Conference is held under the auspices 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

Socio-Economic Modelling and Forecasting

No.1 (2018)

The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings

Edited by

Monika Papież and Sławomir Śmiech

MAY 8-11, 2018

ZAKOPANE, POLAND

Editors: Monika Papież, Sławomir Śmiech

All published papers have been reviewed by two anonymous referees before publishing

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

Publishing of the proceedings is founded by Polish Academy of Sciences and Cracow University of Economics

Suggested citation:

AUTHOR, A. (2018). Title of the paper. 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, pp. xx-xx. ISBN: 978-83-65907-20-2 (HTML).

© Copyright: Proceedings are licensed under Creative Commons Attribution 3.0 License

e-ISSN
ISBN: (HTML) 978-83-65907-20-2
DOI: 10.14659/SEMF.2018.01

Year of publishing: 2018
Number of pages: 611

Publishing House:
Foundation of the Cracow University of Economics
Rakowicka 27, 31-510 Cracow, Poland

Programme Committee:

Chairman: **Prof. dr hab. Józef Pociecha (Chairman)**
Department of Statistics
Cracow University of Economics, Poland

Members: **Prof. dr hab. Andrzej S. Barczak**
Department of Econometrics
University of Economics in Katowice, Poland

Prof. dr hab. Czesław Domański
Department of Statistical Methods
University of Łódź, Poland

Prof. dr hab. Krzysztof Jajuga
President of the Committee of Statistics and
Econometrics of The Polish Academy of Sciences
Department of Financial Investments and Risk
Management Wroclaw University of Economics, Poland

Prof. dr hab. Jacek Osiewalski
Vice President of the Committee of Statistics and Econometrics
of The Polish Academy of Sciences
Department of Econometrics and Operations Research
Cracow University of Economics, Poland

Prof. RNDr. Viera Pacakova, PhD
Institute of Mathematics and Quantitative Methods
University of Pardubice, Czech Republic

Prof. Anatoliy I. Pilyavskyy
Department of Mathematics
Lviv Academy of Commerce, Ukraine

Doc. Ing. Eva Sodomova, PhD
Department of Statistics
University of Economics in Bratislava, Slovakia

Prof. dr hab. Mirosław Szreder
Department of Statistics
University of Gdansk, Poland

Organizing Committee:

Chairman: **Prof. dr hab. Józef Pociecha (Chairman)**
Department of Statistics
Cracow University of Economics, Poland

Vice-chairman: **Prof. UEK dr hab. Anna Malina**
Department of Statistics
Cracow University of Economics, Poland

Prof. UEK dr hab. Barbara Pawełek
Department of Statistics
Cracow University of Economics, Poland

Members: **Jadwiga Kostrzewska, PhD (Scientific Secretary)**
Department of Statistics
Cracow University of Economics, Poland

Prof. UEK dr hab. Monika Papież
Department of Statistics
Cracow University of Economics, Poland

Prof. UEK dr hab. Sławomir Śmiech
Department of Statistics
Cracow University of Economics, Poland

Katarzyna Frodyma, PhD
Department of Statistics
Cracow University of Economics, Poland

Roman Huptas, PhD
Department of Statistics
Cracow University of Economics, Poland

Artur Lipieta, PhD
Department of Statistics
Cracow University of Economics, Poland

Journal website:
http://www.semf.pl

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

Organizer's website:
http://www.katstat.uek.krakow.pl/en/

List of Referees:
Adam P. Balcerzak, Nicolaus Copernicus University in Toruń, Poland
Paweł Baran, University of Szczecin, Poland
Barbara Batóg, University of Szczecin, Poland
Tomasz Bartłomowicz, Wrocław University of Economics, Poland
Katarzyna Bech, SGH Warsaw School of Economics, Poland
Jacek Białek, University of Łódź, Poland
Beata Bieszk-Stolorz, University of Szczecin, Poland
Justyna Brzezińska, University of Economics in Katowice, Poland
Anita Csesznák, Budapest Business School, Hungary
Małgorzata Ćwiek, Cracow University of Economics, Poland
Marek A. Dąbrowski, Cracow University of Economics, Poland
Grażyna Dehnel, Poznań University of Economics and Business, Poland
Sabina Denkowska, Cracow University of Economics, Poland
Krzysztof Dmytrów, University of Szczecin, Poland
Kamil Fijorek, Cracow University of Economics, Poland
Iwona Foryś, University of Szczecin, Poland
Katarzyna Frodyma, Cracow University of Economics, Poland
Ewa Genge, University of Economics in Katowice, Poland
Justyna Góral Institute of Agricultural and Food Economics - National Research Institute in Warsaw, Poland
Krzysztof Guzik, Cracow University of Economics, Poland
Sergiusz Herman, Poznań University of Economics and Business, Poland
Maciej Jewczak, University of Łódź, Poland
Alina Jędrzejczak, University of Łódź, Poland
Paweł Kawa, Cracow University of Economics, Poland
Karol Korczak, University of Łódź, Poland
Daniel Kosiorowski, Cracow University of Economics, Poland
Maciej Kostrzewska, Cracow University of Economics, Poland
Nataliia Kovtun, Taras Shevchenko National University of Kyiv, Ukraine
Marta Kuc, Gdansk University of Technology, Poland
Tomasz Kwarciński, Cracow University of Economics, Poland
Łukasz Kwiatkowski, Cracow University of Economics, Poland
Łukasz Lenart, Cracow University of Economics, Poland
Paweł Lula, Cracow University of Economics, Poland
Radosław Mącik, Maria Curie-Sklodowska University, Poland
Justyna Majewska, University of Economics in Katowice, Poland
Iwona Markowicz, University of Szczecin, Poland
Błażej Mazur, Cracow University of Economics, Poland
Dominik Mielczarek, AGH University of Science and Technology, Poland
Artur Mikulec, University of Łódź, Poland
Małgorzata Misztal, University of Łódź, Poland
Iwona Müller-Frączek, Nicolaus Copernicus University in Toruń, Poland
Sabina Nowak, University of Gdańsk, Poland
Joanna Olbryś, Bialystok University of Technology, Poland
Jacek Osiewalski, Cracow University of Economics, Poland
Anna Pajor, Cracow University of Economics, Poland
Monika Papież, Cracow University of Economics, Poland
Dorota Pekasiewicz, University of Łódź, Poland
Marcin Pełka, Wrocław University of Economics, Poland
Michał Bernard Pietrzak, Nicolaus Copernicus University in Toruń, Poland
Mateusz Pipień, Cracow University of Economics, Poland
Włodzimierz Rembisz Institute of Agricultural and Food Economics - National Research Institute in Warsaw, Poland
Elżbieta Rogalska, University of Warmia and Mazury, Poland
Elżbieta Roszko-Wójtowicz, University of Łódź, Poland
Michał Rubaszek, SGH Warsaw School of Economics, Poland
Aneta Rybicka, Wroclaw University of Economics, Poland
Jerzy P. Rydlewski, AGH University of Science and Technology, Poland
Marcin Salamaga, Cracow University of Economics, Poland
Éva Sándorné Kriszt, Budapest Business School, Hungary
Katerina Shestaeva, Plekhanov Russian University of Economics Moscow, Russia
Victor Shevchuk, Cracow University of Technology, Poland
Elena Sibirskaya Plekhanov Russian University of Economics Moscow, Russia
Agata Sielska, SGH Warsaw School of Economics, Poland
Iwona Skrodzka, Bialystok University of Technology, Poland
Aneta Sobiechowska-Ziegert, Gdansk University of Technology, Poland
Erik Šoltés, University of Economics in Bratislava, Slovakia
Małgorzata Stec, University of Rzeszów, Poland
Jacek Stelmach, University of Economics in Katowice, Poland
Adam Suchecki, University of Łódź, Poland
Sławomir Śmiech, Cracow University of Economics, Poland
Michael Thrun University of Marburg, Germany
Grażyna Trzpiot, University of Economics in Katowice, Poland
Paweł Ulman, Cracow University of Economics, Poland
Marta Wajda-Lichy, Cracow University of Economics, Poland
Marek Walesiak, Wrocław University of Economics, Poland
Grzegorz Wałęga, Cracow University of Economics, Poland
Stanisław Wanat, Cracow University of Economics, Poland
Małgorzata Wosiek, University of Rzeszów, Poland
Justyna Wróblewska, Cracow University of Economics, Poland
Wojciech Zieliński, Warsaw University of Life Sciences, Poland
Zinaida Palian, Taras Shevchenko National University of Kyiv, Ukraine

CONTENTS

Preface 11

Paweł Baran, Iwona Markowicz
Analysis of intra-Community supply of goods shipped from Poland 12

Jacek Białek
Comparison of Jevons and Carli elementary price indices 22

Beata Bieszk-Stolorz, Krzysztof Dmytrów
Application of the survival trees for estimation of the influence of determinants on probability of exit from the registered unemployment 30

Justyna Brzezińska, Aneta Rybicka, Marcin Pełka
Multivariate statistical analysis of environmental data 40

Nataliia Cherkas, Myroslava Chekh
Economic effects of production fragmentation and technological transfer: the evidence from CEE countries 50

Anna Chudzicka-Bator, Mateusz Pipień
Modelling intensity of EUR/PLN high frequency trade – a comparison of ACD-type models 60

Marta Chylińska
Aluminium price discovery on the London Metal Exchange, 2007-2017 70

Małgorzata Ćwiek
Determinants of digital divide in Polish households 80

Marek A. Dąbrowski, Łukasz Kwiatkowski, Justyna Wróblewska
Sources of real exchange rate variability in Poland – Evidence from a Bayesian SVAR model with Markov Switching Heteroscedasticity 90

Marek A. Dąbrowski, Monika Papież, Sławomir Śmiech
Is there a trade-off between monetary independence and exchange rate stability in Central and Eastern European economies? 100

Grażyna Dehnel, Łukasz Wawrowski
Robust estimation of revenues of Polish small companies by NACE section and province . 110

Czesław Domański, Robert Kubacki
Application of differential evolution algorithm to group bank's individual clients 120

Iwona Foryś, Barbara Batóg
Analysis of transaction prices on one of the housing estates in Szczecin: Do the buyers differentiate prices on local markets with respect to the kind of the right to own? 130

Katarzyna Frodyma, Monika Papież, Sławomir Śmiech
Efficiency of investments in solar power in the EU countries 140

Justyna Góral, Włodzimierz Rembisz
Spatial differentiation of the remuneration and labour productivity in Polish agriculture 150

Sergiusz Herman
Application of Generalized Data Envelopment Analysis on Warsaw Stock Exchange 160

Mateusz Jankiewicz
Changes of the consumption structure in European countries considering its modernization process 170

Alina Jędrzejczak, Dorota Pekasiewicz, Wojciech Zieliński
Comparison of parameter estimators for the Dagum distribution 180

Maciej Jewczak, Jadwiga Suchecka, Karol Korczak, Marek Melaniuk
Spatial analysis and assessment of effectiveness of selected social services 190

Daniel Kosiorowski, Dominik Mielczarek, Jerzy P. Rydlewski
New proposal of robust classifier for functional data 200

Daniel Kosiorowski, Dominik Mielczarek, Jerzy P. Rydlewski
Outliers in functional time series – challenges for theory and applications of robust statistics 209

Jadwiga Kostrzewska, Maciej Kostrzewski
The logistic regression in predicting spike occurrences in electricity prices 219

Marta Kuc, Aneta Sobiechowska-Ziegert
Standard of living in Poland at regional level - classification with Kohonen self-organizing maps 229

Łukasz Lenart
Bayesian inference for deterministic cycle with time-varying amplitude 239

Łukasz Lenart, Justyna Wróblewska
Nonlinear stochastic cycle model 248

Paweł Lula, Renata Oczkowska, Sylwia Wiśniewska
Automatic identification of competences expected by employers with the use of exploratory text analysis 256

Radosław Mącik
Identification of factors affecting the popularity of cross-border online shopping in the European Union countries 266

Justyna Majewska, Grażyna Trzpiot
The health transition and an ageing society – example of selected European countries 276

Błażej Mazur
Cyclical fluctuations of global food prices: a predictive analysis 286

Błażej Mazur, Mateusz Pipień
A family of non-standard bivariate distributions with applications to unconditional modelling in empirical finance 296

Artur Mikulec, Małgorzata Misztal
Firms' duration analysis in Lodzkie Voivodship using a regression tree approach 306

Paweł Miłobędzki, Sabina Nowak
The accuracy of trade classification rules for the Warsaw Stock Exchange 316

Iwona Müller-Frączek
Dynamic measure of development 326

Joanna Olbryś, Michał Mursztyn
Assessing accuracy of trade side classification rules. Methods, data, and problems 335

Jacek Osiewalski, Anna Pajor
A hybrid MSV-MGARCH generalisation of the t-MGARCH model 345

Monika Papież, Sławomir Śmiech, Katarzyna Frodyma
Is development of the renewable energy sector crucial for the electricity consumption-growth nexus in the EU countries? 355

Aleksandra Pawłowska
Evaluation of investment support from Common Agricultural Policy with propensity score matching method 365

Michał Bernard Pietrzak
The composition of agricultural macroregions in Eastern Poland - an empirical example of the Aggregation Problem 375

Dominika Polko-Zając
Comparing populations based on squared Euclidean distance between sets of variables 385

Elżbieta Rogalska
Entrepreneurship conditions in Poland at NUTS 3 level. Application of taxonomic measure of development based on median vector Weber 395

Elżbieta Roszko-Wójtowicz, Jacek Białek
Application of the AG mean index for CPI substitution bias reduction 405

Michał Rubaszek
A note on the private housing rental market in Poland 413

Marcin Salamaga
An Application of conjoint analysis to study determinants of Polish direct investment located in European countries 423

Katerina Shestaeva, Elena Sibirskaya, Paweł Lula, Evgenia Progonova
Anatomy of entrepreneurial activity and challenges of the new century 433

Victor Shevchuk, Roman Kopych, Marianna Golynska
Fiscal and monetary effects in Ukraine: SVAR approach 443

Agata Sielska
Armed conflicts and multicriterial evaluation of the development of the country in a long-term perspective 453

Iwona Skrodzka
International technology transfer and smart growth of EU countries in 2010-2015 463

Erik Šoltés, Tatiana Šoltésová
Comparison of income poverty and social exclusion in the EU in 2008 and 2016 473

Małgorzata Stec, Małgorzata Wosiek
Evaluation of the socio-economic situation of European Union countries, taking into account accuracy of statistical data 483

Jacek Stelmach
On stationarity of changes in the trends of selected refining variables 493

Adam Mateusz Suchecki
Spatio-temporal decomposition of the growth of communal budgets expenses on culture in Poland in 2003-2016. Implementation of the SSANOVA model. 503

Sławomir Śmiech, Aleksandra Mucha
Are uncertainty measures powerful predictors of real economic activity in the Euro Area?. 513

Michael C. Thrun
Cluster analysis of the World Gross-Domestic Product based on the emergent self-organization of a swarm 523

Michael C. Thrun, Alfred Ultsch
Effects of the pay-out system of income taxes to municipalities in Germany 533

Paweł Ulman, Tomasz Kwarciński
The quality of life in Poland and Germany 543

Marta Wajda-Lichy, Paweł Kawa, Kamil Fijorek, Sabina Denkowska
Trade openness and financial development: Granger causality analysis for new EU member states 553

Marek Walesiak, Grażyna Dehnel
Evaluation of economic efficiency of small manufacturing enterprises in districts of Wielkopolska province using interval-valued symbolic data and the hybrid approach 563

Agnieszka Wałęga, Grzegorz Wałęga
Does debt improve housing conditions? Evidence from Polish households 573

Stanisław Wanat, Krzysztof Guzik
Estimation of VaR bounds under dependence uncertainty and their use for the SCR calculation in Solvency II 583

Antonina M. Yerina, Zinaida O. Palian
Estimation of indirect demographic losses in Ukraine due to armed conflict 593

Wojciech Zieliński, Alina Jędrzejczak, Dorota Pekasiewicz
Estimation of quantile ratios of the Dagum distribution 603

PREFACE

Quantitative approach to investigation of contemporary socio-economic processes is the unique way to formulate proposals to resolve global as well as regional economic problems. Statistical and econometric methods from both, theoretical and empirical point of views, will be discuss at The 12th International Conference in honour of Professor Aleksander Zeliaś, which take place in Zakopane at 8-11 of May 2018. Participants of the conference are both well known and young scientists from Poland, Czech Republic, Germany, Italy, Slovakia, Russia and Ukraine.

The volume presented here contains selected conference proceedings, independently revised by two anonymous reviewers, among almost 80 submitted studies. The papers present a current stage of statistical and econometric knowledge, interesting applications, econometric modelling technics and data analysis applications in a variety of areas of economic processes. Conference presentations concentrate on financial issues, analysing the macro- and micro-levels, threats of contemporary world, modelling and forecasting economic processes, computational tools for statistical and econometric analyses, spatial and regional modelling, risk analysis, statistical methods for business investigations.

We hope that Readers find in the collection of papers original ideas, flashes of inspiration, useful methods and interesting results of empirical investigations of important socio-economic problems related to both Central-East European countries and global economy as well.

Józef Pociecha

Analysis of intra-Community supply of goods shipped from Poland

Paweł Baran[1], Iwona Markowicz[2]

Abstract

After ceasing all customs duties at the borders between EU member states, the Community lost a viable source of data on international trade. This is why the Intrastat system was introduced. Poland's accession to EU imposed new duties on every entity selling goods to or buying them from other EU member states. Such businesses are required to submit INTRASTAT declarations to National Revenue Administration. Statistical data on international trade collected in the process are then combined at Eurostat. Such data are often incompatible. An example is the difference between two datasets: one containing data on intra-Community supplies (ICS) dispatched from Poland and the other containing data on intra-Community acquisitions (ICA) originating in Poland. The authors have examined such differences on Combined Nomenclature chapter (2-digit) level for both total figures and divided by country. The next part of the survey was to classify countries by the structure of ICS from Poland. The goals of the article were pointing out the CN chapters with the largest differences between ICS and ICA from Poland and what follows – that choice of the source of data on foreign trade may result in different outcomes and conclusions. We need to stress out that we will base our whole work on public statistics only. The very same data serve as the basis for all knowledge on EU intra-Community trade.

Keywords: *Intra-Community supply, cluster analysis, public statistics*
JEL Classification: C38, F14
DOI: 10.14659/SEMF.2018.01.01

1 Introduction

One result of creation of the EU (and its predecessor, the EC) is that all customs duties at the borders between EU countries were revoked. At the same time all customs clearances stopped and Simple Administrative Documents (SADs) are no longer in use between EU members. Thus, the Community has been deprived of a viable source of data on international trade. It became necessary to introduce a new, common system of statistics of trade in goods. This is why on January 1st, 1993 the Intrastat system was introduced in the whole area of the European Single Market.

In Poland these regulations became effective on May 1st, 2004, i.e. Poland's accession to EU imposed new duties on every entity selling goods to or buying them from other EU member states. An entity trading in goods with other member states of the EU is required to

[1] Corresponding author: University of Szczecin, Department of Econometrics and Statistics, ul. Mickiewicza 64, 71-101 Szczecin, Poland, e-mail: Pawel.Baran@usz.edu.pl.
[2] University of Szczecin, Department of Econometrics and Statistics, ul. Mickiewicza 64, 71-101 Szczecin, Poland, e-mail: Iwona.Markowicz@usz.edu.pl.

submit INTRASTAT declarations on intra-community supplies and acquisitions to Revenue Administration Regional Office in Szczecin. Effective March 1st 2017 National Revenue Administration has taken over the collection process of data gathered through INTRASTAT declarations as well as data management and the process of creating a dataset for Central Statistical Office with the use of its own resources. Deploying the INTRASTAT system after the accession was a huge project and it still needs maintaining. A similar project, started in 2012 in Croatia, has lately been described in detail by Erceg (2015).

Statistical data on international trade collected from individual declarations are then combined at Eurostat (European Statistical Office) together with other countries' data. There are still works underway aimed at getting the datasets from different national statistical offices fully comparable and compatible. These works are of great importance since huge discrepancies still exist. An example is the difference between two datasets – first of them containing data on intra-Community supplies (ICS) dispatched from Poland (collected at national level) and the second containing data on intra-Community acquisitions (ICA) originating in Poland (aggregated by Eurostat from other EU members data) that will be addressed later in the article. The above-mentioned differences between datasets are hard to explain in terms of exchange rate or late collection of data. The authors have examined such differences on Combined Nomenclature chapter (2-digit) level for both total figures and divided by country. The next part of the survey was to classify countries by the structure of ICS from Poland. The goal of the article is to point out the CN chapters with the largest differences between ICS and ICA from Poland and what follows – that choice of the source of data on foreign trade even from the same database may result in different outcomes and conclusions. We need to stress out that we have based our whole work on public statistics only. The very same data serve as the basis for all knowledge on EU intra-Community trade.

2 Public statistics of foreign trade

EU member states data on international trade are collected within two parallel systems of collecting public statistics. These are: INTRASTAT system – a system of public statistics containing data on intra-community trade, which is based on data collected from Intrastat declarations, EXTRASTAT system – a system of public statistics containing all trade with third parties (that is countries other than EU members). Data obtained from these two parallel systems constitute homogenous set of statistical data on foreign trade turnover.

On June 1st 2016 Polish Ministry of Finance has deployed a new computer system AIS/INTRASTAT dedicated to process INTRASTAT declarations. Intrastat declaration form

contains the most important data on intra-Community transactions. For minimising overall burden of statistical reporting put on small businesses, only turnover above specified threshold needs to be registered. In 2016 statistical basic threshold was 3,000,000 PLN for arrivals, and 1,500,000 PLN for dispatches**.**

It is only for the last couple of years that Polish exports exceed imports in net balance of foreign trade. Namely, overall foreign trade net balance over the period 2004-2014 was negative (Fig. 1), and it became positive in 2015 for the first time. However, intra-Community trade was quite different. Polish exports to other EU member states have exceeded imports from them since the accession in 2004. In 2016 all Polish exports reached the net value of €184,842.9m (while imports were worth €180,924.6m), from which exports to Europe – €162,963.0m (88,2%), and to EU member states €147,563.6m (79,8%) (Ministry of Economic Development, 2017).

