	UK	UK
	ACD	AT	AT, CZ	AT, CZ	AT, CZ, CY, DE	CZ, DE	AT, CZ	CZ	CZ
	BCD	EE	IE	IE	IE	SI, BE	SI	SI, FR, LU, BE	FR, BE

Class		2010	2011	2012	2013	2014	2015	2016	2017
III	AB	CY	CY	CY				EE	EE
	AC	CZ				LV			
	AD	PT, DE	PT, DE	DE			DE	DE	PT, HU, DE
	BC	PL, LT	LT		PL	CY, PL	CY, PL, EL	CY, PL, EL	CY, LU, PL, IE, EL
	BD	IE	BE	BE			BE		
IV	A	HU						MT, HU	
	B	ES	ES	ES	ES	ES	ES	ES	ES
	C	SK, HR	SK, HR, PL	SK, HR, PL	SK, HR	SK, HR, EL	SK, HR	SK, HR	SK, HR
	D			PT	HU	HU	IT, HU	IT, PT	IT
V		BG, RO, EL, IT, LV, MT	BG, RO, EL, IT, LV, HU, MT	BG, RO, EL, IT, LV, HU, MT	BG, RO, EL, IT, MT, PT	BG, RO, IT, MT, PT	BG, RO, MT, PT	BG, RO	BG, RO, MT

Table 4. Assignment of countries to groups – national goals – in the years 2010–2017

Class		2010	2011	2012	2013	2014	2015	2016	2017
I (ABCD)		SE, DK	NL, DK, CY	FI, NL, DK	SE, DK	AT, DK	AT, DK	AT	
II	ABC	SI	SE, FI, LT, EE	SE, LT, EE	NL, LT, LV, EE	SE, LT, LV	SE, LT, LV	LT, LV	LT, LV
	ABD	FI, NL, CY		CY				SE, CZ	SE, PL, CZ
	ACD	LU, HR, DE, CZ	HR, CZ	HR, DE, CZ	DE, CZ	HR, IE, DE, CZ	HR, DE	HR	HR
	BCD		SI	SI	SI, EL, CY	SI, CY, EL	CY, IT EL	NL, CY, IT, EL, DK	AT, NL, CY, IT, EL, DK

Class		2010	2011	2012	2013	2014	2015	2016	2017
III	AB	UK	UK	UK	UK, FI	UK, RO, EE	UK, EE	UK, EE	UK, MT, EE
	AC	AT	AT, LU	AT, LU	AT, LU	LU	LU, IE	LU, IE	LU, IE
	AD	IE	IE, DE	IE	IE		PL, CZ	SK, DE	SK, DE
	BC	LT, EE		LV			SI	FI, SI	SI
	BD	IT, ES, EL	IT, ES, EL	IT, ES, EL	HU, IT	NL, HU	NL, HU		
	CD	SK, HU	SK	SK	SK	IT		BE	BE
IV	A	PT			MT	MT	MT	PL, MT	RO
	B	LV, BE	LV, BE	BE	ES	FI, ES	FI		FI
	C	PL, BG	PL, BG		FR	SK, FR	SK, FR	FR	FR
	D		HU	HU	HR, BE	BE	BG, BE	HU, BG	HU
V (-)		RO, MT, FR	RO, PT, MT, FR	RO, PT, PL, MT, FR, BG	RO, PT, PL, BG	PT, PL, BG	RO, PT, ES	RO, PT, ES	PT, ES, BG

10 out of 28 countries (35.7%) in the analysed period belonged to the same classes: Finland, Sweden, Denmark and the Netherlands – Class I, Great Britain – Class II ABD, Spain – Class IV B, Slovakia and Croatia – Class IV C, Bulgaria and Romania – Class V.

As regards the evaluation of the implementation of national goals – in the results of the classification only Great Britain remained in the same class for the entire period (III AB). The assignment of Poland to the classes in both considered variants is as follows:

– for differences in relation to the Europe 2020 strategy goals: the year 2010 and the period 2012–2017 (class III BC – differences for EMPLO and BR_GDP variables are less favourable than the median), the years 2011–2012 (IV C – only the EARLY difference is more favourable than the median),
– for differences regarding national targets in 2010–2011 class IV C (only the difference between the value of the EARLY characteristic and the value of the NRP goal is equal to or more favourable than the median), in the consecutive three years all differences were less favourable than the median. In 2015 Poland was in class III AD, in 2016 in class IV A, in 2017 it was in class II ABD (differences of three variables were more favourable than the median – except for EARLY).

Conclusions

Taking into account (the first variant) national goals, the most numerous group includes countries with one or two characteristics approaching the adopted strategic goals according to the criterion ‘value more favourable than the median of difference’ (75% of the countries in the last year of the survey), and only three, two, one country over eight years of research is not in a situation where all the characteristics meet this criterion. The stability of the size of the other two groups indicates that the pace of achieving the goals by the countries was maintained. The structure of countries' classification in terms of European strategic goals is insignificantly different. Slightly more countries (6%) than in the case of achieving national goals are in the first group comprising states closest to the adopted target, taking into account all the variables used in the study. In the last analysed year in the II and III group there are 15 the EU countries (53% of the total number), which implement the European strategic goals, achieving for two and three variables more favourable values in comparison to the adopted median criterion. Poland is in the III or IV class of this classification. The decrease in the value of median of differences between the strategic goal and the value of the variables indicates a consistent pursuit of goals. Dynamic aspects of the implementation of strategic goals, both in the European and national dimension, should be subject to further attention and analysis.

Acknowledgements

The paper is financed by the National Science Centre: 2015/17/B/HS4/01021 and from the sources awarded to the Department of Management at WSB University in Dąbrowa Górnicza as a grant for maintaining research potential.

References

Europe 2020. A strategy for smart, sustainable and inclusive growth (2010). Communication from the Commission, European Commission, Brussels, 3.3.2010 COM(2010) 2020.

Fura, B., Wojnar J., & Kasprzyk, B. (2017). Ranking and classification of EU countries regarding their levels of implementation of the Europe 2020 strategy. *Journal of Cleaner Production*, 165, 968–979.

Hudrliková, L. (2013). Composite indicators as a useful tool for international comparison: The Europe 2020 example. *Prague Economic Papers,* 4, 459–473.

Markowska, M., & Strahl, D. (2003). Statystyki pozycyjne w klasyfikacji porównawczej. In: Jajuga, K., Walesiak M. (eds.), *Klasyfikacja i analiza danych – teoria i zastosowania*. *SKAD*, Prace Naukowe Akademii Ekonomicznej nr 988, Taksonomia 10, AE, Wrocław, 299–308.

Markowska, M. (2019). Wielokryterialny monitoring i ocena realizacji celów strategii Europa 2020 w zakresie inteligentnego rozwoju. Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław (in print).

Pasimeni, F., & Pasimeni, P. (2016). An Institutional Analysis of the Europe 2020 Strategy, *Social Indicators Research,* 127(3), 1021–1038

Pasimeni, P. (2013). The Europe 2020 Index. *Social Indicators Research*, 110, 613–635.

Stec, M. & Grzebyk, M. (2018). The implementation of the Strategy Europe 2020 objectives in European Union countries: the concept analysis and statistical evaluation, *Quality&Quantity,* 52(1), 119–133.

Strahl, D. (2002). *Klasyfikacja regionów z medianą.* In: Dziechciarz, J. (ed.), *Zastosowania metod ilościowych*, Prace Naukowe Akademii Ekonomicznej nr 950, Ekonometria 10, AE, Wrocław, 11–18.

Bayesian interpretation of quasi-Bayesian inference in a normal hierarchical model

Jacek Osiewalski[1]

Abstract
In modern parametric statistics and its applications latent variables and random effects are widely used, and their estimation or prediction is of interest. Under some prior assumptions, Bayes formula can be used to obtain their posterior distribution. However, on the sampling-theory grounds, the unknown constants which appear in the prior distribution are estimated by means of the data being actually modelled. We call such approaches quasi-Bayesian; *parametric empirical Bayes* is an important example. In this paper we propose theoretical framework that enables the Bayesian interpretation of incoherent, quasi-Bayesian inference techniques. Our framework amounts to establishing a formal Bayesian model that justifies a quasi-Bayesian "posterior" (resulting from some data-based "prior") as a valid posterior distribution. From such Bayesian model, i.e. the joint distribution of observations and other quantities, one can deduce the true sampling model, that is the conditional distribution of observations, and the true prior (or marginal) distribution of the remaining quantities – latent variables or parameters. Since analytical derivations are possible in very specific cases, this paper presents only a simple, illustrative example based on a normal hierarchical model. It clearly shows that quasi-Bayesian approaches can lead to posterior distributions, which formally correspond to sampling models and prior distributions different than the assumed (declared) ones.

Keywords: *Bayesian statistics, coherent inference, empirical Bayes procedures, random parameters, shrinkage estimation*
JEL Classification: *C11, C18, C51*

1. Introduction

Bayes Theorem, usually used as Bayes formula for density functions of continuous random variables, is a central, important tool of Bayesian statistics, but it is not the only characteristic of this mode of statistical modelling and inference. There are two defining features of the Bayesian approach to statistics. Probabilistic representation of uncertainty about observations (available, missing, future), latent variables (or random parameters) and classical parameters (unknown constants) is the main feature of *Bayesian modelling*, and treating all "unknowns" as random variables is closely related to the concept of subjective probability. Obeying rules of probability calculus is then the main characteristic of *Bayesian inference*. Obviously, Bayes formula is one of these rules (and a very useful one), but following it in isolation from other rules does not mean conducting Bayesian inference.

In modern statistics and its numerous subject areas (like econometrics) latent variables, random effects and other unobservable random quantities are frequently used and their estimation or prediction is usually of particular interest. Bayes formula can be used to obtain their posterior distribution, given appropriate distributional assumptions. Then, the posterior mean can be used

[1] Cracow University of Economics (Faculty of Management, Dept. of Econometrics and Operations Research); 27 Rakowicka St., 31-510 Cracow, Poland, e-mail: eeosiewa@cyf-kr.edu.pl.

as the estimator (or predictor), even within a non-Bayesian approach to statistics. However, the posterior distribution (and the posterior mean) of a latent variable depends on unknown constants (parameters) of the assumed class of marginal (or prior) distributions of this variable. The purely Bayesian solution amounts to treating all unknowns probabilistically and using probability rules on each level of the hierarchical model. On the sampling-theory grounds, however, the unknown constants in the prior distribution are estimated on the basis of the data being actually modelled. Such approach has been popular since 1970s under the name (*parametric*) *empirical Bayes*; see Efron and Morris (1972), Morris (1983) and Casella (1985). Empirical Bayes (EB) methods can be described as incoherent by an orthodox Bayesian, who by coherency means following basic rules of probability. While such description is formally exact and true, it does not provide us with a deeper Bayesian understanding of incoherent inferences that are practically useful and frequently adopted in empirical research.

We propose theoretical framework that enables the purely Bayesian interpretation of incoherent, quasi-Bayesian inference techniques such as EB. Our framework amounts to establishing such formal Bayesian model that justifies a quasi-Bayesian "posterior" (resulting from some data-based "prior") as a valid posterior. From this Bayesian model, i.e. the joint distribution of observations and other quantities, which justifies the posterior in question, one can deduce (at least in principle) the true sampling model, that is the conditional distribution of observations, and the true prior (or marginal) distribution of the remaining quantities – latent variables and parameters. Since analytical derivations are possible only in very specific cases, we present a simple, illustrative example in this paper. However, it clearly shows that incoherence of quasi-Bayesian approaches can lead to posterior distributions, which formally correspond to sampling models and prior distributions different than the assumed (declared) ones. In the next section our general framework is presented. Section 3 is devoted to the Bayesian and quasi-Bayesian (EB type) approaches to dealing with random effects in a normal hierarchical model. Again, our analysis is kept as simple as possible in order to enable fully analytical derivations. Section 4 concludes.

2. Bayesian interpretation of "posteriors" resulting from data-based "priors"

By quasi-Bayesian inferences we mean approaches where Bayes formula is used mechanically, outside the fully probabilistic Bayesian context that guarantees coherence. Consider the conditional density of observations $p(y|\omega) = g(y; \omega)$, corresponding to some parametric statistical model, and the prior density $p(\omega)$ from some parametric family. Then the Bayesian inference relies on the posterior density function $p(\omega|y) \propto p(y|\omega)p(\omega)$ that, again, belongs to some parametric family. Assuming that instead of specifying the prior hyper-parameters, one estimates them using the actual data and inserts to the formula for the posterior density. It leads to the "posterior" density $p^*(\omega|y)$ that amounts to using the data-based "prior" $p^*(\omega) = f(\omega; y)$, which cannot be the marginal distribution of parameters (and other unknowns, like latent variables). Thus, the density $p^*(\omega|y) \propto p(y|\omega)p^*(\omega) \propto g(y;\omega)f(\omega;y)$ is not the posterior density in the original model with initially assumed $p(y|\omega)$ and $p(\omega)$. However, such $p^*(\omega|y)$ is a member of the same parametric family as $p(\omega|y)$, so it is a well-defined probability density function and it can be the true, formal posterior in a completely different Bayesian model. The main question we pose here is as follows: what are the true building blocks (sampling model and prior) coherently justifying $p^*(\omega|y)$ that was initially obtained from a data-based "prior"? It would be useful to know hidden assumptions underlying formal Bayesian inference based on $p^*(\omega|y)$.

Specifying the marginal distribution of the parameters with the use of the actual data is a fundamental form of incoherence, and $p^*(\omega|y)$ obtained through Bayes formula corresponds then to some statistical model and prior assumptions, which have to be discovered. In order to obtain $p^*(\omega|y)$ as the posterior density, we consider the joint distribution of observations and parameters that is characterised by the density function:

$$\tilde{p}(y,\omega) \propto p(y|\omega)p^*(\omega) \propto g(y;\omega)f(\omega;y),$$

which can be decomposed in two ways:

$$\tilde{p}(y,\omega) = \tilde{p}(\omega|y)\tilde{p}(y) = \tilde{p}(y|\omega)\tilde{p}(\omega).$$

Note that $\tilde{p}(\omega|y)$ and $\tilde{p}(y|\omega)$ are probability density functions if and only if $\tilde{p}(y)$ and $\tilde{p}(\omega)$ are densities of σ-finite measures. Then $\tilde{p}(\omega|y) = p^*(\omega|y)$ by construction, since

$$\tilde{p}(\omega|y) \propto \tilde{p}(y,\omega) \propto g(y;\omega)f(\omega;y).$$

Also note that the joint density $\tilde{p}(y,\omega)$ represents the Bayesian model corresponding to the sampling density $\tilde{p}(y|\omega)$ and the prior density $\tilde{p}(\omega)$. In this Bayesian model we obtain $p^*(\omega|y)$ as the formal (true) posterior density.

3. Quasi-Bayesian and Bayesian analysis of hierarchical models

Now we consider a statistical model with hierarchical structure which can be the starting point for explanation and justification of the EB approach. However, in order to use purely analytical tools and obtain closed-form solutions, only normal distributions with known variances and covariances are examined here. More general priors that correspond to the ones in the basic EB literature are left for future research. Hierarchical Bayesian estimation of a more general random parameters regression type model is presented in Greene (2008, section 18.8); however, it cannot be examined analytically.

The hierarchical structure of a statistical model amounts to assuming the conditional distribution of observations $p(y|\theta) = g(y;\theta)$ ($y \in Y$, $\theta \in \Theta$) where the parameters are in fact latent random variables with some distribution dependent on deeper parameters (treated as unknown constants on non-Bayesian grounds); its density is denoted as $f_0(\theta;\alpha)$, $\alpha \in A \subseteq \mathbb{R}^s$. Then the joint distribution of observations and latent variables (with α fixed) can be written and decomposed in the following way:

$$p(y|\theta)f_0(\theta;\alpha) = g(y;\theta)f_0(\theta;\alpha) = f_1(\theta|y;\alpha)\,h(y;\alpha),$$

where $h(y; \alpha)$ and $f_1(y|\theta; \alpha)$ are the densities of the marginal distribution of observations and the conditional distribution of latent variables, respectively. Of course, Bayes formula describes the relation between all four density functions:

$$f_1(\theta|y;\alpha) = g(y;\theta)f_0(\theta;\alpha)/h(y;\alpha) \propto g(y;\theta)f_0(\theta;\alpha).$$

Thus, in order to make inferences on latent variables (given observations) Bayes formula is used. However, within non-Bayesian approaches, like EB, no prior distribution is assumed for the deeper parameters α, which are estimated using (for example) the maximum likelihood principle applied to the density $h(y; \alpha)$, which is considered as a function of α (for any given y). Then the estimate of α, e. g.,

$$\hat{\alpha} = \hat{\alpha}_{ML} = \arg\max L(\alpha; y) = \arg\max h(y;\alpha), \quad \alpha \in A,$$

is inserted into the posterior density of latent variables. So such quasi-Bayesian approach uses

$$p^*(\theta|y) = f_1(\theta|y,\hat{\alpha}) \propto p(y|\theta)f_0(\theta;\hat{\alpha}),$$

i.e. the “posterior” of θ corresponding to the “prior” with hyper-parameter based on y. This cannot be a formal Bayesian approach, although Bayes formula has been used at an earlier stage. Thus, it is called quasi-Bayesian.

Now let us consider the Bayesian hierarchical model (BHM)

$$p(y, \omega) = p(y, \theta, \alpha) = p(y|\theta)\, p(\theta|\alpha)\, p(\alpha),$$

where $\omega = (\theta, \alpha)$, $p(\alpha)$ is the prior density for $\alpha \epsilon A$, and conditional independence: $y \perp \alpha|\theta$ (characteristic for hierarchical models) leads to $p(y|\omega) = p(y|\theta)$. We use the same notation $p(y|\theta) = g(y; \theta)$ and $p(\theta|\alpha) = f_0(\theta; \alpha)$ as in the quasi-Bayesian case. Basic rules of probability calculus lead to the decomposition of our Bayesian model; this decomposition serves making inferences on $\omega = (\theta; \alpha)$:

$$p(y, \theta, \alpha) = p(y)\, p(\theta, \alpha|y) = p(y)\, p(\alpha|y)\, p(\theta|y, \alpha),$$

where

$$p(\theta|y,\alpha) = \frac{p(y|\theta)p(\theta|\alpha)}{p(y|\alpha)} = \frac{g(y;\theta)f_0(\theta;\alpha)}{h(y;\alpha)} = f_1(\theta|y;\alpha),$$

$$p(\alpha|y) = \frac{p(y|\alpha)p(\alpha)}{p(y)} = \frac{h(y;\alpha)p(\alpha)}{p(y)}, \quad p(y) = \int_A p(y|\alpha)p(\alpha)\,d\alpha.$$

Note that Bayes formula has been used twice: for latent variables θ (given parameters α) and for parameters α themselves. Again, according to probability rules, the marginal density of latent variables is a continuous mixture

$$p(\theta|y) = \int_A f_1(\theta|y;\alpha)p(\alpha|y)\,d\alpha,$$

which is the basis of Bayesian inference on θ. Clearly, uncertainty about α is now fully taken into account, contrary to the quasi-Bayesian case, where unknown α is simply replaced by its point estimate.

In order to provide a coherent Bayesian interpretation of the quasi-Bayesian "posterior" $p^*(\theta|y) = f_1(\theta|y, \hat{\alpha})$ in any particular case, we should consider the joint density (Bayesian model) $\tilde{p}(y, \theta)$ that formally leads to $\tilde{p}(\theta|y) = p^*(\theta|y)$. Then the sampling density $\tilde{p}(y|\theta)$ and the prior density $\tilde{p}(\theta)$ Bayesianly justify quasi-Bayesian inference based on $p^*(\theta|y)$. Now we use a simple normal hierarchical model as a purely analytical example of our approach. Firstly, we present the strict (coherent) Bayesian analysis. Secondly, quasi-Bayesian results are given and their Bayesian interpretation is derived. Let θ_i denote an unobservable characteristic, randomly distributed over n observed units ($i = 1,\ldots, n$). Since θ_i is specific for the i-th unit, it can be called an individual effect. Let $x_{ij} \sim iiN(\theta_i, c_0)$ (for $j = 1,\ldots, m$) denote m independent measurements of θ_i, with c_0 known. Let e_m denote the $m \times 1$ vector of ones. Then $y_i = \frac{1}{m}e'_m x_i = \bar{x}_i$ is a sufficient statistic (for given θ_i) and n average measurements grouped in $y = (y_1 \ldots y_n)'$ are independent and normally distributed given $\theta = (\theta_1 \ldots \theta_n)'$, i.e. $y_i|\theta \sim iiN(\theta_i, c)$, where $c = c_0/m$. As it has been assumed, the unobserved parameters or individual effects are random (thus they are latent variables), independent and normally distributed with the same unknown mean α and the same known variance d, i.e. $\theta_i \sim iiN(\alpha, d)$. Thus in this example

$$p(y|\theta) = f_N^n(y|\theta, cI_n), \quad f_0(\theta; \alpha) = f_N^n(\theta|\alpha e_n, dI_n).$$

where $f_N^k(.|b, C)$ denotes the density function of the k-variate normal distribution with mean vector and covariance matrix C.

We can decompose the product $p(y|\theta) f_0(\theta; \alpha)$ into $f_1(\theta|y; \alpha)\, h(y; \alpha)$, where

$$h(y; \alpha) = \int_{\mathbb{R}^n} p(y|\theta) f_0(\theta; \alpha)\, d\theta = f_N^n(y|\alpha e_n, (c + d)I_n)$$

is the density function of the marginal distribution of the observation vector (given α), and

$$f_1(\theta|y; \alpha) = f_N^n\left(\theta \left| \frac{d^{-1}}{c^{-1} + d^{-1}}\, \alpha e_n + \frac{c^{-1}}{c^{-1} + d^{-1}} y,\ \frac{1}{c^{-1} + d^{-1}} I_n \right.\right)$$

is the posterior density of the vector of random effects (given α), with the mean

$$E(\theta|y; \alpha) = w \cdot \alpha e_n + (1 - w) \cdot y, \quad w = \frac{d^{-1}}{c^{-1} + d^{-1}} = \frac{c}{c + d} \in (0, 1).$$

Note that the posterior precision (the inverse of posterior variance) is the sum of the sample precision c^{-1} and the prior precision d^{-1}, and the posterior mean is a weighted average of the vector of prior means and the observation vector – with weights equal to the share of prior or sample precision in the posterior (or final) precision. Thus $E(\theta|y; \alpha)$ is a point (in $\Theta = \mathbb{R}^n$) that lies on the line segment between $(\alpha\ \alpha\ \ldots\ \alpha)'$ and $(y_1\ y_2\ \ldots\ y_n)$.

It is worth stressing that the conditional density $f_1(\theta|y;\alpha)$ follows from Bayes formula for any fixed α, so to this point the presented approach obeys coherence. However, the deeper parameter (the prior mean α) is unknown, so there are two possible ways of treating it. On the sampling theory grounds (like in the EB approach) some point estimate $\hat{\alpha}$ is inserted into $f_1(\theta|y;\alpha)$, which results in $p^*(\theta|y) = f_1(\theta|y, \alpha = \hat{\alpha})$. In our example we get

$$p^*(\theta|y) = f_N^n\left(\theta|\hat{\theta}_{EB}, \frac{1}{c^{-1}+d^{-1}}I_n\right),$$

where

$$\hat{\alpha} = \bar{y} = \frac{1}{n}e_n'y, \quad \hat{\theta}_{EB} = w\bar{y}e_n + (1-w)y.$$

Three points are worth mentioning. Firstly, uncertainty about α is not fully taken into account. Secondly, $p^*(\theta|y)$ is not the posterior density, because the conditional prior mean has been replaced by the sample average. Thus, coherency is violated despite the use of Bayes formula at the initial step of this statistical procedure (which can be called quasi-Bayesian). Thirdly, within the sampling theory approach, $\hat{\theta}_{EB}$ is a natural point estimate of the vector of random effects. It has the "shrinking" property, since the measurement average y_i corresponding to θ_i is "shrunk" towards the overall average of measurements. So, in $\hat{\theta}_{EB}$ all observations are used to estimate θ_i, not only observations related to this particular effect. While incoherence is a crucial deficiency from the Bayesian point of view, shrinkage estimators have interesting sampling properties. It is therefore important to provide Bayesian interpretation and justification of conducting inference on the basis of $p^*(\theta|y)$.

We seek for the sampling density $\tilde{p}(y|\theta)$ and the prior density $\tilde{p}(\theta)$ that lead to the Bayesian model $\tilde{p}(y,\theta) = \tilde{p}(y|\theta)\tilde{p}(\theta)$ characterised by the joint density of the form

$$\tilde{p}(y,\theta) = p(y|\theta)p(\theta|\alpha = \hat{\alpha})$$

$$\propto g(y;\theta)f_0(\theta;\hat{\alpha}) = f_N^n(y|\theta, cI_n)f_N^n(\theta|\bar{y}e_n, dI_n)$$

which results in

$$p^*(\theta|y) = f_1(\theta|y, \alpha = \hat{\alpha}) \propto g(y;\theta)f_0(\theta;\hat{\alpha})$$

as the true posterior $\tilde{p}(\theta|y)$. Elementary calculations show that

$$\tilde{p}(y,\theta) \propto f_N^n\left(y|\theta, \left(\frac{1}{c}I_n + \frac{1}{dn}e_ne_n'\right)^{-1}\right)\exp\left(-\frac{1}{2d}\theta' M\theta\right),$$

$$\tilde{p}(\theta) = \int_Y \tilde{p}(y,\theta)\,dy \propto \exp\left(-\frac{1}{2d}\theta' M\theta\right), \quad M = I_n - \frac{1}{n}e_ne_n',$$

$$\tilde{p}(y|\theta) = \frac{\tilde{p}(y,\theta)}{\tilde{p}(\theta)} = f_N^n\left(y|\theta, c\left(I_n - \frac{c}{n(c+d)}e_ne_n'\right)\right).$$

Since M is an idempotent singular matrix, the true prior is improper, but σ-finite. It is informative, as it favours approximate equality $\theta_1 \approx ... \approx \theta_n$. In fact, this prior deserves more attention. Consider the non-singular linear transformation of θ into $(\bar{\theta}, \eta)$, where $\bar{\theta} = e_n'\theta/n$ and $\eta_i = \theta_i - \bar{\theta}$ ($i = 1,\ldots, n-1$). Since $\theta' M\theta = \eta'(I_{n-1} + e_{n-1}e_{n-1}')\eta$, $\tilde{p}(\theta)$ leads to $\tilde{p}(\bar{\theta}, \eta) = \tilde{p}(\bar{\theta})\tilde{p}(\eta)$ with $\tilde{p}(\bar{\theta})$ constant and $\tilde{p}(\eta) = f_N^{n-1}\left(\eta|0, d\left(I_{n-1} - \frac{1}{n}e_{n-1}e_{n-1}'\right)^{-1}\right)$. The prior of $\bar{\theta} \in \mathbb{R}$ is improper uniform, but η, the vector of $n-1$ deviations $\theta_i - \bar{\theta}$, is *a priori* normally distributed around 0. The sampling density $\tilde{p}(y|\theta)$ is different from $p(y|\theta)$. The true conditional distribution is normal, like the initially declared one, but it assumes that the observations are equally correlated (instead of being independent). The true sampling covariance matrix leads to the same correlation coefficient for each pair of observations:

$$\widetilde{Corr}(y_i, y_j|\theta) = -\frac{c}{(n-1)c + nd} \quad (i \neq j),$$

which tends to zero when n increases; $\tilde{p}(y|\theta)$ practically coincides with $p\,(y|\theta)$ when n is sufficiently large. However, we cannot use the standard Bayesian asymptotic argument to say that the prior does not matter when n is large, because θ is of dimension n, which is not fixed. Thus the full Bayesian justification of $p^*(\theta|y)$, even asymptotic, requires considering the prior specification as well. Intuitively, $\tilde{p}(\theta)$ is a very reasonable prior. It explains shrinking through giving equal random effects the highest prior chance without introducing any prior information about the average value of all n random effects. In order to show how such prior distribution of individual effects can appear within the fully Bayesian approach, we now consider the Bayesian normal hierarchical model, which introduces one more level – the normal prior distribution of α.

If we assume that

$$p(y|\theta) = f_N^n(y|\theta, cI_n), p(\theta|\alpha) = f_N^n(\theta|\alpha e_n, dI_n), p(\alpha) = f_N^1(\alpha|a, v),$$

we can write

$$p(\theta) = \int_{-\infty}^{+\infty} p(\theta|\alpha)p(\alpha)\, d\alpha = f_N^n(\theta|ae_n, dI_n + ve_ne_n'),$$

$$p(\theta|y) \propto p(y|\theta)p(\theta) = f_N^n(y|\theta, cI_n)p(\theta)$$

or, equivalently,

$$p(\theta|y) = \int_{-\infty}^{+\infty} p(\theta|y,\alpha)p(\alpha|y)\, d\alpha = \int_{-\infty}^{+\infty} f_1(\theta|y;\alpha)p(\alpha|y)\, d\alpha,$$

where

$$p(\alpha|y) = f_N^1\left(\alpha \left|\left(\frac{n}{c+d} + \frac{1}{v}\right)^{-1}\left(\frac{n}{c+d}\bar{y} + \frac{a}{v}\right), \left(\frac{n}{c+d} + \frac{1}{v}\right)^{-1}\right.\right).$$

Since we use the Bayesian approach, where all unknown elements of the statistical model are random variables, it is important to distinguish fixed and random (individual) effects in a Bayesian sense. We use the definition of Koop, Osiewalski and Steel (1997), who call effects *fixed* if they are marginally independent and *random* if they are not. In our example effects θ_i are only conditionally independent (given α), but they are marginally dependent – as it is clear from the non-diagonal covariance matrix of the marginal prior $p(\theta)$. This definition can be extended to σ-finite measures (improper priors). If $p(\theta) = p(\theta_1)p(\theta_2) \dots p(\theta_n)$, then θ_i are fixed effects, otherwise they are random. Note that $\tilde{p}(\theta) \propto \exp\left(-\frac{1}{2d}\theta' M\theta\right)$ cannot be presented as a product of σ-finite measures of individual θ_i, therefore this improper prior describes random effects – similarly as the proper prior $p(\theta)$ does.

Finally, we obtain the marginal posterior density of random effects:

$$p(\theta|y) = f_N^n\left(\theta|(1-w)y + w\left(\frac{n}{c+d}+\frac{1}{\nu}\right)^{-1}\left(\frac{n}{c+d}\bar{y}+\frac{a}{\nu}\right)e_n,\right.$$
$$\left.\frac{cd}{c+d}I_n + w^2\left(\frac{n}{c+d}+\frac{1}{\nu}\right)^{-1} e_n e_n'\right).$$

If $\nu^{-1} = 0$, then $p\,(\alpha) \approx const$, the marginal prior of θ is $p(\theta) \approx \tilde{p}(\theta)$ and the posterior is:

$$p(\theta|y) \approx f_N^n\left(\theta|\hat{\theta}_{EB}, \frac{c}{c+d}\left(dI_n + \frac{c}{n}e_n e_n'\right)\right).$$

Note that the flat prior of α leads to the informative improper marginal prior $\tilde{p}(\theta)$ obtained earlier. However, the above presented posterior $p(\theta|y)$ has a different covariance matrix than $\tilde{p}(\theta|y) = p^*(\theta|y) = f_N^n\left(\theta|\hat{\theta}_{EB}, \frac{cd}{c+d}I_n\right)$; the additional term $\frac{c^2}{n(c+d)}e_n e_n'$ in the posterior covariance matrix reflects uncertainty about α. If both ν and n are large enough, then our fully Bayesian and quasi-Bayesian posteriors approximately coincide: $p(\theta|y) \approx p^*(\theta|y)$ and, thus, small-sample incoherence does not matter.

4. Concluding remarks

In order to somehow validate the use of data-based "priors" (met in incoherent, quasi-Bayesian approaches), a formal method has been proposed in this paper. It amounts to defining and examining the Bayesian model that coherently generates the same posterior distribution as the "posterior" obtained in the original sampling model coupled with the data-based "prior". Although the method is general, only in simple cases it can lead to the closed forms of the true sampling model and the true prior. Therefore, our example includes a normal sampling model with a normal prior structure, always with known variance. The unknown (estimated) prior variance case, fundamental for the parametric empirical Bayes, is left for future research.

As a by-product of our research, we have obtained an important prior structure that favours parameters' equality. This prior is completely uninformative (improper uniform) about the parameters' average, but it is quite informative (proper normal) about the deviations from the average.

Acknowledgement
The author acknowledges support from research funds granted to the Faculty of Management at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

Casella, G. (1985). An introduction to empirical Bayes data analysis. *The American Statistician*, 39, 83–87.

Efron, B. & Morris, C. (1972). Limiting the risk of Bayes and empirical Bayes estimators – Part II: the empirical Bayes case. *Journal of the American Statistical Association*, 67, 130–139.

Greene, W.H. (2008). *Econometric Analysis* (Sixth Edition), Pearson, Upper Saddle River NJ.

Koop, G., Osiewalski, J. & Steel, M.F.J (1997). Bayesian efficiency analysis through individual effects: Hospital cost frontiers. *Journal of Econometrics*, 76, 77–105.

Morris, C. (1983). Parametric empirical Bayes inference: Theory and applications (with discussion). *Journal of the American Statistical Association*, 78, 47–65.

Assessment of the impact of socio-economic situation on health status of inhabitants in the European Union countries

Viera Pacáková[1], Lucie Kopecká[2]

Abstract

Good health status is a priority for citizens of each country and a precondition for economic prosperity. Significant differences in health status exist between European countries and regions. Health status and its supposed determinants are multidimensional categories, specified by a number of indicators. Health indicators based on reliable and mutually comparable data are a prerequisite for establishing and monitoring the implementation of strategies and policies aimed at improving Europeans' health. The European Commission is currently publishing various sets of health indicators that are, and will continue to be, used and disseminated in order to achieve the targets set under the Europe 2020 Strategy.

There is a number of major studies which have demonstrated a clear link between socio-economic background (such as income or occupation) and health. The goal of this article is to assess and quantify inequalities in health status of inhabitants depending on socio-economic situation in European union countries based on selected health, social and economic indicators by using multidimensional methods, namely factor and cluster analysis.

Keywords: *health status, socio-economic situation, inequalities, multidimensional statistical methods*
JEL Classification: *C38, I14, I15*

1. Introduction

Health is important for the wellbeing of individuals and society, but a healthy population is also a prerequisite for economic productivity and prosperity. Our aim was to analyse the socio-demographic and economic factors associated with different levels of health status in European Union countries.

Diseases of the heart and circulatory system (CVD) are the leading cause of mortality in Europe as a whole, responsible for over 3.9 million deaths a year, or 45% of all deaths. CVD is also the leading cause of mortality in the EU, where it causes just over 1.8 million deaths each year – around 800,000 deaths in men and 1 million deaths in women. The main forms of CVD are ischaemic heart disease (IHD) and stroke. As in Europe, IHD and stroke are the first and second most common single causes of death in the EU (Wilkins et al., 2017).

Comparing the CVD mortality burden across individual European countries reveals substantial variation, with a higher burden typically found in Central and Eastern European countries compared to that in Northern, Southern and Western countries (Jindrová and Kopecká, 2017; Pacáková et al., 2016). The incidence of a disease describes the number of new cases that develop within a population over a specified period of time. In 2015, there were just fewer than 11.3 million new cases of CVD in Europe. In the EU there were 6.1 million new cases of CVD

[1] Corresponding author: University of Pardubice, Faculty of Economics and Administration, Institute of Mathematics and Quantitative Methods, 95 Studentská, 532 10 Pardubice, Viera.Pacakova@upce.cz.

[2] University of Pardubice, Faculty of Economics and Administration, Institute of Mathematics and Quantitative Methods, 95 Studentská, 532 10 Pardubice, Lucie.Kopecka1@student.upce.cz.

in 2015. Half of these new CVD cases were due to IHD, while around 10% of new CVD cases were due to stroke (OECD/EU, 2018). CVD has major economic costs as well as human costs for Europe. Overall CVD is estimated to cost the EU economy €210 billion a year. Of the total cost of CVD in the EU, 53% (€111 billion) is due to direct health care costs, 26% (€54 billion) to productivity losses and 21% (€45 billion) to the informal care of people with CVD (Wilkins et al., 2017).

By comparison, cancer – the next most common cause of death – accounts for just fewer than 1.1 million deaths (24%) in men and just fewer than 900,000 deaths (20%) in women. Large variations exist in cancer incidence across EU countries. These variations in incidence rates reflect not only variations in the real number of new cancers occurring each year, but also differences in national policies regarding cancer screening to detect different types of cancer as early as possible, as well as differences in the quality of cancer surveillance and reporting (ECIS, 2019).

There is a large literature examining socio-economics determinants and income inequality in relation to health (Marmot, 2002; Pacáková and Kopecká, 2018; Pickett and Wilkinson, 2015). The social determinants of health are closely inter-linked. Indeed, it makes it hard to empirically disentangle the individual effects of different factors on health. But what is evident is that these factors will, in general, reinforce each other. For example, the better educated are also likely to be richer, live in healthier environments, and be less likely to smoke. Furthermore, some researchers argue that large income differences not only cause health inequalities, but may also be detrimental to population health (Pickett and Wilkinson, 2015).

The goal of this article is to assess and quantify inequalities in health status of inhabitants depending on socio-economic situation in European union countries based on selected health, social and economic indicators by using multidimensional methods, namely factor and cluster analysis. Unlike the most publications for assessing the health status, indicators of incidence were used instead of indicators of mortality due to serious illnesses.

2. Indicators, data and methods

The most of publications in assessing the impact of socio-economic situation on the health status of the population use indicators of mortality of selected diseases. In our opinion, the values of mortality indicators are the result of the impact not only of the socio-economic situation and the quality of life in individual countries but also depend on the level of health expenditure and health care quality. Therefore, to achieve the goal of the article, indicators of incidence of cardiovascular and oncological diseases have been used instead of mortality indicators in line with the Introduction. These indicators are complemented by indicators of the healthy life expectancy of men and women in EU countries.

The used indicators of the socio-economic situation in the EU countries were selected using pairwise Spearman rank correlation coefficients with the selected health indicators. Surprisingly, low dependence of health indicators with unemployment, income inequality indicators and health risk indicators such as alcohol consumption and obesity has been demonstrated, although these indicators are clearly seen as important determinants of health in healthcare publications. The significant impact of risk factors on health may arise if the statistical units are directly people – the inhabitants of countries and not the countries. Unfortunately, indicators of this type are currently not available, the hope for the future can be e-health.

In accordance with the objectives 15 indicators have been selected. The first 8 indicators together characterise the state of health and the last 7 indicators characterize socio-economics situation in EU countries. The values of the indicators have been taken from ECIS (2019), Eurostat (2018), OECD (2016, 2017 and 2018), OECD/EU (2018) and Wilkins et al. (2017).

List of selected indicators:

CVD_I	Incidence of cardiovascular disease, per 100 000, both sexes, 2015
IHD_T	Incidence of ischaemic heart disease, per 100 000, both sexes, 2015
Stroke_T	Incidence of stroke, per 100 000, both sexes, 2015
Cancer_I	Estimated cancer incidence, per 100 000, both sexes, all sites but non-melanoma skin, all ages, 2018
HLY_W	Healthy life years, women, 2016 (or nearest year)
HLY_M	Healthy life years, men, 2016 (or nearest year)
HLY_65_W	Healthy life years at 65, women, 2016 (or nearest year)
HLY_65_M	Healthy life years at 65, men, 2016 (or nearest year)
GDP	GDP per capita in EUR PPP, 2017
Mean	Mean equivalised net income, 2017 or latest available year, Euro
Median	Median equivalised net income, 2017 or latest available year, Euro
Poverty	At risk of poverty rate, 2016
Deprivation	Material deprivation, 2016
Smoking	% of population who smokes daily
Edu_high	Share of the population (%) age 25–54 with tertiary educational attainment

Two multidimensional methods were used to achieve the goal of the article, Factor and Cluster Analysis.

Factor analysis is a statistical approach that can be used to analyse interrelationships among a large number of variables and to explain these variables in terms of their common underlying factors. The general purpose of factor analytic techniques is to find a way of condensing (summarising) the information contained in a number of original variables into a smaller set of new composite factors with a minimum loss of information. Numerous variations of the general factor model are available. The component model is used when the objective is to summarise most of the original information (variance) in a minimum number of factors. An important concept in factor analysis is the rotation of factors. In practice, the objective of all methods of rotation is to simplify the rows and columns of the factor matrix to facilitate interpretation. The *Varimax* criterion centres on simplifying the columns of the factor matrix. With the Varimax rotation approach, there tend to be some high loadings (i.e., close to –1 or +1) and some loadings near 0 in each column of the matrix. The *factor loadings* show the correlation between the original variables and the factors and they are the key to understanding the nature of a particular factor. The *Factor Scores* in output of Factor analyse procedure display the values of the rotated factor scores for each of n cases, in our analysis for each of 28 European Union countries (Hair et al., 2007).

Cluster Analysis procedure is designed to group observations (countries) into clusters based upon similarities between them. A number of different algorithms is provided for generating clusters and are described in detail in many statistical publications. We used the agglomerative algorithm, starting with separate clusters for each observation and then joining clusters together based upon their similarity. To form the clusters, the procedure began with each observation in a separate group. It then combined the two observations which were closest together to form a new group. After re-computing the distance between the groups, the two groups then closest together are combined. This process is repeated until only one group remained. The results of the analysis are displayed in a *dendogram* (Hair et al., 2007).

3. Results and discussion

To assess the suitability of indicators for the factor analysis, we applied the Kaiser-Meyer-Olkin measure (KMO). The KMO = 0.709747 shows suitability of the source variables for factor analysis. In this case, 3 factors were extracted, since 3 factors had eigenvalues greater than or equal to 1. Together they account more than 78 % of the variability in the original data (Table 1).

Table 1. Factor analysis – number of factor extracted

Factor number	Eigenvalue	Percent of Variance	Cumulative Percentage
1	8.21701	54.780	54.780
2	2.57545	17.170	71.950
3	1.03643	6.910	**78.859**
4	0.955858	6.372	85.232
⋮	⋮	⋮	⋮
15	0.00094131	0.006	100.000

Factor loadings present the correlation between the original variables and the extracted factors and they are the key to understanding the nature of a particular factor. After varimax rotation factor loadings shown in Table 2 have been obtained. Rotation is performed in order to simplify the explanation and naming of the factors.

Based on Factor loadings in Table 2 we found out that the *Factor 1* has the strong negative correlation with the indicators of CVD incidence and strong positive correlation with income indicators. The *Factor* 2 demonstrates strong positive correlation with the indicators of healthy life years. The *Factor* 3 shows strong positive correlation with the indicators of poverty, material deprivation and smoking and significant negative correlation with incidence of cancer. So the extracted factors were interpreted as follows: Factor 1 – *factor of CVD incidence and income level*, Factor 2 – *factor of healthy life years* and Factor 3 – *factor of social exclusion and cancer incidence*.

Table 2. Factor loading matrix after Varimax Rotation

Variables	Factor 1	Factor 2	Factor 3
CVD_I	**–0.8589**	–0.3604	0.1395
IHD_T	**–0.8675**	–0.3410	0.0906
Stroke_T	**–0.7209**	–0.4165	0.3555
Cancer_I	0.1439	–0.0415	**–0.7179**
HLY_W	0.0319	**0.95255**	0.1174
HLY_M	0.2224	**0.9566**	0.0349
HLY_65_W	0.3021	**0.8682**	–0.3021
HLY_65_M	0.4319	**0.8461**	–0.2402
GDP	**0.7705**	0.1152	–0.3083
Mean	**0.8121**	0.2282	–0.4401
Median	**0.7972**	0.2301	–0.4620
Poverty	–0.4724	–0.0668	**0.7321**
Deprivation	–0.2259	–0.0246	**0.7796**
Smoking	–0.3279	–0.2841	**0.6045**
Edu_high	**0.6076**	–0.0099	–0.2603

Table 3 shows the factor scores for each monitored country. The Factor Scores displays the values of the rotated factors for each country. Graphical display of countries in a two-dimensional coordinate system with the axes of the extracted factors allows us to quickly assess the observed situation in each country and also to compare the situation in different countries.

In the coordinate system of the factors *Factor 1* and *Factor 2* (Fig. 1) three main groups of countries were created. The first one with high values of both factors, including all the old EU member countries, the second one with low values of both factors, including the new EU countries and the third one with the middle level of the first and the low to medium level of the second factor. In countries with a high level of income and education and a low incidence of CVD there is a high level of healthy life years.

Fig. 2 illustrates the high inequalities in social inclusion and incidence of cancer and healthy life in EU countries. This inequality can also be caused by the fact that high values of *Factor 3* indicate high social exclusion and a high proportion of smokers, and paradoxically a relatively low incidence of cancer. Unfortunately, the Fig. 2 shows again a low level of healthy years of life and a high level of social exclusion in former socialist countries.

Table 3. Table of factor scores

Country	Code	Factor 1	Factor 2	Factor 3
Austria	AT	2.619	–1.802	–0.979
Belgium	BE	4.542	2.783	–3.274
Bulgaria	BG	–10.076	–1.715	8.572
Croatia	HR	–7.947	–6.363	4.415
Cyprus	CY	3.238	2.685	0.992
Czech Republic	CZ	–3.021	–0.751	–0.699
Denmark	DK	6.796	2.691	–5.980
Estonia	EE	–4.394	–5.200	1.939
Finland	FI	4.262	–0.147	–3.879
France	FR	4.788	2.399	–3.164
Germany	DE	2.667	3.994	–2.221
Greece	EL	–3.512	–0.809	4.525
Hungary	HU	–6.668	–4.574	2.139
Ireland	IE	8.872	6.791	–4.883
Italy	IT	–0.032	3.119	1.114
Latvia	LV	–8.507	–8.932	4.475
Lithuania	LT	–7.345	–6.303	4.049
Luxembourg	LU	12.081	3.840	–6.339
Malta	MT	3.484	7.977	–0.754
Netherlands	NL	5.554	1.583	–4.365
Poland	PL	–4.575	–1.529	2.453
Portugal	PT	–1.912	–2.346	1.697
Romania	RO	–9.923	–5.353	7.836
Slovakia	SK	–4.433	–5.832	1.329
Slovenia	SI	–1.241	–2.298	–1.166
Spain	ES	2.343	3.131	0.989
Sweden	SE	7.193	10.027	–5.061
United Kingdom	UK	5.144	2.936	–3.761

The factor analysis based on principal component method resulted in 3 mutually independent factors. These factors are appropriate for the cluster analysis.

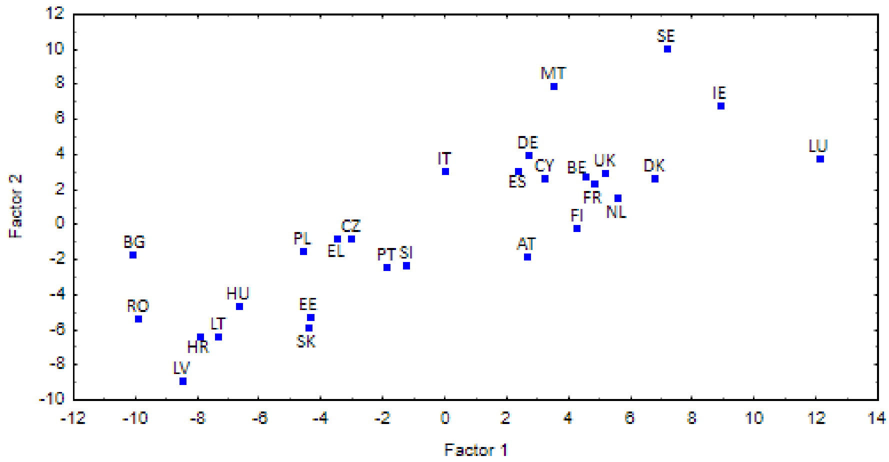


Fig. 1. Location of EU countries according to Factor 1 and Factor 2

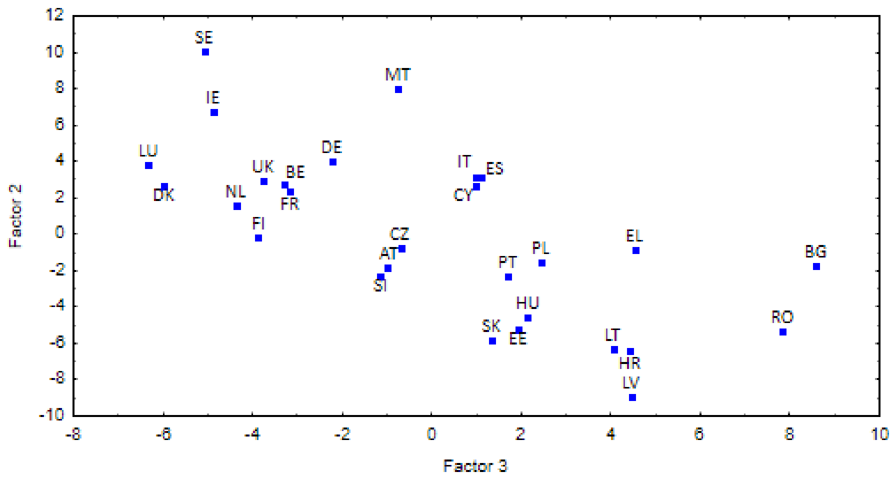


Fig. 2. Location of EU countries according to Factor 3 and Factor 2

The cluster analysis results presented by the dendrogram in Fig. 3 are consistent with the results of factor analysis. In particular, it presents a very large gap between Western EU countries and EU countries in Central, Eastern and Southern Europe by selected health and socio-economic indicators.

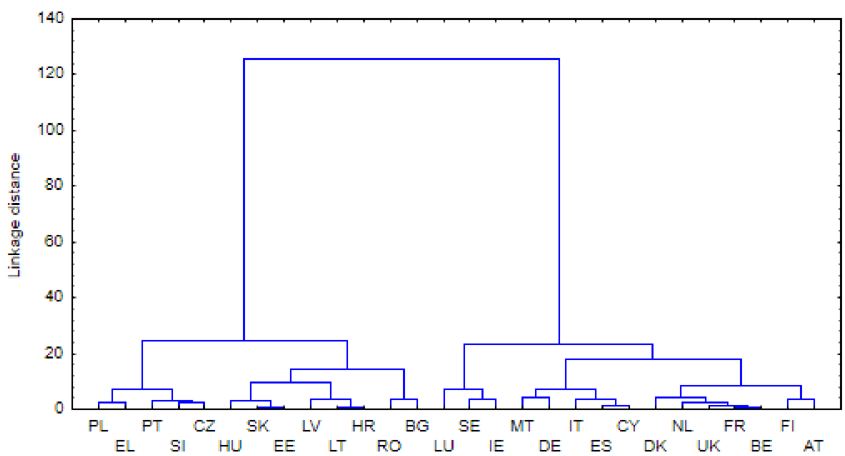


Fig. 3. Cluster analysis results

Conclusions

Large disparities exist between countries in CVD and cancer incidence and healthy life years and their significant dependence on the social and economic situation in the EU countries has been confirmed. The results of factor and cluster analysis showed significant health and social disparities in the EU countries, mainly a large gap between Western EU countries and EU countries in Central, Eastern and Southern Europe. The objectives of the European Commission concerning the gradual reduction of these inequalities, which have been declared in European Commission (2013) and OECD (2015), apparently not successful.

A wide range of policies is required to reduce inequalities. These include greater efforts to prevent health problems starting early in life, promote equal access to care for the whole population, and better manage chronic health problems when they occur to reduce their disabling effects (OECD, 2017).

Reducing inequalities in health in the EU countries necessarily requires reducing inequality in income levels and social exclusion of the population.

Acknowledgements

This article was prepared with the support of the research project SGS_2019_018, funded by the University of Pardubice.

References

ECIS (2019). *European Cancer Information System*. Retrieved from: https://ecis.jrc.ec.europa.eu.

European Commission (2013). *Improving health for all EU citizens*. Luxembourg: Publications Office of the European Union. Retrieved from: https://ec.europa.eu/health//sites/health/files/health_policies/docs/improving_health_for_all_eu_citizens_en.pdf.

Eurostat (2018). *People at Risk of Poverty or Social exclusion*. Retrieved from: https://ec.europa.eu/eurostat/statistics%20explained/index.php/People_at_risk_of_poverty_or_social_exclusion.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. & Tatham, R.L. (2007). *Multivariate Data Analysis*. New Jersey: Pearson Education.

Jindrová, P. & Kopecká, L. (2017). Assessment of risk factors of serious diseases in OECD countries. 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, 123–132.

Marmot, M.G. (2002). The Influence of Income on Health: Views of an Epidemiologist. *Health Affairs* 21, 31–46.

OECD (2015). *Health Inequalities.* Retrieved from http://www.oecd.org/health/inequalities-in-health.htm.

OECD (2016). *Society at a Glance 2016: OECD Social Indicators*. OECD Publishing, Paris.

OECD (2017). *Health at a Glance 2017: OECD Indicators*. OECD Publishing, Paris.

OECD/EU (2018). *Health at a Glance: Europe 2018: State of Health in the EU Cycle*. OECD Publishing, Paris/EU, Brussels.

Pacáková, V., Jindrová, P. & Zapletal, D. (2016). Comparison of Health Care Results in Public Health Systems of European Countries. *European Financial Systems 2016,* Vol. XIII, 534–541.

Pacáková, V. & Kopecká, L. (2018). Inequalities in Health Status Depending on Socio-economic Situation in the European Countries*. Ekonomie a Management,* 21(2), 4–20.

Pickett, K.E., & Wilkinson, R.G. (2015). Income inequality and health: a causal review. *Social Science & Medicine*, 128, 316–326.

Wilkins, E., Wilson, L., Wickramasinghe, K., Bhatnagar, P., Leal, J., Luengo-Fernandez, R., Burns, R., Rayner, M.R & Townsend, N. (2017). *European Cardiovascular Disease Statistics 2017.* European Heart Network, Brussels.

The missing data problem in statistical surveys on the example of the SOF-1 report (Report on the activities of foundations, associations and similar social organisations)

Barbara Pawełek[1], Tomasz Sekuła[2]

Abstract

The problem of missing data is often encountered when carrying out statistical research. To properly address this issue, the factors contributing to the lack of data must be investigated. The analysis is based on data from a study conducted for 2014 via the SOF-1 Reporting Form. The purpose of this research is to establish the level of economic and social resources of the non-profit organisations operating in Poland and to provide a description of their activities. The aim of this work is to present the results of empirical research into existing correlations between the missing answers in the SOF-1 report and the characteristics of the research subjects. The added value of the work is to demonstrate that the method used to visualise the distribution of missing data in a database and the association rules can be successfully applied to the analysis of correlations between the missing answers in the SOF-1 report and the profile of the respondents. The identified correlations may help researchers achieve a higher level of completeness for the SOF-1 report in terms of answers given. The potential benefits of this approach include reduced study costs and more accurate generalisations. The findings of this analysis can be used to identify mechanisms which contribute to generating missing data and thus to select the right method for estimating missing information.

Keywords: *non-profit organisations, missing data, visualisation methods, association rules*

JEL Classification: *L310, C810, C100*

1. Introduction

The problem of missing data is often encountered when carrying out statistical research. A number of different reasons for missing information in statistical databases are reported by the literature (Eaton et al., 2005). The incomplete nature of data sets poses a major obstacle to conducting an effective statistical analysis, assessing the status and dynamics of the studied phenomenon and making rational decisions. For this reason, the public institutions involved in collecting statistical data strive to ensure a high level of completeness of the data sets intended for use by state and local government authorities, businesses and research centres.

The present study involving the SOF-1 Reporting Form, *Report on the activities of foundations, associations and similar social organisations*, is part of a wider research into the non-profit sector carried out by the Statistics Poland (*Główny Urząd Statystyczny*, GUS). The SOF-1 report is intended to address the increasing importance of civil society, provide data to assess the progress made in implementing public policies which support social economy and social capital and to monitor the activities of non-profit organisations in Poland.

[1] Corresponding author: Cracow University of Economics, Department of Statistics, 27 Rakowicka St., 31-510 Cracow, Poland, e-mail: barbara.pawelek@uek.krakow.pl.

[2] Statistical Office in Krakow, Research Centre for Social Economy, 3 Kazimierza Wyki St., 31-223 Cracow, Poland, e-mail: t.sekula@stat.gov.pl.

Once the entities (selected for research) which gave incomplete answers in the SOF-1 report are identified, we contact them to obtain the missing answers. Contact is made by phone or via e-mail. One of the objectives of the research undertaken by the Statistical Office in Cracow and the Department of Statistics at Cracow University of Economics is to define the profile of the entities/respondents who failed to provide certain answers in the survey conducted using the SOF-1 Reporting Form.

The aim of this work is to present the selected results of empirical research into existing correlations between the missing answers in the SOF-1 report and the characteristics of the research subjects. The added value of the work is to demonstrate that the method used to visualise the distribution of missing data in a database and the association rules can be successfully applied to the analysis of correlations between the missing answers in the SOF-1 report and the profile of the respondents.

To our knowledge, there is no previous record of the association rules being applied to the analysis of correlations between the missing answers and the profile of the respondents. Likewise, we are not aware of any reported case of the visualisation methods for distribution of missing data in a database and the association rules being used to examine the activities of foundations, associations and similar social organisations.

2. Data

Empirical data for analysis were collected via the SOF-1 Reporting Form as part of the study carried out in Poland for 2014 (database status as of May 20*th* 2015).

The SOF-1 report was targeted at selected non-profit organisations. According to the definition provided in the United Nations' Handbook on Non-Profit Institutions (*Satellite Account…*, 2018), the non-profit sector includes all entities that are: institutionalised to some extent (i.e. following registration with a competent public office) or have some degree of persistence of goals, structure and activities; institutionally separate from government; non-profit making (do not exist primarily to generate profits and do not return profits generated to their owners, employees, etc.); self-governing (able to control their own activities); involve some meaningful degree of voluntary participation.

The SOF-1 report was used to gather information on the profile of selected non-profit organisations, determine their economic and social potential, development strategies and forms of activity and customer profiles. The form was divided into eight sections:

- field of activity, scope, form and nature of activity the reporting entity's (Section I),
- participation in the projects implemented under the European Social Fund (Section II),
- the type and number of beneficiaries of the activities carried out by the reporting entity and the type of activities undertaken for the benefit of individual consumers (Section III),
- members and volunteers, their social work (Section IV),
- employed on the basis of employment contract and civil law contracts (Section V),
- operating revenues and expenses (Section VI),

- fixed assets held, capital expenditure incurred and in-kind gifts received (Section VII),
- operating conditions (Section VIII),
- contact details.

The Database of Statistical Units (*Baza Jednostek Statystycznych*, BJS) served as a sampling frame for the statistical research involving the selection of non-profit organisations having legal personality, having a legal form which meets the target group criteria for the study and which are recorded in the BJS as having a specific legal and economic status of: an active entity engaged in an activity; an active entity not engaged in an activity, but preparing to engage in one.

The entities were selected for inclusion in statistical records using the following methods:

- purposive sampling was used for a highly varied sample of entities. This type of sampling was applied to: entities with registered offices, entities registered as Public Benefit Organisations, project promoters for the European Social Fund (ESF), entities employing more than 5 people, denominational entities engaged in social activities;
- stratified random sampling with proportional allocation was used for other entities. The strata represented Polish voivodships with a focus on the largest cities: Warsaw, Cracow, Lodz, Wroclaw and Poznan; legal form: foundations, associations and similar social organisations. The following samples were taken from the stratum which represented Polish associations: volunteer firefighters, hunting clubs, sports associations and other organisations.

The research involving the SOF-1 report was conducted in the spring of 2015 by the Statistical Office in Cracow. Data for research were collected from 24 000 entities, out of the 33 500 that qualified for the research. As of December 31st 2015, the level of completeness for the study as whole was recorded at 85%. In addition, the entities that the authors had been unable to contact the time before that took part in a registration survey carried out in September and October 2015 aimed at bringing greater precision to the generalisation of the research data. The results of the registration study, as well as the administrative data on employment from The Polish Social Insurance Institution (*Zakład Ubezpieczeń Społecznych,* ZUS) and the revenue data from the Ministry of Finance, were used to define the sampling weights for each result set, allowing for generalisation of the data collected during proper research.

3. Methods

When identifying the most effective ways of dealing with the problem of missing data in a database, the contributing factors should be investigated. Rubin (1976) defined three kinds of missing data mechanisms: missing completely at random, missing at random and missing not at random. Identifying a mechanism which contributes to generating missing data is not an easy task. One approach is to visualise the distribution of missing data in a database to exclude the possibility of missing not at random (Schafer and Graham, 2002).

The graphical methods for data presentation are widely used in statistical surveys (Kelleher and Wagener, 2011). These methods are also used to visualise the distribution of missing data in a database (Templ et al., 2012).

Another tool used in the study is 'association rule' (Kotsiantis and Kanellopoulos, 2006). Association models represent the co-occurrence of values/variants in a dataset. Association rule is an expression $A \Rightarrow B$ which states that if A occurs, then B occurs with a certain probability. A number of measures can be used to assess the degree of confidence in the association rule, including:

- $\text{support}_{AB} = P(A \cap B)$ – provides information about the probability of co-occurrence of both A and B;
- $\text{confidence}_{AB} = P(B|A) = \frac{P(A \cap B)}{P(A)}$ – provides information about the conditional probability of B in sets that contain A;
- $\text{lift}_{AB} = \frac{P(A \cap B)}{P(A)P(B)} = \frac{\text{support}_{AB}}{P(A)P(B)} = \frac{\text{confidence}_{AB}}{P(B)}$ – provides information about the change in the probability of B given the presence of A.

The study involved applying the methods used to visualise the distribution of missing data in a database and the association rules to define the profiles of the entities/respondents which failed to provide certain answers in the study conducted using the SOF-1 Reporting Form.

The analysis uses the R package *VIM* and the *Statistica* program.

4. Results

The database contained 11 345 records (the number of analysed research subjects) and about 300 variables. The variables represented the questions contained in the SOF-1 form. The following presentation of the results was limited only to questions from Section VIII (Operating conditions). Section VIII contained the following questions:

- Question d8_1_lok: As of December 31st, 2014 – did the entity have access to the premises? d8_1_l1: was the entity the owner (co-owner) of the premises – in m^2; d8_1_l2: did the entity use the premises on a rental basis – in m^2; d8_1_lok_n: did the entity use the premises free of charge (e.g. lending for use) – in m^2.
- Question d8_1_sam: As of December 31st, 2014 – did the entity have access to cars? d8_1_s1: was the entity the owner (co-owner) of cars – specify quantity; d8_1_s2: did the entity use cars on a rental basis – specify quantity; d8_1_sam_n: did the entity use cars free of charge (e.g. lending for use) – specify quantity.
- Question d8_1_kom: As of December 31st, 2014 – did the entity have access to computers? d8_1_k1: was the entity the owner (co-owner) of computers – specify quantity; d8_1_k2: did the entity use computers on a rental basis – specify quantity; d8_1_kom_n: did the entity use computers free of charge (e.g. lending for use) – specify quantity.
- Question d8_2: What were the most serious operating problems encountered by the entity in 2014? The question had 12 variants and more than one answer could be selected.

Table 1 shows the number of missing answers to questions from Section VIII.

Table 1. Number of missing answers to questions from Section VIII

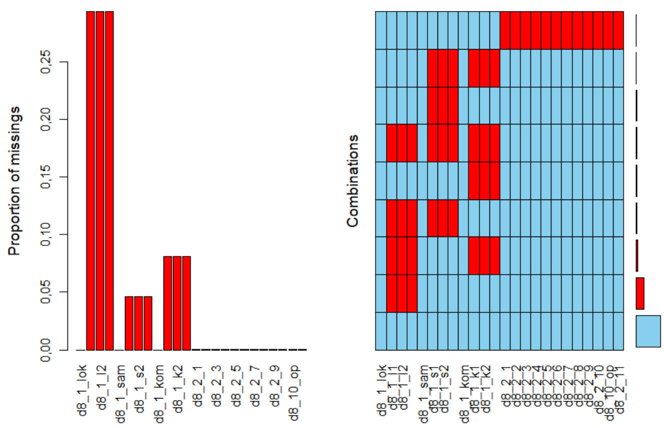
Question	Number of missing answers	Question	Number of missing answers
d8_1_lok	0	d8_2_1	8
d8_1_l1	3332	d8_2_2	8
d8_1_l2	3332	d8_2_3	8
d8_1_lok_n	3332	d8_2_4	8
d8_1_sam	0	d8_2_5	8
d8_1_s1	524	d8_2_6	8
d8_1_s2	524	d8_2_7	8
d8_1_sam_n	524	d8_2_8	8
d8_1_kom	0	d8_2_9	8
d8_1_k1	923	d8_2_10	8
d8_1_k2	923	d8_2_10_op	8
d8_1_kom_n	923	d8_2_11	8

Table 2 shows the combinations of given and missing answers to questions from Section VIII (1 – no answer, 0 – answer given). Data visualisations for respective proportions of specific types of combinations of answers in the total number of analysed research subjects can be found on the right-side graph in Figure 1.