The discussion about the regulation of commercial barriers and the European Community's Value-Added Tax System has been going on for many years (MacLean, 1999; Hart, 1994). According to European Commission (2015), two fundamental issues were identified with the current taxation system. These are: 1 – the additional obligations and costs associated with VAT compliance for businesses engaging in cross-border trade, 2 – the existing levels of VAT fraud within the EU through fraudulent transactions such as MTIC (‘Missing Trader Intra-Community’) fraud (also known as carousel fraud).

There are many reasons, some of them mentioned above, for which generating reliable data on foreign trade isn’t straightforward. Later in the article, there will be some research into discrepancies in public statistics presented.

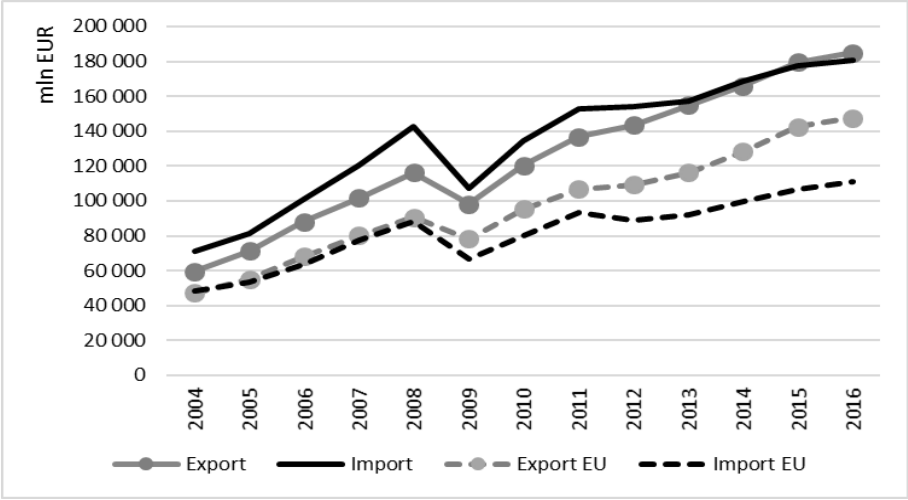


Fig. 1. Polish foreign trade turnover total and with EU member states (Data: SWAID, GUS).

3 Statistical data and research methodology

We used data from Eurostat's COMEXT database. Data on ICS originating from Poland in 2016 were analysed with regard to country and CN chapters. These were values declared by Polish entities mixed together with figures estimated in place of missing data according to methodology provided by Central Statistical Office. On the other hand, we considered data on ICA to EU member states from Poland, declared in 2016 by contractors of Polish businesses and again estimated by national statistical offices respectively. Of course, these data would not be 100% consistent, but it turns out that there exist certain chapters for which there are huge differences, both positive (by positive we mean predominance of exports declared by Polish entities) and negative.

There are several possible causes of differences between ICA and ICS. The main is introducing aforementioned thresholds for the obligatory declaration of foreign trade (these are different for export and for import and differ between member states). Some other are: concealing transactions from taxation (tax evasion), multiplying transactions and declaring fictitious ones (VAT carousels) or simple errors in declarations (e.g. wrong CN code or wrong value of traded goods).

In order to find structural misrepresentation in Comext data, we calculated fractions of ICS from Poland to every member state in all chapters. Then for every pair of member states we compared the structure with a structural similarity index:

$$W_{ij} = \sum_{d=1}^{k} \min_{i \neq j} \left(w_{id}, w_{jd} \right) \tag{1}$$

where:

i, j – EU country, $i, j = 1, \ldots, l$, $l = 27$,

d – CN chapter, $d = 1, \ldots, k$, $k = 97$,

w_{id}, w_{jd} – shares of CN chapter d in structures of trade with countries i and j, respectively.

The above-mentioned index is easy in terms of both computation and interpretation. It takes on values from the interval [0; 1] (Chomątowski and Sokołowski, 1978) and reveals countries that have similar structures of acquisitions from Poland.

A unit of data under consideration contains all information on ICS from Poland to a specified member state as an observation of an object. Features are fractions of CN chapters within total trade hence it is a vector. A different, more complex approach, with exports from many countries (i.e. matrices of data) considered as data units, presents Salamaga (2017) to compare full structures of foreign trade of 18 EU member states.

Another approach to a similar problem is by Landesmann (2000), who refers to structural change in two ways: changes in compositional structures (of output, employment, exports, etc.) and changes in behaviour, that is in the ways in which variables relate to each other.

It is worth noting that this approach to structure similarity is just one of possible choices. A more general approach would be to understand it as a close relative to distance measures widely considered in multivariate statistical methods. Many such measures need data normalization as a prerequisite, yet features are subject to weighting procedures. The two steps are absent in the above method.

Aside of examining similarities for pairs of countries we have also undertaken an attempt to classify EU member states as destinations of Polish foreign trade. In order to show groups of countries with similar structure of goods bought in Poland we used hierarchical clustering. Agglomerative clustering methods have certain advantages, among them: one strict algorithm, results presented in a form of a series of classifications, possibility of graphical presentation with emphasis on sequence of classes generated (Gatnar and Walesiak, 2004). In the analysis, we used Euclidean distance, unitisation of features and Ward's linkage. More on methods and assumptions of classification of objects provide e.g. Anderberg (1973), Kaufman and Rousseeuw (1990), Gordon (1999), Jajuga and Walesiak (2000), Walesiak and Dudek (2010), Markowska et al. (2016).

4 Results of research

The authors calculated the differences between sum of ICS and ICA in 2016 by chapter. They are presented in Fig. 2. As we can see, there are several chapters in which there existed huge differences between figures declared in Poland (ICS dispatched from Poland) and collected from declarations from other member states (ICA originating in Poland). This means that goods from many CN chapters are misrepresented in either of these two datasets. The biggest positive (meaning there was more goods declared as shipped from Poland than those declared as acquired from Poland) and negative (meaning there was more goods declared as acquired from Poland than Polish exporters declared) differences are presented in Table 1.

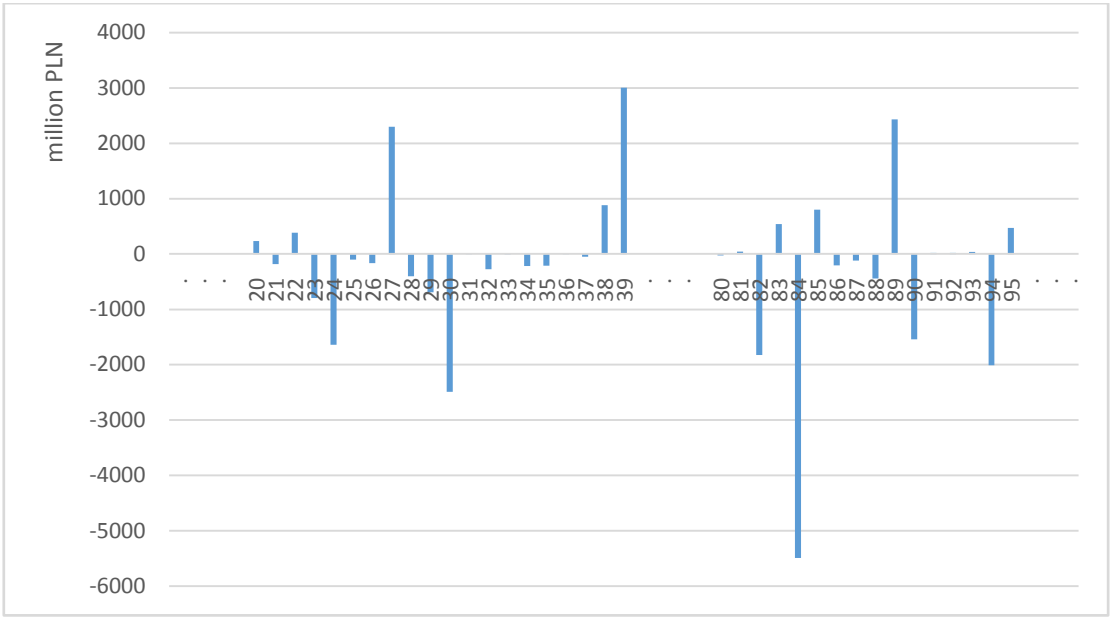


Fig. 2. Differences between ICS and ICA in 2016 by chapter (selection) (Data: Comext).

Table 1. CN Chapters with biggest differences between ICS and ICA.

Number of CN chapter	Description	Difference in bln PLN
39	Plastics and articles thereof	3.01
89	Ships, boats and floating structures	2.43
27	Mineral fuels, mineral oils and products of their distillation; bituminous substances; mineral waxes	2.30
84	Nuclear reactors, boilers, machinery and appliances; parts thereof	−5.49
30	Pharmaceutical products	−2.49
94	Furniture; bedding, mattresses, cushions; lamps and lighting fittings; illuminated signs, nameplates; prefabricated buildings	−2.01
82	Tools, implements, cutlery, spoons and forks, of base metal; parts thereof of base metal	−1.82
24	Tobacco and manufactured tobacco substitutes	−1.64
90	Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments and apparatus; parts and accessories thereof	−1.54

Differences between ICS and ICA turnover have been converted to fractions of the sum of absolute differences, then structural similarity indices were calculated for every pair of EU

member states. Since there exist both positive and negative differences between declared ICS and ICA, the index (1) could not be used directly. Instead, we used a slightly modified version of it – we used absolute values of the differences, and we doubled the set of columns to preserve the negative values from being ruled out of the procedure. The whole matrix is too big to be displayed, thus only a selection of columns is presented in Table 2. There exist countries with similar structures, the two most similar are Greece and Cyprus (W_{ij} = 0.55), followed by Estonia and Latvia (W_{ij} = 0.49), Cyprus and Malta (W_{ij} = 0.49), and Greece and Croatia (W_{ij} = 0.48). Countries like Croatia or Sweden have structures similar to many others. Bulgaria is on the other end of the spectrum, with the structure being least similar to those of other countries.

Table 2. Structural similarities indices for ICS-ICA balance (selection).

	AT	BE	BG	HR	CY	CZ	DK	EE	FI	FR	DE	GR	…
AT	—	0.15	0.06	0.23	0.11	0.27	0.24	0.11	0.24	0.25	0.19	0.25	…
BE	0.15	—	0.16	0.13	0.19	0.15	0.26	0.27	0.13	0.19	0.24	0.19	…
BG	0.06	0.16	—	0.16	0.15	0.17	0.17	0.24	0.15	0.11	0.17	0.14	…
HR	0.23	0.13	0.16	—	**0.31**	**0.29**	0.28	0.26	**0.37**	**0.34**	**0.31**	**0.48**	…
CY	0.11	0.19	0.15	**0.31**	—	0.05	**0.30**	0.19	0.16	0.15	0.11	**0.55**	…
CZ	0.27	0.15	0.17	**0.29**	0.05	—	0.22	0.25	0.22	0.20	0.27	0.17	…
DK	0.24	0.26	0.17	0.28	**0.30**	0.22	—	0.24	**0.35**	0.19	0.25	**0.32**	…
EE	0.11	0.27	0.24	0.26	0.19	0.25	0.24	—	0.17	0.22	0.24	0.16	…
FI	0.24	0.13	0.15	**0.37**	0.16	0.22	**0.35**	0.17	—	0.28	**0.30**	0.25	…
FR	0.25	0.19	0.11	**0.34**	0.15	0.20	0.19	0.22	0.28	—	**0.36**	**0.30**	…
DE	0.19	0.24	0.17	**0.31**	0.11	0.27	0.25	0.24	**0.30**	**0.36**	—	0.19	…
GR	0.25	0.19	0.14	**0.48**	**0.55**	0.17	**0.32**	0.16	0.25	**0.30**	0.19	—	…
HU	**0.35**	0.15	0.14	**0.36**	0.07	0.25	0.27	0.24	0.28	0.25	**0.40**	0.17	…
IE	0.20	0.20	0.14	0.27	**0.31**	0.09	**0.38**	0.26	0.26	0.26	0.25	0.26	…
IT	0.25	**0.42**	0.21	0.18	0.05	**0.39**	0.14	0.19	0.18	0.23	0.28	0.16	…
LV	0.14	0.13	0.23	0.27	0.14	0.24	0.22	**0.49**	0.24	0.15	0.26	0.11	…
LT	**0.32**	0.19	0.23	**0.38**	0.08	**0.30**	**0.34**	**0.30**	**0.31**	0.27	0.28	0.25	…
LU	**0.30**	0.24	0.11	0.21	0.21	0.07	0.24	0.23	**0.33**	**0.42**	0.25	0.27	…
MT	0.19	0.25	0.11	0.17	**0.49**	0.13	**0.31**	0.20	0.17	**0.30**	0.19	**0.39**	…
□	□	□	□	□	□	□	□	□	□	□	□	□	

Because of the differences described above it is important to choose data source and provide research in the area of foreign trade with care. In the second part of the article, we provide an example of classification that is affected by the choice of data source.

First, we classified EU member states according to declared structure of goods sent from Poland (data on ICS from Comext). The results are presented on Fig. 3. There are three clusters of countries visible (cut off at height 5, which was chosen arbitrarily), one of them containing only Germany, the second containing six countries (Czech Republic, United Kingdom, Slovakia, France, Italy, and the Netherlands) and one with all the other member states. Fig. 4 is based on data regarding the same transactions (combined data on ICA from Comext) and reveals similar but significantly different division. The main difference is the absence of Slovakia in the second cluster. Southern neighbour of Poland was replaced by Spain, a country with very different characteristics regarding Poland's foreign trade. There were no changes in method applied, the only difference were the datasets. They were supposed to contain the same data beforehand.

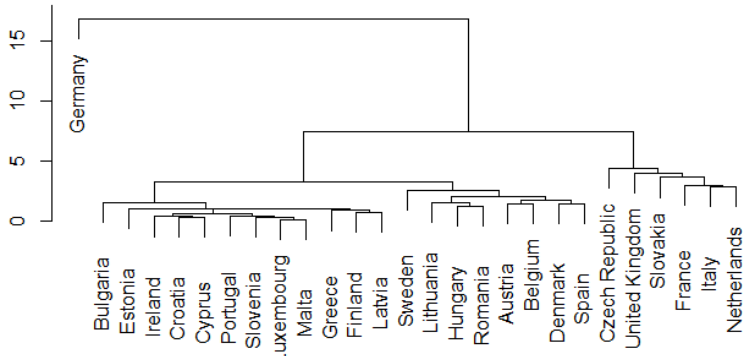


Fig. 3. Classification of EU member states by the structure of ICS from Poland.

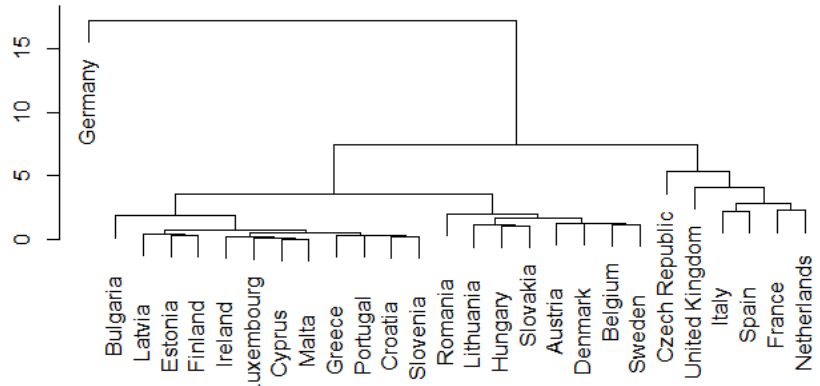


Fig. 4. Classification of EU member states by the structure of ICA from Poland.

Conclusions

Several CN chapters exist in which there are huge differences between figures declared in Poland (ICS) and collected from declarations from other member states (ICA). This means that goods classified in many chapters are misrepresented in either of these two datasets.

The differences tend to have structural nature. One evidence of it is that there are countries with similar structures of such differences.

Researchers need to be cautious with data collected from statistical declarations made by businesses. Since there are virtually no penalties, they may not be reliable. In the second part of the article, we provided an example of classification affected by the choice of data source.

Described differences can be a source of various and possibly vital consequences regarding economic research. They can also affect different aspects of economic policy of the state. These consequences include all possible use of inaccurate public statistics data on foreign trade or wrongly assessed GDP level and/or dynamics. Such a situation where there are no fixed or reliable foreign trade data can also make it hard to perform tax audit as well as to estimate state's tax revenues.

In authors' opinion, it would still be desirable to work on harmonising the system of collecting data on foreign trade, especially trade in goods between EU member states in order to minimise misrepresentation in databases.

References

Chomątowski, S. & Sokołowski, A. (1978). Taksonomia struktur. *Przegląd Statystyczny*, *2*, 217-226.

Erceg, A. (2015). Challenges of Intrastat Implementation in Croatia in Changing Environment. In: Kozarevic, E. & Okicic, J. (eds.). *4th International Scientific Conference Economy of Integration*. Tuzla, 809-822.

European Commission. (2015). *Implementing the 'destination principle' to intra-EU B2B supplies of goods. Feasibility and economic evaluation study.* Final Report.

Gatnar, E. & Walesiak, M. (2004). *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*. Wrocław: Wydawnictwo AE.

Gordon, A. D. (1999). *Classification*. *2nd ed*., Chapman Hall/CRC, London.

GUS. (2017). *Handel zagraniczny. I-IX 2017 r.* Warszawa.

Hart, C. A. (1994). The European Community's Value-Added Tax System: Analysis of the New Transitional Regime and Prospects for Further Harmonization. *Berkeley Journal of International Law 12*(1), 1-62. http://scholarship.law.berkeley.edu/bjil/vol12/iss1/1.

Jajuga, K. & Walesiak, M. (1999). Standardisation of data set under different measurement scales. In: Decker, R. & Gaul, W. (eds.). *Classification and Information Processing at the Turn of the Millennium*. Bielefeld, 105-112.

Kaufman, L. & Rousseeuw, P. J. (1990). *Finding Groups in Data. An Introduction to Cluster Analysis*. Wiley, New York.

Landesmann, M. (2000). Structural change in the transition economies, 1989-1999. *Economic Survey of Europe 2/3*, 95-118.

MacLean, R. M. (1999). The European community's trade barrier regulation takes shape – Is it living up to expectations? *Journal of World Trade 33*(6), 69-96.

Markowska, M., Sokołowski, A. & Strahl, D. (2016). Three-Way Clustering Problems in Regional Science. In: Wilhelm, A. & Kestler, H. (eds.) *Analysis of Large and Complex Data*. Studies in Classification, Data Analysis, and Knowledge Organization. Springer, Cham.

Ministerstwo Rozwoju. (2017). *Syntetyczna informacja o eksporcie i imporcie Polski 2016*.

Salamaga, M. (2017). Ocena podobieństwa struktury towarowej handlu zagranicznego w krajach Unii Europejskiej, *Wiadomości Statystyczne*, *11*, 5-16.

Walesiak, M. & Dudek, A. (2010). Finding Groups in Ordinal Data: An Examination of Some Clustering Procedures. In: Locarek, J. H. & Weihs, C. (eds.). *Classification as a tool for Research.* Dresden, 185-192.

Comparison of Jevons and Carli elementary price indices

Jacek Białek[1]

Abstract

Most of countries use either Jevons or Carli index for the calculation of their Consumer Price Index (CPI) at the lowest (elementary) level of aggregation. The choice of the elementary formula for the inflation measurement does matter and the effect of the change of the index formula was estimated by the Bureau Labor Statistics (2001). It was shown (Hardy et. al, 1934) that the difference between the Carli index and the Jevons index is bounded from below by the variance of the price relatives. In this paper we extend this result comparing expected values of these sample indices under the assumption that prices are described by geometric Brownian motion.

Keywords: *CPI, Jevons index, Carli index*

JEL Classification: E1, E2, E3

DOI: 10.14659/SEMF.2018.01.02

1 Introduction

Elementary price indices are used in inflation measurement on the lowest level of aggregation. Choice of the elementary formula does matter. For instance, in January 1999 the formula used for aggregating price changes for the US consumer price index (CPI) at the lower level of aggregation was changed into a ratio of geometric means of prices (Silver and Heravi, 2007). The effect of this change was researched by the Bureau of Labor Statistics (2001) to reduce the annual rate of increase in the CPI by approximately 0.2 percentage points. As a consequence, it increased a cumulative national debt from over-indexing the federal budget by more than $200 billion per twelve years (Boskin et al., 1996, 1998).

In March 2013, the UK's Office for National Statistics (ONS) started to publish a new inflation index – RPIJ. This index is identical to the Retail Price Index (RPI), except it uses a geometric mean of price relatives (known as Jevons index) rather than an arithmetic mean of price relatives (the Carli index). Moreover, none of the 28 European Union countries makes use of the Carli index in their national price indices. Eurostat regulations do not allow the use of the Carli index in the construction of members' Harmonized Index of Consumer Prices (HICP). There has been a general trend in replacing the Carli index with the Jevons or the Dutot formulas (Evans, 2012). Some countries abandoned the Carli index formula in favour of other price indices over the last few decades, like Canada (in 1978), Luxemburg (in 1996),

[1] University of Lodz, Department of Statistical Methods, Rewolucji 1905 r. St., No. 41, 90-214 Łódź, email: jbialek@uni.lodz.pl.

Australia (in 1998), Italy (in 1999) or Switzerland (in 2000). In 1996, the Boskin Commision in the USA recommended that a Carli-like index that was used in the US CPI should be replaced with the Jevons index (Levell, 2015).

There are many papers that compare the above-mentioned unweighted price index numbers (Silver and Heravi, 2007; Levell, 2015). In this paper we focus on only two elementary price indices, namely we consider Jevons and Carli formulas. It was shown (Hardy et. al, 1934) that the difference between the Carli index and the Jevons index is bounded from below by the variance of the price relatives. In this paper we extend this result comparing expected values of these sample indices under the assumption that price relatives are described by geometric Brownian motion.