Table 2. Combinations of given and missing answers to questions from Section VIII

Combination of answers	Number	Share (in %)
0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0	7677	67.669
0:1:1:1:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0	2407	21.216
0:1:1:1:0:0:0:0:1:1:1:0:0:0:0:0:0:0:0:0:0:0:0:0	516	4.548
0:1:1:1:0:1:1:1:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0	239	2.107
0:0:0:0:0:0:0:0:1:1:1:0:0:0:0:0:0:0:0:0:0:0:0:0	213	1.877
0:1:1:1:0:1:1:1:0:1:1:1:0:0:0:0:0:0:0:0:0:0:0:0	170	1.498
0:0:0:0:0:1:1:1:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0	91	0.802
0:0:0:0:0:1:1:1:0:1:1:1:0:0:0:0:0:0:0:0:0:0:0:0	24	0.212
0:0:0:0:0:0:0:0:0:0:0:0:1:1:1:1:1:1:1:1:1:1:1:1	8	0.071

Note: 1 – no answer, 0 – answer given



Note: missing answers are marked in red, answers given are marked in blue.

Fig. 1. Distribution of the missing answers for questions from Section VIII by question (left graph) and proportions of the combinations of answers and missing answers for questions from Section VIII in the total number of research subjects

The following section of this study shows the results of applying the methods used to visualise the distribution of missing data in a database and the association rules to the analysis of correlations between the missing answers in the SOF-1 report and the characteristics of the respondents, using question d8_1_l1 as an example.

Figure 2 shows the distribution of missing answers to question d8_1_l1 by three selected characteristics of the respondents:

- 'Organisation' variable: 1 – Foundations, 2 – Volunteer firefighters, 3 – Physical culture associations and sports associations, 4 – Hunting clubs, 5 – Typical associations and social organisations, 6 – Denominational entities engaged in social activities;
- 'Revenue class' variable: 1 – no revenue, 2 – with revenue below PLN 1 000, 3 – with revenue of PLN 1 000 to 10 000, 4 – PLN 10 000 to 100 000, 5 – PLN 100 000 to 1 million, 6 – over PLN 1 million;
- 'Urbanisation' variable: *ug* – urban gmina, *cps* – city with powiat status, *urg* – urban-rural gmina, *rg* – rural gmina.

Employees from the Statistical Office in Cracow selected these characteristics drawing on their knowledge and experience. To provide greater clarity of the following graphical presentation of results, only two types of organisations were examined, i.e. foundations (variant 1) and typical associations and social organisations (variant 5). The selected types of organisations represent the largest share of the studied population. The dimensions of the rectangles (width and height) depend on the number of units of a particular type in the total number of analysed research subjects, while the area of the rectangles represents the number of possible answers to question d8_1_l1 (i.e. the number of records in a database). The rectangle area marked in red represents the number of missing answers in the total number of analysed research subjects. The rectangle area marked in blue represents the share of answers given.

Fig. 2. Distribution of missing answers to question d8_1_11 by three selected characteristics of the respondents

Figure 2 shows that the highest level of completeness in terms of answers given can be observed for the reporting forms which have been completed by foundations (variant 1) generating more than PLN 1 million in revenue (variant 6), based in the cities with powiat status (variant *cps).* Whereas the lowest level of completeness in terms of answers given was reported for typical associations and social organisations (variant 5) with a revenue of PLN 100 000 to 1 million (variant 5), which are based in the rural gminas (variant *rg*).

The methods used to visualise the distribution of missing data in a database provided a valuable first insight into existing correlations between the missing answers in the SOF-1 report and the profile of the respondents. A graphical analysis can be used to select the potential characteristics of the respondents that can be helpful in defining the profile of respondents who failed to answer certain questions included in the study conducted using the SOF-1 Reporting Form.

The next step was to apply association rules based on the assumption that the consequent must be a missing answer variable for question d8_1_11, while the antecedent must be variants of three selected characteristics of the examined entities. The minimum support for association rules was set at 5%, and the minimum confidence – at 20%. The results of the analysis are presented in Table 3.

The highest lift value (1.523) was reported for the association rule where the antecedent is variant 6 of the 'Organisation' variable (denominational entities engaged in social activities). The association rule involving variant *rg* of the 'Urbanisation' variable (i.e. rural gmina) was ranked second in terms of lift value (1.476). A lift value of more than 1 was reported also for the association rules with the following antecedents: a revenue of PLN 10 000 to 100 000, a revenue of PLN 100 000 PLN to 1 million, and typical associations and social organisations with a revenue of PLN 100 000 to 1 million.

The results obtained using association rules are consistent with those of the graphical analysis. The reason why entities with the following three characteristics: typical associations and social organisations (variant 5) with a revenue of PLN 100 000 to 1 million (variant 5) which are based in the rural gminas (variant *rg*) did not co-occur in the antecedent – despite this being supported by the visualisation of missing answers – is that the minimum support was adopted for the association rule, i.e. 5%.

Table 3. Association rules for question d8_1_l1

Antecedent	⇒	Consequent	Support (%)	Confidence (%)	Lift
Organisation = 5	⇒	d8_1_l1 = *NA*	13.7	28.2	0.959
Urbanisation = *cps*	⇒	d8_1_l1 = *NA*	11.7	22.5	0.767
Revenue = 5	⇒	d8_1_l1 = *NA*	11.4	30.9	1.051
Urbanisation = *rg*	⇒	d8_1_l1 = *NA*	9.5	43.4	1.476
Revenue = 4	⇒	d8_1_l1 = *NA*	9.0	33.0	1.123
Organisation = 5 Revenue = 5 }	⇒	d8_1_l1 = *NA*	5.6	30.5	1.037
Organisation = 6	⇒	d8_1_l1 = *NA*	5.0	44.7	1.523

By establishing contact with the entities selected on the basis of analysis, we can improve the level of completeness in terms of the answers given to question d8_1_l1.

Conclusions

The methods used to visualise the distribution of missing data in a database and the association rules can be successfully applied to examine the correlations between the missing answers in the statistical research and the profile of the respondents.

The method of analysis outlined in this paper can be performed for all questions in the SOF-1 form. The results of such analysis can be used to increase the level of completeness of the answers given in the SOF-1 report, resulting in reduced study costs, more accurate generalisations, etc.

The research findings can be used to identify data missing mechanisms and thus to select the right method for estimating missing information.

Further research is planned, this time including a breakdown into voivodships. If continued, this research will foster cooperation between the Statistical Office in Cracow and the Department of Statistics at Cracow University of Economics.

Acknowledgements

Publication was financed from the funds granted to the Faculty of Management at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

Eaton, C., Plaisant, C., & Drizd, T. (2005). Visualizing Missing Data: Graph Interpretation User Study. In: Costabile, M.F., Paternò, F. (eds.), *Human-Computer Interaction – INTERACT 2005. INTERACT 2005*, Lecture Notes in Computer Science, vol. 3585, Springer, Berlin, Heidelberg, 861–872.

Kelleher, Ch., & Wagener, T. (2011). Ten guidelines for effective data visualization in scientific publications. *Environmental Modelling & Software*, 26(6).

Kotsiantis, S., & Kanellopoulos, D. (2006). Association Rules Mining: A Recent Overview. *GESTS International Transactions on Computer Science and Engineering*, 32(1), 71–82.

Rubin, D.B. (1976). Inference and missing data. *Biometrika*, 63(3).

Satellite Account on Non-profit and Related Institutions and Volunteer Work (2018), United Nations, New York.

Schafer, J.L., & Graham, J.W. (2002). Missing Data: Our View of the State of the Art. *Psychological Methods*, 7(2), 147–177.

Templ, M., Alfons, A., & Filzmoser, P. (2012). Exploring incomplete data using visualization techniques. *Advances in Data Analysis and Classification*, 6(1), 29–47.

Identification of factors that can cause mobile phone customer churn with application of symbolic interval-valued logistic regression and conjoint analysis

Marcin Pełka[1], Aneta Rybicka[2]

Abstract

The Polish mobile market is a fast growing one: according to forecasts, in 2020 the whole telecommunication market will be worth 47.78 bln PLN. At the end of 2017 the number of active sim cards reached 53.3 mln. At the same time, post-paid services were more popular than pre-paid services. As customers can easily change their current phone operator to another, there arises a key question: what factors can cause customer churn on the Polish mobile phone market? To identify the main factors, conjoint analysis and symbolic logistic regression (with centers approach for model estimation) are used. Both techniques allow different groups of factors to be identified. Conjoint analysis focuses on preferences, while symbolic logistic regression allows main factors of customer loyalty to be identified. Both results complement each other and allow a more comprehensive look at mobile customer churn.

Keywords: *customer churn, conjoint analysis, symbolic interval-valued logistic regression, mobile phone market*

JEL Classification: *C01, C81, D12*

1. Introduction

Customer churn, known also as customer turnover, customer attrition or customer deflection, is a major concern for a number of industries (e.g. banks, internet service providers, insurance companies and telephone service companies). Customer churn is particularly acute in the competitive and quite liberalised mobile telecommunication industry (Keaveney 1995).

The costs of gaining new customers are usually five to even six times higher than the costs of retaining an existing customer (Bhattacharya, 1998). Such costs vary when considering different countries and can vary from 300 USD to 600 USD (Athanassopoulos, 2000).

There are many studies concerning customer churn in general and mobile customer churn (see for example Xia and Jin 2008; Richter, Yom-Tov and Slonim 2010; Neslin et. al. 2006; Burez and Van den Poel 2009; Śmiatacz 2012). But usually these studies are focused only on one side of the customer churn problem – customer loyalty – without taking into account elements of the current offer (especially customer preferences).

As Polish customers can change their mobile phone operator to another one quite easily, the following question arises: what factors can cause churn on the Polish mobile phone market? The main aim of the paper is to identify main factors that can cause mobile phone customer churn. To obtain this goal, conjoint analysis and symbolic logistic regression are used (with the

[1] Wroclaw University of Economics, Faculty of Economics, Management and Tourism, Department of Econometrics and Computer Science, 118/120 Komandorska St., 53-345 Wroclaw, e-mail: marcin.pelka@ue.wroc.pl.

[2] Wroclaw University of Economics, Faculty of Economics, Management and Tourism, Department of Department of Econometrics and Computer Science, 118/120 Komandorska St., 53-345 Wroclaw, e-mail: aneta.rybicka@ue.wroc.pl.

centers method for model estimation). Both techniques allow different groups of factors to be identified. Conjoint analysis focuses on preferences, while symbolic logistic regression allows main factors of customer loyalty to be identified. The results are supplementary to each other and allow us to get a more comprehensive look at mobile customer churn.

2. Symbolic logistic regression and ensemble symbolic logistic regression

Symbolic objects can be described by the following variable types (see e.g. Bock and Diday 2000; Diday and Noirhomme-Fraiture 2008; Billard and Diday 2006; Noirhomme-Fraiture and Brito 2011):

1. Quantitative (numerical) variables:
   a) numerical single-valued variables,
   b) numerical multi-valued variables,
   c) interval-valued variables,
   d) histogram variables.
2. Qualitative (categorical) variables:
   a) categorical single-valued variables,
   b) categorical multi-valued variables,
   c) categorical modal variables.

In general, this kind of data allows objects to be described more precisely, but it requires new, special methods and algorithms. More details about symbolic variables and objects can be found in e.g. Bock and Diday (2000), Billard and Diday (2006), Diday and Noirhomme-Fraiture (2008), Noirhomme-Fraiture and Brito (2011).

In logistic regression for symbolic data, as in logistic regression for classical data, we model binomial (binary, dichotomous) variables (e.g. y – employment status: 1 – employed, 0 – unemployed), while explanatory variables are symbolic interval-valued variables. The general multivariate regression model can be written as:

$$Y_t = b_0 X_{0t} + b_1 X_{1t} + \cdots + b_m X_{mt} + e_t = \sum_{j=0}^{m} b_j X_{jt} + e_t \, , \tag{1}$$

where: Y – dependent variable, $X_0, X_1, \ldots, X_m$ – explanatory (dependent) variables, $b_0, b_1, \ldots, b_m$ – model coefficients, e – model error, $t = 1, \ldots, T$ – observation number, $j = 0, 1, \ldots, m$ – variable number.

In the logit model we assume that we are dealing with a latent variable y^* that cannot be directly observed. However we can observe:

$$y_i = \begin{cases} 1 & \text{if } y^* > 0 \\ 0 & \text{if } y^* \leqslant 0 \end{cases} . \tag{2}$$

The probability that the independent variable y_i will be 0 or 1 is defined as follows:

$$P_i = F(x_i^T b) = \frac{1}{1 + \exp(-x_i^T b)} = \frac{\exp(x_i^T b)}{1 + \exp(-x_i^T b)} . \tag{3}$$

Symbolic interval-valued data is represented by an interval $[\underline{x}_i \ \overline{x}_i]$, where $\underline{x}_i$ – is the lower bound of an interval-valued variable, and $\overline{x}_i$ – is the upper bound of an interval-valued variable.

De Souza et. al. (2011) describe four different approaches of probability estimation (de Souza et. al. 2011, 275–278):

1. The centers method, where each interval is represented by its center (midpoint) $\frac{\underline{x}_i+\overline{x}_i}{2}$. The probability (3) is estimated for centers of explanatory (independent) variables.
2. The lower and upper bounds method, where each interval is represented by its $\overline{x}_i$ lower and upper bound, respectively. The probability (3) can be estimated:
   a) conjointly for lower and upper bounds, so we get $2m$ variables, and the final probability is estimated for them,
   b) separately for lower and upper bounds (so two different models are estimated – one for lower, and one for upper bounds). Final probability is calculated as the mean of these two probabilities.
3. The vertices method, where instead of symbolic interval-valued variables we have a combination of lower and upper bounds that are represented by an **M** matrix. If we have one object and two interval-valued variables $[\underline{x}_{11} \ \overline{x}_{11}]$, $[\underline{x}_{21} \ \overline{x}_{21}]$, then the **M** matrix is:

$$\mathbf{M} = \begin{bmatrix} \underline{x}_{11} & \underline{x}_{21} \\ \underline{x}_{11} & \overline{x}_{21} \\ \overline{x}_{11} & \underline{x}_{21} \\ \overline{x}_{11} & \overline{x}_{21} \end{bmatrix}. \quad (4)$$

The final probability in the vertices method is calculated as the mean, maximum or minimum probability for these combinations (de Souza et. al. 2011, 277).

Usually the centers, both lower and upper bound approaches obtain better results than the vertices method for symbolic interval-valued logistic regression (de Souza et. al. 2011).

In this paper the centers method will be used for model estimation.

In the ensemble model D different models $D = (D_1, ..., D_l)$ are combined to obtain one single (aggregated) model (D^*) that obtains better results (in terms of lower error) than any of the models that are the part of the ensemble.

In this paper, the bagging approach will be used. The initial data set is divided into $U_1, ..., U_l$ subsets (subsamples) that are drawn with replacement from the initial data set. Then each subset is used to build a model (obtain regression results). In the case of this paper, 20 subsets will be taken.

Final results (final probabilities in the case of logistic regression) are obtained by averaging all results (see for example Polikar 2007, 60–61).

3. Conjoint analysis

Conjoint analysis is a powerful market research technique that measures how people make decisions based on certain features of a product or service. This method is well-described in the literature (see for example Orme 2006; Sagan 2013; Gustafsson et. al. 2013; Rao 2014; Wiley et. al. 2014), so only the most important elements will be presented in this paper.

The conjoint method originated in mathematical psychology and was also developed beginning in the mid-sixties by researchers in marketing and business. Conjoint analysis is a statistical method for finding out how consumers make trade-offs and choose among competing products or services. It is also used to predict (simulate) consumers choices for future products or services (see for example: Sagan 2013).

The main aim of conjoint analysis is to estimate part-worth utilities for all attribute levels. The part-worth utilities are estimated for each respondent separately, and as an average values for the whole sample. Estimated utilities allow the following to be estimated: total utilities of a profile for all respondents, average total utilities in the sample, average attribute importance and average total utilities in the segments (clusters, groups) of respondents. In conjoint analysis, attributes (also called factors) are used to describe explanatory variables describing goods or services. Attribute levels describe values of attributes and profiles (stimuli, runs, treatments) that are variants of goods or services.

The most important feature of conjoint analysis based on a full profile method is that the number of attributes taken into consideration is usually limited to six. The profiles are described using all attributes and are presented to the respondents to be assessed. Profiles are generated on the basis of the orthogonal factor system and are maximally and mutually varied. All respondents evaluate the same set of profiles. Conjoint analysis, which represents the decomposition approach, can take into account main effects and the effects of an attribute interaction (Green and Srinivasan 1978).

4. Empirical results

For the purposes of symbolic interval-valued logistic regression, 109 mobile phone users (a convenience sample) from Lower Silesia were asked to answer the following questions (opinions):

y – Would you consider changing your current mobile phone operator (0 – no, 1 – yes),

x_1 – “I use the services of my mobile phone operator as they are the best choice for me”,

x_2 – “If I could, I would make the same choice again” (probability of making the same choice again),

x_3 – “I use the services of the same mobile phone operator over time”,

x_4 – “I would consider using services of another mobile phone operator if the prices of my mobile network rose slightly”,

x_5 – “If had a chance, I would try the services of other mobile phone operator”,

x_6 – “I would consider using the services of another mobile phone operator if the current one had some technical issues”,

x_7 – “My current mobile phone operator provides better services than others”,

x_8 – “In my opinion, the services of my current mobile phone operator are not better than competitors”,

x_9 – "I'm sharing positive opinions about my mobile phone operator with other people",

x_{10} – "I would recommend my current mobile phone operator to other people",

x_{11} – "I feel that I'm emotionally bound to my current mobile phone operator",

x_{12} – "I use the services of my mobile phone operator as I want to (I don't have to)".

All of the questions had the following intervals that reflect the willingness (probability) of the respondent to agree with a question (statement): [0; 15], [7; 25], [25; 70], [45; 80], [60; 100].

Besides that, two open-ended questions were taken into account:

x_{13} – "I usually spend not less than … (insert an amount), but surely not more than … (insert an amount) monthly on my mobile services",

x_{14} – "My mobile phone calls usually last from … (insert an integer) to … (insert an integer) minutes".

The model was estimated using glm function from the stats package of R software. The lrtest function from the lmtest package of R (see Hothron et. al. 2018) software was used to check the model.

The estimated model obtained the results that are shown in the Table 1.

Table 1. Results of estimation

Coefficients	Estimate	Standard error	z-value	p-value	odds ratio
Intecept	0.032936	3.617657	0.009	0.9927	1.0334840
x_1	0.052520	0.029475	1.782	0.0748*	**1.0539238**
x_2	0.046198	0.025659	1.800	0.0718*	**1.0472815**
x_3	−0.053805	0.025669	−2.096	0.0361*	0.9476170
x_4	−0.020768	0.024264	−0.856	0.3920	0.9794458
x_5	−0.021036	0.022383	−0.940	0.3473	0.9791837
x_6	0.045450	0.025999	1.748	0.0804*	1.0464985
x_7	−0.092530	0.047409	−1.952	0.0510*	0.9116219
x_8	0.069309	0.042100	1.646	0.0997*	**1.0717672**
x_9	0.029991	0.030101	0.996	0.3191	1.0304456
x_{10}	−0.016937	0.041599	−0.407	0.6839	0.9832058
x_{11}	−0.006965	0.019349	−0.360	0.7189	0.9930596
x_{12}	0.011584	0.030175	0.384	0.7011	1.0116510
x_{13}	−0.032541	0.024372	−1.335	0.1818	0.9679829
x_{14}	0.027575	0.019164	1.439	0.1502	1.0279585

Variables significant at $\alpha=0.1$ level are marked with "*"
Highest odd ratios for significant variables are in bold.

The null deviance for the centers model was equal to 90.927 (on 108 degrees of freedom), while residual deviance was equal to 51.091 (on 94 degrees of freedom). The Akaike information criterion (AIC) reached 81.091.

Variables x_1, x_2, x_3, x_6, x_7 and x_8, are the only relevant ones ($\alpha = 0.1 > p - value$). Consequently, they are the only ones to be interpreted in the further analysis.

The most important variable, in terms of the odds ratio, is the x_8 ("In my opinion, the services of my current mobile phone operator are not better than competitors"). Increasing midpoint values for this variable raise the chance that a customer will consider changing his/her current mobile phone provider. Variables x_1 ("I use the services of my mobile phone operator as they are the best choice for me"), x_2 ("If I could I would make the same choice again"), x_6 ("I would consider using the services of another mobile phone operator if my current one had some technical issues"), have a slightly lower, but positive impact on the probability that the customer will consider changing his/her current mobile phone provider. Variables x_3 ("I use the services of the same mobile phone operator over time") and x_7 ("My current mobile phone operator provides better services than others") have a negative impact on the chance of changing mobile phone providers. Increasing midpoint values for these variables lower the probability that the customer will consider changing his/her current mobile phone provider.

The regression model was checked if it is relevant with the application of the lrtest function from the lmtest package of R software. As the whole regression is relevant.

All model pseudo R^2 values (Efron's, McFadden's, Nagelkerke's and square of correlation coefficient between empirical and theoretical values) were equal for the single model 0.8098423 (Nagelkerke's pseudo R^2), and all other others were similar, reaching a value of around 0.70. These values were slightly better for ensemble model (0.8728378 in terms of Nagelkerke's pseudo R^2) and around 0.74 in terms of other pseudo R^2 values.

The same 109 respondents also evaluated 17 profiles with the following attributes and levels:

1) phone operator with the following levels – T-Mobile, Orange, Play, Other
2) offer type – with phone, without phone,
3) contract type – post-paid, pre-paid, duet,
4) monthly fee (rate) – up to 30 PLN, 30–50 PLN, 50–100 PLN, more than 100 PLN,
5) supplementary promotions – all services without limitations, some services are unlimited.

The customers were asked to evaluate profiles using a scale from 1 to 10, where 1 signifies, "I would not choose this" and 10, "I'll absolutely choose this."

In the case of this paper the output variable differs slightly from the typical cases, as it reflects a negative, not positive, action that can be performed by customers.

The conjoint model was estimated using the conjoint package of R software. The average importance of all factors is presented in Figure 1.

When considering all factors, the monthly fee was the most important one (36.07% with the highest utilities for two first levels (fee up to 30 PLN, fee 30 to 50 PLN) then mobile phone operator (28.01% with positive values for Plus and Play and negative for others). The least important was the offer type (9.72% with positive values for offer with phone, while the offer without the phone has a negative value).

Positive values were also recorded for services without limitations (in supplementary promotions factor) and Duet and pre-paid services (in terms of contract type). All part-worth utilities for all attribute levels are shown in the Table 2.

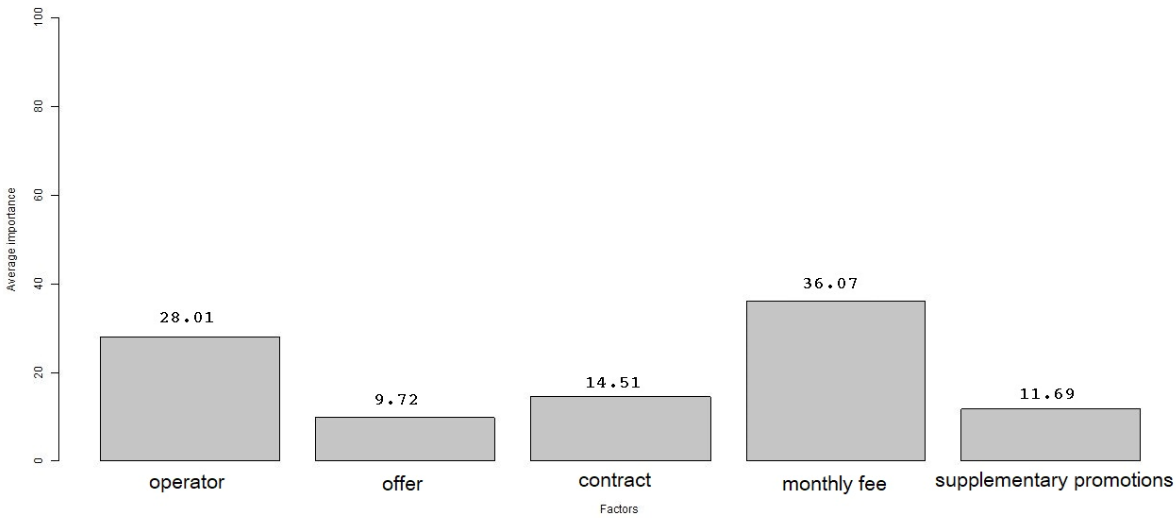


Fig. 1. Average importance of all factors

Table 2. Part-worth utilities

Levels	Utilities	Levels	Utilities
Intercept	3.6456	Duet	0.0825
Plus	0.4944	up to 30 PLN	1.2696
T-Mobile	−0.1389	30–50 PLN	0.5294
Orange	−0.1496	50–100 PLN	0.4480
Play	0.3071	more than 100 PLN	−1.3510
Other	−0.5130	all services without limitations	0.2525
with phone	0.2852		
without phone	−0.2852	some services without limitations	−0.2525
post-paid	−0.1191		
pre-paid	0.0365		

Conclusions

Symbolic logistic regression discovered that customer opinion about services when compared to other competitors ("In my opinion, the services of my current mobile phone operator are not better than competitors") is the most important factor. The increasing midpoint values for this variable raise the chance that a customer will consider changing his/her current mobile phone provider. Furthermore, customers were more likely to change their provider depending on their loyalty to their present operator, whether there might be technical problems with their operator or whether another operator made a better offer. Continuous use of the same services, and the feeling that current services are better than those of competitors, decrease the probability of mobile phone provider change.

Conjoint analysis found that when taking into consideration customer preferences, the monthly payment is the most important factor (rising values encourage customers to change mobile phone providers and decrease the likelihood that they will choose a more expensive offer). Other important factors include the phone operator, contract type and supplementary promotions and offer type.

References:

Athanassopoulos, A.-D. (2000). Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of business research*, 47(3), 191–207.

Bąk, A., & Bartłomowicz, T., The conjoint package for R software. www.r-project.org.

Billard, L., & Diday, E. (2006). *Symbolic Data Analysis: Conceptual Statistics and Data Mining.* John Wiley.

Bhattacharya, C.B. (1998). When customers are members: Customer retention in paid membership contexts. *Journal of the academy of marketing science*, 26(1), 31–44.

Burez, J., & Van den Poel, D. (2009). Handling class imbalance in customer churn prediction. *Expert Systems with Applications*, 36(3), 4626–4636.

de Souza, R.M., Queiroz, D.C., & Cysneiros, F.J.A. (2011). Logistic regression-based pattern classifiers for symbolic interval data. *Pattern Analysis and Applications*, 14(3), 273.

Diday, E., & Noirhomme-Fraiture, M. (Eds.). (2008). *Symbolic data analysis and the SODAS software*. John Wiley & Sons.

Diday, E., & Bock, H.H. (2000). Analysis of symbolic data: Exploratory methods for extracting statistical information from complex data.

Green, P.E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: issues and outlook. *Journal of consumer research*, 5(2), 103–123.

Gustafsson, A., Herrmann, A., & Huber, F. (Eds.). (2013). *Conjoint measurement: Methods and applications*. Springer Science & Business Media.

Hothron, T., Zeileis, A., Farebrother, R.-W., Cummins, C., Millo, G., & Mitchell, D. (2018), The lmtest package for R software. www.r-project.org.

Keaveney, S.-M. (1995). Customer switching behavior in service industries: An exploratory study. *The Journal of Marketing*, 71–82.

Neslin, S.A., Gupta, S., Kamakura, W., Lu, J., & Mason, C.-H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. *Journal of marketing research*, 43(2), 204–211.

Noirhomme-Fraiture, M., & Brito, P. (2011). Far beyond the classical data models: symbolic data analysis. *Statistical Analysis and Data Mining: the ASA Data Science Journal*, 4(2), 157–170.

Orme, B.K. (2006). *Getting started with conjoint analysis: strategies for product design and pricing research*.

Polikar, R. (2007). Bootstrap-inspired techniques in computation intelligence. *IEEE signal processing magazine*, 24(4), 59–72.

Rao, V.R. (2014). *Applied conjoint analysis*. New York, NY: Springer.

Richter, Y., Yom-Tov, E., & Slonim, N. (2010, April). Predicting customer churn in mobile networks through analysis of social groups. In: *Proceedings of the 2010 SIAM international conference on data mining*, 732–741. Society for Industrial and Applied Mathematics.

Sagan, A. (2013). Market research and preference data. *The SAGE handbook of multilevel modeling. London: SAGE Publications Ltd*, 581–99.

Śmiatacz, K. (2012). *Badanie satysfakcji klientów na przykładzie rynku usług telefonii komórkowej w Polsce*. Wydawnictwo Uczelniane Uniwersytetu Technologiczno-Przyrodniczego w Bydgoszczy.

Wiley, J.B., Raghavarao, D., & Chitturi, P. (2010). *Choice-based conjoint analysis: models and designs*. Chapman and Hall/CRC.

Xia, G.E., & Jin, W.D. (2008). Model of customer churn prediction on support vector machine. *Systems Engineering-Theory & Practice*, 28(1), 71–77.

A note on the accuracy of commodity prices forecasts based on futures contracts

Marek Kwas[1], Michał Rubaszek[2]

Abstract
In this study we focus on the predictive power of futures for prices of six key global commodities. For that purpose, we use a comprehensive database of individual contracts to create continuous futures with weekly maturities ranging from one to fifty two weeks. We use this database to check how accurate futures-based forecasts in comparison to the random walk model are. We show that the futures curve does deliver accurate forecasts, which confirm the reliability of the common practice of financial institutions to use futures contracts in forecasting.

Keywords: *Commodity prices, futures contracts, forecasting*
JEL Classification: *G13, G17, Q02, Q47*

1. Introduction

Commodities play an important role in the world economy and commodity prices are important drivers of economic activity, inflation or trade balances. For that reason, understanding the dynamics and the ability to formulate reliable forecasts for commodity prices are important to many economic agents in decision making process. A question arises on whether it is possible to forecast commodity prices accurately. This question is the subject of a long-standing debate in the economic literature. On top of that, since commodity prices are more volatile than stock prices or exchange rates (Fratzscher et al., 2014), constructing a method that delivers accurate forecasts is considered to be a real challenge.

In this study we focus on the predictive power of futures for prices of six key global commodities: crude oil (West Texas Intermediate – WTI – as well as Brent), natural gas, gold, silver and copper. We analyse the accuracy of forecasts based on commodity futures, as they are regularly used in central banks and other policy institutions. This analysis supplements the earlier literature, which delivers ambiguous results. For instance, Alquist and Kilian (2010), Alquist et al. (2013) reported that futures-based forecasts of oil prices are no better than the no-change forecast from the random walk model. On the other hand, Chinn and Coibion (2014) showed that futures of energy commodities are essentially unbiased predictors of spot prices, providing good point and directional forecasts, whereas for industrial and precious metals this is not the case and futures perform relatively poorly.

We contribute to the literature by creating a comprehensive database of individual futures contracts, which are subsequently used to create continuous futures at weekly maturities ranging from one week to one year. We show that the futures curve does deliver accurate forecasts,

[1] SGH Warsaw School of Economics, Econometrics Institute, 162 Niepodległości, 02-554 Warsaw, Poland, marek.kwas@sgh.waw.pl.

[2] Corresponding author: SGH Warsaw School of Economics, Econometrics Institute, 162 Niepodległości, 02-554 Warsaw, Poland, michal.rubaszek@sgh.waw.pl.

which confirm the reliability of the common practice of financial institutions to use futures contracts in forecasting.

2. Data

We use weekly data (daily closing prices at the end of each week) for hundreds of individual futures contracts that were quoted between January 4th, 2009 and September 23rd, 2018. We consider prices of the following commodities: WTI and Brent crude oil, natural gas, gold, silver and copper. The data for individual contracts was downloaded from the CME Group, with an exception of Brent futures prices which were sourced from ICE. The time series of continuous weekly futures are derived as weighted averages of the prices of the two nearest active futures contracts, with the formula (Pindyck, 2001):

$$f_{t,h} = \left(\frac{h_2 - h}{h_2 - h_1}\right) f_{t,h_1} + \left(\frac{h - h_1}{h_2 - h_1}\right) f_{t,h_2}$$

where $f_{t,h}$ is the log price of a continuous futures expiring at horizon h observed in time t, and f_{t,h_1} and f_{t,h_2} denote the logs of prices of the two nearest active futures with $h_1 < h_2$. Following Pindyck (2001) we calculate the synthetic spot price as $s_t = f_{t,0}$. All futures analysed here are physically settled, with an exception of Brent futures which are settled financially. Thus, they all provide benchmarks for spot prices of respective commodities, derived in a similar way as described above.

The dynamics of time series for spot commodity prices are presented in Fig. 1. One can observe a qualitative difference of the natural gas market dynamics from the remaining five markets, the former characterised by rapidly changing trends and higher volatility. It is also possible to observe the post Great Financial Crisis decoupling of oil and natural gas markets, which is described in detail e.g. by Zhang and Ji (2018). Apart from the natural gas, the remaining markets exhibit relatively high degree of similarity in the analysed period.

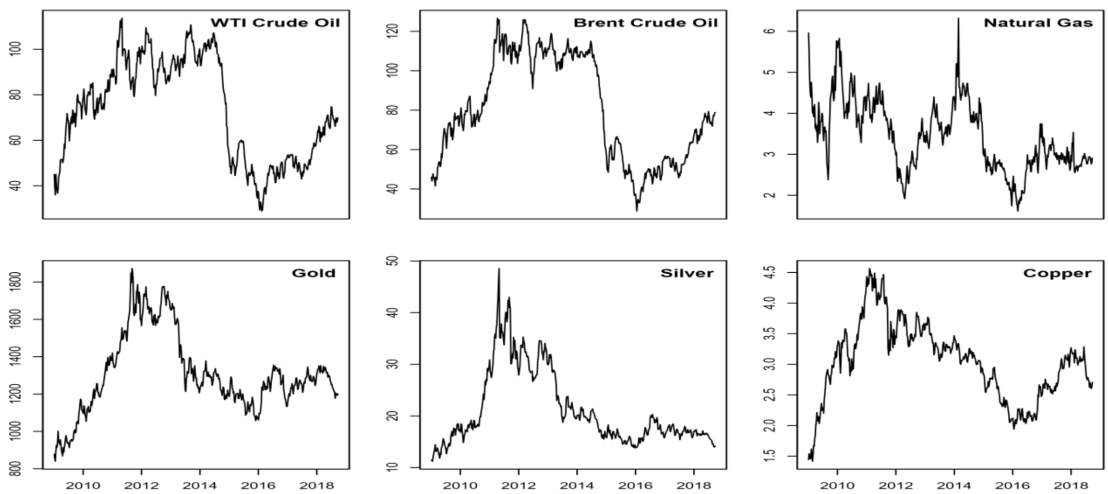


Fig. 1. The dynamics of spot commodity prices

The descriptive statistics for all commodities are presented in Table 1. The table shows gold and copper prices are the least volatile, whereas the variability of natural gas prices is the highest. In the case of skewness, the distributions of natural gas, WTI and copper prices turned out to be symmetric, whereas gold, Brent oil and silver prices were negatively skewed. Furthermore, the excess kurtosis statistics indicate heavy tails, which are the most pronounced for silver. The above is confirmed by the results of Jarque-Bera test, which show non-Gaussian distribution for all series. Let us point out that both WTI and Brent prices, although strongly correlated, show noticeable statistical differences. It illustrates the divergence of the US and European oil markets, with Brent oil gradually substituting WTI as the global benchmark (Manescu and Van Robays, 2014).

Table 1. Descriptive statistics

	Mean	**SD**	**Min.**	**Max.**	**Skew.**	**Kurt.**	**JB**
WTI	0.09	4.46	–15.91	23.01	–0.06	1.91	0.00
Brent	0.11	4.05	–15.26	13.97	–0.23	1.25	0.00
Nat. Gas	–0.14	6.44	–30.98	30.17	0.06	1.67	0.00
Gold	0.06	2.26	–10.13	6.78	–0.26	1.11	0.00
Silver	0.04	4.29	–31.98	13.37	–1.37	10.35	0.00
Copper	0.12	3.35	–18.05	13.94	0.07	2.18	0.00

Notes: The table presents the descriptive statistics for the weekly log-changes (×100) of analysed commodities. JB refers to the p value of the Jarque-Bera normality test.

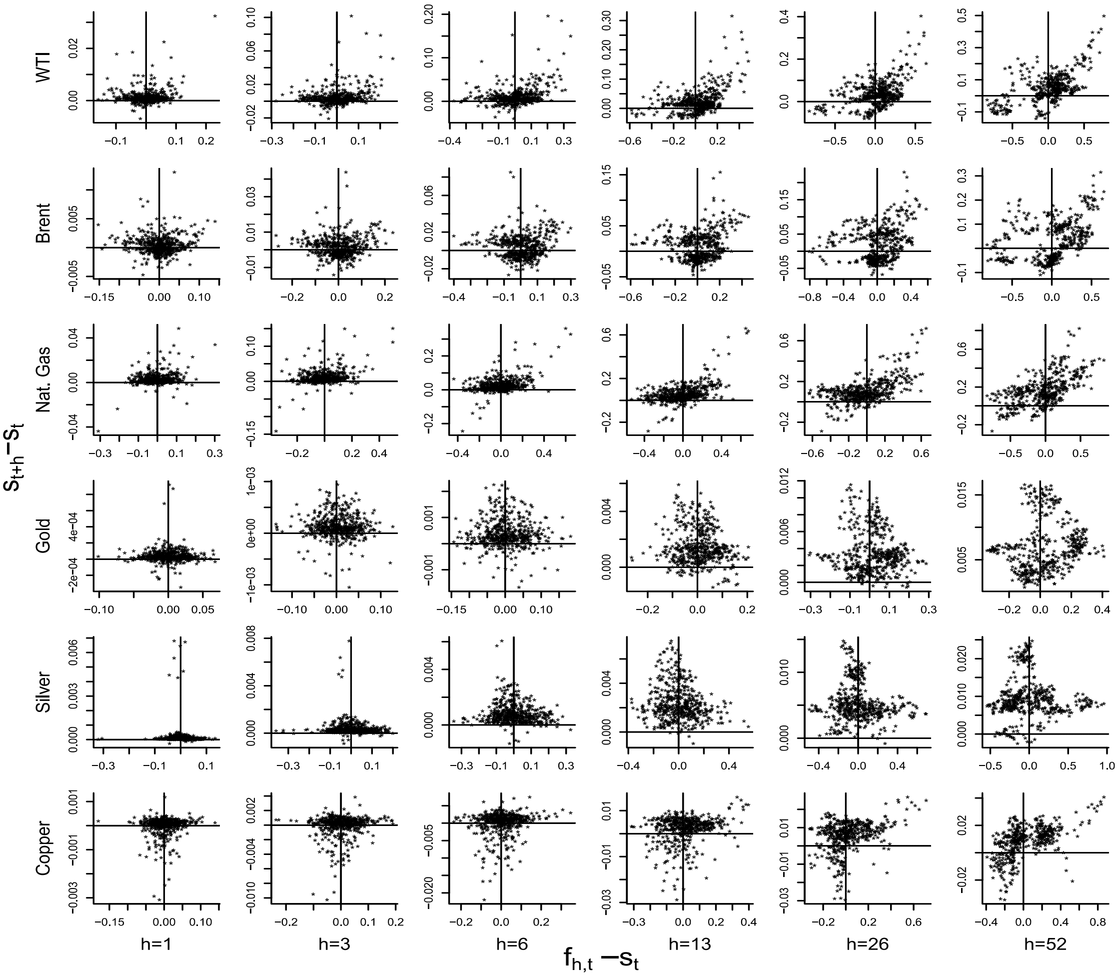
3. In-sample evidence

In this section we analyse how the spot price adjusts to the basis which is defined as the difference between futures and spot prices (Fama and French, 1987). Specifically, we estimate the parameters of regressions:

$$s_{t+\mathrm{h}} - s_t = \alpha_h + \beta_h(f_{h,t} - s_t) + \varepsilon_{t+h},$$

where s_t and $s_{t+\mathrm{h}}$ are the spot prices at times t and $t + h$ respectively, and $f_{h,t} - s_t$ is the basis at h weeks ahead horizon. Positive β_h means that the spot price adjusts to the basis. Moreover, if futures are effective predictors for the spot price the coefficients of the above regression should equal to $\alpha_h = 0$ and $\beta_h = 1$.

The relation between the basis and subsequent changes in spot prices over various horizons is illustrated by Fig. 2. It shows that the correlation is rather weak, which is confirmed by the estimates of the linear regression for six commodities and various horizons (see Table 2). Even if the estimates of are usually positive, the fit of the models is rather poor. The values of show that the model explains only a small fraction of spot price dynamics of investigated commodities. The poor fit of regression (1) might be due to two reasons: the predictive content of the basis is low or the relationship between the basis and the subsequent changes in the spot price is non-linear, which is suggested e.g. by Gulley and Tilton (2014) and Fernandez (2016).



Notes: The horizontal axes present the basis for weeks ahead horizon and the vertical axes describe the subsequent adjustment (log change) of spot prices over weeks horizon.

Fig. 2. The basis and commodity price adjustment

Table 2. The adjustment in the spot price to the basis

	WTI		**Brent**		**Nat. Gas**		**Gold**		**Silver**		**Copper**	
h	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2	$\hat{\beta}_h$	R^2
1	2.65	0.03	0.78	0.00	2.91	0.07	2.77	0.00	-2.27	0.00	5.37	0.00
3	1.90	0.06	0.85	0.01	1.75	0.12	-2.45	0.00	-9.36	0.01	3.94	0.01
6	1.69	0.12	0.67	0.01	1.40	0.21	-2.55	0.00	-14.02	0.01	3.44	0.02
13	1.91	0.20	1.06	0.03	1.18	0.28	-6.55	0.01	-15.12	0.02	4.41	0.05
26	1.78	0.22	1.14	0.06	1.20	0.32	-3.47	0.01	-7.99	0.01	6.79	0.12
52	2.05	0.36	1.57	0.16	1.22	0.37	6.18	0.02	-1.42	0.00	11.24	0.30

Notes: The table presents the estimates of regression $s_{t+h} - s_t = \alpha_h + \beta_h(f_{h,t} - s_t) + \varepsilon_{t+h}$. The parameters were estimated with weekly data covering the period between January 4th, 2009 and September 23rd, 2018.

4. Forecasting contest results

We compare the accuracy of forecasts from two models. The first one is a widely used benchmark, i.e. the naive random walk (RW). From the perspective of a practitioner, there is nothing more conservative than assuming that the price will remain constant over the forecast horizon. The forecast for horizon h formulated in period t is:

$$s_{t+h,t}^{\mathrm{RW}} = s_t.$$

The second method, which we test in this study, is based on continuous futures. For this method the value of forecast formulated at period t for horizon h is:

$$s_{t+h,t}^{\mathrm{Fut}} = f_{h,t}.$$

where $f_{h,t}$ is the price of the continuous futures contract maturing at horizon h.

We evaluate the accuracy of forecasting for horizons ranging from one week to one year (h = 52 weeks). The first set of forecasts is formulated for the period between January 4th, 2009 and December 27th, 2009 and the last one for the period between September 16th, 2018 and September 15th, 2019. Given that our sample ends on September 16th, 2018, we assess the quality of forecasts using 507 forecast errors for one-week horizon, whereas for horizon the number of observations is equal to 508 − h.

We begin our analysis by measuring the forecasting performance of the competing methods with the root mean squared and mean absolute forecast errors statistic (RMSFE and MAFE). Table 3 reports the values of RMSFE and MAFE for RW, whereas for the futures-based forecast it presents ratios relative to RW. Thus, the values below unity indicate the outperformance of the RW benchmark. We also test the null of equal forecast accuracy with the one-sided Coroneo and Iacone (2015) version of the Diebold and Mariano test, which offers relatively good finite sample size and power performance (Harvey et al., 2017).

Table 3. RMSFE and MAFE of RW and futures-based forecasts

	Random walk forecasts						**Futures-based forecasts**					
	1	**3**	**6**	**13**	**26**	**52**	**1**	**3**	**6**	**13**	**26**	**52**
							RMSFE					
WTI	0.05	0.08	0.11	0.17	0.23	0.32	0.991**	0.974**	0.949**	0.918*	0.908*	0.870*
Brent	0.04	0.07	0.10	0.16	0.23	0.32	0.999	0.997	0.996	0.981	0.966	0.931
Nat. Gas	0.07	0.11	0.14	0.19	0.25	0.32	0.982**	0.960**	0.926*	0.929	0.968	1.012
Gold	0.02	0.04	0.05	0.07	0.11	0.16	1.000	1.000	1.000	1.000	0.997	0.989
Silver	0.04	0.08	0.10	0.14	0.20	0.30	1.000	1.000	1.000	1.001	1.001	1.000
Copper	0.03	0.06	0.08	0.12	0.17	0.24	0.999**	0.998**	0.994**	0.988*	0.980*	0.970**

							MAFE					
WTI	0.03	0.06	0.08	0.13	0.17	0.24	0.991**	0.982**	0.943**	0.910*	0.893*	0.837*
Brent	0.03	0.06	0.08	0.12	0.17	0.25	0.999	1.001	0.999	0.963	0.941	0.908
Nat. Gas	0.05	0.08	0.11	0.15	0.20	0.25	0.992**	0.978**	0.956*	0.911	0.936	1.003
Gold	0.02	0.03	0.04	0.06	0.09	0.13	1.000	1.000	1.001	0.998	0.993	0.989
Silver	0.03	0.06	0.08	0.11	0.15	0.24	1.000	1.001	1.002	1.002	1.007	1.009
Copper	0.03	0.04	0.06	0.09	0.13	0.19	1.000**	0.998**	0.995**	0.990*	0.982*	0.979**

Notes: The figures describe the values of RMSFE and MAFE from futures-based forecasts in comparison to the RMSFE and MAFE from RW. Asterisks ***, ** and * denote the 1%, 5% and 10% significance levels of the one-sided Coroneo and Iacone (2015) version of the Diebold-Mariano test with the alternative that a given model performs better than RW.

A look at Table 3 leads to several immediate findings. The first one is that RMSFEs from RW for metal commodities are somewhat lower than those for energy commodities. The second finding is that for WTI, natural gas and copper, futures-based forecasts outperform RW consistently across horizons in both RMSFE and MAFE measures. The only exception is the longest horizon for natural gas prices. Moreover, for WTI and copper the gains are statistically significant, whereas for natural gas it holds for short to medium horizons. The third observation is that, surprisingly, one can observe a substantial difference for WTI and Brent oil prices. Even though for both commodities futures-based forecasts are better than those from the RW, only for WTI the gains are significant. Finally, in case of both precious metals – gold and silver – futures-based forecasts are of comparable accuracy to RW, with insignificant gains for gold for longer horizons. It can be added that our results are more optimistic than those of Alquist and Kilian (2010) and Alquist et al. (2013), who found that futures-based forecasts for oil prices are statistically indistinguishable from those of RW. On the other hand, our results confirm the findings of Chinn and Coibion (2014) or Fernandez (2017), who show that there is a predictive content in the futures prices.

We complement the forecast accuracy analysis with a visual illustration, by plotting the whole sequence of forecasts conducted at different points in time and comparing them to achieved values (see Fig. 3). A first inspection indicates that all models encounter problems to forecast sharp movements in prices, for instance the WTI price decline in 2015. It is mainly because forecasts are excessively conservative, in a sense that they do not deviate substantially from the last spot price.

We conclude the forecasting contest by calculating the percentage of correct sign forecasts together with the independence test of Pesaran and Timmermann (1992). Table 4 indicates that only for natural gas more than half of predictions are of correct sign consistently across horizons. In turn, for WTI, Brent and copper the share is above 50% only for longer horizons. For gold and silver the fractions of correct predictions are either below or insignificantly different from 50%.

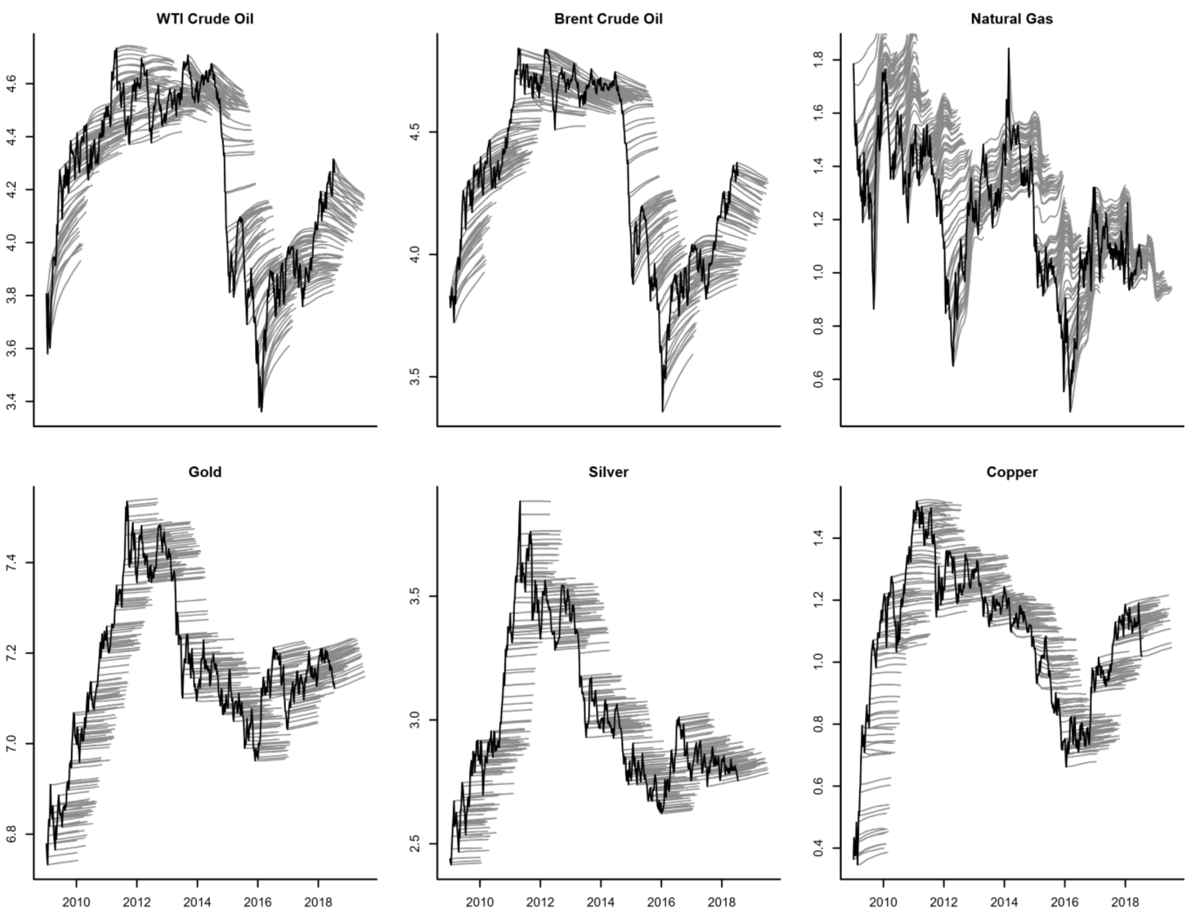


Fig. 3. Rolling forecasts for the log WTI price

Notes: Rolling futures-based forecasts for the logs of prices of six analysed commodities. The first set of forecasts is generated with data ending on December 28th, 2008 and the last set of forecasts is based on series ending on July 15th, 2018, which in total makes forecasts from each model.

Table 4. Fraction of correct sign predictions

	1	**3**	**6**	**13**	**26**	**52**
WTI	0.528	0.538	0.582***	0.604***	0.622***	0.765***
Brent	0.504	0.472	0.510	0.545**	0.600***	0.715***
Nat. Gas	0.516**	0.534***	0.534***	0.556***	0.588***	0.596***
Gold	0.512	0.480	0.496	0.545	0.577	0.564
Silver	0.478	0.480	0.454	0.501	0.459	0.456***
Copper	0.502	0.508	0.524	0.598***	0.631***	0.634***

Notes: The figures represent the fraction of futures-based forecasts that correctly predict the sign of the change. Asterisks ***, ** and * denote the rejection of the null of the independence test by Pesaran and Timmermann (1992) at the 1%, 5% and 10% significance levels.

Conclusions

A common practice of policy making institutions is to rely on futures prices while formulating predictions for commodity prices. It raises a question whether futures contracts contain information that could be extracted to precisely forecasts spot price movements. In this paper, we addressed this question by comparing the accuracy of futures-based forecasts for six main global commodities to those from the random walk model.

Using weekly data from the period 1999–2018, we found that futures-based forecasts are superior or comparable to the naive ones. It turned out that the gains are higher for longer rather than shorter horizons. Moreover, we showed that using futures is most advisable for energy commodities as well as copper, whereas the results for gold and silver are more mixed. A broad conclusion that emerges from our analysis is that it is justified to use forecasting methods that exploit information content of the futures contract prices rather than rely on no-change random walk forecast.

Acknowledgements

The project is financed by the Polish National Center of Science (Narodowe Centrum Nauki). Grant number 2017/25/B/HS4/00156.

References

Alquist, R., & Kilian, L. (2010). What do we learn from the price of crude oil futures? *Journal of Applied Econometrics* 25(4), 539–573.

Alquist, R., Kilian, L., & Vigfusson, R. J. (2013). Forecasting the price of oil. In: Elliott, G., Timmermann, A. (eds.), *Handbook of Economic Forecasting,* 2, 427–507.

Chinn, M.-D., & Coibion, O. (2014). The predictive content of commodity futures. *Journal of Futures Markets* 34(7), 607–636.

Coroneo, L., & Iacone, F. (2015). Comparing predictive accuracy in small samples. *Discussion Papers* 15/15, Department of Economics, University of York.

Fama, E.F., & French, K.R. (1987). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. *The Journal of Business* 60(1), 55–73.

Fernandez, V. (2016). Spot and futures markets linkages: Does contango differ from backwardation? *Journal of Futures Markets* 36(4), 375–396.

Fernandez, V. (2017). A historical perspective of the informational content of commodity futures. *Resources Policy* 51(C), 135–150.

Fratzscher, M., Schneider, D., & Van Robays, I. (2014). Oil prices, exchange rates and asset prices. *ECB Working Paper Series* 1689, European Central Bank.

Gulley, A., & Tilton, J.E. (2014). The relationship between spot and futures prices: An empirical analysis. *Resources Policy* 41(C), 109–112.

Harvey, D.I., Leybourne, S.J., & Whitehouse, E J. (2017). Forecast evaluation tests and negative long-run variance estimates in small samples. *International Journal of Forecasting* 33(4), 833–847.

Manescu, C., & Van Robays, I. (2014). Forecasting the brent oil price: Addressing time-variation in forecast performance. *ECB Working Paper Series* 1735, European Central Bank.

Pesaran, M.H., & Timmermann, A. (1992). A Simple Nonparametric Test of Predictive Performance. *Journal of Business & Economic Statistics* 10(4), 561–565.

Pindyck, R.S. (2001). The dynamics of commodity spot and futures markets: A primer. *The Energy Journal* 22(3), 1–29.

Zhang, D., & Ji, Q. (2018). Further evidence on the debate of oil-gas price decoupling: A long memory approach. *Energy Policy* 113(C), 68–75.

Monetary policy transmission mechanisms in Ukraine

Victor Shevchuk[1]

Abstract

This paper explores monetary policy transmission mechanisms in Ukraine, within a framework of the IS-MP-IA model. We use a six-variable structural VAR model, for the period 2002Q1-2018Q2, encompassing monetary and fiscal variables, to investigate the effects of the National Bank of Ukraine (NBU) interest rate policy. Following an increase in the NBU reference rate, major macroeconomic effects are as follows: (i) a significant decline in inflation, but with a substantial lag of 6 to 8 quarters, (ii) a temporary decrease in the output gap, with the restricting effect disappearing within a year, (iii) an increase in the lending rate, with a slow convergence to the initial level, (iv) improvement in the budget balance, with a two quarters lag, (v) a short-lived depreciation of the real exchange rate (RER) above its equilibrium trend. Among other results, improvement in the budget balance brings about a drop in the inflation rate combined with a temporary appreciation of the RER, while being expansionary in respect to the output gap. It is interesting that budget surplus is associated with a decrease in the NBU reference rate, while not affecting the economy's lending rate. As implied by the IS-MP-IA model, the output gap is inflationary and instrumental in determining of the RER and the NBU reference rate. In accordance with the Taylor rule, the NBU responds to acceleration of inflation on impact, but monetary policy response to the output gap is rather moderate and slow. No monetary policy reaction to the RER undervaluation is detected.

Keywords: *inflation, output gap, budget balance, central bank reference rate, Ukraine*
JEL Classification: *C5, E5, H6*

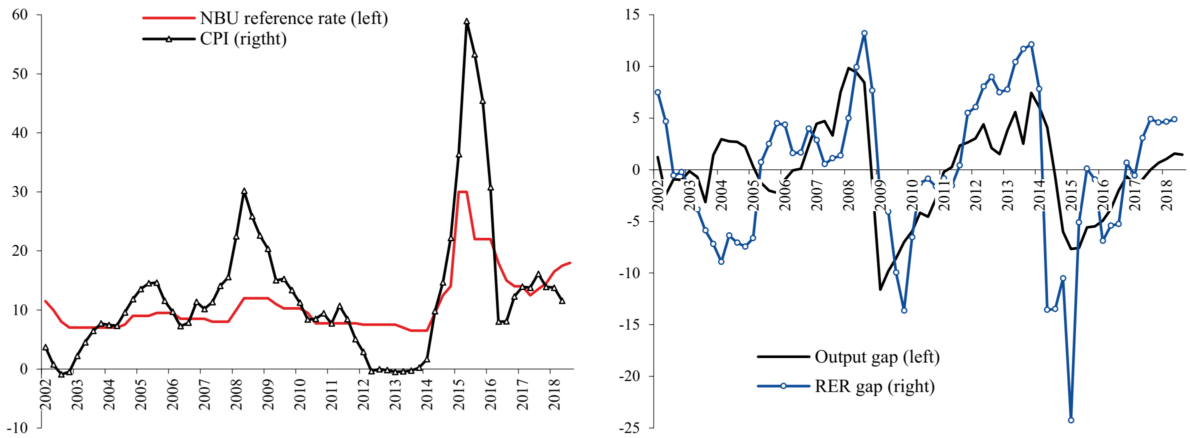
1. Introduction

Since the beginning of the 2000s, the Ukrainian economy has experienced two periods of acceleration of consumer price inflation (CPI) (Fig. 1a). As measured by the level of the National Bank of Ukraine (NBU) reference rate, loose monetary stance was observed in 2007–2009, while monetary tightening was implemented in 2014–2016. The NBU adopted the inflation targeting regime in 2015. Initially, it was a success, but since the beginning of 2016 inflation has accelerated. Even though CPI rates have decreased since the middle of 2017, slow disinflation remains a concern for the monetary authorities. Although higher interest rates could be blamed for a slow recovery from the 2014–2015 financial crisis, the resulting exchange rate appreciation is likely to bring about an opposite expansionary effect. Fig. 1b suggests that the movements of the GDP and relative price gaps are synchronized, with the periods of economic boom coinciding with appreciation of the real exchange rate (RER) and recessions being observed against the backdrop of steep currency depreciations.

Though a mere correlation between output and RER and/or NBU reference rate and CPI is not sufficient to determine the direction of the relationship between the variables that they are all endogenously determined, visual analysis of data series provides a hint on possible causal links between them. As implied by the IS-MP-IA (or Taylor-Romer) model, inflation target can

[1] Corresponding author: Cracow University of Technology, Institute of Economy, Sociology and Philosophy, 24 Warszawska St., 31-155 Cracow, Poland, e-mail: vshevchuk@pk.edu.pl.

be achieved by setting the central bank policy rate at the level of the natural rate of interest, with both output gap and expected inflation accounted for. However, the monetary transmission mechanisms used to be dependent upon the exchange rate effects, with the fiscal policy also playing its part.



a) NBU reference rate and CPI (%)

b) output gap and the inverted RER gap (deviation from trend, %)

Note: GDP and RER are de-trended with the Hodrick-Prescott filter. An increase in the inverted RER gap reflects real appreciation of the exchange rate.
Source: IMF *International Financial Statistics*, National Bank of Ukraine

Fig. 1. Ukraine: selected macroeconomic indicators, 2002–2018

The aim of this study is to examine monetary policy transmission mechanisms in Ukraine, in the presence of fiscal and exchange rate effects. We implement structural vector auto-regression (SVAR) approach for modeling the inter-dependencies between monetary and fiscal policies, output gap and CPI.

The rest of the paper is organized as follows. Section 2 provides a brief outline of analytical issues. Section 3 describes data and outlines the structure of SVAR model. Section 4 discusses empirical results and Section 5 offers some concluding remarks.

2. Analytical framework

As proposed by Romer (2000), the IS-MP-IA model is considered as a simple but informative tool applied in the analysis of the inflation-output relationship and the monetary policy effects by focusing on the interest rate rather than on money supply. While traditional IS and Phillips curves are kept, the LM curve is replaced with a Taylor-type interest rate (Taylor, 2000). If extended by fiscal variables (Bofinger et al., 2006; Clark and Hsing, 2005), the modeling framework enables the analysis of fiscal policy effects as well. For open economies, it is suggested to include real exchange rate (RER) into the reaction function (Ball, 1999; Caporale et al., 2018; Heipertz et al., 2017; Nojković and Petrović, 2015), such approach is also criticised (Leitemo and Söderström, 2005). As the RER affects both aggregate demand and inflation, it further complicates monetary policy.

The extended IS-MP-IA model is presented below:

$$y = \alpha_0 - \alpha_1(i - p^e) - \alpha_2 b + \alpha_3 q, \quad (1)$$

$$i = \bar{r} + p^e + y_1(p - \bar{p}) + y_2 y + y_3 q, \quad (2)$$

$$p = p^e + \beta_1 y + \beta_2 q, \quad (3)$$

where y is the real output gap, q is the RER gap (an increase in the value of q means undervaluation of the real exchange rate), b is the budget surplus, i is the central bank reference rate, r is the 'natural' rate of interest, p and p^e are actual and expected inflation rates, respectively and p is the inflationary target.

The first equation is the IS curve, characterising the inverse relation between real interest rate and output. The budget surplus and the RER overvaluation are expected to be restrictionary. Equation (2) presents a Taylor-type monetary policy rule that implies the response of the central bank reference rate to inflation, output and RER gaps. Although accounting for the exchange rate is not required in the case of the developed economies, it might be of importance in the emerging economies (Caporale et al., 2018). In Equation (3), the aggregate supply is given by the positive short-run open economy Phillips curve. Lags of the variables entering the model could be added.

Policy implications of the IS-MP-IA model in general and the Taylor rule in particular are extensively tested. For example, it was empirically established that over the pre-crisis period of 1989–2007 interest rate forecasts were consistent with Taylor-type rules for the G7 countries (Fendel et al., 2011). Earlier it was found that the interest rate in Germany and the euro area could be described by a Taylor rule with the interest rate smoothing (Peersman and Smets, 1998). For the CEE countries, a clear shift of the interest rate setting in favour of targeting inflation is found in the Czech Republic, Hungary and Poland, with slightly weaker results for Slovenia and Romania (Frömmel et al., 2011). Similar results were obtained more recently in several other studies (Feldkircher et al., 2016; Ryczkowski, 2016; Wang et al., 2015).

There is empirical evidence that central banks in the Czech Republic, Poland, Hungary and Serbia react to the RER gap, while in Romania and Albania there is response to the changing rate of RER (it implies that only accelerated RER developments affect policy decision on interest rate, while the constant rate of change does not trigger any policy shifts) (Nojković and Petrović, 2015). In several emerging countries (Indonesia, Israel, South Korea, Thailand, Turkey) the exchange rate has impact on the reaction function of monetary authorities under the high inflation regime but not under the low inflation regime (Caporale et al., 2018).

3. Data and statistical model

The quarterly series used in the SVAR are the NBU reference rate (%), i_t, lending rate (%), rl_t, CPI (%), p_t, the cyclical components of real output (deviations form trend, %), y_t, RER (deviations form trend, %), q_t, and the budget balance (% of GDP), b_t. The data set for the sample period of 2002Q1:2018Q2 was collected from IMF International Financial Statistics (IFS) and Ukraine's State Committee of Statistics. All data were seasonally adjusted using the Census X12 procedure, except for CPI and RER. Both the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) stationarity tests indicate that all the macroeconomic variables are stationary at the 5% significance level (not reported).