2 Unweighted Jevons and Carli indices

There are several elementary price indices in the literature (Von der Lippe, 2007).
In particular we have the following formulas

- the Carli price index (Carli, 1804)

$$P_C = \frac{1}{N}\sum_{i=1}^{N}\frac{p_i^t}{p_i^0}, \tag{1}$$

- the Jevons price index (Jevons, 1865)

$$P_J = \prod_{i=1}^{N}\left(\frac{p_i^t}{p_i^0}\right)^{\frac{1}{N}}, \tag{2}$$

where the time moment $\tau = 0$ we consider as the basis, N is the number of items observed at times 0 and t, p_i^τ denotes the price of the i-th item at time τ.

The Carli index is an arithmetic mean of price relatives (partial indexes), whereas the Jevons index is a geometric mean. As a consequence, these indices satisfy the classic inequality for arithmetic and geometric means

$$P_J \le P_C. \tag{3}$$

The difference between the Carli index and the Jevons index is bounded from below by the variance of the price relatives (Hardy at al., 1934):

$$P_C - P_J \ge D^2\left(\frac{p_i^t}{p_i^0}\right), \tag{4}$$

and thus the analogical inequality holds for their expected values. From the axiomatic price index theory the Jevons index seems to be better; it satisfies main tests (axioms), whereas the Carli index does not satisfy the time reversal test and circularity (Balk, 1995; Levell, 2015).

3 Comparison of expected values of sample indices

Let us treat price processes as stochastic ones and let the Carli index (1) and the Jevons index (2) be sample indices, where N denotes the sample size. Let us denote by P_C^0 and P_J^0 the following, unknown (a priori) values:

$$P_C^0 = \frac{1}{N}\sum_{i=1}^{N} E(\frac{p_i^t}{p_i^0}). \tag{5}$$

$$P_J^0 = \prod_{i=1}^{N} E(\frac{p_i^t}{p_i^0})^{\frac{1}{N}}. \tag{6}$$

In this section we are going to compare expected values of sample Jevons and Carli price indices in the continuous time stochastic model. We assume that prices are described by the geometric Brownian (Wiener) motion (also known as the exponential Brownian motion), i.e.

$$dp_i^t = \alpha_i p_i^t dt + \beta_i p_i^t dW_i^t, \tag{7}$$

where percentage drifts α_i and percentage volatilities β_i are constant, $\{W_i^t : 0 \le t < \infty\}$ are independent Wiener processes. The solution for the stochastic differential (7) is as follows (Oksendal, 2002)

$$p_i^t = p_i^0 \exp((\alpha_i - \frac{\beta_i^2}{2})t + \beta_i W_i^t), \tag{8}$$

and since we assume that all initial prices satisfy $p_i^0 = 1$ we obtain the following expected values of the price relatives P_i^t, where $i = 1,2,...,N$ (Jakubowski et al., 2003)

$$E(P_i^t) = E(\frac{p_i^t}{p_i^0}) = \exp(\alpha_i t). \tag{9}$$

Obviously, from (3) or (4) we know that $E(P_c) \ge E(P_J)$. Let us notice that it holds

$$P_J = \prod_{i=1}^{N} (P_i^t)^{\frac{1}{N}} = \prod_{i=1}^{N} \exp(\frac{\alpha_i - \beta_i^2/2}{N} t + \frac{\beta_i}{N} W_i^t), \tag{10}$$

or equivalently

$$P_J = \exp((\sum_{i=1}^{N} \frac{\alpha_i}{N} - \frac{1}{2}\sum_{i=1}^{N} (\frac{\beta_i}{N})^2)t + \sum_{i=1}^{N} \frac{\beta_i}{N} W_i^t) \exp(\frac{1}{2}(\sum_{i=1}^{N} (\frac{\beta_i}{N})^2 - \sum_{i=1}^{N} \frac{\beta_i^2}{N})t). \tag{11}$$

Let us denote by $vol(\beta_1, \beta_2, ..., \beta_N)$ a component connected with price volatilities, i.e.

$$vol(t, \beta_1, \beta_2, ..., \beta_N) = \exp(\frac{1}{2}(\sum_{i=1}^{N} (\frac{\beta_i}{N})^2 - \sum_{i=1}^{N} \frac{\beta_i^2}{N})t) = \exp(\frac{1-N}{2N^2}\sum_{i=1}^{N} \beta_i^2 t). \tag{12}$$

From (11) and (12) and under the assumption about independent Wiener processes we can write an expected value of the Jevons price index as follows

$$E(P_J) = vol(t, \beta_1, \beta_2, ..., \beta_N) \prod_{i=1}^{N} E[\exp((\frac{\alpha_i}{N} - \frac{1}{2}(\frac{\beta_i}{N})^2)t + \frac{\beta_i}{N} W_i^t). \quad (13)$$

In analogous way to (8) and (9) we obtain

$$E(P_J) = vol(t, \beta_1, \beta_2, ..., \beta_N) \prod_{i=1}^{N} \exp(\frac{\alpha_i}{N} t) = vol(t, \beta_1, \beta_2, ..., \beta_N) \prod_{i=1}^{N} (E(P_i^t)^{\frac{1}{N}}. \quad (14)$$

In the case of the Carli price index, from (1) and (9) we get

$$E(P_C) = \frac{1}{N} E(\sum_{i=1}^{N} \frac{p_i^t}{p_i^0}) = \frac{1}{N} \sum_{i=1}^{N} E(P_i^t) = \frac{1}{N} \sum_{i=1}^{N} \exp(\alpha_i t). \quad (15)$$

From (14) and (15) we have that $E(P_C) = P_C^0$ and $E(P_J) \neq P_J^0$, where

$$E(P_J - P_J^0) = (vol(t, \beta_1, \beta_2, ..., \beta_N) - 1) P_J^0. \quad (16)$$

Analogously to (3) we have

$$P_C^0 \geq P_J^0, \quad (17)$$

and thus

$$E(P_C - P_J) = P_C^0 - vol(t, \beta_1, \beta_2, ..., \beta_N) P_J^0 \geq P_J^0 (1 - vol(t, \beta_1, \beta_2, ..., \beta_N)). \quad (18)$$

Obviously, if price processes are deterministic, i.e. if $\beta_1 = \beta_2 = ... = \beta_N = 0$, we get trivial conclusion that $vol(t, \beta_1, \beta_2, ..., \beta_N) = 1$ and thus

$$E(P_C - P_J) = P_C - P_J = P_C^0 - P_J^0. \quad (19)$$

The main conclusion from the relation described in (18) is that the difference between expected values of the sample Carli index and the sample Jevons index depend on number of items, volatilities of price relatives and values of arithmetic and geometric means of expected values of sample price relatives. In particular, the inequality in (18) states that the higher the inflation is, the bigger differences between expected values of the Carli index and the Jevons index appear.

Remark

The estimation of variances of Carli and Jevons indices in the stochastic model would exceed the limited size of this paper and thus it is omitted. However, we calculate these statistics numerically in the simulation study (see Section 4).

4 Simulation study

Let us take into consideration a group of $N = 4$ items, the time horizon of observations $T = 1$ and the following parameters of price processes described in (7).

Case 1 (small volatilities)

$\alpha_1 = 0.02$, $\beta_1 = 0.05$, $\alpha_2 = 0.03$, $\beta_2 = 0.06$, $\alpha_3 = 0.05$, $\beta_3 = 0.015$,

$\alpha_4 = 0.06$, $\beta_4 = 0.01$.

Case 2 (medium volatilities)

$\alpha_1 = 0.02$, $\beta_1 = 0.25$, $\alpha_2 = 0.03$, $\beta_2 = 0.26$, $\alpha_3 = 0.05$, $\beta_3 = 0.15$,

$\alpha_4 = 0.06$, $\beta_4 = 0.1$.

Case 3 (big volatilities)

$\alpha_1 = 0.02$, $\beta_1 = 0.85$, $\alpha_2 = 0.03$, $\beta_2 = 0.76$, $\alpha_3 = 0.05$, $\beta_3 = 0.75$,

$\alpha_4 = 0.06$, $\beta_4 = 0.6$.

Without loss of generality we assume that $p_i^0 = 1$ for each $i \in \{1,2,3,4\}$. Some realizations of price relatives from Case 1 (for $t \in [0,1]$) are presented in Fig.1. Fig.2 presents $K = 10000$ sample realizations of each P_i^1 in Case 1. Basic statistics for generated K values of Jevons and Carli indices depending on the considered case are presented in Table 1 - 3.

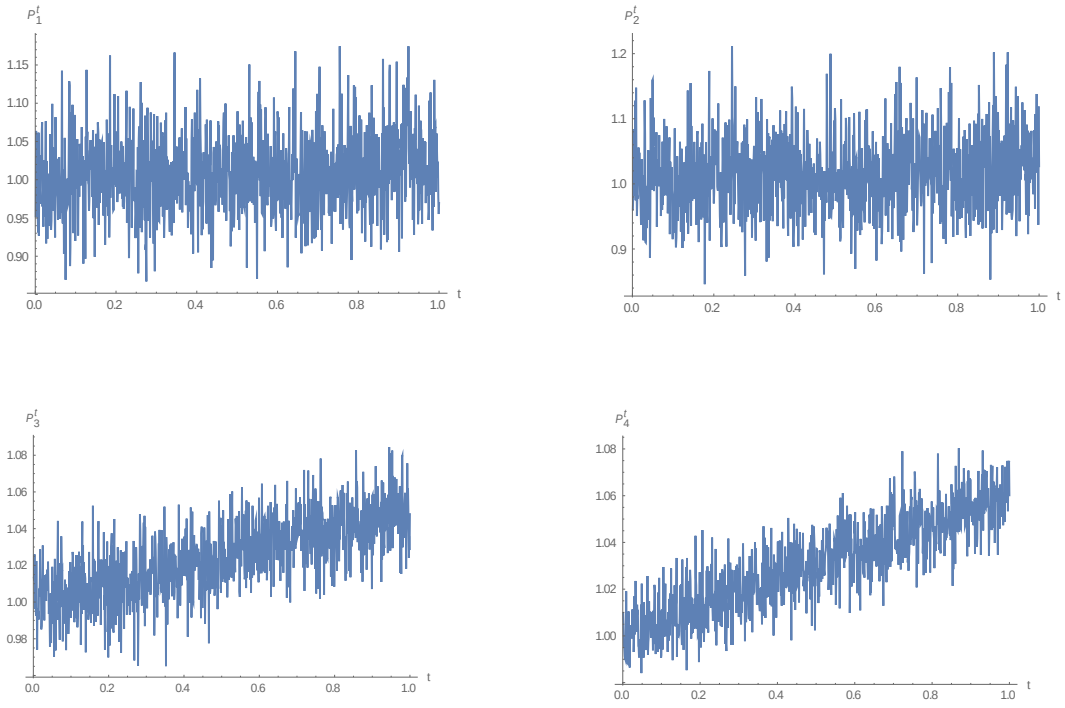


Fig. 1. Some realizations of price relatives processes for Case 1 and $t \in [0,1]$.

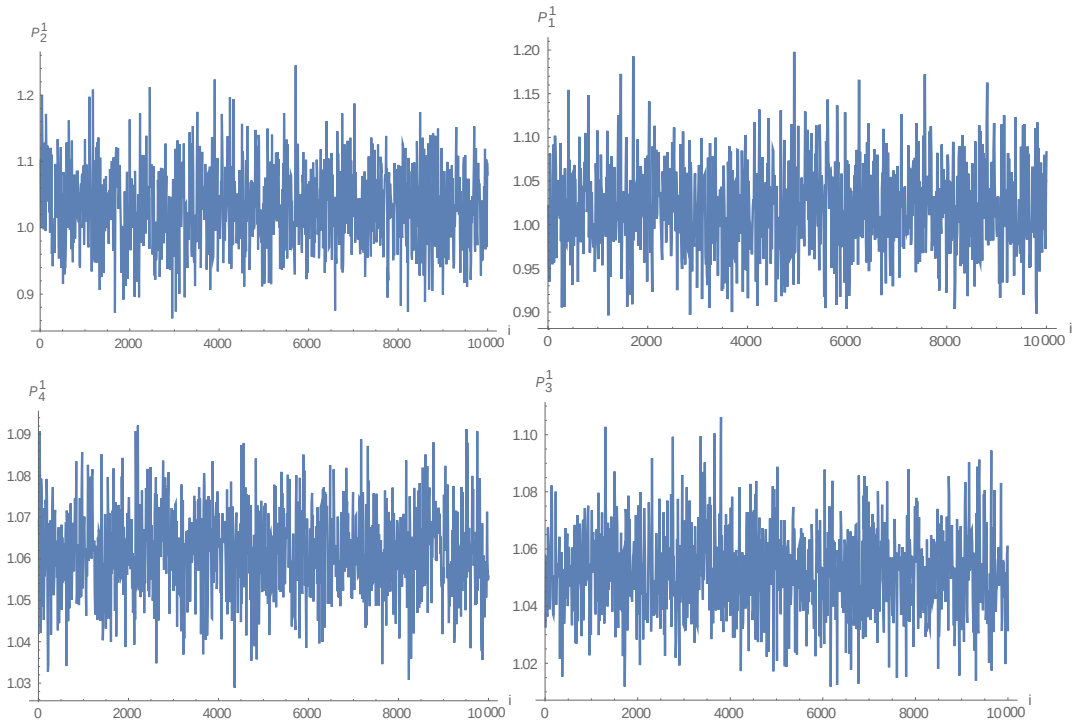


Fig. 2. Some realizations of price relatives processes for Case 1 and $t = T = 1$.

Table 1. Basic statistics for generated Jevons and Carli price indices (Case 1).

Basic statistics	Jevons index	Carli index
Mean	1.0411	1.0412
Standard Deviation	0.0204	0.0206
Volatility coefficient	0.0196	0.0198

Table 2. Basic statistics for generated Jevons and Carli price indices (Case 2).

Basic statistics	Jevons index	Carli index
Mean	1.0265	1.0415
Standard Deviation	0.1044	0.1059
Volatility coefficient	0.1017	0.1017

Table 3. Basic statistics for generated Jevons and Carli price indices (Case 3).

Basic statistics	Jevons index	Carli index
Mean	0.8435	1.0414
Standard Deviation	0.3232	0.4567
Volatility coefficient	0.3832	0.4385

Conclusions

There are several sources of the CPI bias including the elementary index bias (White, 1999). As it was mentioned, a choice of the elementary formula does matter in final inflation calculations. There has been a general trend in replacing the Carli index with the Jevons or the Dutot formulas and most of papers recommend the Jevons index rather than the Carli index. In the paper we show some similarities and differences in practical using of these indices. First of all, in our simulation study we observe that the expected (mean) value of generated values of the Jevons formula depends strongly on price volatilities whereas the mean value of generated values of the Carli index does not react on price fluctuations. In the case of strong price fluctuations, the differences between the expectations of Jevons and Carli price indices increase. In particular, we obtain the following $vol(t,\beta_1,\beta_2,...,\beta_N)$ function values: 0.999 (Case 1 with small price volatilities), 0.984 (Case 2 with medium price volatilities) and 0.812 (Case 3 with high price volatilities). Thus, the differences between expected values of sample Jevons and Carli indices are the strongest in the Case 3. Moreover, we show analytically that the difference between expected values of the sample Carli index and the sample Jevons index depend on number of items, volatilities of price relatives and values of arithmetic and geometric means of expected values of sample price relatives. We also can observe (from the inequality in (18)) that the higher the inflation is, the bigger differences between expected values of sample Carli and Jevons indices appear. It is quite interesting that volatilities of these generated (in the simulation) indices, measured by their standard deviations and volatility coefficients, seem to be comparable although they still depend on price dispersions.

Acknowledgements

The author of the paper would like to thank the National Science Centre in Poland for financing this publication (grant no. 2017/25/B/HS4/00387).

References

Balk, M. (1995). Axiomatic Price Index Theory: A Survey. *International Statistical Review, 63*, 69-95.

Boskin, M. S. (1996). (Chair) *Advisory Commission to Study the Consumer Price Index. Towards a more accurate measure of the cost of living*. Washington DC.

Boskin, M. S., Delberger, E. R., Gordon, R. J., Griliches, Z., & Jorgenson, D. W. (1998). Consumer prices in the consumer price index and the cost of living. *Journal of Economic Perspectives*, *12*(1), 3-26.

Bureau of Labor Statistics (BLS). (2001). The experimental CPI using geometric means (CPI-U-XG). *http://www.bls.gov/cpi/cpigmtoc.htm.*

Carli, G. (1804). Del valore e della proporzione de'metalli monetati. In: *Scrittori Classici Italiani di Economia Politica*, *13*, 297-336.

Evans, B. (2012). *International comparison of the formula effect between the CPI and RPI. Office for National Statistics*. Newport.

Hardy, G. H., Littlewood, J. E, & Polya, G. (1934). Inequalities. *Cambridge: Cambridge University Press.*

Jakubowski, J., Palczewski, A., Rutkowski, M., & Stettner, Ł. (2003). *Matematyka finansowa. Instrumenty pochodne*. Warszawa: Wydawnictwa Naukowo-Techniczne.

Jevons, W. S. (1865). The variation of prices and the value of the currency since 1782. *J. Statist. Soc. Lond*., *28*, 294-320.

Levell, P. (2015). Is the Carli index flawed?: assessing the case for new retail price index RPIJ. *J. R. Statist. Soc. A*, *178*(2), 303-336.

Oksendal, K. (2002). Stochastic Differential Equations: An Introduction with Applications. *Springer.*

Silver, H., & Heravi, S. (2007). Why elementary price index number formulas differ: Evidence on price dispersion. *Journal of Econometrics*, *140* (2007), 874-883.

White, A. G. (1999). Measurement Biases in Consumer Price Indexes. *International Statistical Review*, *67*(3), 301-325.

Von der Lippe, P. (2007). *Index Theory and Price Statistics*. Peter Lang.

Application of the survival trees for estimation of the influence of determinants on probability of exit from the registered unemployment

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

Abstract

Survival trees are very useful regression tools for modelling of relations between the survival time and the vector of covariates. They are the example of the recursive binary partitioning. The aim of this method is creation of homogeneous subsets with respect to analysed response variables. Tree based methods, due to their non-parametric nature and flexibility, have become very popular in the last decades as an alternative to the traditional proportional hazard model. In the presented research, the conditional inference trees were used. It is the non-parametric class of regression trees that can be applied for all regression types. The goal of the analysis was the assessment of the influence of gender, age and education on the probability of exit from the registered unemployment. Due to the existence of censored observations, survival analysis methods were applied. The Kaplan-Meier estimator was used for estimation of the survival function of homogeneous groups in each terminal node. The splitting criterion was the statistic of the log-rank test, which is used for comparison of survival distribution for various groups. The two most numerous forms of de-registration were considered – finding a job and removal from the register for reasons attributable to the unemployed person. These forms helped to distinguish the subgroups of persons with the highest and lowest probability of de-registration to work, and resignation from the mediation of the labour office.

Keywords: *survival trees, Kaplan-Meier estimator, log-rank test, registered unemployment*
JEL Classification: C38, C41, J64
DOI: 10.14659/SEMF.2018.01.03

1 Introduction

Labour market analyses generally focus on the persons that leave the unemployment by finding a job. The statistical data collected in the poviat labour offices is the rich source of information about other causes of de-registration. There are about fifty of them: retirement, receiving pension, going abroad for period longer than 30 days, change of residence, death, granting pre-retirement allowance and many others. Finding a job is the most frequent cause of de-registration. Removal due to lack of readiness to work is the second one. It happens in

[1] Corresponding author: University of Szczecin/Institute of Econometrics and Statistics, Department of Operations Research and Applied Mathematics in Economics, ul. Mickiewicza 64, 71-101 Szczecin, beatus@wneiz.pl.
[2] University of Szczecin/Institute of Econometrics and Statistics, Department of Operations Research and Applied Mathematics in Economics, ul. Mickiewicza 64, 71-101 Szczecin, krzysztof.dmytrow@usz.edu.pl.

case of refusal of accepting proposed employment or absence in the labour office in due time. In years 2008-2014 removal constituted from 27% do 32% of all de-registrations in Poland (Fig. 2). Many unemployed people do not inform the labour office about finding a job thinking that it is their employer's responsibility. Formally, they should do it within the week since finding a job. The labour office pays premiums for the unemployed before removing them from the register. The trial to decrease the scale of this occurrence is punishing the unemployed people that were removed because of their fault. The punishment is the difficulty of reclaiming the status of the unemployed person, hence the right to the health insurance and benefit.

The goal of the research was the assessment of the influence of gender, age and education on the probability of exit from the registered unemployment. The two most numerous forms of de-registration were considered – finding a job and removal from the register for reasons attributable to the unemployed person. Due to the existence of censored observations, survival analysis methods were applied.

Situation on the Polish labour market in last years has improved. It can be observed by decreasing unemployment rate (Fig. 1). In the era of globalisation it is connected to the general situation on the world market and particularly in the European Union. As the analyses show, processes on the Polish labour market are similar to these on the Slovak and Hungarian markets (Hadaś-Dyduch et al., 2016). Unemployment is influenced by the social policy of the country, realised among the other things, by the labour offices. Activisation activities addressed to groups of people threatened by the unemployment influence the labour market positively. However, expanded system of benefits for the unemployed people may increase the time of job searching (Bieszk-Stolorz and Markowicz, 2015).

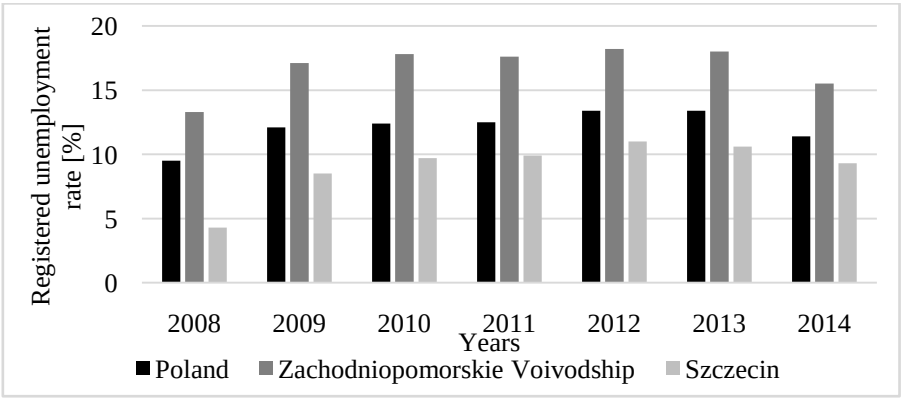


Fig. 1. Registered unemployment rate in Poland, Zachodniopomorskie Voivodship and Szczecin in years 2008-2014.