Structural VARs enable us to separate out systematic responses to changes in interest rates from exogenous monetary policy shocks. Omitting details of a general specification for the economy described by a structural form equation of a linear, stochastic dynamic form, our SVAR presents as follows (in terms of the contemporaneous innovations):

$$b = u_1 + a_{12}\gamma + a_{15}rl, \tag{4}$$

$$\gamma = u_2 + a_{24}i + a_{26}q, \tag{5}$$

$$p = a_{31}b + a_{32}\gamma + u_3, \tag{6}$$

$$i = a_{42}\,\gamma + a_{43}\,p + u_4, \tag{7}$$

$$r\,l = a_{53}\,p + a_{54}\,i + u_5, \tag{8}$$

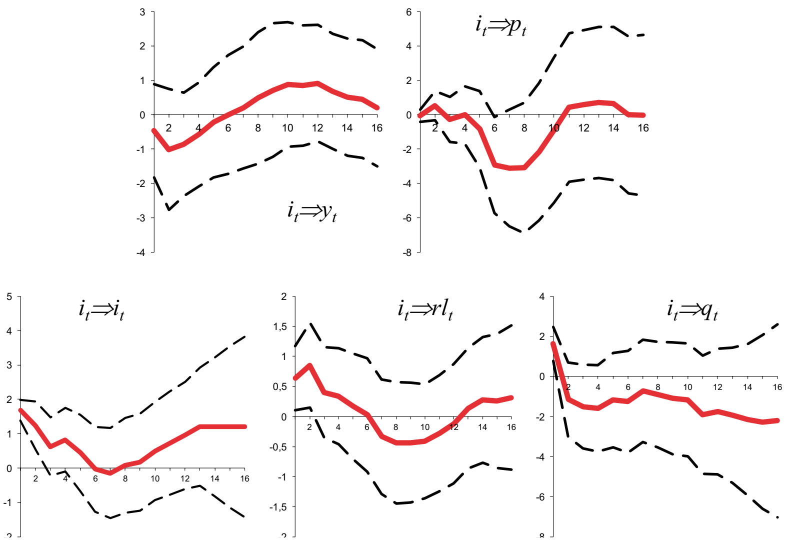
$$q = a_{61}\,b + a_{62}\,\gamma + a_{63}\,p + a_{64}\,i + a_{65}\,r\,l + u_6. \tag{9}$$

All variables in equations (4)-(9) represent the first stage VAR residuals. It is assumed that the budget balance responds to changes in the output gap and lending rate (equation (4)). As implied by the IS curve, the output gap is influenced by the NBU reference rate and the RER gap (equation (5)). Inflation in the current period is affected by fiscal policy and output gap (equation (6)). Thus, it is assumed that monetary policy exerts its inflationary effects through its impact upon the output gap. As argued by Giordani (2004), using the output gap instead of the level of real output helps to avoid the price puzzle when monetary tightening does not bring about a deceleration of inflationary dynamics.

The NBU reference rate is a function of output and CPI shocks (equation (7)). The lending rate reacts to changes in inflation and the NBU reference rate (equation (8)). It is assumed that the correction of the RER misalignment is not among the central bank priorities in the short run. Finally, the RER gap is affected by all other endogenous variables in the current period (equation (9)). Among exogenous variables, our SVAR includes a dummy variable to control for the financial turmoil of 2014–2016. In estimation, we use five lags of each endogenous variable, as implied by most of the lag length criteria. It is worth mentioning that using of terms-of-trade (TOT) or foreign direct investments (FDI) as exogenous variables does not change the results significantly.

4. Estimation results

The impulse response functions to an unexpected increase in the NBU reference rate are presented in Fig. 2. The horizontal axis indicates quarters after shock, and the vertical axis represents percentage changes of endogenous variables. Output gap declines temporarily, with the downward effect reaching a peak in the second quarter but the effect is not significant. For more than year inflation does not react to a higher NBU reference rate and only thereafter it starts to decline. The lending rate increases significantly on impact but then tends to decline to its initial level, in line with weakening of the monetary shock. Though there is a short-lived depreciation of the RER, it is reversed after a quarter. An increase in i_t seems to be neutral in respect to the budget balance.



Note: Solid lines are the point estimates of the impulse-response mean. Dashed lines are the point estimates ± 2 standard deviations.

Fig. 2. Macroeconomic effects of higher NBU reference rate

A combination of restrictionary impact on the output gap in the short run and slow disinflation effect, accompanied by increase in the lending rate and the exchange rate depreciation, can be easily (and mistakenly) interpreted as monetary policy "inefficiency" in the disinflation process. Consequently, it may lead to intense interest group pressure for the reversal of monetary policy tightening. After 6 quarters the NBU reference rate returns to its initial level, which is not viable in the long run. Therefore, it is necessary to return to an increase in i_t.

As seen in Fig. 3, the budget surplus increases the output gap, but the effect is only significant on impact. The anti-inflationary effect of b_t becomes significant after two quarters, much faster in comparison to the monetary shock. The improvement in the budget balance is associated with the RER appreciation on impact and decline in the NBU reference rate, with 3 quarters lag. Only the lending rate does not respond in the expected way.

Undervaluation of the RER reduces the output gap and improves the budget balance (Fig. 4). The effect of q_t on inflation peaks after 4 quarters, but inflationary pass-through is not persistent. There is no evidence of significant NBU reaction to a RER shock, although there is an increase in i_t after 3 quarters, with a downward correction after 9 quarters. The exchange rate instability can help reduce price volatility (Ball, 1999) or lead to welfare gains (Heipertz et al., 2017). However, this approach is also criticised. For example, it is argued that the Taylor rule without exchange rate seems more robust in modeling uncertainty in the open economy (Leitemo and Söderström, 2005).

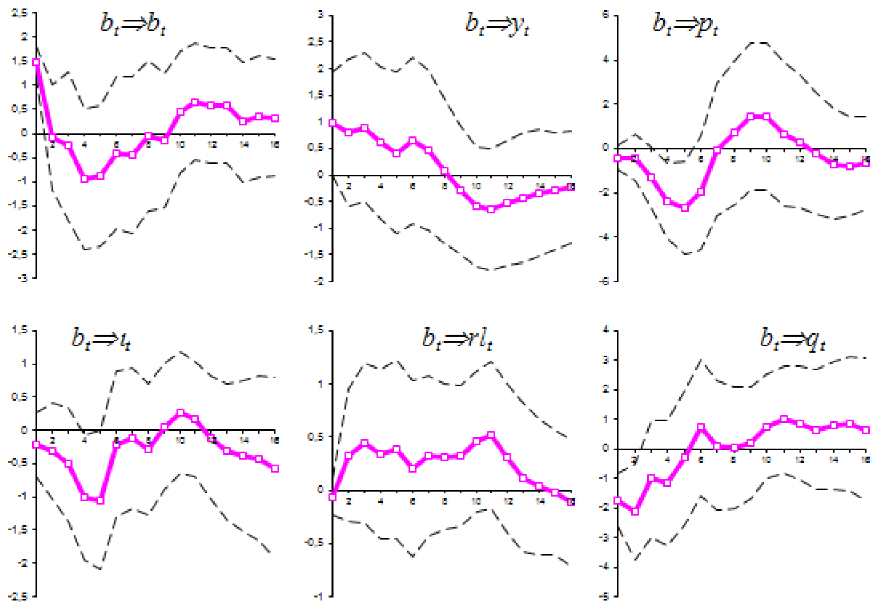


Fig. 3. Macroeconomic effects of the budget surplus

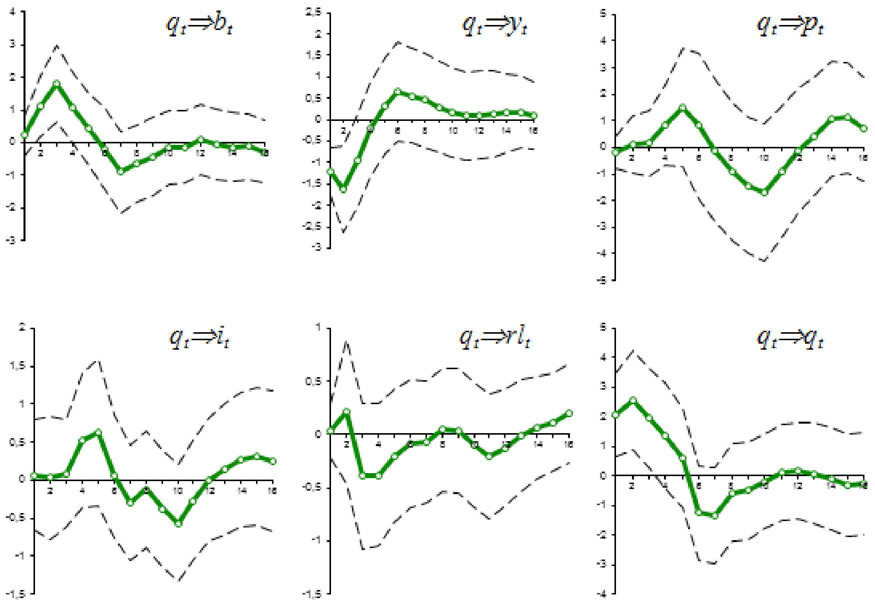


Fig. 4. Macroeconomic effects of the RER gap

The reaction of NBU reference rate to other macroeconomic shocks is reported in Fig. 5. As expected, there is an immediate increase of i_t in response to inflation, but the reaction is rather weak and it reverses after two years. As for the output gap, the reaction appears to be slow, with an increase of i_t four quarters after a shock, and quite weak (this effect is not statistically significant). At the same time, there is an asymmetric reaction to an increase in the lending rate, in 7 quarters after the shock. The similarity of our results to those found in CEE country studies in respect to the central bank policy rate effects (Fig. 2) does not contradict the possibility of successful inflation targeting in Ukraine, but the response to inflation and output gap should be much stronger. Also, a monetary policy shock should be more persistent. It cannot be ruled out that the above mentioned weaknesses of the NBU response to adverse macroeconomic shocks stand behind slow reaction of inflation to monetary tightening. Assuming significant real effects of the RER undervaluation, it is open to discussion whether the NBU should not react to such shocks. Also, it seems that fiscal policy plays some role in inflationary dynamics.

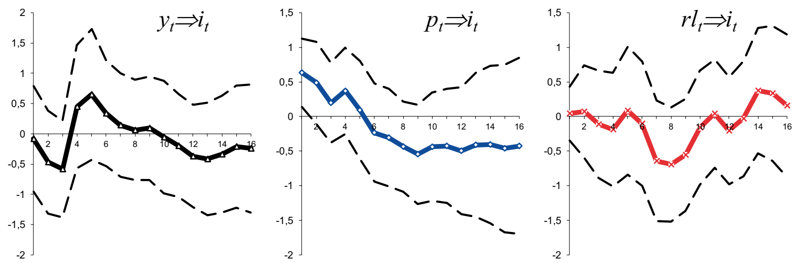


Fig. 5. Response of the NBU reference rate to macroeconomic shocks

The forecast error variance decomposition (FEVD) confirms that the contribution of both output gap and inflation shocks to the NBU reference rate is rather limited and comparable with the contribution of the fiscal shock (Table 1). The NBU reference rate determines up to 16% of the output gap but its share in the FEVD of CPI gradually increases to 27% over time. At the end of the second year the contribution to RER volatility is below 20% and it increases to 36% till the end of the sample period. Interest rate shocks and RER together account for more than a half of output variability at any horizon. In the short run prices are quite sticky but after two years and onwards monetary factors determine above 40% of CPI variability. Shocks to fiscal policy account for 16–18% of the fluctuations in the output gap and 20–25% of the fluctuations in CPI and RER.

Conclusions

The IS-MP-IA model implies that inflation target can be achieved with the response of the central bank policy rate to the output gap and inflationary pressure. Our findings seem to support such policy option but the reaction of inflation to an increase in the NBU reference rate seems to be very slow, as it takes 6 to 8 quarters for the inflation rate to decline. On the other hand, there is a negative impact on the output gap on impact, with an increase in the lending rate and a short-lived depreciation of the RER above its equilibrium trend. In the presence of a slow disinflation effect, it can be mistakenly interpreted as monetary policy "inefficiency" in the disinflation process and cause pressure for the reversal of monetary policy tightening. No monetary policy reaction to the RER undervaluation is found, despite its negative effects upon inflation and output. A positive link between weak currency and the budget balance can justify such policy stance, but it could weaken efforts aimed at permanent improvement in the budget balance. Contrary to standard predictions of the IS curve, the budget surplus is expansionary, while contributing to a decline in both inflation and the RER undervaluation. As the budget surplus is associated with a decrease in the NBU refinancing rate (not affecting the economy's lending rate), an increase in the NBU reference rate is likely to bring about an improvement in the budget balance.

Table 1. Forecast error variance decomposition of selected endogenous variables

Responses of	Innovations in	Forecast horizons			
		4	**8**	**12**	**16**
y	*b*	16	18	17	18
(real output)	*y*	26	26	23	22
	p	3	6	5	5
	i	13	11	15	16
	rl	13	13	17	17
	q	29	26	22	21
p	*b*	25	21	20	20
(inflation)	*y*	1	15	14	13
	p	60	21	17	16
	i	1	27	26	25
	rl	11	10	14	14
	q	3	5	10	12
i	*b*	17	22	17	14
(the NBU	*y*	8	11	9	7
reference rate)	*p*	9	9	13	12
	i	62	44	45	55
	rl	1	8	8	6
	q	4	7	8	6
q	*b*	25	21	20	18
(RER)	*y*	8	7	8	7
	p	2	4	6	6
	i	17	19	25	36
	rl	8	7	7	7
	q	41	41	33	26

References

Ball, L. (1999). Policy rules for open economies. In: Taylor, J.B. (ed.), *Monetary Policy Rules*, University of Chicago Press, 127–156.

Bofinger, P., Mayer, E., & Wollmershauser, T. (2006). The BMW Model: A New Framework for Teaching Monetary Economics. *Journal of Economic Education*, 37(1), 98–117.

Caporale, G.M., Helmi, M.H., Çatık, A.N., Ali, F.M., & Akdeniz, C. (2018). Monetary policy rules in emerging countries: Is there an augmented nonlinear Taylor rule? *Economic Modelling*, 72(C), 306–319.

Clark, D.P., & Hsing, Y. (2005). Application of the IS-MP-IA Model and the Taylor Rule to Korea and Policy Implications. *The Journal of the Korean Economy*, 6(2), 297–311.

Fendel, R., Frenkel, M., & Rülke, J.C. (2011). 'Ex-ante' Taylor rules – Newly discovered evidence from the G7 countries. *Journal of Macroeconomics*, 33(2), 224–232.

Frömmel, M., Garabedian, G., & Schobert, F. (2011). Monetary Policy Rules in Central and Eastern European Countries: Does the Exchange Rate Matter? *Journal of Macroeconomics*, 33(4), 807–818.

Giordani, P. (2004). An alternative explanation of the price puzzle. *Journal of Monetary Economics*, 51(6), 1271–1296.

Heipertz, J., Mihov, I., & Santacreu, A. M. (2017). The Exchange Rate as an Instrument of Monetary Policy. *FRB of St. Louis Working Papers*, 2017–28, Federal Reserve Bank of St. Louis.

Leitemo, K., & Söderström, U. (2005). Simple monetary policy rules and exchange rate uncertainty. *Journal of International Money and Finance*, 24(3), 481–507.

Nojković, A., & Petrović, P. (2015). Monetary policy rule in inflation targeting emerging European countries: A discrete choice approach. *Journal of Policy Modeling*, 37(4), 577–595.

Peersman, G., & Smets, F. (1998). Uncertainty and the Taylor rule in a simple model of the euro-area economy. Ghent University Working Paper.

Romer, D. (2000). Keynesian Macroeconomics without the LM Curve. *Journal of Economic Perspectives,* 14(2), 149–169.

Ryczkowski, M. (2016). Poland as an inflation nutter: The story of successful output stabilization. *Zbornik radova Ekonomskog fakulteta u Rijeci,* 34(2), 363–392.

Taylor, J. (2000). Reassessing Discretionary Fiscal Policy. *Journal of Economic Perspectives,* 14(3), 21–36.

Wang, Y., Jiang, C., Chang, H.-L., & Su, C.-W. (2015). Are Taylor Rules Valid in Central Eastern European Countries? *Ekonomický časopis*, 63(7), 665–685.

Assessment of the severity of armed conflicts based on the changes in the quality of life

Agata Sielska[1]

Abstract

Armed conflicts have a significant impact on the economy and quality of life. Their severity and intensity are usually measured based on war expenditures and losses, especially direct ones. It is difficult to assess the related opportunity costs which result i.e. from the weakening trade ties if the number of states involved in the conflict is growing. The paper attempts to assess the severity of international conflicts based on their characteristics as well as indirect effects which were approximated by changes in the multi-criteria assessment of the quality of life in an affected country. Quality of life assessments is based on economic, social and political factors, all of which are reflected in current quality of life rankings.

The approach suggested in the paper allows for a comprehensive and synthetic assessment of the impact of the chosen conflict on the economy and quality of life in affected country. The paper also attempts to examine to what extent the addition of the criteria related to the quality of life changes the assessment of the severity of the conflict made only on the basis of characteristics of the conflict itself. An important element of the analysis is the length of the considered time period (from the nineteenth century). Multicriteria rankings are the main method used in the paper. .

Keywords: *armed conflicts, quality of life, multicriteria rankings*

JEL Classification: *I310, C440, F510*

1. Introduction

Conflicts affect the countries' economy and quality of life in many dimensions. First of all they change conditions under which economics operates which in turn modifies the allocation of resources, uncertainty and risk levels. Consumption of goods and services, trade (Feldman and Sadeh, 2018), education (Lai and Tyne, 2007), politics are healthcare (Lai and Tyne, 2007) are affected. The overall quality of life decreases. On the other hand it is mentioned that due to those changing conditions requiring specific actions from economic agents conflicts may also contribute to the increase of efficiency (Kang and Meernik, 2005) or development of new technologies (Ruttan, 2006). It can be shown that there is a negative relation between quality of life and the intensity of internal conflicts and a positive relationship in case of external ones (Sielska, 2018a).

The aim of the paper is to assess the severity of international (interstate) conflicts based on their characteristics as well as on indirect effects which were approximated by changes in the multi-criteria assessment of the quality of life in an affected country. The paper consists of 4 parts. In the first part data and criteria used for evaluating both conflicts and quality of life are presented. The second one describes the outranking approach and weights elicitation methods used in the study. In the last part results are presented and discussed.

[1] Corresponding author: Warsaw School of Economics, Department of Applied Economics, 162 Niepodległości, 02-554 Warsaw, e-mail: asiels@sgh.waw.pl.

2. Data and criteria

Data for the analysis comes from the Correlates of War database (Sarkees and Wayman, 2010). We use definition provided by Sarkees, according to which interstate wars 'take place between or among (…) members of the interstate system' including additional criteria.

Some of the conflicts were removed from the study due to the lack of corresponding data which would allow the author to evaluate the quality of life in a participating country. As a result, the analysis was conducted on the basis of 44 interstate conflicts listed in table 1.

Due to the nature of the data the severity of each war was assessed from a given participants' perspective. Therefore the notation used in the next parts of the paper is as follows: state_warcode. For example GBR_65 means the conquest of Egypt evaluated from the perspective of the Great Britain.

Two war characteristics (battle-related combatant fatalities suffered by the participating state and binary variable reflecting if the participant won the conflict) were always included in the criteria set. Quality of life assessments are based on the Clioinfra data (https://www.clio-infra.eu) and (van Zanden et al., 2014) and are taken from the study by Sielska (2018b). Because the data frequency in the most cases is 10 years, all conflicts that occur in the same decade are grouped. In such cases the codes of additional conflicts appear in the war symbol after the semicolon. Four different approaches were considered for quality of life assessment. The first one (further denoted by P) is based on the criteria referring to the political state of the country including polity2 index (for years 1820–1980) and democracy index (1820–1980). Demographic approach (D) is based on the homicide rate (1820–1980), average height in population (1820–1980) and life expectancy at birth (1870–1980). Economic assessment (E) considers growth rate of GDP per capita (1820–1980), income inequality (1820–1980) and real wages of construction workers (1910–1980). Multidimensional approach (T) uses: numeracy index (1810–1970), inflation (1810–1990), GDP per capita (1810–1990), GDP per capita growth rate (1900–1990), urbanization ratio (1810–1990), average height in population (1810–1990), numeracy inequality (1830–1900), real wages of construction workers (1900–1990), average years of education (1900–1990), homicide rate (1930–1990), life expectancy (1900–1990), education inequality (1910–1990), income inequality (1950–1990), ratio of female to male life expectancy (1940–1990).

Table 1. Analysed conflicts

War code (1)	War name (1)	War code (2)	War name (2)
1	Franco-Spanish War	83	Sino-Russian
4	First Russo-Turkish	85	Russo-Japanese
7	Mexican-American	94	Second Spanish-Moroccan
10	Austro-Sardinian	97	Italian-Turkish
13	First Schleswig-Holstein	106	World War I

War code (1)	War name (1)	War code (2)	War name (2)
16	Roman Republic	107	Estonian Liberation
19	La Plata	108	Latvian Liberation
22	Crimean	109	Russo-Polish
25	Anglo-Persian	116	Franco-Turkish
28	Italian Unification	118	Manchurian
31	First Spanish-Moroccan	127	Conquest of Ethiopia
40	Franco-Mexican	133	Changkufeng
46	Second Schleswig-Holstein	136	Nomonhan
49	Lopez	139	World War II
52	Naval War	142	Russo-Finnish
55	Seven Weeks	145	Franco-Thai
58	Franco-Prussian	151	Korean
61	Second Russo-Turkish	155	Sinai War
65	Conquest of Egypt	156	Soviet Invasion of Hungary
67	Sino-French	158	IfniWar
79	Spanish-American	163	Vietnam War, Phase 2
82	Boxer Rebellion	170	Second Laotian, Phase 2

3. TOPSIS method

Multicriteria ranking are constructed using the TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) method (Hwang and Yoon, 1981), which is also used to assess the quality of life (Sahin and Yapici Pehlivan, 2017; Sanda and Krupka, 2016; Sanda and Mandys, 2017). The first step of constructing a ranking is to evaluate all alternatives on the basis of all decision criteria taking into account criteria weights. Characteristic feature of the TOPSIS method is that the final rank of the alternative is based on its relative distances from ideal solution and negative ideal solution . Those reference points are defined according to formulas (1–4)

$$T^{+} = (t_1^{+}, t_2^{+}, \dots, t_n^{+}) \tag{1}$$

where

$$t_j^{+} = \begin{cases} \max\limits_i t_{ij} & \text{for maximised criteria} \\ \min\limits_i t_{ij} & \text{for minimised criteria} \end{cases} \tag{2}$$

t_{ij} – evaluation of the alternative i based on j-th criterion (considering criterion importance).

$$T^{-} = (t_1^{-}, t_2^{-}, \dots, t_n^{-}) \tag{3}$$

where

$$t_j^- = \begin{cases} \max\limits_i t_{ij} & \text{for maximised criteria} \\ \min\limits_i t_{ij} & \text{for minimised criteria} \end{cases} \tag{4}$$

The closer the alternative is to the ideal solution, the better it is considered. Similarly, the closer the alternative is to the negative ideal, the worse it is considered. The final rank is calculated based on the following relative distance:

$$D_p(a_i) = \frac{d_p^-(a_i)}{d_p^-(a_i) + d_p^+(a_i)} \tag{5}$$

Where $d_p^{-(a_i)}$ denotes to the distance to the negative ideal, and $d_p^{+(a_i)}$ denotes distance to the ideal solution.

Criteria weights were calculated using ROC (Rank Order Centroid) approach which is one of the most popular (see for example Sielska, 2015; Alfares and Duffuaa, 2016) methods which allows for eliciting criteria weights based only on the ranking of criteria. All possible sets of weights were considered, i.e. for 3 criteria we build 6 rankings, for 4–24, for 5–120 and in the last case (6 criteria) 720 rankings were constructed. Because of the limited volume of the paper we do not present the entire rankings. We focus on the leading positions and overall similarity of the rankings instead.

4. Results

The following tables present the results of ranking based on the core set of two war characteristics with addition of other criteria representing the change in the quality of life after the conflict. Because of the limited volume of the paper only median ranks are presented. Numbers in the parentheses denote median ranks obtained for the core dataset (two war characteristics only) while columns refer to the type of approach to quality of life assessment.

Table 2. Most severe conflicts. Evaluation based on 3 criteria

Conflict	P	T	E	D
ARG_49	3.5 (17.0)			
BRA_49	4.5 (21.0)			
ESP_158				1.0 (4.0)
ESP_31				6.5 (9.5)
ESP_52			2.0 (26.0)	
ESP_94			4.5 (8.0)	6.0 (7.0)
FRA_1		1.0 (37.0)		3.0 (5.0)

Conflict	P	T	E	D
FRA_139;145	1.0 (36.0)			
FRA_16		4.0 (36.0)	4.0 (7.0)	6.0 (6.0)
FRA_40	2.0 (40.0)			
FRA_82			3.0 (2.0)	
GBR_22;25		2.0 (25.0)		
GBR_82				2.0 (2.0)
MEX_40	4.0 (41.0)			
NLD_139		3.0 (30.0)		
NLD_151			1.0 (25.0)	
USA_79		5.0 (34.0)		

It should be noted that evaluations obtained by TOPSIS method are relative. Therefore conflicts that involve only one of the analysed states are more likely to be ranked higher than conflicts involving most of the states and to the similar degree. In case of the results presented in table 2 it can be seen that the additional criteria change the first positions. For example Franco-Spanish war for France (FRA_1) is considered the most severe for the general point of view, while its median rank is 37 if assessed only on the basis of core criteria. This is the most distinct example, but it can be clearly seen that the changes in the quality of life (measured by any approach) provide additional information. Including additional criteria, connected to different spheres of the quality of life leads to different conclusions. For example the impact of Ifni War on the demographic and social situation of Spain is clearly visible (ESP_158), as well as the effect of Korean war on the economics of Netherlands (NLD_151). The effect of both II World War and Franco-Thai War on the political situation of France was assessed as significant.

Table 3. Most severe conflicts. Evaluation based on 4 criteria

Conflict	P+T	E+P	E+T	E+D	D+P	D+T
ARG_49	6.0 (10.0)					
ESP_158				6.0 (4.0)		
ESP_31					5 (9)	
ESP_52				1.0 (19.0)		
ESP_94				3.5 (7.0)		
FRA_1	2.0 (5.0)		1.5 (5.0)	6.0 (5.0)		1.0 (4.0)
FRA_139;145	2.5 (20.0)	1.5 (24.5)			1.0 (18.5)	
FRA_16			5.0 (6.0)	5.0 (6.0)		3.0 (5.0)

Conflict	P+T	E+P	E+T	E+D	D+P	D+T
FRA_22;28		6.5 (15.5)			4.0 (11.0)	
FRA_40	5.0 (22.0)	3.0 (26.5)			2.0 (20.5)	
FRA_67						5.0 (6.0)
FRA_82				2.0 (1.0)		
GBR_22;25	2.0 (12.0)	5.0 (14.0)	2.5 (12.0)		4.0 (10.0)	2.0 (8.0)
GBR_65					7.0 (3.0)	
ITA_127						4.0 (7.0)
NLD_139	3.5 (9.0)		3.0 (10.0)			
NLD_151		2.0 (21.5)	2.5 (20.0)			
USA_79			6.0 (8.0)			

In the case of 4 criteria evaluation (two war characteristics and two different approaches to quality of life measurement) similar conclusions may be drawn. Top positions in rankings change depending on the criteria taken into account (table 3). Evaluation based on 5 criteria (table 4) leads to similar conclusions. It should be noted, however, that in this case Franco-Spanish war for France (FRA_1) once again can be assessed as critical. Boxer Rebellion (82) and Crimean war (22) are ranked high considering their impact on different countries. In general approach if whole set of criteria is taken into account Korean (151), Sinai (155) and Ifni (158) wars are ranked highest (table 5). In all cases (3, 4 and 5 criteria) 2 groups of conflicts can be distinguished. For some wars, severity is rated significantly higher if the quality of life is taken into account, while for the other group additional criteria do not affect the rank. In case of 3 criteria first group includes i.e. Franco-Mexican war (40) assessed from French perspective and naval war (ESP_52) while both Franco-Spanish War (1) and Boxer Rebellion for France (82) belong to the other group. In case of 4 criteria respective examples are: Korean War (NLD_151) and Boxer Rebellion (FRA_82).

However, it should be also noted that the approach to the quality of life cannot be neglected. For example in the case of Franco-Mexican war (FRA_40) taking into account political, demographic and total set of criteria (P+D+T) do not provide much additional information, while other approaches do, which results in the change of ranks.

Table 4. Most severe conflicts. Evaluation based on 5 criteria

Conflict	E+D+T	E+D+P	P+D+T	P+E+T
ESP_52		3.0 (16.5)		
FRA_1	1.0 (4.0)		1.5 (14.0)	2.0 (5.0)
FRA_139;145		2.0 (18.5)	2.5 (5.5)	3.0 (19.0)

Conflict	E+D+T	E+D+P	P+D+T	P+E+T
FRA_16	4.0 (5.0)			
FRA_22;28		6.5 (11.0)		
FRA_40	3.0 (15.5)	2.0 (20.5)	4.0 (4.5)	5.0 (20.0)
FRA_67			6.0 (12.0)	
FRA_82	3.0 (1.0)	6.0 (1.0)		
GBR_139	6.0 (14.5)			
GBR_22;25	3.0 (8.0)	5.0 (10.0)	2.0 (11.0)	3.0 (11.0)
GBR_65			6.0 (15.0)	
ITA_127	7.0 (7.0)			
NLD_139				4.0 (9.0)
NLD_151				3.0 (18.0)
USA_79				8.0 (8.0)

Table 5. Evaluation based on 6 criteria

Conflict	Median rank (including effects on the quality of life)	Median rank (not including effects on the quality of life)
FRA_1		4.0
FRA_139	8.0	
FRA_151;155;158	1.0	
FRA_16	8.0	5.0
FRA_22;28	6.0	
FRA_67	4.0	
FRA_82		1.0
GBR_151;155	7.0	
GBR_65	7.0	3.0
GBR_82	4.0	2.0

In the final step we assessed the similarity of core and more complex rankings that take into account the quality of life. Spearman rank correlation coefficients are presented in fig. 1. In general, rankings with additional criteria are not similar to the original ones, build only on basis of war characteristics. The greatest similarity can be observed in case of demographic and economic changes. It is worth pointing out that about half of the relations are insignificant (at significance level 0.01) which is represented on the plot by the blank bars.

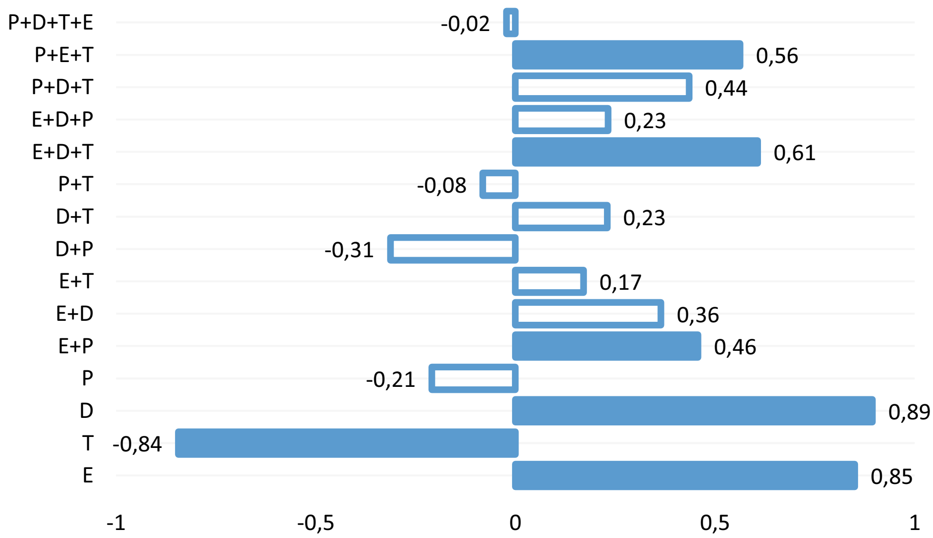


Fig. 1. Rank correlation coefficients

Conclusions

Severity of 44 interstate conflicts was assessed on the basis of both conflict characteristics and its' impact on the relative quality of life in a participating country. The results show that for all approaches to the quality of life measurement additional criteria change positions in the rankings. Secondly, those changes depend on the criteria and approach used. It means that even if changes in one sphere of well-being are reflected by the changes in the other (e.g. economic conditions influence society and demography and vice versa) none of them should be neglected while building a ranking. In general, some sets of criteria connected to the quality of life do not provide additional information even if some of the individual criteria, included in respective sets do. Due to this fact we may conclude that it is not only possible to assess the severity of each conflict, but to describe those spheres of wellbeing which are not significantly affected by the conflict.

The overall similarity of rankings built based on war characteristics only and more complex rankings is high only if demographic or economic approaches are considered.

Acknowledgements

Paper is a part of the project "Wielokryterialna analiza konfliktów zbrojnych z uwzględnieniem ich efektów" (Multi-criteria analysis of armed conflicts including their effects) KZiF/BMN/23/18 at Warsaw School of Economics.

References

Alfares, H.K., & Duffuaa, S.O. (2016). Simulation-Based Evaluation of Criteria Rank-Weighting Methods in Multi-Criteria Decision-Making. *International Journal of Information Technology & Decision Making,* 15(1), 43–61.

Clio Infra | Reconstructing Global Inequality. Retrieved July 20th, 2017, from https://www.clio-infra.eu/

Feldman, N., & Sadeh, T. (2018). War and Third-Party Trade. *Journal of Conflict Resolution,* 62(1), 119–142.

Hwang, C.L., & Yoon, K. (1981). Multiple Attribute Decision Making: Methods and Applications. New York: Springer-Verlag.

Kang, S., & Meernik, J. (2005). Civil War Destruction and the Prospects for Economic Growth. *The Journal of Politics*, 67(1), 88–109.

Lai, B., & Thyne, C. (2007). The Effect of Civil War on Education, 1980–97. *Journal of Peace Research*, 44(3), 277–292.

Ruttan, V.W. (2006). Is war necessary for economic growth? Military procurement and technology development. Oxford: Oxford University Press.

Sahin, A., & Yapici Pehlivan, N. (2017). Evaluation of life quality by integrated method of AHP and TOPSIS based on interval type-2 fuzzy sets. *Hacettepe Journal of Mathematics and Statisctics,* 46(3), 511–523.

Sanda, M., & Krupka, J. (2016). Quality of Life Evaluation in Visegrad Group and Progression of Evaluation in Years 2008–2014. In: Stejskal, J., Krupka, J. (eds.) *Proceedings of the 11th International Scientific Conference Public Administration 2016,* 251–259.

Sanda, M., & Mandys, J. (2017). Quality of Life Evaluation in NUTS3 Regions of the Czech Republic and Progression of Evaluation in Years 2000–2015. In: Matejova, L. (ed.), *Proceedings of the 21st International Conference Current Trends in Public Sector Research*, 113–121.

Sarkees, M.R. Inter-state Wars (Version 4.0): Definitions and Variables (PDF document). Retrieved from Correlates of War Online Website: http://www.correlatesofwar.org/data-sets/COW-war/inter-state-wars-codebook/at_download/file.

Sarkees, M.R., & Wayman, F. (2010). Resort to War: 1816–2007. Washington DC: CQ Press. http://www.correlatesofwar.org/data-sets/COW-war.

Sielska, A. (2015). The impact of weights on the quality of agricultural producers' multicriteria models. *Operations Research and Decisions,* 25(4), 51–69.

Sielska, A. (2018a). Armed Conflicts And Multicriterial Evaluation Of The Development Of The Country In A Long-Term Perspective. In: Papież, M., Śmiech, S. (eds.) *The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena : conference proceedings*, 453–462.

Sielska, A. (2018b). Konflikty zbrojne a jakość życia w ujęciu wielokryterialnym. Report of the project "Konflikty zbrojne a jakość życia w ujęciu wielokryterialnym" (Armed conflicts and quality of life – multicriteria approach) KZiF/BMN/24/17 at Warsaw School of Economics. Unpublished.

Zanden, J.L., & Baten, J. (2014). How was life?: global well-being since 1820. Paris: OECD Publ.

An example of application of optimal sample allocation in a finite population

Dominik Sieradzki[1], Wojciech Zieliński[2]

Abstract
The problem of estimating a proportion of objects with particular attribute in a finite population is considered. This paper shows an example of the application of estimation fraction using new proposed sample allocation in a population divided into two strata. Variance of estimator of proportion which uses proposed sample allocation is compared to variance of the standard one.

Keywords: *survey sampling, sample allocation, stratification, estimation, proportion*
JEL Classification: *C83, C99*

1. Introduction

In microeconomics, the main subject of interest is human as a managing individual, whereas macroeconomics places the greatest emphasis on households and enterprises (Bartkowiak, 2003). Such objects frequently form multi-million populations. Due to amount of costs it is impossible to subject the population of interest to exhaustive sampling, even for Statistical Office. In economics populations consist of a finite number of units. Survey sampling deals with finite populations. Therefore a sample is drawn from the population. When sampling, two types of errors can be distinguished: sampling error and non-sampling error. Non-sampling error is associated with the non-response problem. Proposal on how to deal with such an issue can be found in Hansen and Hurwitz (1946) or Chaudhuri et al. (2009). This article is focused on sampling error, hence it is assumed that responses were obtained from all of the chosen units in the sample. The sampling error, among others, depends on sampling scheme. In the next part of this paper an example of application of sample allocation proposed in Sieradzki and Zieliński (2019) is presented.

In economics the aim of the research is often to inference about dychotomus occurences, for example support for a particular party or candidate in elections (Szreder, 2010), unemployment rate (Hadaś-Dyduch, 2015), farmers' decision about production credit and EU measures (using these funds or not) (Roszkowska-Mądra and Mańkowski, 2010) or deciding on ecological farming (Sieradzki and Stefańczyk, 2017). Consider a problem of support for a particular candidate in the elections. The main issue to consider in the study is a population $U = \{u_1, ..., u_N\}$ which contains a finite number of N people who may vote. In this population a number of people who support a particular candidate is observed. All the units in this population could be considered as a vector $\boldsymbol{Y} = (Y_1, ..., Y_N)^T$ where $Y_k = 1$ if k-th person supports a candidate and $Y_k = 0$ if k-th

[1] Corresponding author: Warsaw University of Life Sciences, Department of Econometrics and Statistics, 159 Nowoursynowska St., Warsaw, Poland, dominik_sieradzki@sggw.pl.

[2] Warsaw University of Life Sciences, Department of Econometrics and Statistics, 159 Nowoursynowska St., Warsaw, Poland, wojciech_zielinski@sggw.pl.

person doesn't support a candidate, for $k = 1, \ldots, N$. Hence $\sum_{k=1}^{N} Y_k$ stands for an unknown number of people in the population who support a candidate. Let us denote this number as M. The aim of the study is to estimate an unknown proportion (fraction) $\theta = \frac{M}{N}$. A sample of size n is drawn using simple random sampling without replacement scheme. In the sample number of people who support a candidate is observed and this number is a random variable. Let ξ denote this random variable. The random variable ξ has hypergeometric distribution (Barnett, 1974; Greene and Wellner, 2017) and its probability distribution function is

$$P_{\theta,N,n}\{\xi = x\} = \frac{\binom{\theta N}{x}\binom{(1-\theta)N}{n-x}}{\binom{N}{n}}, \tag{1}$$

for integer x from set $\{\max\{0, n-(1-\theta)N\}, \ldots, \min\{n, \theta N\}\}$. Unbiased estimator with minimal variance of the parameter θ is $\hat{\theta}_c = \frac{\xi}{n}$ (Cochran, 1977; Steczkowski, 1995). Variance of that estimator equals $D_\theta^2(\hat{\theta}_c) = \frac{1}{n^2} D_\theta^2 \xi = \frac{\theta(1-\theta)}{n} \frac{N-n}{N-1}$ for all θ. It is obvious that the worst variance occurs when θ equals $\frac{1}{2}$.

2. Stratified estimator

In some cases, the population of the study is strongly variable and support for a particular candidate depends on e.g. region or gender of voters. Therefore, the sample is drawn due to simple random sampling without replacement scheme, so it can contain only people who support a candidate. To avoid this, stratified random sampling is used. This method of sampling assumes a division of the population among disjoint strata. After such division of the population, random sample is taken in each strata (Cochran, 1977). Let us divide the population into two disjoint strata U_1 and U_2, $U = U_1 \cup U_2$ of N_1 and N_2 people, respectively. The details of division of the population in stratification can be found in (Horgan, 2006; Hidiroglou and Kozak, 2017). For example, support in elections can depend on dominant political option at the time. In each strata fraction of people who support a candidate equals θ_1 and θ_2, respectively. The aim of the study is still to estimate the overall proportion θ, not θ_1 and θ_2. Let w_1 denote a contribution of the first strata, i.e. $w_1 = \frac{N_1}{N}$. Obviously the overall proportion equals

$$\theta = w_1\theta_1 + w_2\theta_2 \tag{2}$$

where $w_2 = 1 - w_1$ is a contribution of the second strata. It seems intuitively obvious to take as our estimate of θ,

$$\hat{\theta}_w = w_1\frac{\xi_1}{n_1} + w_2\frac{\xi_2}{n_2}, \tag{3}$$

where n_1 and n_2 denote sample sizes from the first and the second strata, respectively. Now, there are two random variables describing the number of units with a particular attribute in samples drawn from each strata:

$$\xi_1 \sim H(N_1, \theta_1, N_1, n_1),\ \xi_2 \sim H(N_2, \theta_2, N_2, n_2) \tag{4}$$

Let us consider costs of sampling. Suppose that cost of sampling from the first strata equals c_1 and from the second one c_2. Funds for the poll are limited. Cost function is of the form:

$$C = c_1 n_1 + c_2 n_2 \tag{5}$$

The main goal is to estimate the overall fraction θ, not fraction in each strata. The parameter θ_1 will be considered as a nuisance one. This parameter will be eliminated by appropriate averaging. Note that for a given $\theta \in [0, 1]$, parameter θ_1 is a fraction M_1/N_1 (it is treated as the number, not as the random variable) from the set

$$A = \left\{ a_\theta, a_\theta + \frac{1}{N_1}, \dots, b_\theta \right\}, \tag{6}$$

where

$$a_\theta = \max\left\{0, \frac{\theta - w_2}{w_1}\right\} \text{ and } b_\theta = \min\left\{1, \frac{\theta}{w_1}\right\} \tag{7}$$

and let L_θ be cardinality of A.

It is facile to prove that estimator $\hat{\theta}_w$ is an unbiased estimator of fraction θ (Sieradzki and Zieliński, 2017). Hence, it is necessary to compare variances of both estimators. Averaged variance of estimator $\hat{\theta}_w$ having regard to cost, could be as follows:

$$D_\theta^2 \hat{\theta}_w = \frac{1}{L_\theta} \sum_{\theta_1 \in A} \left(\frac{w_1^2}{n_1} \theta_1 (1 - \theta_1) \frac{N_1 - n_1}{N_1 - 1} + \frac{w_2^2}{(C - c_1 n_1)/c_2} \frac{\theta - w_1\theta_1}{w_2} \left(1 - \frac{\theta - w_1\theta_1}{w_2}\right) \frac{N_2 - (C - c_1 n_1)/c_2}{N_2 - 1} \right) \tag{8}$$

Detailed analysis of variance $D_\theta^2 \hat{\theta}_w$ can be found in (Sieradzki and Zieliński, 2017; Sieradzki and Zieliński, 2019). In further steps: firstly, finding 'the worst' situation, i.e. such value of proportion for which variance $D_\theta^2 \hat{\theta}_w$ takes on its maximal value is needed. Then it is necessary to find such (n_1^{opt}, n_2^{opt}) that minimises this maximal variance. The optimal allocation of the sample is $(n_2^{opt} = (C - c_1 n_1^{opt})/c_2)$:

$$n_1^{opt} = \begin{cases} \frac{C\sqrt{(N_2-1)}w_1}{c_1\sqrt{(N_2-1)}w_1 - \sqrt{c_1 c_2 w_2 (N(w_1^2 - 3w_1 + 1.5) - w_1)}} & \text{for } w_1 \leqslant w_1^* \\ \text{numerical solution available} & \text{for } w_1 \leqslant w_1^* \end{cases} \tag{9}$$

where w_1 equals about 0.46 (Sieradzki and Zieliński, 2018).

In order to compare effectiveness of both estimators, it is necessary to determine sample size for the classical estimator $\hat{\theta}_c$. Let n_c denote a sample size for estimator $\hat{\theta}_c$. Sample size could be described as follows (Sieradzki and Zieliński, 2019):

$$n_c = \frac{C}{w_1 c_1 + w_2 c_2}. \tag{10}$$

Example of application of this sample allocation will be considered in the next section.

3. Example

Suppose that the aim of the research is to estimate support for a candidate (it will be referred to as a candidate “A”) in second round of presidential elections in Poland. In Poland there are more than 30 milions people who are entitled to vote (due to offcial statistics, in 2015 there were $N = 30\,709\,281$ voters). The standard way of estimation θ is to take a sample of size n_c due to the scheme of simple sampling without replacement. In the sample the number of answers “yes, I will vote for candidate A” is counted. Let us denote this number as ξ. Obviously the standard estimator of the support is $\frac{\xi}{n_c}$.

In 2015 some party which is linked with candidate ‘A’ won in 7 of 16 voivodeships. In those voivodeships there were 14 526 524 people who may vote. In the remaining ones there were 16 182 757 voters. Hence, let us divide the population of electorate into two strata: the first one of the weight $w_1 = 14\,526\,524/30\,709\,281 = 0.47$ and the second one of the weight $w_2 = 16\,182\,757/30\,709\,281 = 0.53$. Suppose that costs of sampling from the first and the second strata equal $c_1 = 3$ and $c_2 = 1$, respectively. Funds for the sampling for this poll equal e.g. $C = 1200$. These are exemplary values of these magnitudes, but for all values sample allocation is calculated in the same way. Sample size n_c equals 618. The optimal division (n_1^{opt}, n_2^{opt}) of the sample for this numerical case could be calculated. After some calculations (which can be done in e.g. Mathematica) optimal sample allocation is obtained: $n_1 = 242$, $n_2 = 474$.

Suppose that in the whole sample 100 ’yes’ answers were obtained. The point estimate of the support with classical estimator equals $\hat{\theta}_c = 100/618 = 16.18\%$ and its estimated variance equals

$$\hat{v}_c(100) = \frac{\hat{\theta}_c(1-\hat{\theta}_c)}{618}\frac{30709281-618}{30709281-1} = 0.00021946, \tag{11}$$

Suppose that in the sample of size n_1 from the first strata there were 10 ’yes’ answers and the number of ’yes’ answers in the sample of size n_2 equals 128. The point estimate of the support would be $\hat{\theta}_w = 16.14\%$. The estimated variance of the estimator $\hat{\theta}_w$ equals

$$\begin{aligned}\hat{v}(10,128) = &\left(\frac{14\,526\,524}{30\,709\,281}\right)^2 \frac{\frac{10}{242}\left(1-\frac{10}{242}\right)}{616}\frac{14\,526\,524-24}{14\,526\,524-1} \\ &+\left(\frac{16\,182\,757}{30\,709\,281}\right)^2 \frac{\frac{128}{474}\left(1-\frac{128}{474}\right)}{474}\frac{16\,182\,757-474}{16\,182\,757-1} = 0.00001516.\end{aligned} \tag{12}$$

The relative reduction of estimated variance equals

$$reduction = \left(1-\frac{\hat{v}_w(10,128)}{\hat{v}_c(100)}\right)\cdot 100\% = 30.94\%. \tag{13}$$

Table 1 shows other possible results of the poll, assuming that the overall 'yes' answers equal to 100, total funds equal 1200, costs of sampling from the first and the second stratum equal 3 and 1, respectively.

Table 1. Possible results for $\xi = 100$, $\hat{v}_c(100) = 0.000$, $c_1 = 3$, $c_2 = 1$, $C = 1200$, $n_1 = 242$, $n_2 = 474$, $n_c = 618$, $\hat{\theta}_w = 16.18\%$

ξ_1	ξ_2	**support**	**variance**	**reduction**
10	128	16.14%	0.000 151 6	30.94%
20	111	16.22%	0.000 175 0	20.28%
30	93	16.19%	0.000 193 0	12.08%
40	75	16.16%	0.000 206 1	6.10%
50	57	16.13%	0.000 214 3	2.33%
60	40	16.21%	0.000 218 8	0.31%
70	22	16.18%	0.000 217 4	0.94%
80	4	16.15%	0.000 211 2	3.77%

In Tables 2 and 3 possible results are given assuming that the overall positive answers are 300 and 400 respectively.

Table 2. Possible results for $\xi = 200$, $\hat{v}_c(200) = 0.000$, $c_1 = 3$, $c_2 = 1$, $C = 1200$, $n_1 = 242$, $n_2 = 474$, $n_c = 618$, $\hat{\theta}_w = 32.36\%$

ξ_1	ξ_2	**support**	**variance**	**reduction**
10	274	32.31%	0.000 178 8	49.53%
20	257	32.39%	0.000 215 0	39.29%
30	239	32.36%	0.000 246 6	30.37%
40	221	32.33%	0.000 273 3	22.83%
50	204	32.41%	0.000 295 4	16.6%
60	186	32.38%	0.000 312 5	11.78%
70	168	32.35%	0.000 324 7	8.32%
80	150	32.32%	0.000 332 1	6.24%
90	133	32.40%	0.000 335 2	5.37%
100	115	32.37%	0.000 332 9	6.01%
110	97	32.33%	0.000 325 8	8.02%
120	80	32.41%	0.000 314 6	11.17%
130	62	32.38%	0.000 297 9	15.89%
140	44	32.35%	0.000 276 3	21.99%

ξ_1	ξ_2	**support**	**variance**	**reduction**
150	26	32.32%	0.000 249 8	29.46%
160	9	32.40%	0.000 219 7	37.97%

Table 3. Possible results for $\xi = 300$, $\hat{v}_c(300) = 0.000$, $c_1 = 3$, $c_2 = 1$, $C = 1200$, $n_1 = 242$, $n_2 = 474$, $n_c = 618$, $\hat{\theta}_w = 48.54\%$

ξ_1	ξ_2	**support**	**variance**	**reduction**
10	421	48.59%	0.000 094 7	76.57%
20	403	48.56%	0.000 144 7	64.19%
30	385	48.53%	0.000 189 9	53.01%
40	367	48.5%	0.000 230 3	43.03%
50	350	48.58%	0.000 265 2	34.4%
60	332	48.55%	0.000 295 9	26.8%
70	314	48.52%	0.000 321 7	20.41%
80	297	48.6%	0.000 342 4	15.29%
90	279	48.57%	0.000 358 6	11.28%
100	261	48.54%	0.000 369 9	8.47%
110	243	48.51%	0.000 376 4	6.87%
120	226	48.59%	0.000 378 1	6.45%
130	208	48.55%	0.000 375	7.22%
140	190	48.52%	0.000 367	9.2%
150	172	48.49%	0.0003541	12.38%
160	155	48.57%	0.000 336 8	16.66%

Tables 4–6 contain proper columns, assuming that cost of sampling in the second strata is greater than in the first strata, i.e. $c_1 = 1$ and $c_2 = 3$.

Table 4. Possible results for $\xi = 100$, $\hat{v}_c(100) = 0.000$, $c_1 = 1$, $c_2 = 3$, $C = 1200$, $n_1 = 405$, $n_2 = 265$, $n_c = 582$, $\hat{\theta}_w = 17.18\%$

ξ_1	ξ_2	**support**	**variance**	**reduction**
10	81	17.22%	0.000 234 2	4.23%
20	75	17.2%	0.000 237 2	2.98%
30	69	17.19%	0.000 238 5	2.45%
40	63	17.17%	0.000 238 1	2.63%

ξ_1	ξ_2	support	variance	reduction
50	57	17.16%	0.000 235 9	3.52%
60	51	17.14%	0.000 232	5.13%
70	45	17.12%	0.000 226 3	7.45%
80	39	17.11%	0.000 218 9	10.49%

Table 5. Possible results for $\xi = 200$, $\hat{v}_c(200) = 0.000$, $c_1 = 1$, $c_2 = 3$, $C = 1200$, $n_1 = 405$, $n_2 = 265$, $n_c = 582$, $\hat{\theta}_w = 34.36\%$

ξ_1	ξ_2	support	variance	reduction
10	157	32.28%	0.000 264 5	25.31%
20	152	32.46%	0.000 280 5	20.79%
30	146	32.44%	0.000 295 5	16.56%
40	140	32.43%	0.000 308 8	12.82%
50	134	32.41%	0.000 320 3	9.58%
60	128	32.4%	0.000 33	6.82%
70	122	32.38%	0.000 338	4.56%
80	116	32.36%	0.000 344 3	2.79%
90	110	32.35%	0.000 348 8	1.52%
100	104	32.33%	0.000 351 6	0.74%
110	98	32.32%	0.000 352 6	0.45%
120	92	32.3%	0.000 351 9	0.65%
130	86	32.28%	0.000 349 4	1.35%
140	80	32.27%	0.000 345 2	2.54%
150	75	32.45%	0.000 341	3.73%
160	69	32.44%	0.000 333 4	5.86%

Table 6. Possible results for $\xi = 300$, $\hat{v}_c(300) = 0.000$, $c_1 = 1$, $c_2 = 3$, $C = 1200$, $n_1 = 405$, $n_2 = 265$, $n_c = 582$, $\hat{\theta}_w = 51.55\%$

ξ_1	ξ_2	support	variance	reduction
10	254	51.49%	0.000 054 8	87.23%
20	248	51.48%	0.000 088 6	79.36%
30	242	51.46%	0.000 120 6	71.89%
40	237	51.64%	0.000 147 9	65.54%

ξ_1	ξ_2	**support**	**variance**	**reduction**
50	231	51.63%	0.000 176 6	58.85%
60	225	51.61%	0.000 203 6	52.56%
70	219	51.6%	0.000 228 9	46.67%
80	213	51.58%	0.000 252 4	41.2%
90	207	51.56%	0.000 274 1	36.13%
100	201	51.55%	0.000 294 1	31.46%
110	195	51.53%	0.000 312 4	27.21%
120	189	51.52%	0.000 328 9	23.36%
130	183	51.5%	0.000 343 7	19.92%
140	177	51.49%	0.000 356 7	16.88%
150	171	51.47%	0.000 368	14.25%
160	165	51.45%	0.000 377 5	12.03%

Conclusions

In the article an example of application of averaged sample allocation was presented. The classical estimator and stratified estimator were compared with respect to their estimated variances. The variance of estimator depends strongly on costs of sampling c_1, c_2 and limited funds C. In the numerical study it was shown that whatever value c_1, c_2 and C have, the estimated variance of is smaller than estimated variance of . The reduction of the variance is up to 90%. It is worth noting that proposed sample allocation does not need any preliminary investigation, which is necessary in the case of Neyman allocation.

References

Barnett, V. (1974). *Elements of Sampling Theory*, The English Universities Press Ltd.

Bartkowiak, R. (2003). *Historia myśli ekonomicznej*, Warszawa: Polskie Wydawnictwo Ekonomiczne.

Chaudhuri, A., Christofides, T.C., & Saha, A. (2009). Protection of privacy in efficient application of randomized response techniques, *Statistical Methods and Applications,* 18, 389–418.

Cochran, W.G. (1977). *Sampling Techniques* (3rd ed.), New York: John Wiley.

Greene, E., & Wellner, J.A. (2017). Exponential bounds for the hypergeometric distribution, *Bernoulli,* 23, 1911–1950.

Hadaś-Dyduch, M. (2015). Polish macroeconomic indicators correlated-prediction with indicators of selected countries, In: Papież, M. & and Śmiech, S. (Eds.). *9th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings. Cracow: Foundation of the Cracow,* 68–76.

Hansen, M.H. & Hurwitz, W.N. (1946). The problem of non-response in sample surveys, *Journal of the American Statistical Association*, 41, 236, 517–529.

Hidiroglou, M.A., & Kozak, M. (2017). Stratification of Skewed Populations: A Comparison of Optimisation-based versus Approximate Methods, *International Statistical Review,* 86, 87–105.

Roszkowska-Mądra, B., & Mańkowski D.R. (2010). Determinanty decyzji rolników o korzystaniu z funduszy unii europejskiej i kredytów na działalność rolniczą: przykład dla rolnictwa z rozwiniętym systemem produkcji mlecznej w województwie podlaskim, *Roczniki Nauk Rolniczych. Seria G, Ekonomika Rolnictwa*, 97, 1427.

Sieradzki, D., & Stefańczyk J. (2017). The conversion of the area of ecological crops in the selected EU states. *Economic Science for Rural Development*, 44, 190–196.

Sieradzki, D., & Zieliński W. (2017). Sample allocation in estimation of proportion in a finite population divided into two strata. *Statistics in Transition new series*, 18(3), 541–548, 10.21307.

Sieradzki, D., & Zieliński, W. (2019). Cost Issue in Estimation of Proportion in a Finite Population Divided Among Two Strata. *arXiv preprint arXiv:1903.09935*.

Steczkowski, J. (1995). *Metoda reprezentacyjna w badaniach zjawisk ekonomiczno-społecznych*. Warszawa: PWN.

Szreder M. (2010). *Metody i techniki sondażowych badań opinii*. Kraków: Wydawnictwo UEK..

Diversification of the level and structure of research and development expenditure in the European Union countries

Elżbieta Sobczak[1], Małgorzata Markowska[2], Petr Hlaváček[3]

Abstract

The purpose of the study is to assess the degree of the European Union (EU) countries' diversification in terms of the research and development intensity and the structure of R&D expenditure by the source of funding. The analysis includes R&D expenditure as % of the gross domestic product and the structure of R&D expenditure broken down by the following sectors: business enterprise, government, higher education, private non-profit and the "rest of the world", as well as changes in this area in the years 2008–2016. The research used multidimensional statistical analysis with particular focus on classification methods. The study results allowed separating the relatively homogeneous classes of the EU countries in terms of R&D intensity and structure of research and development expenditure by their funding sources.

Keywords: *research and development expenditure, R&D intensity, source of R&D funds, the EU countries*
JEL Classification: *C19, F63, O52*

1. Introduction

Nowadays innovation is considered one of the most important factors having impact on international competition and socio-economic development of countries and regions. It is the feature of economies presenting the capacity to create, implement and absorb innovations. "The research and development activity, defined as creative work, approached in a methodical way to increase the resources of knowledge (...) and to develop new applications for the existing knowledge, remains one of many types of innovative activities" (Manual, 2015, p. 28) and can constitute the source of innovation. R&D can be described taking into account its two correlated aspects: scale and results. It allows analysing both the size and the intensity of expenditure invested in research and development, the funding sources and the R&D effects (Dworak and Grzelak, 2010; Ostaszewska and Tylec, 2016). R&D activities represent an important economic problem analysed by many researchers. In addition, the subject literature discusses issues related to the cyclical nature of R&D (Mand, 2019; Ouyang, 2011), the effectiveness of incurred expenditure (Sawulski, 2018; McGrath and Romeri, 1994) as well as financing of R&D activities (Hall, 2009; Howell, 2017).

The European Union, recognising the importance of innovation, adopted the Europe 2020 development strategy, which defines the targets allowing the Member States to ensure smart development based on knowledge and innovation. Achieving this goal requires considerable expenditure on R&D and applying mechanisms facilitating the transfer of knowledge and tech-

[1] Corresponding author: Wroclaw University of Economics, Department of Regional Economy, 118/120 Komandorska St., 53-345 Wroclaw, Poland, elzbieta.sobczak@ue.wroc.pl.

[2] Wroclaw University of Economics, Department of Regional Economy, 118/120 Komandorska St., 53-345 Wroclaw, Poland, malgorzata.markowska@ue.wroc.pl.

[3] Jan Evangelista Purkyně University in Ústí nad Labem, Department of Regional Development and Public Administration, 1 Pasteurova, 400 96 Ústí nad Labem, Czech Republic, Petr.Hlavacek@ujep.cz.

nology as well as absorbing innovation by the economy and enterprises. It is essential to monitor the realisation level of the targets set in the EU development strategy. The empirical research which was carried out follows this trend, focusing on the assessment of the diversification degree of the EU countries in terms of the intensity of R&D activity and the structure of research and development expenditure by their funding sources. The conducted analysis also covered the problem of changes occurring in this area.

2. Scope and methods of research

The subject of the analysis is the intensity and structure of expenditure on R&D activities by its funding sources in the European Union countries. Gross national expenditure on research and development covers all expenditure on R&D carried out on the territory of a given country in a given reporting period (GERD), including current and investment expenditure on R&D activity, but it does not include these funds' depreciation. The intensity of expenditure on R&D is defined as the ratio of internal expenditure on R&D against GDP (%).

The international research analysing expenditure incurred on R&D applies the classification of funding sources following the institutional classification in accordance with the Frascati Manual (Manual, 2015). There are 5 main funding sources of research and development activities, i.e.: business enterprise sector (BES), government sector (GOV), higher education sector (HES), private non-profit sector (PNP) and the rest of the world sector which covers institutions outside the territory of the country where R&D is carried out, including the EU institutions and bodies, international and supranational organizations regardless of the physical location of their offices or places of running a business. All the aforementioned sectors are, at the same time, the executors of R&D activities.

The data necessary to assess the intensity diversification and structure of R&D expenditure by its funding sources in the EU countries were retrieved from the Eurostat database. Based on the availability of statistical information, the research period covers the years 2008–2016.

The classifications of the EU countries were carried out using the following steps (the review of information on normalization methods, distance measures and classification methods can be found, e.g. in the studies by Hartigan (1975), Kukuła (2000)): determining the diversification among the analysed countries using the Euclidean Squared distance, hierarchical classification of countries using Ward's clustering method, identifying the number of classes based on the results of classifications presented on a dendrogram and the graph of fusion distance in the fusion stages, classification of countries using *k*-means method and the characteristics of the identified classes. In the case of the EU countries classification in terms of expenditure intensity on R&D, the output data were subjected to min-max normalization (Kukuła, 2000).

3. Diversification of the European Union countries in terms of the intensity of research and development activities

The strategic goal included in the Europe 2020 strategy is to allocate 3% of the EU GDP to investments in R&D. Fig. 1 presents the development of R&D intensity ratio in the EU as well as the minimum and maximum values in the EU countries in the years 2008–2016. The average intensity of expenditure on R&D in the EU, in the analysed period, increased slightly, i.e. from 1.43% in 2006 to 1.55% in 2016, not reaching the target value. The lowest share of expenditure on R&D against GDP (approx. 0.40%) was recorded in 2008–2012 and in 2015 in Cyprus, in 2012–2013 in Romania and in 2016 in Latvia. The countries characterised by the highest intensity of expenditure on R&D include Finland (2008–2012, 2014), Sweden (2013, 2015–2016), where the share of expenditure exceeded 3%, ranging from 3.17 to 3.75 %.

The European Union countries were characterised by a very high diversification in the intensity of expenditure on research and development (see Tab. 1). However, in the analysed period the diversification intensity of R&D expenditure declined from approx. 62% to approx. 56%. The empirical range of the analysed indicator showed a right-sided symmetry, as the majority of countries presented lower than average shares of expenditure on R&D in GDP.

The composition of classes of the EU countries identified by the intensity of R&D expenditure and its characteristics are presented in Table 2 and on Figure 2. As a result of applying the *k*-means method 3 classes of countries, characterised by a relatively similar share of expenditure on R&D in GDP in 2008–2016, were identified. Class 1, including countries with the highest intensity of expenditure on R&D reaching, on average, approx. 2.74% of GDP was the least numerous. The class 3 covered 10 EU countries characterised by the mean share of expenditure on R&D in GDP (approx. 1.48%). From among the EU10 it included the Czech Republic, Estonia, Hungary. The class of countries featuring the lowest intensity of expenditure on R&D (approx. 0.69% with a slight upward trend) included 10 countries with only Greece from EU15.

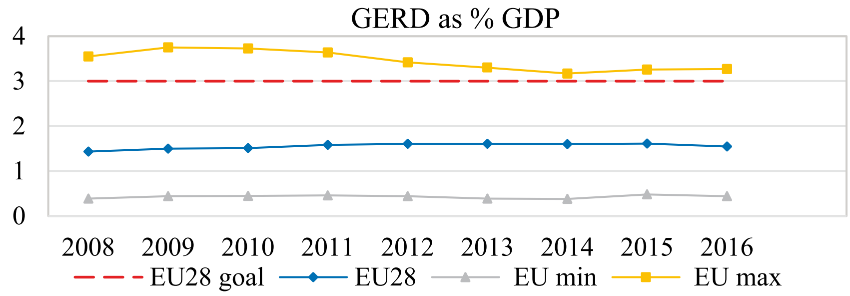


Fig. 1. The intensity of expenditure on R&D in the European Union countries in the years 2008–2016

Table 1. Descriptive parameters of the share of R&D expenditure in GDP in the EU28 in the years 2008–2016

Descriptive parameters	2008	2009	2010	2011	2012	2013	2014	2015	2016
Min	0.39	0.44	0.45	0.46	0.44	0.39	0.38	0.48	0.44
Max	3.55	3.75	3.73	3.64	3.42	3.30	3.17	3.26	3.27
Range	3.16	3.31	3.28	3.18	2.98	2.91	2.79	2.78	2.83

Descriptive parameters	2008	2009	2010	2011	2012	2013	2014	2015	2016
Mean	1.43	1.50	1.51	1,58	1.61	1.61	1.60	1.61	1.55
Median	1.29	1.38	1.43	1.46	1.34	1.36	1.35	1.31	1.27
CV (%)	62.43	62.80	59.98	57.92	56.55	55.63	54.13	52.47	56.87

where: CV – coefficient of variation

Class 1 of countries presenting high expenditure intensity on R&D in 2008 was characterised by the most extensive internal diversification (CV = 27.13%), which showed a declining tendency, reaching the value of 16.38% in 2016. In 2008 class 3 of the EU countries was characterised by the lowest dispersion (CV = 15.85%). However, in the analysed period the diversification of countries included in this class was continuously increasing, reaching 19.59% variability in 2018. Class 2 of countries with the lowest share of expenditure on R&D in GDP showed large diversification in 2008 (CV = 25.99%), indicating an increasing trend, in 2016 the CV coefficient reached the value of 27.76%. In 2016 it was the most diversified class of countries in terms of expenditure intensity on R&D.