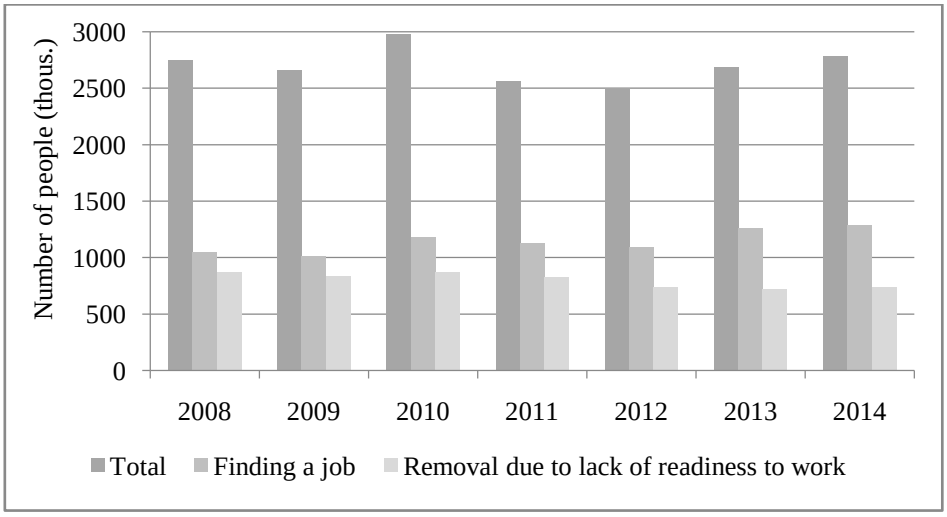


Fig. 2. Number of de-registrations from labour offices in Poland in years 2008-2014.

2 Data used in the research

In the study, anonymous individual data obtained from the Poviat Labour Office (PUP) in Szczecin (Poland) was used. The study covered 22078 unemployed individuals registered in 2013 and observed by the end of 2014. The event that terminated each observation was the moment of de-registration form the labour office list. Time *T* since the moment of registration until the moment of de-registration because of specific cause was analysed.

The two types of events terminating observations were considered: finding a job and removal due to lack of readiness to work. De-registrations due to other causes and observations that did not end with event before the end of 2014 (1856 observations) were considered as censored data. In the analysed period almost 44% of registered unemployed people found a job. They constituted the most numerous group. Slightly smaller group (almost 41%) were people that were absent in the labour office in due time or did not accept the job offer (removal). The size of each group is presented in Table 1.

The job-finding (Job) consists of three main subgroups: finding a job or another form of employment, taking up a government subsidised form of employment and entrepreneurial activity. The Removal from register category includes the unemployed individual's reluctance to cooperate with the labour office and have been removed from the register through their own fault or on their own request. The remaining causes of de-registration are less numerous and, as previous research showed, each of them had a marginal effect on the probability of de-registration (Bieszk-Stolorz, 2017).

Table 1. Structure of analysed unemployed people.

Group	Total	of which	
		Job	Removal
Total	22078	9678	8965
Gender			
Women (*K* or 1)	9770	4836	3264
Men (*M* or 0)	12308	4842	5701
Age			
18-24 (W_1)	4148	1506	2257
25-34 (W_2)	7356	3614	2966
35-44 (W_3)	4259	1869	1734
45-54 (W_4)	3497	1642	1214
55-59 (W_5)	2185	837	629
60-64 (W_6)	633	210	165
Education			
At most lower secondary (S_1)	5123	1410	2932
Basic vocational (S_2)	5016	1968	2220
General secondary (S_3)	2859	1226	1223
Vocational secondary (S_4)	4086	1943	1415
Higher (S_5)	4994	3131	1175

3 Research methodology

Survival analysis, commonly applied in demography and medicine, is more and more often applied in the social and economic phenomena, e.g. in labour market analysis. In this manner, the economic activity of the population (Landmesser, 2013) or duration of unemployment (Bieszk-Stolorz, 2013; Bieszk-Stolorz and Markowicz, 2012) can be analysed. The duration time at given state (duration of the company activity, duration of unemployment, duration of debt payment) is the random variable *T*. The basis for such analysis is the survival function, defined as follows:

$$S(t) = P(t > T) = 1 - F(t) \tag{1}$$

where *T* represents duration and *F*(*t*) – the cumulative distribution function of the random variable *T*. The most widely used estimator of the survival function is the Kaplan-Meier estimator (Kaplan and Meier, 1958):

$$\hat{S}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{n_j}\right) \quad (2)$$

where d_j is the number of events at the moment t_j and n_j is the number of individuals at risk by the moment t_j. The survival function $S(t)$ specifies the probability that the event will not occur at least by the time t. Depending on the defined event, sometimes it is more convenient to analyse the cumulative distribution function $F(t)$ which expresses the probability for the event to occur at most by the time t. When the event is defined as de-registering then the survival function estimator specifies the probability of remaining in the labour office register, while the estimator of the cumulative distribution function designates the probability of de-registering. In this case, d_j was the number of de-registrations due to particular cause at the moment t_j (finding a job or removal due to lack of readiness to work). In case of de-registration to work it is desirable that the survival curves are low-lying, while for the removal it is the opposite.

Analysed community can be divided into groups with respect to specific features, survival function for each group and the significance of differences between these functions can be estimated. Because distributions of the duration were unknown, the non-parametric tests, based on the rank order, were used. Unfortunately, there are no commonly accepted methods of selection of test at given situation. Most of them yield reliable results only for large samples, while effectiveness of these tests for small samples is less recognised. For comparison of two survival curves, the log-rank test is commonly used (Kleinbaum and Klein, 2005). It is used for verification of hypothesis H_0: $S_1(t) = S_2(t)$ stating that the survival curves for both groups are the same versus the hypothesis H_1: $S_1(t) \neq S_2(t)$ stating that they are not the same. Assuming that the null hypothesis is true, the test statistics is chi-square distributed with one degree of freedom. This test has the highest power, when the difference between the hazard functions for single subgroups is constant in time (Landmesser, 2013). Initial analysis with use of the function $\ln(-\ln S(t))$ and certain limitations resulting from assumptions for other tests confirmed the validity of application of the log-rank test in the research. In order to divide the analysed community into homogeneous groups with respect to shape of the survival curves, the survival trees are very useful tools. They are the subgroup of the so-called conditional inference trees. The idea of binary partitioning is used in construction of these trees. They has recently become popular in comparison with other methods (for example the discriminant analysis) because less assumptions are required and they can deal with various data structures (Al-Nachawati et al., 2010; Bou-Hamad et al., 2009;. Zhou and McArdle, 2015; LeBlanc and Crowley, 1993). Construction of any tree is

connected with two aspects (Cappelli and Zhang, 2007): partitioning the data, or tree growing and pruning the tree in order to make it shorter and increase the clarity of results.

Data partitioning is connected with separation of homogeneous with respect to analysed covariates groups. Splitting criterion can be based on the impurity measure or on the value of the log-rank test statistics. Partitioning occurs until the stopping criterion is reached. The necessity of the tree pruning is caused by the fact that data partitioning makes the tree very large and the overfitting occurs. Generally, the partitioning stops if the empirical significance level of the log-rank test statistics exceeds assumed value. However, for large sample size this approach is not always effective. The other criterion is defining the minimum group size, for which the partitioning may occur or the minimum group size in the terminal node. Also, the maximum tree depth may by defined (Mudunuru, 2016). Presented in the article survival trees were constructed by using the `ctree` function in the `partykit` package in R language. They were the conditional inference trees. Every observed unemployed person was described by the following triplet: $\{y_i, \delta_i, \mathbf{x}_i\}$, where y_i was the duration of registration, δ_i indicated whether the observation is censored or not (1 – uncensored, 0 – censored) and the $\mathbf{x}_i$ vector contained three covariates: gender, age and education. The duration of registration was the numerical continuous variable, censoring was the dichotomic variable. The covariates were the categorical variables. In the `ctree_control` function, two default parameters were changed: the `mincriterion` was set at 0.99 in order to set the significance level at 0.01 and by means of the `maxdepth` parameter the tree was pruned at the third level.

4 Analysis of time to de-registration to work or to removal

The analysis was conducted in two stages. The first one consisted in selection of homogeneous groups of unemployed people with respect to the probability of exit from unemployment to work. The Fig. 3 shows that gender was not the significant splitting criterion. In the first step the unemployed people were divided with respect to education – into persons with higher education and the remaining ones. In the next step, persons with at most secondary education were further divided into persons with at most lower secondary and secondary. The unemployed persons were further divided with respect to age. Finally, seven terminal nodes were obtained, their specifications are presented in Table 2. The lowest probability of exit from unemployment to work was observed for persons at the age of 60 and older with at most lower secondary education. On the other hand, the highest probability was observed for young persons (up to 35 years old) with higher education. The distribution of

time to de-registration to work is presented on the Fig. 4. It was extremely positively skewed. The largest number of people (1999 – 20.65%) found a job within first month since registration.

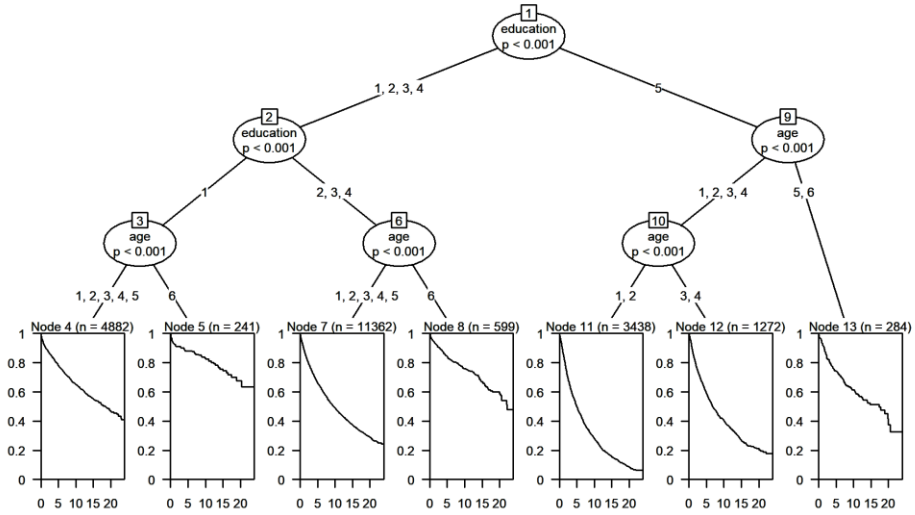


Fig. 3. Survival tree for de-registration to work.

Table 2. Homogeneous groups of unemployed persons – de-registration to work.

Specification	Number of the terminal node						
	4	**5**	**7**	**8**	**11**	**12**	**13**
education	S_1	S_1	S_2-S_4	S_2-S_4	S_5	S_5	S_5
age	W_1-W_5	W_6	W_1-W_5	W_6	W_1, W_2	W_3, W_4	W_5, W_6

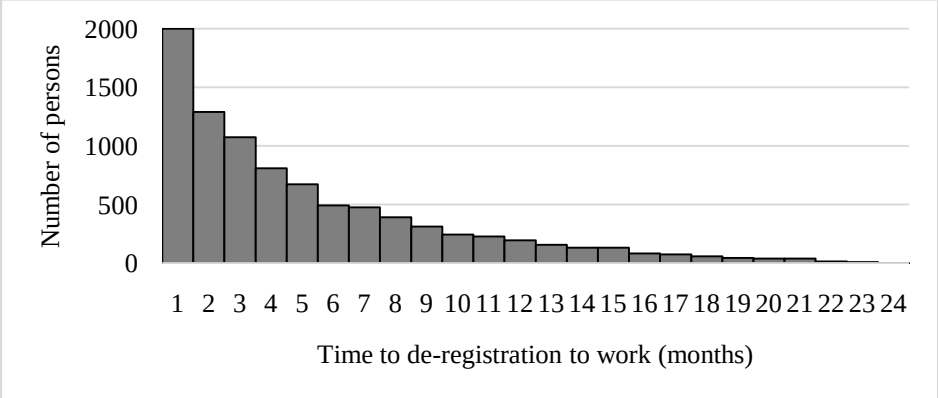


Fig. 4. Distribution of time to de-registration to work.

The second stage of the research consisted in selection of homogeneous groups of unemployed persons with respect to the probability of removal from the register. As seen on

the Fig. 5, unemployed persons removed from the register were in the first step divided with respect to age – into persons at the age up to 24 years and older. In the second step, the youngest persons were further divided with respect to education and gender. On the other hand, the oldest persons were not further divided according to age or education. It can be observed that the largest number of unemployed persons was removed from the register within the first month (values of the survival curve decreased most rapidly) and the second largest – in the fourth month. Finally, eight terminal nodes were obtained, their specifications are presented in Table 3. The lowest probability of removal from the register was observed at the age of 60 and more and the highest – the youngest males with at most lower secondary education.

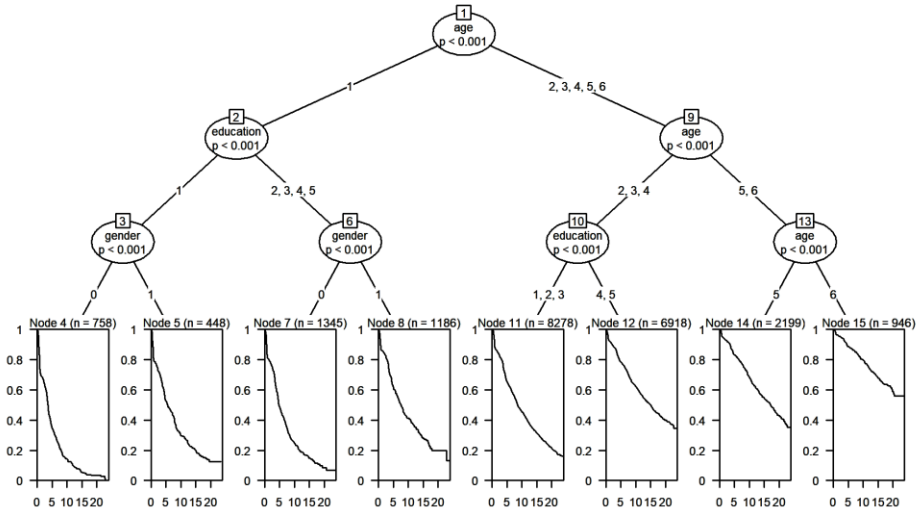


Fig. 5. Survival tree for removal from the register.

Table 3. Homogeneous groups of unemployed persons – removal from the register.

Specification	Number of the terminal node							
	4	**5**	**7**	**8**	**11**	**12**	**14**	**15**
education	S_1	S_1	S_2-S_5	S_2-S_5	S_1-S_3	S_4, S_5	S_1-S_5	S_1-S_5
age	W_1	W_1	W_1	W_1	W_2-W_4	W_2-W_4	W_5	W_6
gender	*M*	*K*	*M*	*K*	*K*, *M*	*K*, *M*	*K*, *M*	*K*, *M*

The distribution of time to removal from the register is presented on Fig. 6. It is worth noting that this distribution differs from the distribution of time to de-registration to work. It is bimodal – the largest probability of removal (2290 persons – 25.5%) was within the first

month since registration. The second largest probability of removal was in the fourth month (1148 persons – 12.8%).

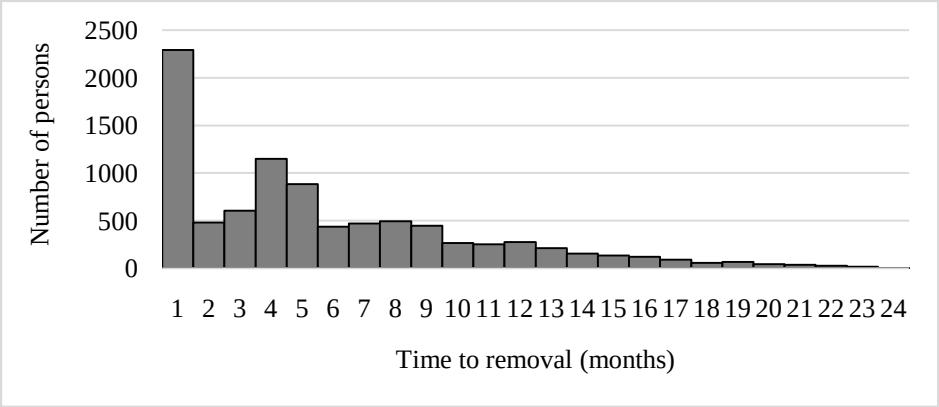


Fig. 6. Distribution of time to removal from the register.

Conclusions

On the basis on conducted analysis, the influence of gender, education and age on the probability of finding a job or removal due to lack of readiness to work was assessed. Results obtained in this research confirm the results obtained by means of other methods of survival analysis referring to the registered unemployment in Szczecin in years 2007-2011 (Bieszk-Stolorz, 2013). Generally, gender hardly influenced the probability of finding a job. On the other hand, education and age of the unemployed people were the strong determinants. In the analysed period, young people with higher education had the largest probability of finding a job. One of goals of labour market analysis is indication of groups particularly threatened by the unemployment and development of activisation programmes for them. Many of these programmes are directed to the young people. However, young and poorly educated persons were more quickly removed from the register. It indicates a lack of interests in the labour office offers such, as: participation in traineeships, trainings, providing additional equipment of workstations, etc. It is worth noting that gender of the unemployed person was important determinant of removal from the register. Males were much more often removed from the register than females.

References

Al-Nachawati, H., Ismail, M. & Almohisen, A. (2010). Tree-structured analysis of survival data and its application using SAS software. *Journal of King Saud University (Science)*, *22*, 251-255.

Bieszk-Stolorz, B. (2013). *Analiza historii zdarzeń w badaniu bezrobocia*. Volumina.pl Daniel Krzanowski, Szczecin.

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. & Markowicz, I. (2012). *Modele regresji Coxa w analizie bezrobocia*. CeDeWu, Warszawa.

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.

Bou-Hamad, I., Larocque, D., Ben-Ameur, H., Mâsse, L. C., Vitaro, F. & Tremblay, R. E. (2009). Discrete-time survival trees. *Canadian Journal of Statistics - Revue Canadienne De Statistique*, *37*(1), 17-32.

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.

Hadaś-Dyduch, M., Pietrzak, M. B. & Balcerzak, A. P. (2016). Wavelet Analysis of Unemployment Rate in Visegrad Countries. In: Kliestik, T. (ed.). *16th International Scientific Conference on Globalization and its Socio-Economic Consequences*. Zilina: University of Zilina, 595-602.

Kaplan, E. L. & Meier, P. (1958). Non-parametric estimation from incomplete observations. *Journal of American Statistical Association*, *53*, 457-481.

Kleinbaum, D. & Klein, M. (2005). *Survival Analysis. A Self-Learning Text*. Springer, New York.

Landmesser J. (2013). *Wykorzystanie metod analizy czasu trwania do badania aktywności ekonomicznej ludności w Polsce*. Wydawnictwo SGGW, Warszawa.

LeBlanc M. & Crowley J. (1993). Survival trees by goodness of split. *Journal of the American Statistical Association; 88*(422), 457-467.

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 (10.12.2017).

Zhou, Y. & McArdle, J. J. (2015). Rationale and Applications of Survival Tree and Survival Ensemble Methods. *Psychometrika*, *80*(3), 811-833.

Multivariate statistical analysis of environmental data

Justyna Brzezińska[1], Aneta Rybicka[2], Marcin Pełka[3]

Abstract

One of the characteristics of environmental data is that many of them are mostly described by complex and large number of variables. To understand this phenomena it is necessary to analyze the relationship and association between them. In this paper we apply multivariate statistical methods for the analysis of environmental problems. The main aim of the paper is to present an application of the linear ordering with multidimensional scaling for results visualization in the environmental data (green growth) analysis. The main contribution of this paper is the empirical part of this paper will that presents the application of linear ordering several multivariate methods and graphical presentation using modern and advanced visualizing tools based on datasets and reports from the Organization for Economic Cooperation and Development (OECD). Presented analysis may be used in all types of environmental practice and real life solutions. All calculations will be conducted done in R software using.

Keywords: *environmental data, multivariate statistical analysis, R software*
JEL Classification: C01, C38, C39
AMS Classification: 62H30, 62H86
DOI: 10.14659/SEMF.2018.01.04

1 Introduction

Most environmental data involve a large degree of complexity and uncertainty. Environmental Data Analysis is created to provide modern quantitative tools and techniques designed specifically to meet the needs of environmental sciences and related fields. Statistics is an indispensable means of environmental research. It is used to analyze and to interpret the increasing flood of vast data from environmental areas, which are often of heterogeneous nature and show high variability. Many important results and statements concerning the environment are based on statistical investigations, such as changes of the ozone layer, climate changes, global warming, air quality etc. But also less spectacular results about the influence of various human activities on certain environmental parameters which are not

[1] University of Economics in Katowice, Faculty of Management, Department of Economic and Financial Analysis e-mail: justyna.brzezinska@ue.katowice.pl.
[2] Wroclaw University of Economics, Faculty of Economics, Management and Tourism, Department of Department of Econometrics and Computer Science e-mail: aneta.rybicka@ue.wroc.pl.
[3] Wroclaw University of Economics, Faculty of Economics, Management and Tourism, Department of Department of Econometrics and Computer Science e-mail: marcin.pelka@ue.wroc.pl

obvious at first glance and superimposed by considerable random variations are important findings of statistical analysis.