Table 2. Classification of the European Union countries in terms of the share of R&D expenditure in GDP in the years 2008–2016 using *k*-means method

Class No.	GERD/GDP (%)	Countries (codes)*
1.	high	DK, DE, AT, FI, SE, BE, FR, SI
2.	low	BG, EL, HR, CY, LV, LT, MT, PL, RO, SK
3.	middle	CZ, EE, IE, ES, IT, LU, HU, NL, PT, UK

* codes for EU countries: Austria (AT), Belgium (BE), Bulgaria (BG), Cyprus (CY), Czech Republic (CZ), Germany (DE), Denmark (DK), Estonia (EE), Greece (EL), Spain (ES), Finland (FI), France (FR), Croatia (HR), Hungary (HU), Ireland (IE), Italy (IT), Lithuania (LT), Luxembourg (LU), Latvia (LV), Malta (MT), Netherlands (NL), Poland (PL), Portugal (PT), Romania (RO), Sweden (SE), Slovenia (SI), Slovakia (SK), United Kingdom (UK)

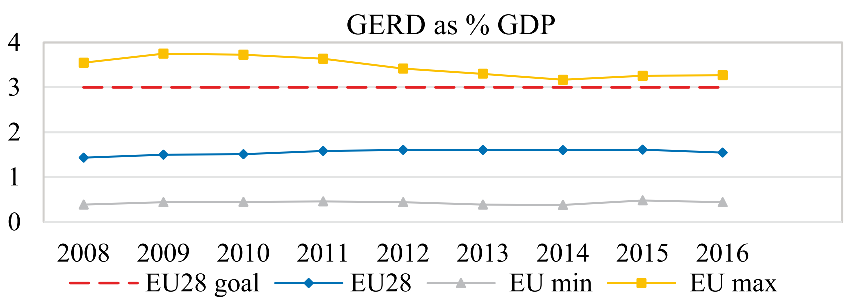


Fig. 2. Mean values of the share of R&D expenditure in GDP in the classes of the European Union countries identified using *k*-means method

In Poland, from 0.60% of GDP in 2008 up to 0.96% of GDP in 2016 was allocated to R&D in the analysed period. Therefore, Poland was assigned to the class of the EU countries characterised by low intensity of expenditure on research and development in the gross domestic product. In addition, as the national target of Europe 2020 strategy Poland declared reaching 1.7% level of R&D expenditure, which still remains a difficult challenge.

4. Classification of the European Union countries in terms of major funding sources of R&D activities

Tables 3 and 4 present the classification results of the European Union countries by the structure of expenditure on research and development by the funding sources in 2008 and in 2016, respectively. Fig. 3 shows dendrograms prepared using Ward's hierarchical clustering method and mean shares of R&D expenditure by the funding sources in both analysed periods, respectively. Based on the analysis of dendrograms and graphs of fusion distance in the fusion stages, the number of classes of the EU countries in 2008 and 2016 was determined as the basis for applying *k*-means method. Finally, the division of the European Union countries into 3 relatively homogeneous classes in 2008 and 4 classes in 2016 was performed.

Table 3. Classification of the European Union countries in terms of R&D expenditure structure by funding sources in 2008 prepared using *k*-means method

Class No.	GERD by source of funds	Countries	Mean share of R&D expenditure by funding sources (%)				
			BES	GOV	HES	PNP	Rest of the world
1.	Dominant share of BES	BE, DK, DE, LU, MT, SI, FI, SE	63.80	26.19	0.50	0.96	8.54
2.	Dominant share of GOV	BG, EL, CY, LV, LT, PL, RO, SK	27.80	58.95	1.86	0.43	10.94
3.	Balanced share of BES and GOV	CZ, EE, IE, ES, FR, HR, IT, HU, NL, AT, PT, UK	45.76	41.54	1.33	1.33	10.05

The conducted empirical research shows that in 2008 BES and GOV sectors were definitely most involved in R&D activity funding in the EU countries. Their average total shares in R&D expenditure funding amounted to 89.99% in the first, 86.75% in the second and 87.30% in the third class of countries, respectively. The involvement of the rest of the world sector was also similar in all identified classes and ranged from over 8.50% in the first class to almost 11% in the second class. Another joint feature of the identified classes of countries was the marginal funding of R&D expenditure by HE and PNP sectors.

The main difference between the classes of countries identified in 2008 consists in a different degree of involvement of both business enterprise sector and government sector in R&D funding (see Fig. 3). The first class included 8 countries characterised by a dominant share of business enterprise sector in R&D expenditure funding (on average 63.80%), including only two countries from the EU10 group: Malta and Slovenia. Class 2 included countries where the government sector was the dominant one in R&D funding (58.95%). In this also 8-element class all countries except Greece belong to the EU10 group, the so-called countries of the new EU enlargement. These countries are characterised by low GDP per capita, among which there are Bulgaria, Romania and Poland. The characteristic feature of the countries included in class 3 is a relatively balanced share of both business enterprise and government sectors in R&D expenditure funding (45.76% and 41.54% respectively). It is the largest 12-element class in which EU15 countries predominate, but it also includes the Czech Republic, Estonia, Croatia and Hungary.

Table 4. Classification of the European Union countries in terms of R&D expenditure structure by funding sources in 2016 prepared using *k*-means method

Class No.	GERD by source of funds	Countries	Mean share of R&D expenditure by funding sources (%)				
			BES	GOV	HES	PNP	Rest of the world
1.	Dominant share of BES	BE, DK, DE, FR, HU, MT, NL, AT, SI, FI, SE, UK	57.51	28.59	0.63	1.68	11.55
2.	Dominant share of BES with high involvement of Rest of the world sector	BG, IE,	46.00	23.30	0.80	0.35	29.55
3.	Dominant share of GOV with high involvement of BES and Rest of the world sector	CZ, CY, LV, LT,	33.75	40.90	2.68	0.25	22.40
4.	Balanced share of BES and GOV	EE, EL, ES, HR, IT, LU, PL, PT, RO, SK	46.92	40.42	2.33	0.56	9.80

In 2016, 4 classes of countries featuring a different structure of R&D expenditure funding were identified. Similarly to the situation in 2008, all classes were characterised by a clear dominance of R&D expenditure funding by business enterprise and government sectors and the lowest share of higher education and private non-profit sectors. However, in 2016 the classes of countries differed significantly between one other also regarding the involvement degree of the rest of the world sector. In 2016, as in 2008, the class of countries with a dominant share of the business enterprise sector (class 1) and a balanced share of the business enterprise and government sector in R&D expenditure funding (class 4) was distinguished. Class 1 included countries assigned to this class type in 2008 (excluding Luxembourg) and was extended by the following countries: France, Hungary, the Netherlands, Austria and the United Kingdom, which previously were characterised by a balanced involvement degree of BES and GOV sectors. Class 4 was enlarged by Poland, Romania and Slovakia, which in 2016 belonged to the group of countries with the dominant R&D funding from the government sources. In 2016 the class of countries characterised by the dominant share of BES (46%) and the large involvement of the rest of the world sector (29.55%) was identified and included only Bulgaria and Ireland (class 2). Class 3 was made up of the countries where the dominant source of funding was the GOV sector (40.9%), along with strong involvement of BES (33.75%) and the rest of the world sector (22.4%). In 2016 the Czech Republic, Cyprus, Latvia and Lithuania were characterised by such funding structure of R&D expenditure and, except for the Czech Republic, these countries also featured dominant GERD funding from the GOV sector in 2008, whereas the rest of the world was of minor importance (10.94%).

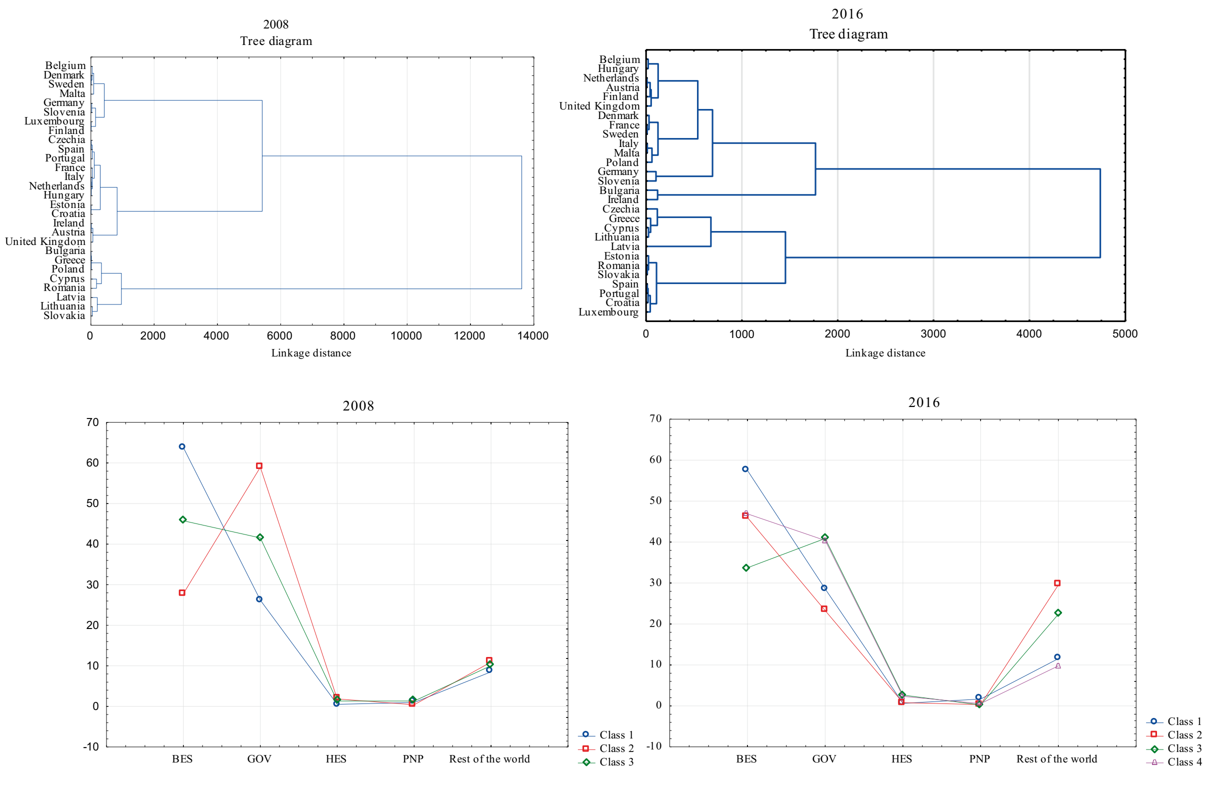


Fig. 3. Dendrogram prepared using Ward's method and mean shares of R&D expenditure by funding sources in the classes identified using *k*-means method in 2008 and 2016

Conclusions

The following conclusions can be presented on the basis of the conducted research:

1. The EU countries showed an extensive diversification throughout the entire analysed period in terms of R&D expenditure intensity, exceeding 50% variation. The tendency towards its reduction should be assessed positively (from approx. 62% to almost 56%). The discussed changes resulted from a slight decline in the intensity of R&D expenditure in the countries featuring the highest values (Finland, Sweden) and an increased share of R&D expenditure in GDP in the countries ranked as last in this respect (Cyprus, Romania, Latvia). In 2008–2016 the number of countries characterised by a lower than average intensity of R&D expenditure prevailed.
2. In the analysed period 3 separate groups of countries characterised by a different intensity of R&D expenditure were distinguished. The least numerous group included countries with a high share of expenditure on R&D in GDP. In the group of 10 countries featuring low intensity of R&D expenditure only Greece was included from the EU15 group, the remaining countries were the so-called new enlargement ones, including Poland.
3. The classification of the EU countries changed significantly in terms of R&D expenditure structure by the funding sources. In 2008 3 classes of countries with the dominant role of BES and GOV sectors and also their balanced share were distinguished, along with the similar significance of the other sectors. In 2016 two specific groups of countries were identified, which along with the dominance of BES sector (Bulgaria, Ireland) or GOV sector (Czech Republic, Cyprus, Latvia, Lithuania) were characterised by a large, amounting to 29.55% and 22.40% share of the rest of the world sector in funding R&D expenditure. More developed EU countries were characterised by higher intensity of R&D expenditure and greater involvement in R&D funding by the business enterprise sector.

Having recognised the importance of expenditure on R&D activities for the innovation and competitiveness of the EU countries, it is necessary to perform the ongoing monitoring of the occurring changes. In the long-term perspective it is worth carrying out in-depth research on the occurrence of club convergence, the crowding-out effect and the complementarity of private and public expenditures.

Acknowledgements

The paper is financed by the National Science Centre: 2015/17/B/HS4/01021, Grant No. 8-11299S: The development of trajectories of traditional industries in the old industrial regions.

References

Dworak, E., & Grzelak, M.M. (2010). Nakłady na działalność badawczo-rozwojową a PKB w krajach Unii Europejskiej. *Gospodarka Narodowa*, 7(21), 107–120.

Hall, B.H., & Lerner, J. (2010). The financing of R&D and innovation. *Handbook of the Economics of Innovation* (1), 609–639.

Hartigan, J.A. (1975). *Clustering Algorithms.* New York, NY: John Wiley & Sons. Inc.

Howell, S.T. (2017). Financing innovation: evidence from R&D grants. *American Economic Review*, 107(4), 1136–64.

Kukuła, K. (2000). *Metoda unitaryzacji zerowanej*. Warszawa: Wydawnictwo Naukowe PWN.

Mand, M. (2019). On the cyclicality of R&D activities. *Journal of Macroeconomics*, 59, 38-???

Manual, F. (2015). The Measurement of Scientific, Technological and Innovation Activities: Guidelines for Collecting and Reporting Data on Research and Experimental Development. *Organization for Economic Co-operation and Development (OECD). Paris: OECD Publishing*.

McGrath, M.E., & Romeri, M.N. (1994). The R&D effectiveness index: a metric for product development performance. *Journal of Product Innovation Management: an International Publication of The Product Development & Management Association*, 11(3), 213–220.

Ostraszewska, Z., & Tylec, A. (2016). Nakłady wewnętrzne na działalność badawczo-rozwojową w Polsce i źródła jej finansowania w sektorze przedsiębiorstw. *Zeszyty Naukowe Politechniki Częstochowskiej. Zarządzanie*, (24), 30–42.

Ouyang, M. (2011). On the Cyclicality of R&D. *Review of Economics and Statistics*, 93(2), 542–553.

Sawulski, J. (2018). *Efektywność wydatków na badania i rozwój w Polsce na tle innych państw Unii Europejskiej*. Warszawa: Difin.

Statistical evaluation of research and development activity of the EU countries with regard to the accuracy of statistical data

Małgorzata Stec[1]

Abstract

The aim of this paper is to evaluation of research and development activity of the EU countries, including the accuracy of statistical data. 8 diagnostic variables describing this economic phenomenon in years 2010 and 2017 were used for the empirical study. Min-max normalisation was employed to perform a linear ordering of objects. Because a quality of obtained results depends, among others, on a quality of data used for calculations, the study also contains an evaluation ofinfluence of the accuracy of statistical data on a result of the linear ordering of the EU countries with regard to the level of R&D. Due to the fact that the problem of influence of the accuracy of diagnostic data on the results of the taxonomic analyses does not have any thorough methodology in the economic research yet, an original approach for the analyses of the subject was proposed. For this purpose, the uncertain theory of measurements, which is used in technical sciences, was employed and adjusted to the specificity of methods of multidimensional comparative analyses. In view of measuring scales used in taxonomy (among others, permissibility of mathematical operations on these scales), the Monte Carlo method was employed in order to determine the uncertainty ranges of a synthetic measure.

Keywords: *R&D, European Union countries, synthetic measure, Monte Carlo method*

JEL Classification: *O32, O52, C15*

1. Introduction

The characteristic trait of the contemporary highly developed countries is an economy based on knowledge and new technologies (Knowledge-based Economy). Intellectual potential, knowledge and new technologies are the factors that are decisive forcountries and regions to have chances for development and competitiveness. Research and development activity have a crucial role in ability to create knowledge and to remold it into new technologies, products and services (Bilbao-Osorio and Rodríguez-Pose, 2004; Bravo-Ortega and Marin, 2011; Cetenak and Oransay 2017; Coccia, 2012; Cunningham and Link, 2016; Falk, 2007; Grzebyk and Stec, 2015; Hall et al., 2010; Rodríguez-Pose, 2001).

Therefore, construction of ranking of the EU countries in terms of the level of progression of research and development activity, is an interesting research problem. While undertaking a statistical evaluation of the EU countries in matters of the researched phenomenon and other complex phenomena[2] by means of widely available statistical data, one should pay attention to a problem of their "accuracy". Statistical information used in the research will decide about the final results. The way of obtaining statistical data by the institutions collecting such data (e.g.

[1] Corresponding author: University of Rzeszów, Faculty of Economics, Department of Quantitative Methods and Economic Informatics, 2 M. Ćwiklińskiej St., 35-601 Rzeszów, Poland, e-mail: malgorzata.a.stec@gmail.com.

[2] Complex economic phenomena are the phenomena not subjected to the direct measurement of, e.g. the socio-economic development of countries, regions, etc., standard of living, job market situation, financial situation of local government unit, companies, banks, etc. (Pawełek, 2008).

Eurostat and statistical offices of each country) results inproducing errorsthat are impossible to omit. Thus, one should be aware of their existence and, whenever possible, should include their impact on the results of undertaken research and drawn conclusions. The aim of this paper is to evaluate research and development activity of the EU countries, including the accuracy of statistical data. 8 diagnostic variables describing this economic phenomenon in years 2010 and 2017 were used for the empirical study. The linear ordering of EU countries was done by the min-max normalisation. Moreover, an influence of uncertainty of measurement of diagnostic variables on the values of synthetic measure was investigated. Monte Carlo method was adopted for this purpose.

2. Theoretical basis for statistical data accuracy

The quality concept of official statistics is based on the definition of the European Statistical System definition of the quality and defined on the basis of the following 6 criteria: relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability, coherence. Accuracy denotes the closeness of computations or estimates (after collecting, processing, imputation, estimation of data and the like) to the exactor true values. The difference between these two values is the error (Vademecum of quality in official statistics, 2012).

The way of obtaining statistical data has an impact on their accuracy. The most accurate data are gathered in official national registers, while less accurate data are collected by means of sample surveys. In yearbooks, there is no information about estimated values of uncertainty that burden a given statistical variable. Therefore, the author made such estimations on the basis of the available knowledge about the way of obtaining data used to calculate a precise variable.

In case of statistical data, the real value of measured variable is usually unknown. Then, the accuracy of statistical data (similarly as in technical sciences) can be equated with the measurement uncertainty (JCGM/WG 1, 2008).

The formal definition of the term ‘uncertainty of measurement’ is as follows: uncertainty (of measurement) parameter, associated with the result of a measurement, that characterises the dispersion of the values that could reasonably be attributed to the measurand. The parameter a uncertainty of measurement may be, for example, a standard deviation called standard measurement uncertainty (or a specified multiple of it), or the half-width of an interval, having a stated coverage probability (Balazs, 2008).

A variable read from a yearbook or downloaded from a data base is taken as a nominal value of a variable X_n. Estimated value of uncertainty u_{Bc} determines the upper and lower limit of a nominal value in which, with an established probability, a real value of this variable can be found:

$$X_R = X_n \pm u_{Bc} \tag{1}$$

where: X_R – limits of the range in which the real value of a variable is; X_n – variable’s nominal value; u_{Bc} – estimated value of variable’s uncertainty.

Assuming that the respective variables are burdened by uncertainty (real values are unknown, only expected values – estimates – and distribution are known), one should analyse if the ranges are not too wide to "blur" the difference between researched objects. In the Fig. 1 a graphical representation of a nature of comparison of assigned values to variables is presented for two objects (countries) whose ranges of uncertainty "overlap". Such case will occur if for a variable-booster, an object's upper limit of uncertainty occupying a lower position in a ranking has a higher value, than the lower limit of uncertainty of an object positioned higher.

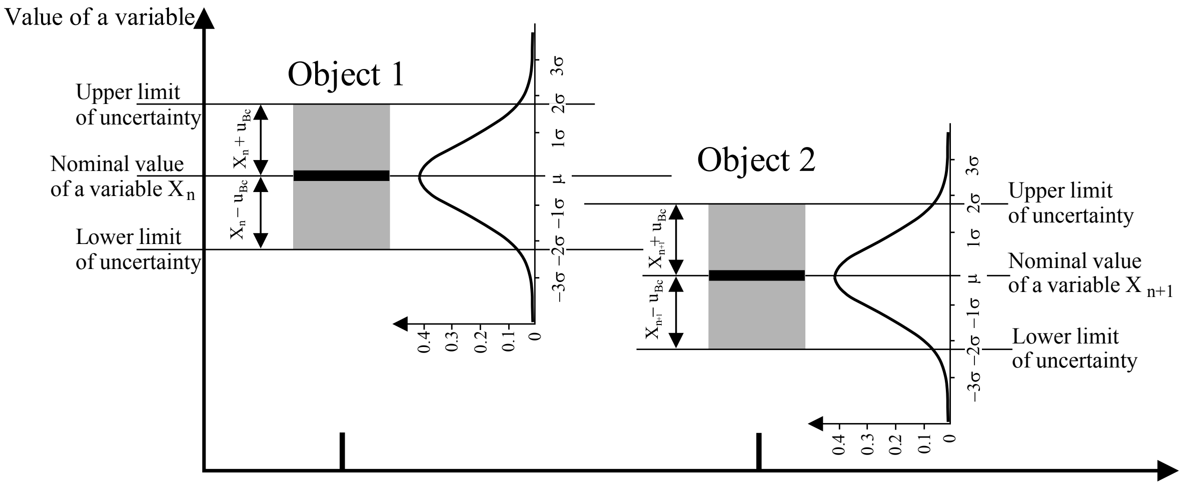


Fig. 1. A nature of estimation of uncertainty of statistical value

The overlapping of ranges of uncertainty can take place if assigned valuesof variables of two or more objects (countries) differ little from each other, while the estimated uncertainties are relatively big. With small differentiations of assigned values of variables of objects, there may be a situation in which a few objects are characterised by similar values of one variable, which may hinder an interpretation of real differences between these objects. Analogical situation also refers to synthetic measures.

The range of uncertainty for determining a diagnostic variable can be described by a density function of normal distribution described by a dependence 2.

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-m)^2}{2\delta^2}} \tag{2}$$

Area under a curve of density function *f*(*X*) is a measure of probability of finding an assigned value to variable in a determined range. It is possible to calculate a probability of event, in which, as a result of change of assigned values of variables, a position of researched objects (countries) changes. Fig. 2 presents such case for a diagnostic variable *X*1 (Research and development expenditure (in % of GDP)), in relation to Hungary and Portugal. If a real value of variable *X*1 for Hungary (*X*1=1.35) was nevertheless lower than a range value X_g, whereas for Portugal (*X*1=1.32) higher than X_g; therefore, these countries would exchange their positions in the ranking. Due to the fact that both events are independent, the total probability would be a product of probabilities for respective ranges determined by the density function of normal distribution. Change of positions of objects is also possible in other cases which were not described in this article.

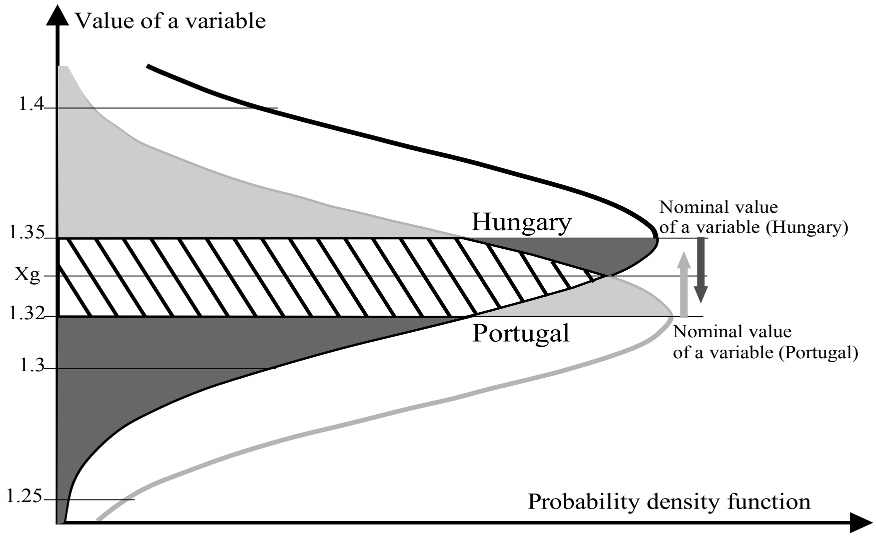


Fig. 2. Density function of normal distribution for Hungary and Portugal for variable *X*1

Analogical situation that takes place for diagnostic variables can also occur in case of an analysis of uncertainty range for a calculated synthetic measure.

3. Methods applied

In this paper, the min-max normalisation was used in order to calculate a value of synthetic measure for all countries in terms of research and development activity in year 2010 and 2017 (Kukuła, 2000).

A normalisation of the variable values was conducted using the following formulas:

$$z_{ij} = \frac{x_{ij} - \min_i\{x_{ij}\}}{R_j} \quad \text{for stimulating factors} \tag{3}$$

$$z_{ij} = \frac{\max_i\{x_{ij}\} - x_{ij}}{R_j} \quad \text{for non-stimulating factors} \tag{4}$$

where: z_{ij} – normalized value of *j*-th variable for the *i*-th object, x_{ij}, – value of *j*-th variable for the *i*-th object, R_j – range for the *j*-th variable.

It should be emphasised that the calculations were done in a dynamic manner, using the so called 'object-periods'. The synthetic measure was calculated as an arithmetic mean of the normalised value of variables:

$$MS_i = \sum_{j=1}^{m} z_{ij} \tag{5}$$

where: MS_i – synthetic measure in *i*-th object, *m* – number of variables.

Due to the fact that the employed method to calculate the synthetic measure leads to change of the measuring scale, the calculation of the uncertainty of the synthetic measure by analytic method would provide false results. That is why, the Monte Carlo[3] method was employed and calculations were performed in the *R* application (Walesiak and Gatnar, 2009).

In order to calculate the value of uncertainty of the synthetic measure, it was concluded that for a sample big enough (the calculations were performed on a set of data counting 1000 for each object), standard deviation can be considered as a measure of distribution identified with a range of uncertainty of the synthetic measure. The following algorithm of procedure was chosen (Stec, 2017; Stec and Wosiek, 2018):

- for each diagnostic variable, 1000 values were drawn for every object (28 countries) which fulfilled the following conditions:
  - value of each drawn variable was comprised in the assumed range of uncertainty created around the nominal value for this variable,
  - drawn values for each variable had normal distribution,
- from the drawn variables, sets of data were created (1000 sets for each object),
- drawn sets of data underwent normalisation,
- on the basis of the normalised set of data, the synthetic measures were calculated (1000 values of partial measures),
- from 1000 set synthetic, partial measures standard deviation was calculated, which constituted the measure of uncertainty of synthetic measure.

The above mentioned procedure allowed for calculation of nominal values of synthetic measures for each object (country) and their uncertainty.

4. Diagnostic variables employed in the research

The evaluation of research and development activity for 28 EU countries was done with an employment of 8 diagnostic variables: *X*1-Research and development expenditure (% of GDP) (S); *X*2-Intramural R&D expenditure (GERD) by source of funds (Business enterprise sector-% of total GERD) (S); *X*3-Share of government budget appropriations or outlays on research and development (% of total) (S); *X*4-Research and development personnel (Full time equivalent-% of the labour force) (S); *X*5-High-tech exports % of exports (S); *X*6-Employment in high- and medium-high technology manufacturing sectors and knowledge-intensive service sectors (% of total employment) (S); *X*7-Human resources in science and technology (HRST) (% of active population) (S); *X*8-Patent applications to the European patent office (EPO) by priority year per 100 thous. population (S)[4].

[3] The Monte Carlo method solves a numerical problem by performing calculations on random variables, it is a tool for solving quantity problems, when analytical methods based on formulas, estimators, etc., fail. (Kopczewska et al., 2016; Liu, 2008; Niemiro, 2013).

[4] (S)-stimulant.

Empirical data describing the states under study were extracted from the Eurostat database[5]. Also, an influence of the uncertainty of diagnostic variables on the results, regarding ordering of objects in terms of values of proposed variables was assessed. Table 1 shows a compilation of investigated values of uncertainty and a number of cases of overlapping (collisions) ranges of uncertainty for each diagnostic variables caused by too small difference between their nominal values in relation to the calculated uncertainty. The uncertainty of data was estimated on the basis of research sample (for Poland) and the uncertainty resulting from rounding. At the same time, a simplifying assumption was taken in order to acknowledge that for the purposes of the article, it is adequate to apply the same uncertainties for all countries for the analysed years.

Table 1. Comparison of estimated uncertainty values of diagnostic variables and overlapping ranges of uncertainty of diagnostic variables (2017)

	Diagnostic variables							
	***X*1**	***X*2**	***X*3**	***X*4**	***X*5**	***X*6**	***X*7**	***X*8**
Overall uncertainty of a variable	2.5%	1.0%	0.8%	1.0%	0.8%	0.7%	0.5%	0.5%
Overall number of conflicts	21	21	11	12	9	13	19	2
Number of conflicts with probability > 0,05	11	12	7	10	5	10	9	1

5. Empirical results

The statistical evaluation of research and development activity of the EU countries in years 2010 and 2017 was done in a traditional way by analysing value of a synthetic measure and by taking into account uncertainties of a given synthetic measure.

Considering only values of synthetic measure and rankings of the EU countries based on them in terms of research and development activity, it can be noticed that:

- in 2010, as regards the R & D activity, the leading positions in the rank of EU countries were taken by: Finland, Germany, Sweden, Denmark and Luxembourg. Last positions were taken by: Bulgaria, Romania, Latvia, Greece and Poland.
- in 2017, the leading positions in the rank of EU countries were taken by: Germany, Denmark, Sweden Austria and Finland. Last positions were taken by: Latvia, Romania, Bulgaria, Cyprus and Greece.

In 2017 in comparison to 2010, the highest advancement in the ranks of the EU countries in terms of R+D was noted for: Poland (from 24th position in 2010 to 19th in 2017), Austria and Ireland (advancement of 3 places). The same position in both years was held by: Estonia,

[5] https://ec.europa.eu/eurostat/data/database (accessed on 07.01.2019).

Hungary, Romania, Slovenia and Sweden. Countries that lost their position were: Luxembourg (from 5th in 2010 to 13th in 2017) and Finland (from 1st to 5th position).

In the second variant of the study, the aim was to verify if the inclusion of uncertainty as to the value of diagnostic variables used in the research, and inclusion of uncertainty for an aggregate measure can have an influence on a change of a position held by the individual EU countries in terms of research and development activity.

Fig. 3 presents results of ordering of the EU countries in regard to the level of research and development activity in 2010; whereas fig. 4 presents the same ordering for 2017.