In official statistics environmental monitoring has become a serious tool for political consulting and economic practice. For environmental research also statistical methods for the design and analysis of experiments play an important role. A specific scientific discipline of environmental statistics does not exist. The whole statistical methodology may be used in investigating environmental questions and a wide range of statistical methods can be applied. Several statistical methods may be applied for the data analysis, such as: time series analysis (see for example: Fu and Weng 2016; Chaudhary et. al. 2015; Proulx et. al. 2015), spatial analysis (see for example: Dale, Fortin 2014; May et. al. 2017) parametric and nonparametric regression analysis (see for example: Rivest et. al. 2016; Cade and Noon 2003; Xu et. al. 2016), exploratory data analysis (see for example: Seifert et. al. 2014), multivariate statistical analysis (see for example: Šmilauer and Lepš 2014).

In environmental research quite often the measurement, collection, storage, processing and analysis of data are not carried out by the same institution or researcher. Only in exceptional cases there is one person who knows about all the details of collection and analysis of the data, who knows about the scientific environmental background and at the same time also about the mathematical procedure and algorithms for the statistical evaluation and the presentation of the results. There is a great complexity in environmental statistics. Heterogeneous data from different sources and collection principles are analyzed simultaneously. There are dependencies between the measured quantities that usually do not follow fixed laws or rules, but reveal random variability. Besides this variability in the nature of the environmental problem there is variability in time and space. And measurements in time are not repeatable. Another problem lies in the abundance of data in environmental statistics. Although modern computer technique can cope with this, problems of compatibility, standardization and data harmonization arise as obstacles. Storing of data is often connected with coding. But the description of the coding principles is sometimes insufficient and this may make it difficult to join data from different sources together. There are several applications of statistical analysis of environmental data (see for example: Fu and Weng, 2016; Chaudhary et. al., 2015; Proulx et. al., 2015; Dale and Fortin, 2014; May et. al., 2017; Rivest et. al., 2016; Cade and Noon, 2003; Xu et. al., 2016; Seifert et. al., 2014; Šmilauer and Lepš, 2014; Crawford et. al., 2018; Ver Hoef, 2018).

The main aim of the paper is to present an application of the linear ordering with multidimensional scaling for results visualization in the environmental data (green growth)

analysis. The main contribution of this paper is the empirical part of this paper that presents the application of linear ordering and graphical presentation using modern and advanced visualizing tools based on datasets and reports from the Organization for Economic Cooperation and Development (OECD). Presented analysis may be used in all types of environmental practice and real life solutions. All calculations will be done in R software.

2 Environmental data analysis

In environmental statistics most data are the result of a measurement process, where the measuring instruments have a certain degree of precision and a limited range of scale. Both have to be taken into account in the analysis of the data, as well as their liability. Unintentional failure of measurement instruments and disturbances of a transmission channel may lead to false or missing values. If the deviations are big enough, the corresponding data are detected as outliers. There are statistical procedures to perform this. Supplementary etiology may even lead to a correction of the values.

Talking about environmental analysis it is crucial to talk about one of the most important area which is green growth. Due to the OECD green growth is a subset of sustainable development. It is narrower in scope, entailing an operational policy agenda that can help achieve concrete, measurable progress at the interface of the economy and the environment. It fosters the necessary conditions for innovation, investment and competition that can give rise to new sources of economic growth that are consistent with resilient ecosystems. Green growth strategies need to pay specific attention to many of the social issues and equity concerns that can arise as a direct result of greening the economy both at the national and international level. This is essential for the successful implementation of green growth policies. Strategies should be implemented in parallel with initiatives focusing on the broader social pillar of sustainable development. Green growth means fostering economic growth and development, while ensuring that natural assets continue to provide the resources and environmental services on which our well-being relies. To do this, it must catalyse investment and innovation which will underpin sustained growth and give rise to new economic opportunities. We need green growth because risks to development are rising as growth continues to erode natural capital. If left unchecked, this would mean increased water scarcity, worsening resource bottlenecks, greater pollution, climate change, and unrecoverable biodiversity loss.

Green growth is a big challenge and huge problem that deserve analysis. Multivariate statistical methods may bring solutions that may solve economic and government problems.

3 Data analysis in R software

The concept of pattern of development and the measure of development was proposed by Professor Zdzisław Hellwig in 1967 (Hellwig 1967). The general procedure in linear ordering of the data set based on pattern object (or anti-pattern object) and metric data requires to choose a complex phenomenon, that cannot be measured directly, for ordering the set of objects. The objects are described by variables. Preferential variables (stimulants, destimulants and nominants) must be identified among variables (see for example Hellwig 1981, for formal definitions of stimulants, destimulants and nominants). Nominants have to be transformed into stimulants. A pattern object and anti-pattern object are added to the data set. Data has to be normalized if the variables are measured on interval or ratio scale. The distance measure between object is calculated. Multidimensional scaling is done. The iterative procedure in the **smacof** algorithm was applied in the study (Borg and Groenen, 2005). Finally, two-dimensional space is obtained. The graphical presentation and interpretation of the results in a two-dimensional (multidimensional scaling results) and one-dimensional space (linear ordering results). In the Fig. a straight line that connects pattern and anti-pattern for the MDS results is added. This line is the so-called axis of the set. Also isoquants of development are added to the MDS plot. These isoquants are determined on the basis of the pattern object, e.g. by dividing the set of axis into four parts. The objects between isoquants present similar development level. The same level of development can be reached by objects placed in different locations on the same isoquant. Such representation expands the interpretation of the results. Finally, normalized distances d_i^+ of *i*-th object from the pattern of development are calculated as follows (Hellwig, 1981):

$$d_i^+ = \frac{\sqrt{\sum_{j=1}^{2}\left(v_{ij} - v_{+j}\right)^2}}{\sqrt{\sum_{j=1}^{2}\left(v_{i+j} - v_{-j}\right)^2}}, \quad d_i^+ \in [0;1], \tag{1}$$

where: $\sqrt{\sum_{j=1}^{2}\left(v_{ij} - v_{+j}\right)^2}$ – Euclidean distance between *i*-th object and the patter object, $\sqrt{\sum_{j=1}^{2}\left(v_{i+j} - v_{-j}\right)^2}$ – Euclidean distance between pattern and anti-pattern object.

The objects are ordered by the growing values of the distance measure (1). Linear ordering results are graphically presented on thure. The empirical study uses the statistical data representing green growth level of 33 OECD member countries in 2010. The data was

selected by using the convenience sampling. The evaluation of the green growth was done by using eight metric variables (measured on a ratio scale):

x_1 – renewable electricity generation (% of total energy produced),

x_2 – production-based CO_2 productivity (GDP per unit of energy-related CO_2 emissions),

x_3 – environmentally related government R&D budget (% of total government R&D),

x_4 – development of environment-related technologies (% of all technologies),

x_5 – population with access to improved sanitation (% total population),

x_6 – mortality from exposure to PM2.5,

x_7 – forests under sustainable management certification (% total forest area),

x_8 – municipal waste generated in kg per capita.

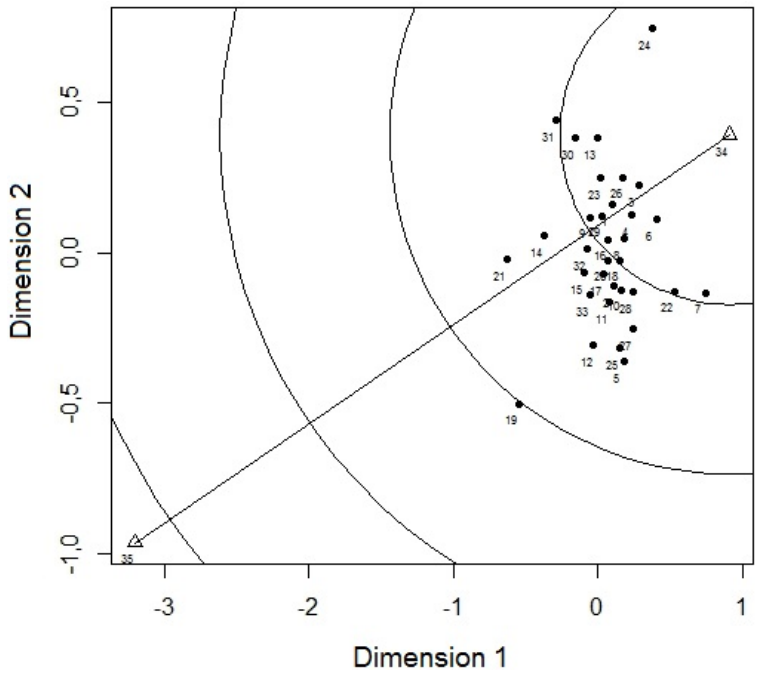


Fig. 1. Graphical presentation of multidimensional scaling results in two-dimensional space of 35 objects containing 33 countries, pattern object (34) and anti-pattern object (35) referring to the OECD countries green growth. Source: authors' compilation using R software.

Variables x_1, x_3, x_4, x_5, x_7 are stimulants, x_2, x_6 and x_8 are destimulants. The data was collected in 2010 from OECD data bank. A pattern and anti-pattern were added to the data set, so data matrix covers 35 objects described by eight variables.

Due to the fact all the variables are metric data was normalized. mdsOpt (Walesiak and Dudek, 2017) package of R software was used to find optimal multidimensional scaling procedure. The best result was obtained using positional standardization, Manhattan distance and ratio multidimensional scaling model. The Stress-1 was equal to 0.137825 and HHI spp (Hirschman-Herfindahl HHI index calculated based on stress per point) 345.2902.

The results of linear multidimensional scaling with the axis of the set, four isoquants are presented on the Fig. 1.

The ordering 33 countries from the pattern object by growing measure value (1) are shown in Table 1.

Table 1. The ordering of 35 objects regarding green growth.

Object no.	Name	Distance
34	Pattern	0.000000
7	Estonia	0.126559
6	Denmark	0.133721
24	Norway	0.147263
22	Netherlands	0.148296
3	Canada	0.150826
4	Chile	0.167198
26	Portugal	0.172901
8	Finland	0.186051
1	Austria	0.194838
28	Slovenia	0.195421
18	Korea	0.199848
23	New Zealand	0.208219
16	Italy	0.209917
10	Germany	0.210028
13	Iceland	0.210512
29	Spain	0.212897
27	Slovak Republic	0.215031
20	Luxembourg	0.216233
2	Belgium	0.219286
17	Japan	0.229008

11	Greece	0.230475
9	France	0.232427
25	Poland	0.240424
5	Czech Republic	0.241395
32	Turkey	0.244379
30	Sweden	0.246673
33	United Kingdom	0.254697
15	Israel	0.256432
12	Hungary	0.270643
31	Switzerland	0.27851
14	Ireland	0.306645
21	Mexico	0.367897
19	Latvia	0.395211
35	Anti-pattern	1.000000

Distances from the pattern object sorted by growing values of measure (1) are presented on the Fig. 2.

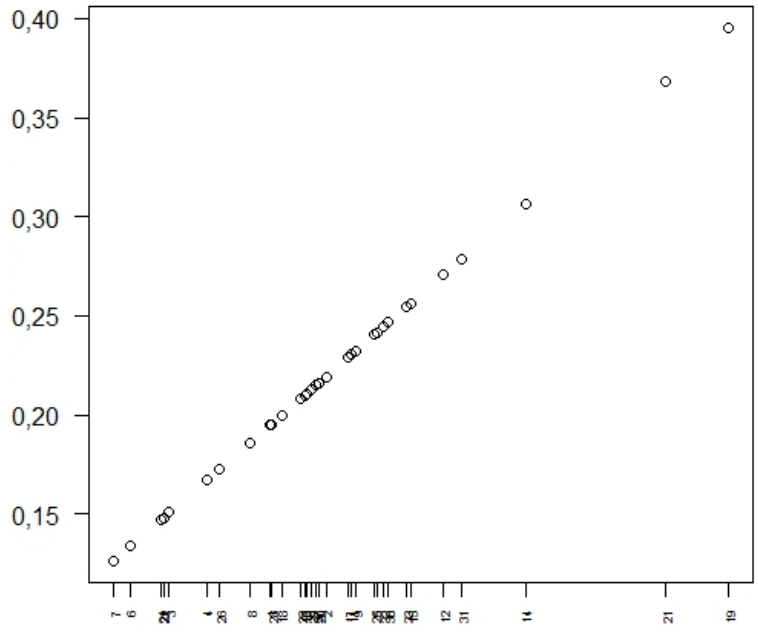


Fig. 2. Graphical presentation of linear ordering of 33 OECD countries referring to their green growth value. Source: authors' compilation using R software.

The top five countries (the best one) when considering values of the measure (1) are Estonia, Denmark, Norway and Netherlands. The five worst countries are Hungary, Switzerland, Ireland, Mexico and Latvia. Poland is the 24-th country, Czech Republic is the 25-th, Germany 15-th. When taking into consideration the location of the countries on the Fig. 1 all countries are within the first three isoquants. Countries with similar level of green growth but different positions on the Fig. 1 are: Estonia, Netherlands, Finland, Chile, Canada, Denmark, Italy, Korea (below the set of the axis), Austria, Portugal, Spain, New Zealand, Iceland, France, Sweden, Norway (above the set of the axis) that are within the first isoquant, Czech Republic, Poland, Slovak Republic, Slovenia, Hungary, Greece, Germany, Belgium, Luxembourg, Japan, United Kingdom, Israel, Turkey (below the axis of the set) and Ireland, Switzerland, Mexico (above the axis of the set) within the second isoquant. In general, we can say that more countries are below the set of the axis than above it. Also in general second isoquant contains mostly post-communist, central European countries.

Conclusions

The paper presents an application of the proposal introduced by Walesiak (2016) that allows the visualization of linear ordering results for the set of objects by applying multidimensional scaling for this task. The concept of isoquants and the path of development proposed by Hellwig (1981) allows to represent objects in two dimensions. The application of the proposed methods allows to show results for more than two variables. The proposed approach was illustrated by an empirical example where green growth data for 33 OECD member countries was used. In general, we can say that almost all OCED countries lie between two first isoquants, and so called "post-communist" countries can be found below the axis of the set. The best country, when considering green growth, is Estonia, then Denmark, Norway and Netherlands. The worst country is Latvia, then Mexico, Ireland and what can be surprising – Switzerland. The aim for the future studies should be the longitudinal analysis of the green growth and its changes during last years.

References

Borg, I. & Groenen, P.J.F. (2005). *Modern Multidimensional Scaling. Theory and Applictions. 2nd Edition*. Springer Science+Business Media, New York.

Cade, B. S. & Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, *1*(8), 412-420.

Chaudhary, S., McGregor, A., Houston, D. & Chettri, N. (2015). The evolution of ecosystem services: a time series and discourse-centered analysis. *Environmental Science & Policy, 54*, 25-34.

Crawford, F. W., Ho, L. S. T., & Suchard, M. A. (2018). Computational methods for birth-death processes. *Wiley Interdisciplinary Reviews: Computational Statistics*.

Dale, M. R. & Fortin, M. J. (2014). *Spatial analysis: a guide for ecologists*. Cambridge Unversity Press.

Fu, P. & Weng, Q. (2016). A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery. *Remote Sensing of Environment, 175*, 205-214.

Hellwig, Z. (1967). *Procedure of Evaluating High-Level Manpower Data and Typology of Countries by Means of the Taxonomic Method*, COM/WS/91, Warsaw, 9 December, 1967 (unpublished UNESCO working paper).

Hellwig, Z. (1981). Wielowymiarowa analiza porównawcza i jej zastosowanie w badaniach wielocechowych obiektów gospodarczych. In: Welfe, W. (ed.), *Metody i modele ekonometryczno-matematyczne w doskonaleniu zarządzania gospodarką socjalistyczną*, PWE, Warszawa, 46-68.

May, F., Gerstner, K., McGlinn, D. J., Xiao, X. & Chase, J. M. (2017). mobsim: An R package for the simulation and measurement of biodiversity across spatial scales. *Methods in Ecology and Evolution* (doi: 10.1111/2041-210X.12986).

Proulx, R., Parrott, L., Fahrig, L. & Currie, D. J. (2015). Long time-scale recurrences in ecology: detecting relationships between climate dynamics and biodiversity along a latitudinal gradient. In: *Recurrence Quantification Analysis* (pp. 335-347). Springer, Cham.

Rivest, L. P., Duchesne, T., Nicosia, A. & Fortin, D. (2016). A general angular regression model for the analysis of data on animal movement in ecology. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 65*(3), 445-463.

Seifert, B., Ritz, M. & Csősz, S. (2014). Application of exploratory data analyses opens a new perspective in morphology-based alpha-taxonomy of eusocial organisms. *Myrmecological News 19*, 1-15.

Šmilauer, P. & Lepš, J. (2014). *Multivariate analysis of ecological data using CANOCO 5*. Cambridge University Press.

Ver Hoef, J. M., Peterson, E. E., Hooten, M. B., Hanks, E. M. & Fortin, M. J. (2018). Spatial autoregressive models for statistical inference from ecological data. *Ecological Monographs, 88*(1), 36-59.

Walesiak, M. (2016), Visualization of Linear Ordering Results for Metric Data with the Application of Multidimensional Scaling. *Ekonometria, 2*(52), 9-20.

Walesiak, M. & Dudek, A. (2017). *The mdsOpt package for R software*, http://r-project.org.

Xu, B., Luo, L. & Lin, B. (2016). A dynamic analysis of air pollution emissions in China: evidence from nonparametric additive regression models. *Ecological indicators, 63*, 346-358.

Economic effects of production fragmentation and technological transfer: the evidence from CEE countries

Nataliia Cherkas[1], Myroslava Chekh[2]

Abstract

Foreign direct investment (FDI) inflow is traditionally considered as an important factor of structural changes and productivity growth in Central and Eastern European countries (CEECs) due to transfer of technologies and active participation in global value chains (GVC). The aim of the study is to estimate the influence of technological transfer on structural changes in CEECs. An empirical analysis of the impact of FDI and other indicators of technological transfer on the export structure was performed. We consider three export groups of technology-intensive manufactures: high-, medium- and low skill and one export group of labor and resource-intensive manufactures.The analysis includes a panel framework covering seven CEECs (Croatia, Czech Republic, Hungary, Poland, Romania, Slovakia and Slovenia) over the period of 2001–2016. OLS with pooled data, panel data with fixed effects and dynamic panel-data model were used as principal methods. Our results mostly reflect the prediction of the Flying Geese Model (FGM) and GVC theory in terms of: (i) stimulating effect of FDI on high-skill and technology-intensive manufactures; (ii) significance of impact of technology transfer through technological import growth for all export sectors; (iii) important contribution of EU integration to technological development of CEECs.

Keywords: *technological transfer, export structure, production fragmentation, FDI, imports*
JEL Classification: F14, F63
DOI: 10.14659/SEMF.2018.01.05

1 Introduction

The most of theoretical approaches suggest that foreign direct investment (FDI) positively affect development and structural changes in host countries due to technology transfer through multinational corporations. However, the benefits for host countries considerably depend on their absorption capacity (Damijan, Kostevc and Rojec, 2013; Salamaga, 2013).According to the theory of endogenous growth FDI inflow is an important channel for technology transfer to host countries (Danakol et al., 2017). International business theory implies that technology is a core type of ownership advantages of foreign investorstransmitted tothe country that accepts investments (Dunning and Lundan, 2008). “Flying geese model” (FG), suggested by Akamatsu (1962) and developed by Ozawa (2007) considers the formation

[1] Corresponding author: Kyiv National Economic University, Department of International Economics, Pr. Peremogy 54/1, 03057 Kyiv, Ukraine, natsanex@gmail.com.
[2] Lviv University of Trade and Economics, Department of International Economic Relations, 10, Tuhan-Baranovskiy Str., 79008 Lviv, Ukraine, mirachekh@gmail.com.

of a new dynamic paradigm of multinational development through technology transfer to recipient country by multinational corporations.

The main features of Akamatsu's FG concept are shown on figure 1. His approach suggests four fundamental stages of the FG pattern that was developed in the historical context of the Euro-American leadership and Asia as a follower (Kojima, 2000; Li, 2017).At the first stage manufactured consumer goods are imported from advanced to less-advanced countries (started from t_1 in Panel a). Such import can lead to negative consequences for the industry of less developed country because of the substitution effect. Second stage describes increasing import from time t_1 to $t_2$and possibility of domestic production to start from t_2. Simultaneously, the host country should import capital goods (Panel b). The competition between imported and domestic consumer goods can be observed at this stage. At the third stage, the internal consumer goods industry develops into the export industry (started from t_3 in Panel a).This stage reflects a successful implementation of the catching-up process of the industry concerned along the consistent way import-production-export (M-P-E) which is the basic pattern of the FG model (Kojima, 2000). At fourth stage it is shown the decline of consumer goods exports (started from t_4 in Panel a), whereas capital goods started exporting (started from t_4 in Panel b). The export reduction ensue as a result of consumer goods production transfer to other less-developed countries (offshore production at panel a), besides it is also possible a reverse import existence (Panel a) (Widodo 2007).But, in terms of the FGM, it is difficult to clarify the catching-up process at more advanced stages of host-country development. The influence of FDI along the lines of the FGM appears mostly in industries at the lower end of the technology scale and less when it comes to industries at the upper end (Damijan and Rojec, 2004).

The recent theoretical approach of global value chains (GVC) economics (Baldwin 2012, 2016) is seen as an adjustment of the FGM to the trends of the 21st century, because globalization's 2nd unbundling means off shoring of production stages, but not industries (as in case of FGM). Damijan, Kostevc and Rojec (2013) prove the importance of GVC concept for export sophistication and growth of labor productivity in Central and Eastern Europe (CEE). Using data for industry-level and accounting for technology intensity, they demonstrated significance but heterogeneity of FDI to export restructuring in the CEECs.While Visegrad group countries managed to increase exports in high-tech industries, non-Visegrad countries couldn't change their export specialization. This points out that direction of FDI flows is crucial. In addition, the results show that export sophistication and

economic specialization caused by FDI during the last two decades in CEECs are very important for their potential and productivity growth in the long run.

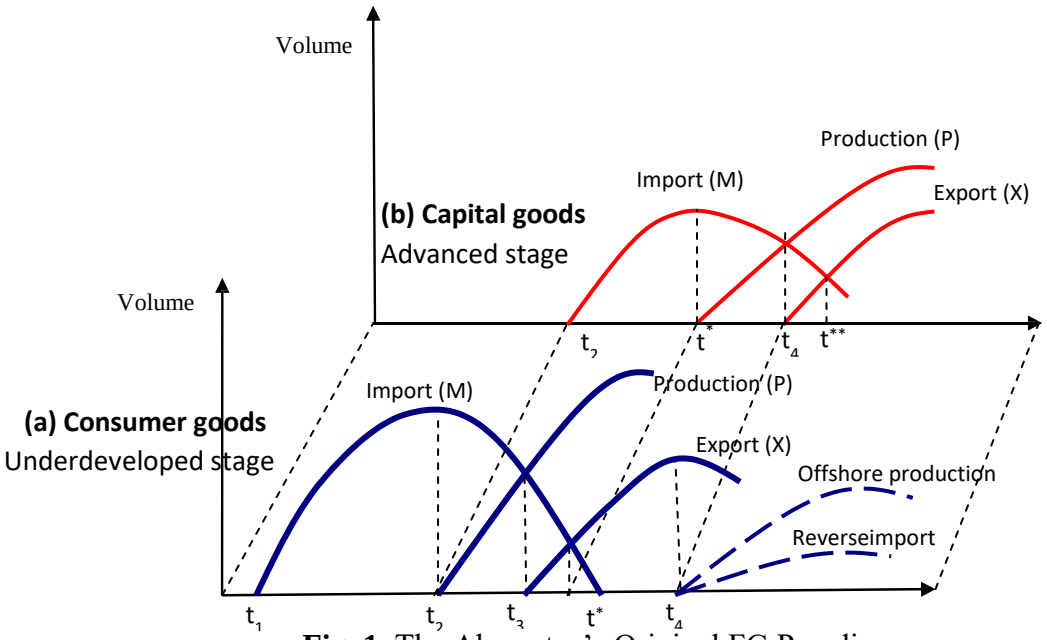


Fig. 1. The Akamatsu's Original FG Paradigm

Source: Widodo (2007).

Nevertheless, in his theoretical approach of GVC Baldwin (2016) claims that, within the "vertical specialization", which is typical for offshore stages of labor-intensive industries, transmitted from headquarter economies to "factory economies", instead of technology transfer we observe the technology lending. Nowadays, manufactured goods export is no longer a sign of economies' competitiveness, but it can simply reflect the position of the nation in global value chains. Such trend could mean the limited impact of FDI on host countries.

Therefore, the aim of the study is to estimate the effects of technological transfer on export structure with the main focus on the impact of high-tech imports, intra-industry trade and FDI on the export groups. Section 2 presents the data. Statistical methodology is provided in Section 3. The estimation of the implied panel methods are interpreted in Section 4 that is followed by the conclusions.

2 Data

The annual data for the period of 2001-2016 is used to study the impact of technological transfer on export structure in CEECs. The analysis includes a sample covering seven CEECs,

namely Croatia (HR), Czech Republic (CZ), Hungary (HU), Poland (PL), Romania (RO), Slovakia (SK) and Slovenia (SI). The data is transformed into logs in order to avoid the influence of outliers. The dependent variables are the following (in US dollars): x_ht_t– exports of high-skill and technology-intensive manufactures; x_ms_t – exports of medium-skill and technology-intensive manufactures; x_ls_t –exports of low-skill and technology-intensive manufactures;x_l_t– exports of labour-intensive and resource-intensive manufactures.

The list of explanatory variables includes: fdi_t – foreign direct investment, inflows (% of GDP);$patents_t$– patent applications of residents;m_ht_t – imports of high-skill and technology-intensive manufactures (US dollars);iit_ht_t – Intra-industry trade, sector of high-skill and technology-intensive manufactures, index; $manuf_va_t$ – manufacturing, value added (% of GDP); $service_va_t$ – services, value added (% of GDP); E_t – price level ratio of PPP (purchasing power parity) conversion factor (GDP) to market exchange rate; EU_t – dummy variable of EU membership (1 – EU member, otherwise – 0); $Crisis_t$ – dummy variable (1 – crisis period from 2009 till 2015, otherwise – 0). Foreign trade data and inward foreign direct investments are obtained from UnctadStat. The data on manufacturing and services, value added, Price level ratio of PPP and patent applications are collected from the World Development Indicators (WDI, 2017). All data are transformed into logarithmic form, except the dummy variables.

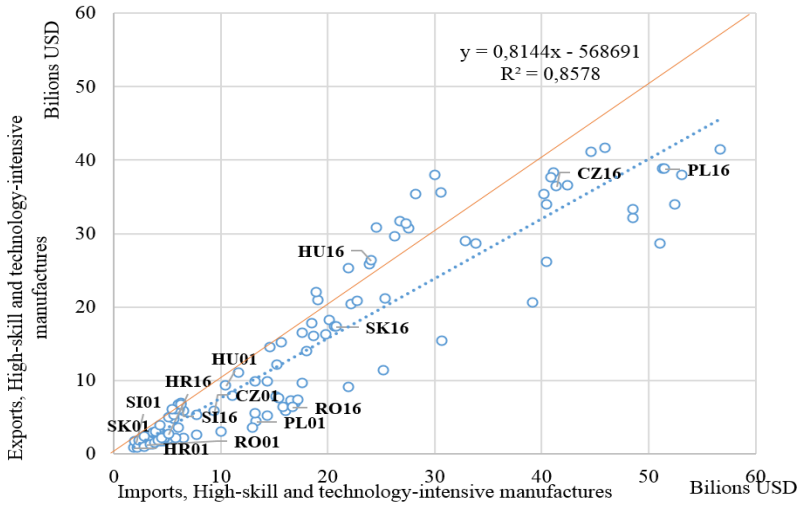


Fig. 2. High-skill and technology-intensive manufactures: exports versus imports (billions, USD), 7 CEE countries (HR, CZ, HU, PL, RO, SK, SI), 2001-2016.

Notes: Red line denotes the 45° line. The countries are marked for the first and last year.

Figure 2 presents the countries' specific positions considering high-skill and technology-intensive manufactures exports versus imports for the sample of CEE countries. These figures give some preliminary evidence about the differences in cross-country data. V-4 economies (especially Poland and Czech Republic) demonstrate increasing over time technology-intensive manufactures exports via imports, while for other CEECs the impact is less evident.

3 Methodology

The empirical analysis is conducted on panel data, which includes 112 observations (seven countries for 16 years). We apply the following techniques: pooled ordinary least squares (OLS), model with fixed effects (FE) and dynamic panel-data model. The selection of FE model in this analysis is confirmed with the Hausman test for all specifications (Hausman, 1978).

The following base model is used to study the relationships following the methodology for panel data estimation (Wooldridge, 2010):

$$\boldsymbol{x_{it}} = a_0 + a_1 fdi_{it} + a_2 patents_{it} + a_3 m_ht_{it} + a_4 iitiit_ht_{it} + a_5 manuf_va_{it} + a_6 service_va_{it} + a_7 e_{it} + a_8 EU_{it} + a_9 Crisis_{it} + \varepsilon_{it}, \quad (1)$$

where $\boldsymbol{x_{it}}$ represent the four groups of manufactured exports: x_ht_t:, x_ht_t;, x_ls_t, and$x_l_{t..}\varepsilon_{it}$ is the error term. The explanatory variables are described in detail in Section 2. Table 1 reflects the correlations results for variables, which are used in logarithms (except dummies).

Table 1. Correlation matrix for the explanatory variables.

	fdi_t	$patents_t$	m_ht_t	iit_ht_t	$manuf_va_t$	$service_va_t$	e_t	EU_t	$CRISIS_t$
fdi_t	1.000	–	–	–	–	–	–	–	–
$patents_t$	0.029	1.000	–	–	–	–	–	–	–
m_ht_t	0.091	0.610	1.000	–	–	–	–	–	–
iit_ht_t	0.046	0.215	0.387	1.000	–	–	–	–	–
$manuf_va_t$	0.088	-0.044	0.228	0.500	1.000	–	–	–	–
$service_va_t$	0.025	0.281	-0.070	0.199	-0.534	1.000	–	–	–
e_t	0.091	-0.039	0.013	0.428	-0.088	0.435	1.000	–	–
EU_t	-0.014	0.163	0.595	0.552	0.212	0.046	0.328	1.000	–
$Crisis_t$	-0.154	0.017	0.144	0.125	-0.211	0.059	0.254	0.214	1.000

4 Results and discussion

As mentioned above, the verification of the main functional dependencies is performed by the following panel methods: OLS (model 1), FE (model 2) and Dynamic panel-data (model 3). Table 2a presents the results for high-and medium-skill technology-intensive manufactures, while table 2b – low-skill technology-intensive, labor- and resource-intensive manufactures.

Our results indicate that inward FDI generally did not support the growth of exports in CEEC during the period of evaluation, which is consistent with the results of Damijan and Rojec (2004), Damijan, Kostevc and Rojec (2013). We observe the positive effect of FDI only on the exports of technology-intensive manufactures in case of OLS. The study of Kalotay (2010) revealed that FDI in CEES had the deepest impact on structural change due to the effective sectoral composition of FDI.

The impact of high-tech imports is highly significant and positive for all sectors and methods used. The results of the impact of intra-industry trade on export performance show different results within four export groups. The impact is positive for industries with higher level of technology that emphasizes the importance of foreign trade exchange. These findings go in line with the results of Jude (2016), where is indicated that the position of a sector in the supply chain is essential for capturing the technology spillovers. An increase of manufacturing value added acts as a stimulating factor for the exports of high-, medium- and low-skill and labor-intensive sectors.

The growth of value added in services contributes positively only to high- and medium-skill technology intensive manufactures. These results can be explained by the fact that services create value added mostly for high-tech industries. According to the estimation results, the strong exchange rate show a positive impact on exports for low-skill and technology intensive manufactures (models 1-3) and for high-skill manufactures (model 3), while the OLS and FE results (models 1-2) for high-technology sector indicate the opposite results. The impact of exchange rate on different export groups of Ukraine was studied in Cherkas (2013) and it was shown that the exports of high value-added goods are strongly dependent on imports but less on exchange rate.

Our data indicate that the impact of EU integration is positive for technology intensive manufactures (high- and medium-skill) and labor-intensive manufactures, but insignificant and even negative (dynamic panel data estimation) for low-skill industries. The influence of another dummy variable, characterizing the impact of global financial crisis is stimulating for high-skill and technology intensive sectors. However, positive effect disappears at the lower level of technology and even turns opposite.

Table 2a. Determinants of $x_ht_t$and x_ms_t.

Explanatory variables	High-skill and technology-intensive manufactures (x_ht_t)			Medium-skill and technology-intensive manufactures (x_ms_t)		
	(1)	(2)	(3)	(1)	(2)	(3)
Constant	-5.978***	-0.762	-1.482**	-10.368**	-16.194***	-7.566***
	(1.21)	(0.94)	(0.47)	(3.32)	(2.59)	(1.30)
Lagged dependent	–	–	0.117***	–	–	0.358***
variable	–	–	(0.02)	–	–	(0.04)
fdi_t	0.419**	0.019	-0.029	-0.797*	-0.224	-0.025
	(0.13)	(0.08)	(0.05)	(0.36)	(0.23)	(0.13)
$patents_t$	-0.022	0.056*	0.022	-0.018	-0.265***	-0.043
	(0.01)	(0.02)	(0.01)	(0.04)	(0.07)	(0.03)
m_ht_t	1.017***	1.078***	0.875***	1.052***	1.054***	0.730***
	(0.02)	(0.03)	(0.02)	(0.05)	(0.07)	(0.06)
iit_ht_t	1.491***	1.036***	1.249***	0.439*	0.653**	0.410***
	(0.07)	(0.08)	(0.05)	(0.19)	(0.22)	(0.09)
$manuf_va_t$	0.368***	-0.048	0.165**	1.901***	1.184***	1.337***
	(0.10)	(0.11)	(0.05)	(0.27)	(0.30)	(0.15)
$service_va_t$	1.114***	-0.229	0.242**	0.993	3.322***	0.640*
	(0.22)	(0.20)	(0.09)	(0.60)	(0.55)	(0.25)
e_t	-0.101*	-0.095*	0.084***	0.105	0.055	0.021
	(0.05)	(0.05)	(0.02)	(0.13)	(0.13)	(0.06)
EU_t	0.031	0.073***	0.021*	0.220**	0.040	-0.031
	(0.03)	(0.02)	(0.01)	(0.07)	(0.05)	(0.03)
$Crisis_t$	0.071**	0.036*	0.018*	-0.004	-0.031	-0.013
	(0.02)	(0.02)	(0.01)	(0.07)	(0.04)	(0.02)
R^2	0.942	0.923	–	0.903	0.901	–
F-test	210.27	183.14	–	298.97	220.94	–
Hausman χ^2 *(Prob>* χ^2*)*		74.88 (0.000)		–	72.71 (0.000)	

Notes: ***, ** and * represent the levels of significance of 1%, 5% and 10% respectively. The values of the standard errors are in parenthesis.

Table 2b. Determinants of x_ls_t and x_l_t.

Explanatory variables	Low-skill and technology-intensive manufactures (x_ls_t)			Labour-intensive and resource-intensive manufactures (x_l_t)		
	(1)	(2)	(3)	(1)	(2)	(3)
Constant	16.671***	1.871	2.292	19.618***	6.729***	7.418***
	(3.74)	(2.45)	(1.59)	(2.87)	(1.84)	(1.60)
Lagged dependent	–	–	0.110**	–	–	0.440***
variable	–	–	(0.04)	–	–	(0.06)
fdi_t	-1.602***	-0.142	-0.095	-1.003**	-0.131	0.014
	(0.41)	(0.22)	(0.16)	(0.31)	(0.17)	(0.14)
$patents_t$	0.210***	0.087	0.010	0.289***	-0.119*	0.058
	(0.04)	(0.06)	(0.04)	(0.03)	(0.05)	(0.05)
m_ht_t	0.770***	0.842***	0.782***	0.532***	0.527***	0.454***
	(0.05)	(0.07)	(0.05)	(0.04)	(0.05)	(0.06)
iit_ht_t	-0.592**	-0.391	0.196	-0.629***	0.520**	-0.473***
	(0.21)	(0.21)	(0.16)	(0.16)	(0.16)	(0.14)
$manuf_va_t$	0.128	0.856**	0.921***	0.158	0.873***	0.132
	(0.30)	(0.29)	(0.15)	(0.23)	(0.22)	(0.18)
$service_va_t$	-3.688***	-0.802	-0.931**	-3.675***	-0.337	-1.700***
	(0.68)	(0.52)	(0.34)	(0.52)	(0.39)	(0.31)
e_t	1.026***	0.509***	0.612***	0.051	0.078	-0.199**
	(0.15)	(0.12)	(0.07)	(0.12)	(0.09)	(0.07)
EU_t	0.140	0.041	-0.121***	0.247***	-0.044	-0.011
	(0.08)	(0.05)	(0.04)	(0.06)	(0.04)	(0.03)
$Crisis_t$	-0.154*	-0.011	-0.029	-0.096	-0.040	-0.048*
	(0.08)	(0.04)	(0.02)	(0.06)	(0.03)	(0.02)
R^2	0.908	0.871	–	0.926	0.568	–
F-test	123.02	131.35	–	155.28	88.09	–
Hausman χ^2 *(Prob>* χ^2*)*		79.51 (0.000)		–	80.44 (0.000)	

Notes: ***, ** and * represent the levels of significance of 1%, 5% and 10% respectively. The values of the standard errors are in parenthesis.

Conclusions

The results of our study are consistent with the prediction of the FG Model (considers a new dynamic paradigm of development through technology transfer) and GVC theory (according to which the CEEC are integrated into EU's supply chains)and can be summarized as follows. First, we point out the importance of technology sophistication of 'implanted' industries for FDI benefits brought to the host country. Second, technological transfer for CEEC takes place rather through the import of technologies. Third, integration into EU is positively correlated with technological development of CEECs. Further research directions include the study of the factors of economic divergence of transitional countries based on the FDI into high-skill and technology intensive sectors.

References

Akamatsu, K. (1962). A historical pattern of economic growth in developing countries. *The Developing Economies, Vol. 1,* 3-25.

Baldwin, R. (2012). Global Supply Chains: Why they emerged, why they matter, and where they are going. *CEPR Discussion Paper No. DP9103*. The Graduate Institute Geneva.

Baldwin, R. (2016). *The Great Convergence*. Harvard University Press.

Cherkas, N. (2013). Structural shifts in the exports and industrial production effects in Ukraine. In: Papież, M. & Śmiech, S. (Eds.), *The 7th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings.* Cracow: Foundation of the Cracow University of Economics, pp. 41-48.

Damijan, J. P., Kostevc, Č. & Rojec, M. (2013). Global Supply Chains at Work in Central and Eastern European Countries: Impact of FDI on Export Restructuring and Productivity Growth. *VIVES Discussion Paper 37.*

Damijan, J. P, & Rojec, M. (2004). Foreign Direct Investment and the Catching-up Process in New EU Member States: Is There a Flying Geese Pattern?*, Research Reports, 310, WIIW.*

Danakol, S. H., Estrin, S., Reynolds, P. & Weitzel, U. (2017). FDI via M&A and domestic entrepreneurship: blessing or curse?. *Small Business Economics, 48*(3), 599-612.

Dunning, J. H.& Lundan, S.M. (2008). *Multinational Enterprises and the Global Economy, Second Edition*. Cheltenham: Edward Elgar.

Hausman, J.A. (1978). Specification tests in econometrics. *Econometrica,Vol. 46*(6), 1251-1271.

Jude, C. (2016). Technology spillovers from FDI. Evidence on the intensity of different spillover channels. *The World Economy, 39*(12), 1947-1973.

Kalotay, K. (2010). Patterns of inward FDI in economies in transition. *Eastern Journal of European Studies, 1*(2), 55-76.

Kojima, K. (2000). The 'Flying Geese' model of Asian economic development: origin, theoretical extensions, and regional policy implications. *Journal of Asian Economics, Vol. 11*, 375-401.

Li, K.-W. (2017). East Asia: Flying Geese Against Wind Currents. *Redefining Capitalism in Global Economic Development.* Academic Press Elsevier Science, 283-297.

Medve-Bálint, G. (2014). The role of the EU in shaping FDI flows to East Central Europe, *JCMS: Journal of Common Market Studies, No. 52(1),* 35-51.

Ozawa, T. (2007). *Institutions, Industrial Upgrading, and Economic Performance in Japan: The Flying-geese Paradigm of Catch-up Growth*. Edward Elgar Publishing.

Salamaga, M. (2013). Modeling the Influence of Foreign Direct Investment on Foreign Trade in the Light of Chosen Economic Theories Using the Example of the Countries of EastCentral Europe, *Scientific Papers, No. 223*, Cracow University of Economics Press, Cracow.

Taglioni, D. & Winkler, D. (2016). *Making Global Value Chains Work for Development*. Washington, DC: World Bank.

Widodo, T. (2007). "Flying Geese" Paradigm: Review, Analytical Tool and Application, *MPRA Paper No. 78218,* University Library of Munich, Germany.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

World Bank (2017). *World Development Indicators*, Washington, DC.

Modelling intensity of EUR/PLN high frequency trade – a comparison of ACD-type models

Anna Chudzicka-Bator [1], Mateusz Pipień[2]

Abstract

In the paper a review of alternative parameterizations of the autoregressive conditional duration (ACD) models is presented. We consider several different specifications of the conditional mean of durations as well as the types of conditional distribution. We discuss relative predictive and explanatory performance of a class of competing specifications on the basis of the series of price and trade durations obtained for EUR/PLN exchange rate. To investigate relative performance we present detailed insight into the goodness of fit of estimated models on the basis of probability integral transformation (PIT). The results show that the Log_1ACD model based on the conditional Burr distribution receives the greatest data support for both price and trade durations under study. The effect of persistence to shocks indicate explosiveness. Although the choice of the Burr distribution results with much more regular histograms of PIT however, the PIT sample is characterized by relatively stronger autocorrelation. In general, models with poor explanatory power exhibit stronger excess of histogram of Z statistics from the uniform distribution but with weaker autocorrelation reported.

Keywords: *high frequency data, exchange rate, durations, ACD models, probability integral transform*
JEL Classification: C52, C53, C58
DOI: 10.14659/SEMF.2018.01.06

1 Introduction

Using the same idea as the one that originated ARCH model for the volatility, (Engle and Russell, 1998) developed the autoregressive conditional duration (ACD) model to describe the evolution of the times between transactions (durations). It was introduced to study transactions data that occur irregularly in time, treating the time between event occurrences as a random process. Most applications of ACD models focus on the analysis of the trading process based on trade and price durations. Initially intensity of trading activity was analysed mainly on equity markets. However there are some exceptions like the paper by Holder et al. (2004) focused on investigation of the price formation process for the futures market. Also Dufour and Engle (2000) analysed transaction data for Treasury Note futures contracts traded at the Chicago Board of Trade. Except of the trading process, some other economic events, such as firm defaults or liquidity traps was subject to research; see Pacurar (2008).

[1] Cracow University of Economics, anna.chudzicka.88@gmail.com
[2] Corresponding Author, Department of Econometrics and Operations Research, Cracow University of Economics, eepipien@cyf-kr.edu.pl

Foreign exchange market was also subject to analysis through the market microstructure perspective. Fisher and Zurlinden (2004) examine trade intensity with relation to interventions done by central banks on foreign exchange markets. Using daily data on spot transactions of the Federal Reserve, the Bundesbank, and the Swiss National Bank on the dollar market, the authors conclude that traditional variables of a central bank's reaction function for interventions do not improve the ACD specification in their sample.

High Frequency Trading (HFT) as a special case of Algorithmic Trading (AT) become very popular during last decade, particularly of FX market. Hence a detailed insight into empirical properties of ACD-type models is necessary not only from in-sample viewpoint but also regarding forecasts. Consequently in this paper we aim at predictive and explanatory performance of a class of competing ACD specifications on the basis of the series of durations obtained for EUR/PLN trade. To investigate relative performance we report BIC score. We also report measures based on the probability integral transform (PIT) as employed for diagnostics on predictive performance of ACD models.