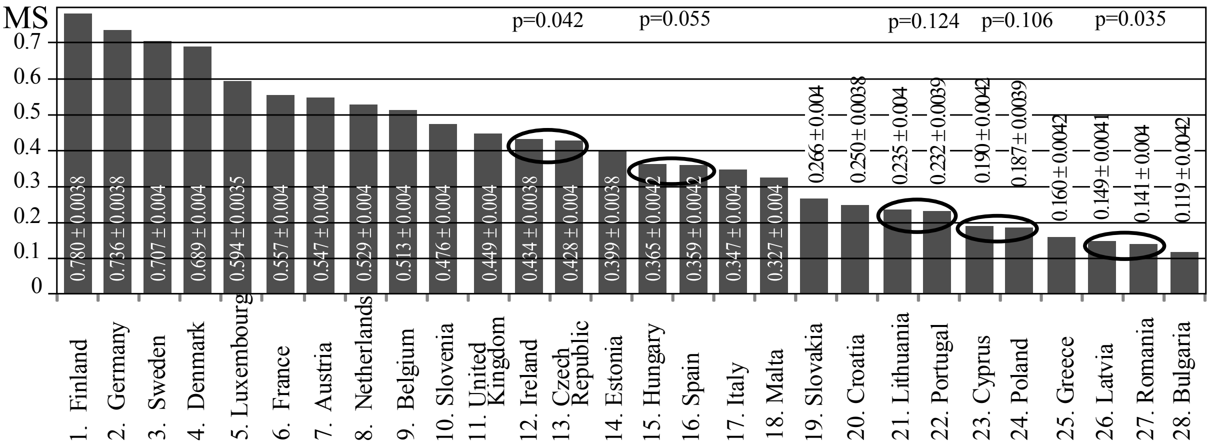


Fig. 3. Results of linear ordering of the EU countries in terms of the level of research and development activity (2010)

The cases in which inclusion of uncertainty resulted in overlapping of uncertainty ranges (collisions) between countries were emphasised. This situation took place in 2010 in: Ireland-Czech Republic, Hungary-Spain, Lithuania-Portugal, Cyprus-Poland, Latvia-Romania. The analysis of probability of position change caused by data errors shows that only in three cases the probability exceeded 0,05.

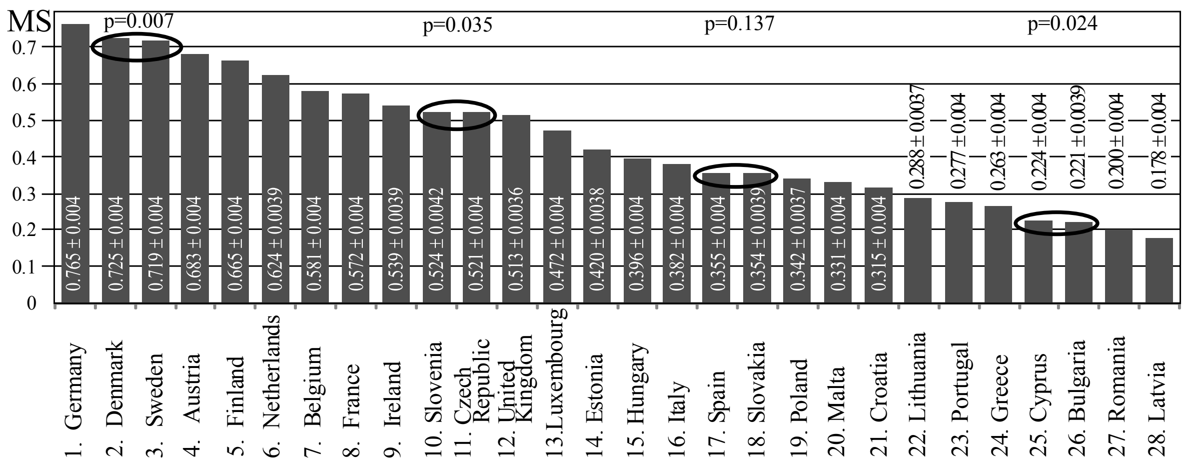


Fig. 4. Results of linear ordering of the EU countries in terms of the level of research and development activity (2017)

For 2017, four collisions were noticed for: Denmark-Sweden, Slovenia-Czech Republic, Spain-Slovakia, Cyprus-Bulgaria. Only in one case, the probability of position change caused by data errors exceeded 0,05.

6. Conclusions

The following conclusions can be drawn based on the carried out research:

- The statistical evaluation of research and development activity of the EU countries was done with an employment of 8 diagnostic variables. The min-max normalisation was employed for the empirical studies.
- The results confirm the diversity of EU countries in terms of research and development activity. In 2017 Germany, Denmark, Sweden Austria and Finland were the leaders interms of research and development activity. The lowest level of the analyzed phenomenon is represented by the following countries: Latvia, Romania, Bulgaria, Cyprus and Greece.
- The conducted research shows that the method of obtaining statistical data influences the value of uncertainty estimation.The method of analysis of diagnostic variables and a synthetic measure proposed in the article, including uncertainty ranges, allows to determine the trust range for the obtained results.The analysis does not change the order of ordering, nor does it verify the actual ordering of objects.Taking into account the uncertainties in the value of synthetic measures may influence the final conclusions resulting from the research.

References

Balazs, A. (2008). International vocabulary of metrology-basic and general concepts and associated terms. *Chemistry International*, 20–1.

Bilbao-Osorio, B., & Rodríguez-Pose, A. (2004). From R&D to innovation and economic growth in the EU. *Growth and Change*, 35(4), 434–455.

Bravo-Ortega, C., & Marin, A.G. (2011). R&D and productivity: A two way avenue?. *World Development*, 39(7), 1090–1107.

Cetenak, O.O., & Oransay, G. (2017). Economic Growth and Dynamic R&D Investment Behavior. In: *Global Business Strategies in Crisis*, 243–259. Cham: Springer.

Coccia, M. (2012). Political economy of R&D to support the modern competitiveness of nations and determinants of economic optimization and inertia. *Technovation*, 32(6), 370–379.

Cunningham, J.A., & Link, A. N. (2016). Exploring the effectiveness of research and innovation policies among European Union countries. *International Entrepreneurship and Management Journal*, 12(2), 415–425.

Eurostat database:https://ec.europa.eu/eurostat/data/database (accessed on 07.01.2019).

Falk, M. (2007). R&D spending in the high-tech sector and economic growth. *Research in Economics*, 61(3), 140–147.

Grzebyk, M., & Stec, M. (2015). Sustainable development in EU countries: concept and rating of levels of development. *Sustainable Development*, 23(2), 110–123.

Hall, B.H., Mairesse, J., & Mohnen, P. (2010). *Measuring the Returns to R&D*. In: *Handbook of the Economics of Innovation* (Vol. 2, pp. 1033–1082). North-Holland.

JCGM/WG 1 2008 Working Group. (2008). Evaluation of measurement data–guide to the expression of uncertainty in measurement. In: *Tech Rep JCGM 100: 2008 (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML*.

Kopczewska, K., Kopczewski, T., & Wójcik, P., (2016). *Metody ilościowe w R: aplikacje ekonomiczne i finansowe*. Warszawa: CeDeWu, 221.

Kukuła, K. (2000). *Metoda unitaryzacji zerowanej*. Warszawa: Wydawnictwo Naukowe PWN.

Liu, J. S. (2008). *Monte Carlo strategies in scientific computing*. Springer Science & Business Media.

Niemiro, W. (2013). *Symulacje stochastyczne i metody Monte Carlo*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.

Pawełek, B. (2008). Metody normalizacji zmiennych w badaniach porównawczych złożonych zjawisk ekonomicznych. *Zeszyty Naukowe/Uniwersytet Ekonomiczny w Krakowie. Seria Specjalna, Monografie*, (187).

Rodríguez-Pose, A. (2001). Is R&D investment in lagging areas of Europe worthwhile? Theory and empirical evidence. *Papers in regional science*, 80(3), 275–295.

Stec, M. (2017). *Taksonomiczna analiza poziomu rozwoju społeczno-gospodarczego województw Polski*. *Studium przypadku-województwo podkarpackie*, Rzeszów: Wydawnictwo Uniwersytetu Rzeszowskiego, 87–99.

Stec, M. & Wosiek, M. (2018). Evaluation of the socio-economic situation of European Union countries, taking into account accuracy of statistical data. In: Papież M. and Śmiech S. (Eds.), *The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomen. ConferenceProceedings*. Cracow: Foundation of the Cracow University of Economics, 483–492.

Vademecum of quality in official statistics. (2012). Warsaw:http://bip.stat.gov.pl/en/activity-of-official-statistics/quality-in-statistics/ (accessed on 04.01.2019).

Walesiak, M., & Gatnar, E. (Eds.). (2009). *Statystyczna analiza danych z wykorzystaniem programu R*. Warszawa: Wydawnictwo Naukowe PWN.

A comparative analysis of rankings of Polish provinces in terms of social cohesion for metric and interval-valued data

Marek Walesiak[1], Grażyna Dehnel[2]

Abstract

The article describes a comparative analysis of rankings of Polish provinces in terms of social cohesion based on metric and interval-valued data between 1st and 3rd quartiles (50% of observations), 1st and 9th deciles (80% of observations) and the minimum and maximum (100% observations). The rankings were obtained using a hybrid approach combining the use of multidimensional scaling (MDS) with linear ordering. Interval-valued variables characterise the objects of interests more accurately than metric data. Metric data are of an atomic nature, i.e. an observation of each variable is expressed as a single real number. In contrast, an observation of each interval-valued variable is expressed as an interval. Interval-valued data were derived by aggregating metric data on social cohesion at the level of districts to the province level. All observations were included in the aggregation and then outliers were omitted.

Keywords: *social cohesion, interval-valued data, multidimensional scaling, composite indicators, outliers*

JEL Classification: *C38, C43, C63*

1. Introduction and motivation

Social cohesion is a multi-faceted phenomenon. When one analyses the conceptualisation of social cohesion, one can note a clear direction of changes, reflecting the growing importance attached to the socio-cultural and political indicators, accompanied by the declining role of the economic dimension (Chan and Chan, 2006). Dickes et al. (2010) and Dickes and Valentova (2013) indicate four dimensions of social cohesion: institutional trust, solidarity, socio-cultural participation and political participation.

The article presents a comparative analysis of the rankings of Polish provinces in terms of social cohesion for metric data and three types of interval-valued data. The rankings were obtained by using a hybrid approach, which combines MDS and linear ordering. Two criteria are proposed as the basis for comparing the rankings. The first one involves cluster analysis, which is used to identify similarities and differences in the ordering of provinces in terms of social cohesion. The second one is based on the analysis of the degree to which different rankings of objects with respect to specific variables correspond to those obtained by using the aggregate measure for 4 datasets (one containing metric data and three with interval-valued data). These two approaches were then used to select a ranking that best represents the level of social cohesion in the provinces of Poland.

[1] Corresponding author: Wrocław University of Economics, Department of Econometrics and Computer Science, 3 Nowowiejska St., 58-500 Jelenia Góra, Poland, marek.walesiak@ue.wroc.pl. ORCiD 0000-0003-0922-2323.

[2] Poznan University of Economics and Business, Department of Statistics, 10 Niepodległości, 61-875 Poznan, Poland, grazyna.dehnel@ue.poznan.pl. ORCiD 0000-0002-0072-9681.

2. An overview of the social cohesion concept

In the EU practice, the level of social cohesion is measured using, among other indicators, the EU Regional Social Progress Index (EU-SPI). The index comprises three dimensions of social progress (Annoni and Dijkstra, 2016, p. 2): basic human needs (nutrition and basic medical care, water and sanitation, shelter (housing), personal safety); foundations of well-being (access to basic knowledge, access to information and communication, health and wellness, environmental quality); opportunity (personal rights, personal freedom and choice, tolerance and inclusion, access to advanced education).

In the article, social cohesion of Polish provinces was analysed on the basis of secondary data. For this reason, the use of the SPI index, which is based on three dimensions (basic human needs, foundations of well-being, opportunity), is considered justified. The final set of variables, used in the study, was selected by the authors of the article (Dehnel et al., 2018). Given this 3-dimensional frame of reference, social cohesion of the Polish provinces was measured using 26 metric variables:

1. **Basic human needs** (7 variables): mean monthly wage (in PLN) – a stimulant, total unemployment rate in % – a destimulant, mean useful floor area of a dwelling per inhabitant in m^2 – a stimulant, average number of persons per room – a destimulant, length of the sewerage network in relation to the length of the water supply network in % – a stimulant, number of doctors and dentists per 10,000 population – a stimulant, crimes reported (criminal offenses, against life and health, against property) per 10,000 population – a destimulant.
2. **Foundations of well-being** (11 variables): people using water treatment services (% of total population) – a stimulant, percentage of all dwellings equipped with central heating – a stimulant, children enrolled in day-care centres per 1000 children up to the age of 3 – a stimulant, children enrolled in nursery schools per 1000 children aged 3–5 – a stimulant, pupils taking obligatory classes of English in primary and intermediate schools (% of all pupils) – a stimulant, number of pupils in secondary schools per class – a destimulant, members of sports club per 1000 population – a stimulant, users of public libraries per 1000 population – stimulant, people participating in cultural events (organised by cultural centres and clubs) per 1000 population – a stimulant, area of public greenspace (parks, residential greenspace) per 10,000 population (in ha) – a stimulant, length of municipal and district improved hard surface roads per 10,000 population (in km) – a stimulant.
3. **Opportunities** (8 variables): persons in households (below the income threshold) using social assistance per 1000 population – a destimulant, age dependency ratio (number of people aged 0–14 and those aged 65 and older per 100 people of working age) – a destimulant, share of women in the labour force in % – a nominant (with the nominal value of 50%), share of youth (up to the age of 25) in the population of registered unemployed in % – a destimulant, share of long-term unemployed (over 12 months) in the population of registered unemployed in % – a destimulant, number of job offers for disabled people per 1000 registered disabled unemployed – a stimulant, places in stationary social welfare facilities per 10,000 population – a stimulant, voter turnout local elections (for municipal authorities and town councils with district rights) in 2014 in % – a stimulant.

The statistical data come from the Local Data Bank maintained by the Central Statistical Office. The reference year is 2016, except for variable “Voter turnout in local elections”, which represents data for 2014 (the last local government elections). The x4 nominant variable was converted into a stimulant. The definitions of stimulant, destimulant and nominant can be found in (Walesiak, 2016, p. 18).

3. Social cohesion of Polish provinces – research methodology

The objects analysed in the study were ranked in terms of social cohesion using a two-step procedure, which makes it possible to visualise results of linear ordering. In the first step the objects of interest undergo MDS, as a result of which they can be visualised in a two-dimensional space. In the second step the objects are linearly ordered to produce a ranking. A description on the procedure can be found in (Walesiak, 2016; Walesiak and Dehnel, 2018).

Datasets used in comparative analysis

A ranking of Polish provinces in terms of social cohesion can be obtained on the basis of metric or interval-valued data. For metric data, an observation for the *j*-th variable for the *i*-th object is expressed as a real number. In the case of interval-valued data, observations for each variable are expressed as intervals $x_{ij} = [x_{ij}^l, x_{ij}^u]$ ($x_{ij}^l \leqslant x_{ij}^u$, x_{ij}^l (x_{ij}^u) denotes the lower bound (the upper bound) of the interval). Studies by (Gioia and Lauro, 2006; Brito et al., 2015) provide different examples of data that in real life are of interval type. In this article we compare two approaches to the assessment of social cohesion in Polish provinces:

1. A classical one-step approach, based on metric data, where the ranking of provinces was created using a matrix consisting of 17 objects (16 provinces plus an average province) described by 26 metric variables.
2. A two-stage approach, based on interval-valued data. Firstly, atomic metric data on social cohesion in Polish districts (LAU units) were collected (380 districts described by 26 variables), which were then aggregated to produce interval-valued data. The lower and upper limit of the interval for each province was determined on the basis of district-level data: the minimum and maximum (100% of observations), 1st and 9th deciles (80% of observations) and 1st (Q_1) and 3rd (Q_3) quartiles (50% of observations). Variable values greater than $\left(Q_3 + 3\frac{Q_3-Q_1}{2}\right)$ and less than $\left(Q_3 - 3\frac{Q_3-Q_1}{2}\right)$ are considered outliers. The decision on the selection of the percentage of the cut-off of outlier observations (quartiles and deciles) was made arbitrarily.

The selection of the optimum multidimensional scaling (MDS) procedure

The problem of selecting an optimum MDS procedure is discussed in (Walesiak and Dudek, 2017). Multidimensional scaling was conducted using the smacofSym function from the smacof R package (Mair et al., 2018). To solve the problem of choosing the optimal MDS procedure two criteria were applied in mdsOpt package (Walesiak and Dudek, 2018b): Kruskal’s *Stress*-1 fit measure and the Hirschman-Herfindahl *HHI* index, calculated based on Stress per point values. For all MDS procedures, for which $Stress\text{-}1_p \leq critical\ stress$, we choose the one for each occurs $\min_p \{HHI_p\}$ (p – MDS procedure number).

For metric data, the optimal MDS procedure was selected after testing 6 normalisation methods (n1, n2, n3, n5, n5a, n12a – see Walesiak and Dudek, 2018a), 5 distance measures (Manhattan, Euclidean, Squared Euclidean, Chebyshev, GDM1[3] – see e.g. Everitt et al., 2011, pp. 49–50), 4 MDS models (ratio, interval, second and third degree polynomial – see Borg and Groenen, 2005; Borg et al., 2018), producing a total of 120 MDS procedures. After applying the optSmacofSym_mMDS function from the mdsOpt package for the R program (R Core Team, 2018) the optimal procedure of MDS was selected, which involves positional standardization (n2), the ratio scaling model and the Manhattan distance.

For interval-valued data, the optimal MDS procedure was selected after testing 6 normalisation methods (n1, n2, n3, n5, n5a, n12a), 4 distance measures (Ichino-Yaguchi, Euclidean Ichino-Yaguchi, Hausdorff, Euclidean Hausdorff – see Billard and Diday, 2006; Ichino and Yaguchi, 1994), 4 MDS models (ratio, interval, second and third degree polynomial), resulting in a total of 96 MDS procedures. After applying the optSmacofSymInterval function from the mdsOpt R package three optimal procedures of MDS were selected for three types of interval-valued data: (a) for an interval between the min and max values: positional standardisation (n2), the ratio scaling model and the Euclidean Ichino-Yaguchi distance, (b) for an interval between 1st and 9th deciles: positional standardisation (n2), mspline 3 scaling model and the Euclidean Ichino-Yaguchi distance, (c) for an interval between 1st and 3rd quartiles: positional standardisation (n2), the ratio scaling model and the Euclidean Hausdorff distance.

After applying the MDS procedures for metric data and three types of interval-valued data, the following results are shown in Figure 1. In each diagram, the anti-pattern (AP) object and the pattern (P) object are connected by a straight line, known as the set axis. Then isoquants of development (curves of equal development) are identified. Objects located between pairs of isoquants represent a similar level of development. Objects located at different points within the same isoquant of development can have the same level of development (as a result of different configurations of variable values).

Ranking of provinces in terms of social cohesion

Based on the results of multidimensional scaling in a two-dimensional space, the provinces can be ranked in terms of social cohesion. Objects are ordered linearly using an aggregated measure (composite indicator) d_i (Hellwig, 1981):

$$d_i = 1 - \sqrt{\sum_{j=1}^{2}(v_{ij} - v_{+j})^2} \Big/ \sqrt{\sum_{j=1}^{2}(v_{+j} - v_{-j})^2}, \tag{1}$$

v_{ij} – j-th coordinate for i-th object in a two-dimensional MDS space, v_{+j} (v_{-j}) – j-th coordinate for the pattern (anti-pattern) object in a two-dimensional MDS space.

[3] See Jajuga et al., 2003.

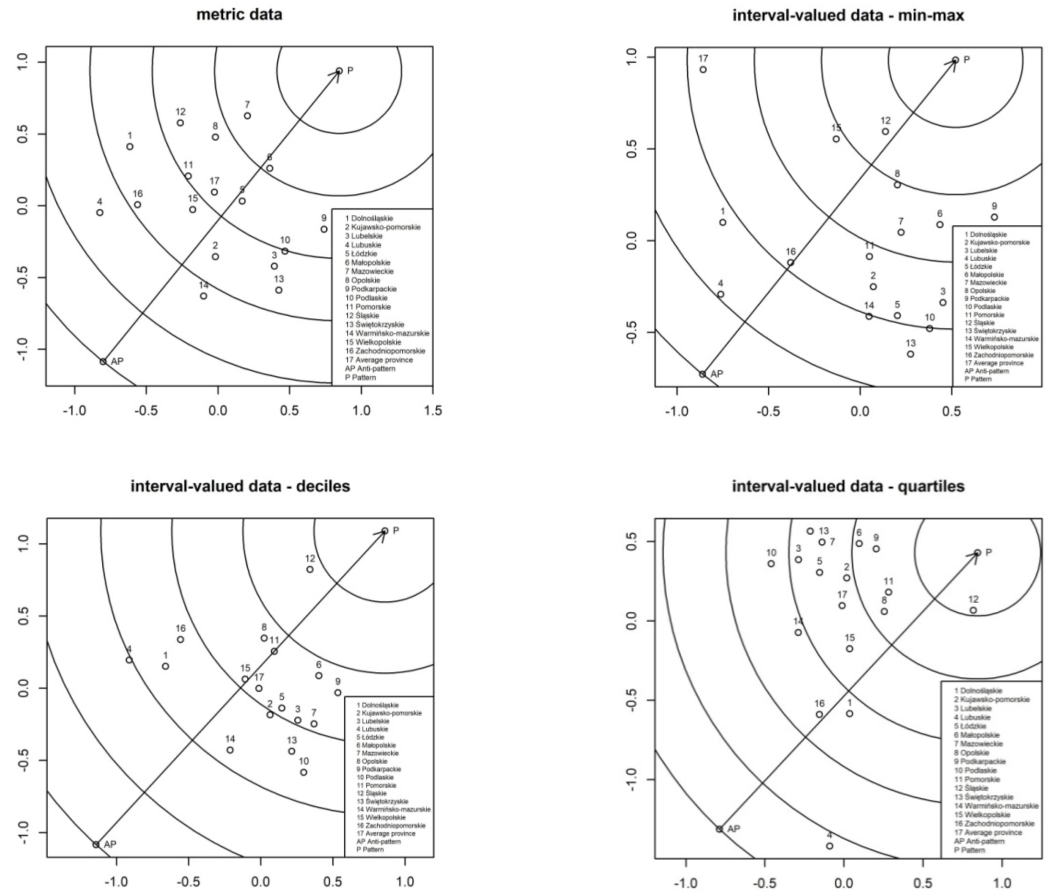


Fig. 1. Results of multidimensional scaling of Polish provinces according to social cohesion

Values of the aggregate measure d_i are included in the interval [0; 1] . The higher the value of d_i, the higher the level of social cohesion of the objects of interest. Table 1 shows the ranking of provinces in terms of social cohesion for 2016.

4. Comparative analysis of the results

The article describes a comparative analysis of rankings (Table 1) of Polish provinces in terms of social cohesion based on metric and interval-valued data between 1st and 3rd quartiles (50% of observations), 1st and 9th deciles (80% of observations) and the minimum and maximum (100% observations).

The comparative analysis was conducted using cluster analysis to identify similarities and differences in the rankings taking the following steps:

1. The four datasets (metric, min-max, deciles, and quartiles) are linearly ordered to produce 4 rankings (see Table 1).
2. The rankings are compared on the basis of Kendall's tau coefficient (Kendall, 1955). For purposes of cluster analysis, values in the rankings are transformed into distances $d = \frac{1}{2}(1 - tau)$ arranged in the form of a distance matrix.

Table 1. Ranking of Polish provinces in terms of social cohesion, based on metric and interval-valued data (3 intervals) for 2016 (values of aggregate measure d_i).

Province	Metric data		min-max		deciles		quartiles	
	d_i	Rank	d_i	Rank	d_i	Rank	d_i	Rank
Dolnośląskie	0.4051	12	0.2961	15	0.3952	15	0.4565	14
Kujawsko-pomorskie	0.4042	13	0.4025	8	0.4916	11	0.6472	6
Lubelskie	0.4507	11	0.3984	9	0.5114	10	0.5252	12
Lubuskie	0.2570	17	0.1777	17	0.3284	17	0.1323	17
Łódzkie	0.5667	5	0.3506	12	0.5198	8	0.5784	9
Małopolskie	0.6806	2	0.5902	5	0.6271	2	0.6856	5
Mazowieckie	0.7269	1	0.5521	6	0.5182	9	0.5869	8
Opolskie	0.6251	3	0.6589	2	0.6214	3	0.7082	4
Podkarpackie	0.5755	4	0.5989	4	0.6052	5	0.7313	3
Podlaskie	0.4975	9	0.3319	13	0.4031	14	0.4524	15
Pomorskie	0.5086	8	0.4685	7	0.6167	4	0.7420	2
Śląskie	0.5534	6	0.7517	1	0.8024	1	0.8473	1
Świętokrzyskie	0.3934	14	0.2629	16	0.4390	13	0.5533	11
Warmińsko-mazurskie	0.2983	16	0.3295	14	0.3713	16	0.4806	13
Wielkopolskie	0.4614	10	0.6452	3	0.5225	7	0.5766	10
Zachodniopomorskie	0.3534	15	0.3522	11	0.4574	12	0.4018	16
Average province	0.5356	7	0.3730	10	0.5275	6	0.6150	7

3. The distance matrix is the basis for cluster analysis, which is conducted using the farthest neighbour method of hierarchical clustering, which can be visualised as a dendrogram (Fig. 2).

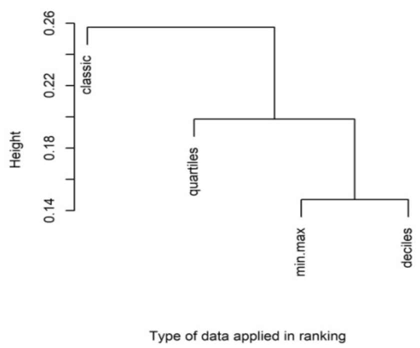


Fig. 2. Dendrogram of rankings of Polish provinces based on 4 types of data

Results of the approach based on interval-valued data are considerably different from those obtained using metric data. The differences increase along with the width of the intervals.

The next analysis focused on the similarity between rankings created with respect to different variables and the ranking based on the aggregate measure, for the four datasets. The following procedure was adopted:

1. The four datasets (metric, min-max, deciles, and quartiles) are linearly ordered according to a set of *m* variables to produce 4 rankings (see Table 1).
2. For each variable ($j = 1, ..., m$) a distance between each object and the pattern object is calculated according to the formula:
   a) for metric data:

$$d_i = 1 - |x_{ij} - x_{+j}|/(x_{+j} - x_{-j}), \quad j = 1,\ldots, m, \tag{2}$$

   where: $j = 1,\ldots, m$ – variable number, x_{ij} – value of j-th variable for i-th object, x_{+j} (x_{-j}) – j-th coordinate of the pattern (anti-pattern) object.

   b) for interval-valued data (Ichino-Yaguchi distance for one variable):

$$d_i = 1 - |\varphi(x_{ij}, x_{+j})|/\varphi(x_{+j}, x_{-j}) \tag{3}$$

   where: $x_{ij} = [x^l_{ij}, x^u_{ij}](x^l_{ij} \leqslant x^u_{ij})$ – interval (min-max, 1st and 9th deciles, 1st and 3rd quartiles), $\varphi(x_{ij}, x_{+j}) = |x_{ij} \oplus x_{+j}| - |x_{ij} \otimes x_{+j}| + 0.5(2 \cdot |x_{ij} \otimes x_{+j}| - |x_{ij}| - |x_{+j}|)$, $|\,|$ – interval length, $x_{ij} \oplus x_{+j} = x_{ij} \cup x_{+j}$; $x_{ij} \otimes x_{+j} = x_{ij} \cap x_{+j}$, x_{+j} (x_{-j}) – the pattern (anti-pattern) object for j-th variable.

   It yields a set of *m* rankings.
3. The general ranking (step 1) is compared with individual rankings (step 2) using Kendall's tau coefficient.
4. Results obtained in step 3 are averaged (see Table 2). A higher average value indicates a higher degree of similarity between the ranking of objects with respect to a given set of variables and the ranking obtained on the basis of the aggregate measure.

Table 2. Assessment of the similarity of rankings of objects with respect to a given set of variables and the ranking obtained on the basis of the aggregate measure

No.	Types of data	Average value of Kendall's tau	Rank
1	Atomic (metric) data	0.0742	4
2	Interval-valued (min-max)	0.1074	2
3	**Interval-valued (1st and 9th deciles)**	**0.1412**	**1**
4	Interval-valued (1st and 3rd quartiles)	0.0861	3

The highest degree of similarity between the rankings of objects with respect to different variables and the ranking obtained on the basis of the aggregate measure is achieved when the intervals are defined by deciles. The decile-based approach can be classified as a robust approach since it reduces the impact of outliers.

Conclusions

The level of social cohesion in Polish provinces was assessed using two approaches: a classical one, based on average metric values, and a symbolic one, based on interval-valued data (min-max, deciles, quartiles). The proposed approach has made it possible to assess social cohesion in provinces not only on the basis of mean values of the variables, but also by taking into account the intervals.

The results of the interval-based approach are considerably different from those obtained using the classical approach (see the dendrogram in Fig. 2). The highest degree of similarity between the rankings of objects with respect to different variables and the ranking obtained on the basis of the aggregate measure is achieved when the intervals are defined by deciles. This approach helps to eliminate the influence of outliers on the assessment of social cohesion in the provinces of Poland.

All the calculations were conducted using scripts written by the authors in the R program.

Acknowledgements

The project is financed by the Polish National Science Centre DEC-2015/17/B/HS4/00905.

References

Annoni, P., & Dijkstra, L. (2016). The EU Regional Social Progress Index: Methodological Note. Brussels: European Commission.

Billard, L., & Diday, E. (2006). *Symbolic Data Analysis: Conceptual Statistics and Data Mining*. Chichester: John Wiley.

Borg, I., & Groenen, P.J.F. (2005). *Modern Multidimensional Scaling. Theory and Applications*. New York: Springer Science+Business Media.

Borg, I., Groenen, P.J.F., & Mair, P. (2018). *Applied Multidimensional Scaling and Unfolding*. Heidelberg, New York, Dordrecht, London: Springer.

Brito, P., Noirhomme-Fraiture, M., & Arroyo, J. (2015). Editorial for special issue on symbolic data analysis. *Advanced in Data Analysis and Classification*, 9(1), 1–4.

Chan, J., To, H., & Chan, E. (2006). Reconsidering social cohesion: Developing a definition and analytical framework for empirical research. *Social Indicators Research*, 75, 273–302.

Dehnel, G., Walesiak, M., & Obrębalski, M. (2018). Comparative analysis of the ordering of Polish provinces in terms of social cohesion. *Argumenta Oeconomica Cracoviensia* (in press).

Dickes, P., & Valentova, M. (2013). Construction, Validation and Application of the Measurement of Social Cohesion in 47 European Countries and Regions, *Social Indicators Research*, 113, 827–846.

Dickes, P., Valentova, M., & Borsenberger, M. (2010). Construct Validation and Application of a Common Measure of Social Cohesion in 33 European Countries. *Social Indicators Research*, 98, 451–473.

Everitt, B.S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. Chichester: Wiley.

Hellwig, Z. (1981). *Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych*. In: Welfe, W. (Ed.), *Metody i modele ekonomiczno-matematyczne w doskonaleniu zarządzania gospodarką socjalistyczną*, 46–68. Warszawa: PWE.

Gioia, F., & Lauro, C.N. (2006). Principal component analysis on interval data. *Computational Statistics*, 21(2), 343–363.

Ichino, M., & Yaguchi, H. (1994). Generalized Minkowski metrics for mixed feature-type data analysis. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(4), 698–708.

Jajuga, K., Walesiak, M., & Bąk, A. (2003). *On the general distance measure,* In: Schwaiger, M., Opitz, O. (Eds.), *Exploratory data analysis in empirical research*, 104–109. Berlin, Heidelberg: Springer-Verlag.

Kendall, M.G. 1955. *Rank correlation methods*. London: Griffin.

Mair, P., De Leeuw, J., Borg, I., & Groenen, P. J. F. (2018). smacof: Multidimensional Scaling, R package ver. 1.10–8. Https://CRAN.R-project.org/package=smacof.

R Core Team (2018). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. Https://www.R-project.org

Walesiak, M. (2016). Visualization of linear ordering results for metric data with the application of multidimensional scaling. *Ekonometria*, 2(52), 9–21.

Walesiak, M., & Dehnel, G. (2018). Evaluation of Economic Efficiency of Small Manufacturing Enterprises in Districts of Wielkopolska Province Using Interval-Valued Symbolic Data and the Hybrid Approach. In: Papież, M. and Śmiech, S. (Eds.), *The 12th Professor Aleksander Zeliaś International Conference on Modelling and Forecasting of Socio-Economic Phenomena*. Conference Proceedings, Foundation of the Cracow University of Economics, Cracow, 563–572.

Walesiak, M., & Dudek, A. (2017). Selecting the optimal multidimensional scaling procedure for metric data with R environment. *Statistics in Transition new series*, 18(3), 521–540.

Walesiak, M., & Dudek, A. (2018a). clusterSim: Searching for Optimal Clustering Procedure for a Data Set. *R package*, version 0.47-3. http://CRAN.R-project.org/package=clusterSim

Walesiak, M., & Dudek, A. (2018b). mdsOpt: Searching for Optimal MDS Procedure for Metric Data. *R package*, version 0.3-3. http://CRAN.R-project.org/package=mdsOpt