2 A review of models describing conditional duration

Let denote by $x_i = t_i - t_{i-1}$ the duration between two consecutive events occurred at the time t_i and t_{i-1} respectively. Let F_{i-1} be the information set available until the time t_{i-1}; i.e a series of x_t up to $\mathrm{t} = \mathrm{i} - 1$. According to Engle and Russel (1998) the duration can be described according to the following formula:

$$x_i = \Psi_i \varepsilon_i, \tag{1}$$

where ε_i is the positive error term with probability density function $f_\varepsilon(\varepsilon_i)$ and $\varepsilon_i \sim IID(1, \sigma_\varepsilon^2)$. In our framework $\Psi_i = \Psi_i(x_i|x_{i-1}, \dots, x_1; \theta) = E(x_i|F_{i-1})$ represents the conditional mean of modelled duration. The original (linear) Autoregressive Conditional Duration model, ACD(p,q) parameterizes the conditional mean duration as follows:

$$\Psi_i = \omega + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \Psi_{i-j}\,. \tag{2}$$

The restrictions: $\omega > 0, \bigwedge_{j=1,2,\dots} \alpha_j \geq 0, \beta_j \geq 0$ ensure positivity of the conditional duration and $\sum_{j=1}^{p} \alpha_j + \sum_{j=1}^{q} \beta_j < 1$ is sufficient for covariance stationarity of the process and existence of the unconditional mean. Here we consider several other specifications of the conditional mean duration Ψ_i

1. Logarithmic ACD_1 model – $Log_1ACD(1,1)$ – Bauwens and Giot (2006):

$$ln\,\Psi_i = \omega + \alpha_1 \ln \varepsilon_{i-1} + \beta_1 \Psi_{i-1}. \tag{3}$$

2. Logarithmic ACD_2 model – $Log_2ACD(1,1)$ – (Bauwens & Giot, 2006):

$$ln\,\Psi_i = \omega + \alpha_1 \varepsilon_{i-1} + \beta_1 \Psi_{i-1}. \quad (4)$$

3. Box-Cox ACD model – BACD(1,1) – (Hautsch, 2001):

$$\Psi_i^{\delta_1} = \omega + \alpha_1 \varepsilon_{i-1}^{\delta_2} + \beta_1 \Psi_{i-1}^{\delta_1}. \quad (5)$$

4. Augmented Box-Cox ACD model – ABACD(1,1) – (Hautsch, 2012):

$$\Psi_i^{\delta_1} = \omega + \alpha_1 \Psi_{i-1}^{\delta_2} [|\varepsilon_{i-1} - \nu| + c_1 |\varepsilon_{i-1} - b|]^{\delta_2} + \beta_1 \Psi_{i-1}^{\delta_1}. \quad (6)$$

5. Additive and Multiplicative ACD model – AMACD(1,1,1) – Hautsh (2011):

$$\Psi_i = \omega + \alpha_1 x_{i-1} + \nu_1 \varepsilon_{i-1} + \beta_1 \Psi_{i-1}. \quad (7)$$

Innovations ε_i are positive thus any distribution with positive support could be applied. We consider only distributions that belong to a family of linear-exponential distributions. This useful family of distributions enables to use a quasi-maximum likelihood estimate (QMLE). The estimation procedure is consistent and asymptotically normal irrespective of the actual distribution of a random components. We consider 4 alternative conditional distributions: exponential, Weibull, Generalised Gamma and Burr. The corresponding density functions (assuming that the error terms have unit expectation) and log-likelihood functions are given by:

1. Exponential; see Engle and Russel (1998):

$$f(x_i|\Psi_i;\theta) = \frac{1}{\Psi_i} f_\varepsilon\left(\frac{x_i}{\Psi_i}\right) = \frac{1}{\Psi_i} \exp\left(-\frac{x_i}{\Psi_i}\right), \quad (8)$$

$$\ln L(\theta) = -\sum_{i=1}^{N} \frac{x_i}{\Psi_i}\left(\frac{x_i}{\Psi_i} + \ln \Psi_i\right). \quad (9)$$

2. Weibull; see Engle and Russel (1998):

$$f(x_i|\Psi_i;\theta) = \frac{1}{\Psi_i/\mu} f_\varepsilon\left(\frac{x_i}{\Psi_i/\mu}\right) = \frac{\gamma}{x_i}\left[\frac{x_i\Gamma\left(1+\frac{1}{\gamma}\right)}{\Psi_i}\right]^{\gamma} \exp\left\{-\left[\frac{x_i\Gamma\left(1+\frac{1}{\gamma}\right)}{\Psi_i}\right]^{\gamma}\right\}, \quad (10)$$

$$\ln L(\theta) = \sum_{i=1}^{N}\left\{\ln\left(\frac{\gamma}{x_i}\right) + \gamma \ln\left[\frac{x_i\Gamma\left(1+\frac{1}{\gamma}\right)}{\Psi_i}\right] - \left[\frac{x_i\Gamma\left(1+\frac{1}{\gamma}\right)}{\Psi_i}\right]^{\gamma}\right\}. \quad (11)$$

3. Generalised Gamma; see Lunde (1999):

$$f(x_i|\Psi_i;\theta) = \frac{\gamma}{x_i\Gamma(\upsilon)}\left[\frac{x_i\Gamma(\upsilon+1/\gamma)}{\Psi_i\Gamma(\upsilon)}\right]^{\upsilon\gamma} \exp\left(-\left[\frac{x_i\Gamma(\nu+1/\gamma)}{\Psi_i\Gamma(\upsilon)}\right]^{\gamma}\right), \quad (12)$$

$$\ln L(\theta) = \sum_{i=1}^{N}\left\{\ln\left(\frac{\gamma}{x_i\Gamma(\nu)}\right) + \nu\gamma \ln\left[\frac{x_i\Gamma\left(\nu+\frac{1}{\gamma}\right)}{\Psi_i\Gamma(\nu)}\right] - \left[\frac{x_i\Gamma\left(\nu+\frac{1}{\gamma}\right)}{\Psi_i\Gamma(\nu)}\right]^{\gamma}\right\}. \quad (13)$$

4. Burr Gramming and Maurer (2000):

$$f(x_i|\Psi_i;\theta) = \frac{1}{\Psi_i/\mu} f_\varepsilon\left(\frac{x_i}{\Psi_i/\mu}\right) = \frac{\kappa(\Psi_i/\mu)^{-\kappa}(x_i)^{\kappa-1}}{[1+\sigma^2(\Psi_i/\mu)^{-\kappa}(x_i)^{\kappa}]^{\frac{1}{\sigma^2}+1}}, \quad (14)$$

$$\ln L(\theta) = \sum_{i=1}^{N} \left[\ln(\kappa\mu^{\kappa-1}) + (1-\kappa)\ln \Psi_i + (\kappa-1)\ln x_i - \left(1+\frac{1}{\sigma^2}\right)\ln(1+\sigma^2(\mu/\Psi_i)^{\kappa} x_i^{\kappa}) \right]. \quad (15)$$

3 Empirical analysis

In this section, a set of competing ACD-type specifications are applied to modelling trade intensity of EUR/PLN exchange rate. We analysed the time series of price and trade durations. The price duration is defined as the time until the unit price has changed by at least 8,27E-05 in absolute value, representing the average change of the price of analysed exchange rate. The trade duration is simply the time between two consecutive transactions. The analysed time series cover the time span of 20 workdays: starting from October 6th to 31st of 2014, when transactions recorded only between 9 AM and 5 PM were considered. Transactions that occurred in the same second were aggregated in one with the price referred to latter trade. We assumed that time between two consecutive days equals 1. This resulted with 44 247 price durations and 79 435 trade durations comprising the subject of the analysis. The data were taken from GAIN Capital Group; see http://ratedata.gaincapital.com/.

As described in the literature, durations exhibit intraday patterns; see (Engle and Russell, 1998; Bauwens and Giot, 2006). Following (Engle and Russell, 1998) we transform data by dividing plain intra-daily durations by estimated periodic component. We use Nadaraya-Watson kernel estimator with quadric kernel, considering every week in the sample separately as suggested by Bauwens and Giot (2006). Such method is commonly used by many authors; see Bauwens and Giot (2006), Bień (2006), Chudzicka (2016), Huptas (2009), Białkoska and Pipień (2015).

We present detailed insight into the goodness of fit of estimated models on the basis of probability integral transformation (PIT). In general the purpose of testing procedure proposed by Diebold et al. (1997) was to check if a sequence of one-step-ahead density forecasts of the ACD model is accurate. We utilise this approach in full description of the data fit of competing models. In this paper the PIT is obtained by taking the cumulative distribution function of the residuals given a particular ACD-type specification. Under a proper specification the PIT series are tested to be independently uniformly distributed on the unit interval $(0,1)$:

$$z_i = \int_{-\infty}^{x_i} f_i\left(u|\Psi_{i-1}, \hat{\theta}\right) du \,, \{z_i\} \sim i.i.d.\, U(0,1). \quad (16)$$

We perform Kolmogorov-Smirnov test to verify the $i.i.d.U(0,1)$ behaviour of $\{z_i\}$. Additionally, we report histograms and correlograms as it is suggested in Diebold et al. (1997).

3.1 Results obtained for price duration

We estimated parameters of 19 models for price durations. According to BIC score presented in Table 1, the conditional Burr distribution is empirically supported as the best choice for all parameterisation of the conditional duration, except of BACD model, the case of the conditional Generalised Gamma distribution receives the best score. Analysing the ranking of competing models with respect to the BIC score the greatest data support seem Log_1ACD with conditional Burr distribution and BACD with Generalised Gamma specification.

In Table 1 we report sum $\alpha_1 + \beta_1$ which represents the effect of persistence to shocks. Among ACD specifications the effect is much stronger in case of exponential and Weibull distribution (the sum $\alpha_1 + \beta_1 = 0.894$ and 0.876 respectively). Much more complicated Burr distribution is able to describe some features of the data in such way the persistence to shocks declines ($\alpha_1 + \beta_1 < 0.7$). Also in case of Log_1ACD the conditional Burr reaches the first place in rank. Surprisingly in case of the best model within analysed subclass this effect indicate explosiveness, as the point estimates of $\alpha_1 + \beta_1$ is greater than 1. Reacher parameterisation of the conditional price duration, proposed as BACD model does not resolve the problem of empirically pervasive effect of persistence to shocks. Close to unity point estimates of underlying sum of parameters is supported in case of three analysed conditional distributions. The best specification, assuming conditional Generalised Gamma distribution is characterised by relatively weaker effect of persistence. For a subclass of ABACD models, for the best model, built on the basis of assumption of the conditional Generalised Gamma distribution the sum is smaller but still close to 1 ($\alpha_1 + \beta_1 = 0.98$). The empirical importance of heavily parameterized conditional distribution in effect of persistence to shocks is also reported in case of AMACD class. Conditionally Burr distribution explains dynamics of the price durations with the use of equation with the weakest effect of persistence as the sum $\alpha_1 + \beta_1$ is smaller than 0.8.

For a detailed insight into differences in explanatory power of the best models (Log_1ACD) we focus on PIT histograms presented in Fig. 1. We calculate Z statistics according to (16) at each data point. Histograms present empirical distribution of Z given ML estimates of parameters. Our results show the role of the conditional Exponential distribution in explaining the distribution of observables can be characterized by huge overestimation of the right tail as the frequencies reach maxima for this region of the distribution function. Also

underestimation of the right tail receives attention, however the nature of this effect is rather different. Histograms exhibit substantial excess from the uniform distribution for the whole region starting from left tail up to the quantile of order 0.25 approximately.

Table 1. Comparison of estimation results and goodness-of-fit.

Model	Distribution	Price durations			Trade durations		
		BIC	$\alpha_1 + \beta_1$	Z	BIC	$\alpha_1 + \beta_1$	Z
ACD	Exponential	65983.75	0.894	0.2049	139080.69	0.925	0.1434
	Weibull	57789.52	0.876	0.0744	135613. 98	0.909	0.0343
	Burr	51706.03	0.659	0.0341	119134.51	0.747	0.0941
Log_1 ACD	Exponential	67223.94	0.959	0.2241	141042.50	0.988	0.0314
	Weibull	57986.59	0.993	0.0944	136471.90	1.006	0.1096
	Burr	49317.34	1.062	0.0358	111772.70	1.057	0.0396
	Generalized Gamma	53467.00	0.837	0.0728	122304.63	0.710	0.0847
Log_2 ACD	Exponential	66381.27	0.986	0.2031	139626.31	1.006	0.1429
	Weibull	58064.70	0.957	0.0784	136083. 00	0.994	0.0900
	Burr	52016.72	0.843	0.0382	119658. 51	0.874	0.0349
BACD	Exponential	65502.92	1.045	0.2023	138525.39	1.055	0.1469
	Weibull	57407.05	1.060	0.0843	135007.03	1.048	0.0969
	Generalized Gamma	51115.53	0.929	0.0740	123116.54	1.088	0.0937
ABACD	Exponential	65351.88	1.004	0.2055	138487.03	1.062	0.1479
	Weibull	57402.57	1.262	0.0837	134969.30	0.968	0.1021
	Burr	NA	NA	NA	114082.90	1.440	0.0425

Model	Distribution	Price durations			Trade durations		
		BIC	$\alpha_1 + \beta_1$	Z	BIC	$\alpha_1 + \beta_1$	Z
AMACD	Generalized Gamma	52255.65	0.983	0.0762	NA	NA	NA
	Exponential	65663.44	0.842	0.2014	138784.53	0.901	0.1413
	Weibull	57628.03	0.827	0.0796	135411.95	0.892	0.0917
	Burr	51568.95	0.760	0.0308	118803.62	0.896	0.0339
	Generalized Gamma	NA	NA	NA	127737.70	0.875	0.0803

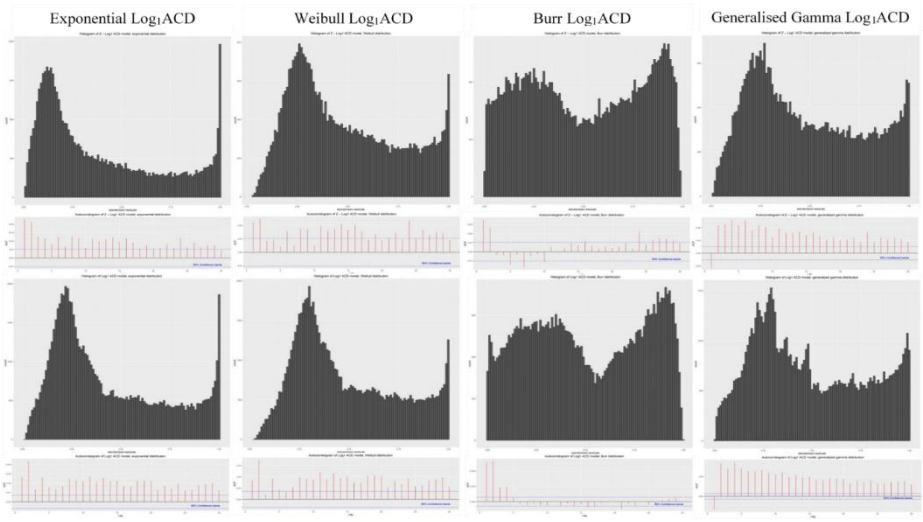
Conditional Weibull distribution also underestimates both tails of the conditional distribution of price durations. But frequencies corresponding to the left tail indicate much more regular coverage as compared to the case analyzed above. Conditional Burr distribution seems the most regular and it confirms result of model comparison discussed previously. The case of conditional Generalized Gamma distribution generates histograms of *Z* statistics very similar to the case of the conditional Weibull distribution. The time pattern of dependence of Z can be analyzed on the basis of correlograms presented in Fig. 1. It seems that autocorrelation is relatively higher in case of models with better explanatory power. Also statistically significant correlations were obtained even at 30 and greater, making the case of independence of *Z* samples improbable in the view of the data. Conditionally Burr distribution, generally supported as the best choice makes histogram more regular however the PIT sample utilized in the procedure is characterized by much stronger autocorrelation.

3.2 Results obtained for trade duration

In case of trade durations we estimated parameters of 20 models. Again, BIC score presented in Table 1 indicates the conditionally Burr Log_1ACD model as the best among all competing specifications. Burr distribution receives the greatest data support for all parameterisation except of BACD model where conditional Generalised Gamma is the most likely.

Just like in case of price durations, when modelling trade durations we report sum $\alpha_1 + \beta_1$ to explain persistence to shocks. Among ACD specifications the conditional Burr distribution receives the strongest data support (the sum $\alpha_1 + \beta_1 < 0.74$). In case of Log1ACD the

conditional Burr reaches the first place in rank again. The sum $\alpha_1 + \beta_1$ is very close to unity for conditionally Exponential model. In case of conditionally Weibull and Burr models this effect indicate explosiveness ($\alpha_1 + \beta_1 > 1$). Equation for conditional trade duration estimated in case of Generalised Gamma distribution can be characterised by the weakest effect of persistence as the point estimated of the sum $\alpha_1 + \beta_1$ slightly crosses 0.8. The class of BACD models is characterised by very strong effect of persistence to shocks. Greater than 1 point estimates of $\alpha_1 + \beta_1$ are obtained in case of all three analysed conditional distributions. For a subclass of ABACD models, for the best model (with Burr distribution), the sum $\alpha_1 + \beta_1$ is much greater than 1 and consequently this case exhibit much stronger effect of persistence to shocks. In case of AMACD class the inference about the sum $\alpha_1 + \beta_1$ is very stable among all four analysed conditional distributions and does not cross the value 0.9. Conditionally Burr distribution explains dynamics of the price durations with the use of equation with the weakest effect of persistence to shocks.



Notes: Top figures relate to price durations, bottom to trade durations.

Fig. 1. Histograms and autocorrelations of Z – the case of Log_1ACD models.

Again we calculated PIT histograms and correlograms presented in Fig. 1. Application of different conditional distributions result with qualitatively the same misspecification measured by excess of the histogram of Z statistics from the uniform case. Overestimation of both tails seems the most important feature of the nature of explanatory power of analysed class of models. Left tail receives maxima of frequencies of Z statistics for conditional Exponential distribution. Conditionally Weibull case does not contribute any important information as the histogram obtained for those class is very similar to the class of

Exponential model. Conditional Burr distribution seems the most regular and it confirms its superiority among all discussed competing specifications. The case of conditional Generalised Gamma distribution generates histogram of Z statistics very similar to the case of the conditional Weibull distribution.

Just like in case of series of price durations autocorrelation is higher for models with better fit. Significant correlations were obtained even at 30 and greater, making the case of independence of Z samples rejected. Analyses bring the same results as in previous case.

Conclusions

The paper presents alternative parameterizations of ACD models. We discuss relative predictive and explanatory performance of a class of competing specifications on the basis of the series of price and trade durations obtained for EUR/PLN exchange rate. According to BIC score the conditional Burr distribution receives the greatest data support for both price and trade durations. It is empirically supported as the best choice for most of parameterisation of the conditional duration. In this case the effect of persistence to shocks, measured by the sum $\alpha_1 + \beta_1$ indicate explosiveness. Application of different conditional distributions result in qualitatively the same misspecification measured by excess of the histogram of Z statistics from the uniform case. Only conditional Burr distribution seems the most regular and it confirms its superiority among all competing specifications. The time pattern of dependence of Z analysed on the basis of correlograms show that autocorrelation is higher for models with better fit. Significant correlations were obtained even at 30 and greater, making the case of independence of Z samples rejected. The stochastic nature of trade duration series may be complex in such way the family of the ACD-type models analysed in the paper is too narrow for a proper description and prediction of observed features of price and trade intensity.

Acknowledgement

This research was supported by the research grant 2017/25/B/HS4/02529 financed by the National Science Centre, Poland

References

Bauwens, L. & Giot, P. (2006). The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks.

Bauwens, L., Giot, P., Grammig, J. & Veredas, D. (2004). A comparison of financial duration models via density forecasts. *Int. J. Forecasting*, *20*(4), 589–609.

Białkowska J. & Pipień M (2015) Analiza czasów trwania pomiędzy zmianami kierunku cen akcji - wpływ uwzględnienia wewnątrzdziennej sezonowości na ranking modeli ACD, *Folia Oeconomica Cracoviensia 56*, 35-80

Bień, K. (2006). Zaawansowane specyfikacje modeli ACD - prezentacja oraz przykład zastosowania. *Przegląd Statystyczny*, *53*(1), 90–108.

Chudzicka, A. (2016). Prognozowanie rozkładu warunkowego czasu trwania pomiędzy transakcjami na kurs EUR/USD - porównanie modeli typu ACD. *Folia Oeconomica Cracoviensia*, (57), 83–104.

Diebold, F. X., Gunther, T. A. & Tay, A. S. (1997). Evaluating Density Forecasts.

Dufour, A. & Engle, R. F. (2000). The ACD Model: Predictability of the Time Between Concecutive Trades. *ICMA Centre Discussion Papers in Finance*.

Engle, R. F. & Russell, J. R. (1998). Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica*, *66*(5), 1127–1162.

Fischer, A. M. & Zurlinden, M. (2004). Are Interventions Self Exciting? *Open Econ Rev.*, *15*(3), 223–237.

Gouriéroux, C., Jasiak, J. & Le Fol, G. (1999). Intra-day market activity. *Journal of Financial Markets*, *2*(3), 193–226.

Grammig, J., & Maurer, K.-O. (2000). Non-monotonic hazard functions and the autoregressive conditional duration model. *Econometrics Journal*, *3*(1), 16–38.

Hautsch, N. (2001). Modelling Intraday Trading Activity Using Box-Cox ACD Models.

Hautsch, N. (2011). *Econometrics of Financial High-Frequency Data*. Springer Science & Business Media.

Hautsch, N. (2012). *Econometrics of Financial High-Frequency Data*.

Holder, M. E., Qi, M. & Sinha, A. K. (2004). The impact of time duration between trades on the price of treasury note futures contracts. *Journal of Futures Markets*, *24*(10), 965–980.

Huptas, R. (2009). Metody szacowania wewnątrzdziennej sezonowości w analizie danych finansowych pochodzących z pojedynczych transakcji. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, *76*(26), 83–96.

Lunde, A. (1999). *A generalized gamma autoregressive conditional duration model. Politics and Public Administration*, Aalbourg University, Denmark.

Pacurar, M. (2008). Autoregressive Conditional Duration Models in Finance: A Survey of the Theoretical and Empirical Literature. *Journal of Economic Surveys*, *22*(4), 711–751.

Aluminium price discovery on the London Metal Exchange, 2007-2017

Marta Chylińska[1]

Abstract

The aim of the paper is to extend the analysis of aluminium price discovery on the London Metal Exchange (LME) beyond the 2007-2008 global financial crisis. To this end a VEC DCC-MGARCH model on the weekly sampled price series of spot and 3-month aluminium futures in the period 3/10/2007–27/09/2017 is estimated (10 years, 522 observations). The results of the study reveal that both prices exhibit a common stochastic trend and their spread have co-integrating properties. The hypothesis stating that they equally quickly revert to the long-run equilibrium relationship is rejected. An increased conditional volatility of their returns is observed during the crisis and after that a slightly decreasing albeit very close to unity their conditional correlation coefficient. Nevertheless, a constant conditional correlation hypothesis (CCC-MGARCH) is rejected. More interestingly, the term premium is likely to be proportional to the exchange rate of US dollar into British pound.

Keywords: *aluminium futures, London Metal Exchange, VEC DCC-MGARCH*
JEL Classification: G13, Q02
DOI: 10.14659/SEMF.2018.01.07

1 Introduction

The beginning of the XXI century is a period of increasing role of the aluminium industry. Major primary aluminium producers are located in China, Russia, Canada, the Middle-East, Australia, Brazil and India. From 1998 to 2016 world smelter production of aluminium increased by approximately 161%. It is a result of rising demand for commodities from emerging markets – particularly China (40% share of aluminium production) and Russia (9%) (Nappi, 2013, p. 20).

Aluminium futures contracts are traded on a small number of specialized markets: the London Metal Exchange (LME), the Commodity Exchange of New York (COMEX) and the Shanghai Futures Exchange (SFHE). The LME competes with the SHFE to dominate in the aluminium price discovery. The total volume of aluminium futures traded on the LME in 2016 was more than 1.3 billion tonnes (53.1 mln lots), which makes the LME a market in which 90% of the world aluminium futures are conducted. Furthermore, from the late 90s aluminium has been the most heavily traded non-ferrous metal on the LME (Figuerola–Ferretti and Gilbert, 2005).

[1] Corresponding author: University of Gdansk, Department of Econometrics, Armii Krajowej 101, 81–866 Sopot, Poland, e-mail: marta.chylinska@ug.edu.pl

The aluminium futures trading on the LME was introduced in October 1978 and actual – the high grade primary aluminium contract (AH) – in August 1987. From the mid of 1980s aluminium has been world-wide sold on the basis of LME quotations (Figuerola–Ferretti, 2005). The primary aluminium is sold in 25 tonnes lots (with a tolerance of +/-2%) in the shape of ingots, T-bars and sows. Price quotation is in US dollars per tonne (USD/t). All aluminium deliverable against LME contracts must be of an LME-approved brand. Premium futures contracts for aluminium were introduced on 23/11/2015. They enable market participants to take delivery of readily available material in non-queued warehouses.

The price setting process at the LME is a subject of intensive research. The recent papers on the issue are listed in Table 1. Most of the papers are focused on or before the 2007-2009 world financial crisis. The main findings include those that spot and futures prices are integrated of order one variables, co-integrated and their returns variances and covariances are time-varying. Some papers focus on correlations among the LME, the COMEX and the SHFE (Figuerola-Ferretti and Gilbert, 2005; Gong and Zheng, 2016), and other on correlations between aluminium and copper futures (Vu-Nhat, 2004; Figuerola-Ferretti and Gilbert, 2008).

Table 1. Literature review.

Study	Time period	Data freq.	Model	Main conclusions
Vu-Nhat (2004)	07/1995-07/2002	weekly	VAR	Spot, 3- and 27-month futures prices of aluminium and copper are co-integrated, aluminium and copper are likely to be substitutes
McMillan (2005)	01/1989-07/2003	daily	GARCH, GARCH-X	Returns variances and covariances are time-varying
Figuerola-Ferretti, Gilbert (2005)	01/1979-12/2003	monthly	VECM	LME aluminium price is more informative than Metal Bulletin price, LME price leads COMEX one (Granger causality), price discovery is on COMEX (permanent-transitory methodology)
Watkins, McAleer	02/1986-09/1998	daily	VAR	3 structural breaks, long-run relationship between spot and futures prices

(2006)				
Figuerola-Ferretti, Gilbert (2008)	10/1982-12/2005	daily	FIGARCH-VECM	Spot and 3-month aluminium and copper volatilities follow long memory processes, they exhibit a common degree of fractional integration, the processes are symmetric
Figuerola-Ferretti, Gonzalo (2010)	01/1989-10/2006	daily	VECM	Spot and futures prices are co-integrated, price discovery takes place in futures market
Gong, Zheng (2016)	04/1995-04/2013	daily	FISC[2]	LME and SHFE aluminium futures markets are more correlated in downturns than upturns

The aim of the paper is to extend the analysis of aluminium futures price discovery beyond the 2007-2008 global financial crisis and focus on the intra LME aluminium price setting in the recent decade. To this end a combined vector error correction and dynamic conditional correlation multivariate GARCH model (VEC DCC-MGARCH) on the weekly sampled price series of spot and 3-month aluminium futures in the period 3/10/2007–27/09/2017 (10 years, 522 observations) is estimated and several hypotheses related to their dynamics are tested. Weekly time series (Wednesdays) are used to avoid the day-of-the-week effects. Computations are performed using Microfit 5 and Stata 14 SE. The data comes from the Thomson Reuters. The remainder of the paper proceeds as follows. In section 2 the methodology is described. In section 3 the data and the empirical results are discussed. The last section briefly concludes.

2 Methodology

Assuming that the logs of futures and spot prices are co-integrated, on the base of risk premium model (Watkins and McAleer, 2006), a VEC DCC-MGARCH model (Johansen, 1995; Tse and Tsui, 2002) is specified:

$$\Delta la3_t = \sum_{k=1}^{p-1} \alpha_k^{(1)} \Delta la3_{t-k} + \sum_{k=1}^{p-1} \beta_k^{(1)} \Delta la0_{t-k} + \sum_{k=1}^{p-1} \gamma_{j,k}^{(1)} \Delta lex_{t-k} + \delta^{(1)} e_{t-1} + \xi_t^{(1)} \quad (1a)$$

[2] FISC stands for fractionally integrated stochastic copula model.

$$\Delta la0_t = \sum_{k=1}^{p-1} \alpha_k^{(2)} \Delta la0_{t-k} + \sum_{k=1}^{p-1} \beta_k^{(2)} \Delta la3_{t-k} + \sum_{k=1}^{p-1} \gamma_{j,k}^{(2)} \Delta lex_{t-k} + \delta^{(2)} e_{t-1} + \xi_t^{(2)} \quad (1b)$$

$$e_t = la3_t - \phi_0 - \phi_1 la0_t - \phi_2 lex_t \quad (1c)$$

$$\xi_t = H_t^{0,5} v_t \quad (2a)$$

$$H_t = D_t^{0,5} R_t D_t^{0,5} \quad (2b)$$

$$R_t = diag(Q_t)^{-0,5} Q_t diag(Q_t)^{-0,5} \quad (2c)$$

$$Q_t = (1 - \lambda_1 - \lambda_2)R + \lambda_1 \Psi_{t-1} + \lambda_2 Q_{t-1} \quad (2d)$$

where : $la0_t, la3_t$ – log price of spot and 3-month futures contracts, x_t – exchange rate of US dollar into British pound, $\xi_t^{(i)}$ – error term, H_t – Cholesky factor of the time-varying conditional covariance matrix, v_t – vector of i.i.d innovations, D_t – diagonal matrix of conditional variances in which each element σ_{kt}^2 evolves according to a univariate GARCH(p_k, q_k) processes $\sigma_{kt}^2 = s_i + \sum_{j=1}^{p_k} \alpha_j^{(k)} \xi_{j,t-j} + \sum_{j=1}^{q_k} \beta_j^{(k)} \sigma_{k,t-j}^2$, R_t – matrix of means to which the dynamic process in Eq. (2d) reverts, Ψ_t – rolling estimator of the correlation matrix $\hat{\xi}_t$, λ_1, λ_2 – parameters that govern the dynamics of conditional correlations such that $0 \leq \lambda_1 + \lambda_2 < 1$.

The model is estimated in two steps. First, the Johansen procedure is employed to identify co-integrating vectors. Second, the residuals from co-integrating relations are used to estimate a full VEC DCC-MGARCH with the maximum likelihood method. Then it is validated by testing:

1. Constant conditional correlations, VECM CCC-GARCH vs. VECM DCC-GARCH (H_0: $\lambda_1 = \lambda_2 = 0$), VC_1 – Wald test statistic under H_0 distributed as $\chi^2(2)$,
2. No return of conditional variances to their mean levels (H_0: $\lambda_1 + \lambda_2 = 1$), VC_2 – t test statistic under H_0 distributed as $N(0,1)$ in large samples,
3. GARCH(1,1) vs. IGARCH(1,1), IG – Wald test statistic under H_0 distributed as $\chi^2(2)$,
4. IGARCH in the variance equation for the price of contract maturing at time $t + k$, IG_k – Wald test statistics under H_0 distributed as $\chi^2(1)$,

Of the particular interest are hypotheses stating whether:

5. The price of futures contracts departures from their long run equilibrium relationship do not affect the current price of contract maturing at time $t+k$ ($H_0: \delta^{(i)} = 0$, $i = 1,2$), W_k – Wald test statistics under H_0 distributed as $\chi^2(1)$,
6. The price of futures contracts departures from their long run equilibrium relationship equally quickly revert to the long-run equilibrium relationship ($H_0: \delta^{(1)} = \delta^{(2)}$), W – Wald test statistics under H_0 distributed as $\chi^2(1)$.



Fig. 1. Log aluminium prices, la3, la0 (left axis) and log exchange rate, lex (right axis), 10/2007-9/2017.

Table 2. ADF-GLS and KPSS tests results.

Variable	**Test**							
	ADF-GLS				KPSS			
	Level	Lag	Trend	Lag	Level	Lag	Trend	Lag
$la0$	-1.90	13	**-3.17**	13	0.87	18	**0.90**	18
$la3$	-1.71	12	**-3.15**	13	0.93	18	**0.09**	18
lex	-0.19	18	-2.13	18	1.14	18	0.23	18
$\Delta la0$	**-2.37**	12	**-3.90**	12	**0.07**	18	**0.05**	18
$\Delta la3$	**-2.13**	12	**-3.73**	12	**0.07**	18	**0.05**	18
Δlex	**-3.70**	12	**-3.80**	17	**0.09**	18	**0.07**	18

The 5% critical values in the ADF-GLS test are obtained using

the response surface approach with a proper augmentation to solve for autocorrelation of random errors if necessary: –1.96 (h=12), –1.96 (h=13), –1.95 (h=18) (level) and –2.84 (h=12), –2.84 (h=13), –2.82 (h=17), –2.82 (h=18) (trend), h – augmentation lag (see Cheung and Lai (1995)). Critical values in the KPSS test are: 0.15 (trend), 0.46 (level).

3 Data and empirical results

The empirical research starts with the analysis of logarithmic price series of spot and 3-month futures contracts ($la0$, $la3$) and exchange rate of US dollar into British pound (lex) against time (see Fig. 1.). As demonstrated the prices sharply fall in September 2008 due to the world's financial crisis originated by the collapse of Lehman Brothers Holdings Inc. Both prices rarely pass through their mean levels which suggests they are not stationary. The results of the ADF-GLS and KPSS tests are gathered in Table 2. They indicate that the log prices and log of exchange rate are integrated of order one variables.

Table 3. Maximal eigenvalue and trace tests estimation results.

Test									
Maximal eigenvalue					Trace				
Hypothesis		Test	Crit. value		Hypothesis		Test	Crit. value	
H_0	H_A	statistic	5%	10%	H_0	H_A	statistic	5%	10%
r=0	r=1	44.64	14.79	12.83	r=0	r≥1	45.18	17.79	15.83
r≤1	r=2	0.54	8.13	6.49	r≤1	r≥2	0.54	8.13	6.49

Co-integration with no intercepts or trends in the VAR.

The lag order $p = 1$ of the VAR system is set using AIC information criterion. Next, based on the maximal eigenvalue and trace test statistics, the existence of one co-integrating vector is identified (see Table 3). So the log prices of spot and 3-month aluminium futures at LME follow a common stochastic trend. Then, setting over identifying restrictions on the parameters of co-integrating vector, aluminium price spread are found to have co-integrating properties. Since the estimate of relevant likelihood ratio test statistic $LR(1)$ = 0.07 this hypothesis cannot be rejected at the 5 % significance level. Under H_0 the LR test statistic is asymptotically distributed as $\chi^2(1)$. Its 95 per cent bootstrap critical value is 4.55. The

hypothesis stating that $\phi_2 = 0$ is rejected ($LR(1)$ = 35.54, under H_0 the LR test statistic is asymptotically distributed as $\chi^2(1)$, its 95 per cent bootstrap critical value is 7.39). The term premium is likely to be proportional to the exchange rate of US dollar into British pound. The estimated co-integrating vector is: $e_t = la3_t - la0_t - 0.030837lex_t$ and the VEC model consists of two equations: $\Delta la3_t = \delta^{(1)} e_{t-1} + \xi_t^{(1)}$ and $\Delta la0_t = \delta^{(2)} e_{t-1} + \xi_t^{(2)}$, where e_t are the residuals from co-integrating vector.

Table 4. VECM DCC-MGARCH estimation and validation results.

Variable/	**Equation**			
Test stat.	$\boldsymbol{\Delta la3_t}$		$\boldsymbol{\Delta la0_t}$	
	Coef.	Std. err.	Coef.	Std. err.
Estimation results				
e_{t-1}	-0.32	0.17	-0.20	0.17
ξ_{t-1}	0.07	0.04	0.08	0.05
σ_{t-1}	0.83	0.10	0.80	0.11
cons	0.00	0.00	0.00	0.00
Validation results				
Statistic	Estimate	$\boldsymbol{p}$-value	Estimate	$\boldsymbol{p}$-value
$\boldsymbol{IG_k}$	2.50	0.11	3.21	0.07
$\boldsymbol{W_k}$	3.53	0.06	1.30	0.25
Residuals				
$\boldsymbol{LB(1)}$	0.19	0.66	0.42	0.52
$\boldsymbol{LB(4)}$	2.58	0.63	3.09	0.54
$\boldsymbol{LB(13)}$	12.20	0.51	12.36	0.50
$\boldsymbol{LB(26)}$	17.70	0.89	18.60	0.85
Sq. of residuals				
$\boldsymbol{LB(1)}$	0.45	0.50	0.42	0.52
$\boldsymbol{LB(4)}$	5.06	0.28	4.47	0.35
$\boldsymbol{LB(13)}$	10.48	0.65	9.95	0.70
$\boldsymbol{LB(26)}$	21.06	0.74	17.68	0.89

IG = 4.15, VC_1 = 671.85, VC_2 =16.14, W = 19.97

The VEC DCC-MGARCH is estimated with maximum likelihood method (using AIC informational criterion the DCC-MGARCH(1,1) is set). The estimation and validation results are gathered in Table 4. They indicate that the mean and variance equations of VEC DCC-MGARCH model are properly specified as the Ljung-Box portmanteau test applied on standardized residuals from Eq. (1a)-(1b) and their squares shows that they are non-autocorrelated processes up to the 26th order (see the estimates of LB(k) – Ljung-Box portmanteau test statistic for autocorrelation of order up to k, under H_0 distributed as χ^2(k)). Second, a DCC-MGARCH is more likely than a CCC-MGARCH (VC_1 = 671.85). The hypothesis of integrated MGARCH is rejected only for the first equation at the 10% significance level (see the estimates of IG and IG_k test statistics in T4.). More interestingly, the hypothesis stating that the price departures from their long run equilibrium relationship do not affect the time t price of 3-month aluminium futures contract is rejected at the 10% significance level (see the estimates of W_3 test statistic in Table 4.). However, the same hypothesis cannot be rejected for spot contract (see the estimates of W_0 test statistic in Table 4.). The hypothesis stating that they equally quickly revert to the long-run equilibrium relationship is rejected (see the estimates of W test statistic in Table 4.).

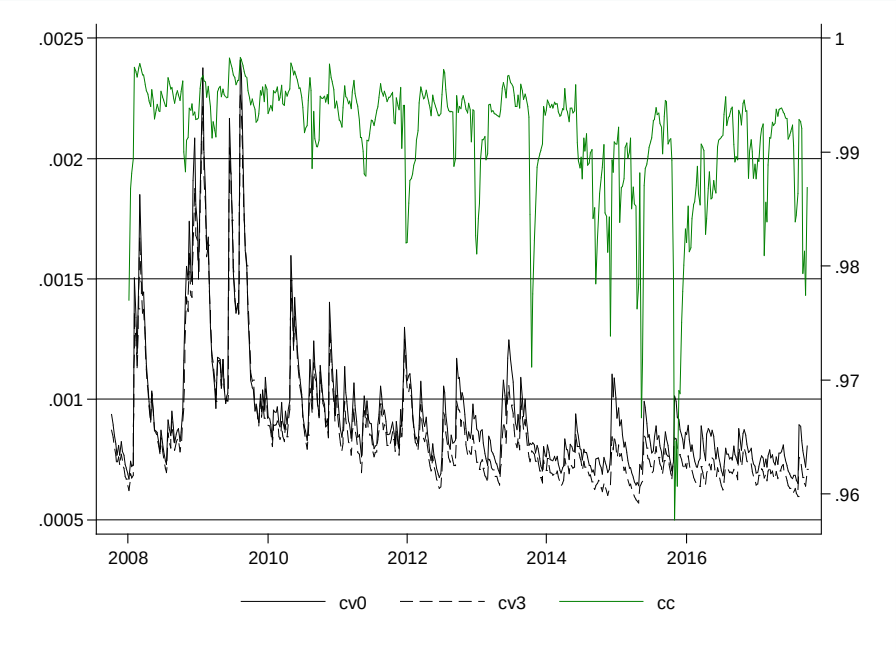


Fig. 2. Conditional correlation, cc (right axis) and conditional variances, cv0, cv3 (left axis) of the weekly log rates of return on aluminium futures prices at LME, 10/2007-9/2017.

Finally, the conditional correlation coefficient and conditional variances for the log rates of returns on spot and 3-month aluminium futures at LME are plotted in Fig. 2. There are 3 periods of an increased conditional volatility of the log returns between 2008-2010 resulting from stages of the crisis on financial markets. After 2010 the levels of conditional volatilities are smaller, but still sudden, though moderate, increases are locally observed. At all times their conditional correlation remain almost stable and is close to one, however two of its slightly different levels are observed before and after 2014.

Conclusions

In the paper the aluminium price discovery on the LME was analysed in the period 2007-2017. The results of the study reveal that the price series of spot and 3-month futures exhibit a common stochastic trend and their spread have co-integrating properties. Based on VEC DCC-MGARCH model the hypothesis stating that they equally quickly revert to the long-run equilibrium relationship is rejected. Moreover, 3 periods of an increased conditional volatility of their returns are observed during the crisis. After 2010 levels of conditional volatilities are smaller, but still sudden, though moderate, increases are locally observed. At all times their conditional correlation remain almost stable and is close to one, however two of its slightly different levels are observed before and after 2014. Nevertheless a constant conditional correlation hypothesis (CCC-MGARCH) is rejected. More interestingly, it is showed that the term premium is likely to be proportional to the exchange rate of US dollar into British pound.

References

Cheung, Y. W. & Lai, K. S. (1995). Lag order and critical values of the augmented Dickey–Fuller test. *Journal of Business & Economic Statistics, 13(3),* 277-280.

Figuerola-Ferretti, I. & Gilbert, C. L. (2005). Price discovery in the aluminum market. *Journal of Futures Markets, 25(10),* 967-988.

Figuerola-Ferretti, I. & Gilbert, C. L. (2008). Commonality in the LME aluminum and copper volatility processes through a FIGARCH lens. *Journal of Futures Markets, 28(10),* 935-962.

Figuerola-Ferretti, I. & Gonzalo, J. (2010). Modelling and measuring price discovery in commodity markets. *Journal of Econometrics, 158(1),* 95-107.

Figuerola–Ferretti, I. (2005). Prices and production cost in aluminium smelting in the short and the long run. *Applied Economics, 37(8),* 917-928.

Gong, Y. & Zheng, X. (2016). Long memory in asymmetric dependence between LME and Chinese aluminum futures. *Journal of Futures Markets, 36(3),* 267-294.

Hua, R. & Chen, B. (2007). International linkages of the Chinese futures markets. *Applied Financial Economics, 17(16),* 1275-1287.

Johansen, S. (1995). Likelihood-based inference in cointegrated vector autoregressive models. *Oxford University Press,* Oxford.

McMillan, D. G. (2005). Time-varying hedge ratios for non-ferrous metals prices. *Resources Policy, 30(3),* 186-193.

Nappi, C. (2013). The global aluminium industry 40 years from 1972. *World Aluminium,* 1-27.

Tse, Y. K. & Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics, 20(3),* 351-362.

Vu-Nhat, N. (2004). Common trends in metal futures curves and inventories: evidence from the London Metal Exchange. *Journal of Derivatives & Hedge Funds, 10(2),* 114-130.

Watkins, C. & McAleer, M. (2006). Pricing of non-ferrous metals futures on the London Metal Exchange. *Applied Financial Economics, 16(12),* 853-880.

Determinants of digital divide in Polish households

Małgorzata Ćwiek[1]

Abstract

No access to information and communication technologies and no ability to use them are seen as potential barriers for individuals to participate in the information society what can lead them to the digital divide. The study shows the selected aspects of the problem of digital inequality in Poland. The purpose of this paper is to identify the socioeconomic factors that are conducive of digital exclusion of Polish households. In order to extract the qualitative factors, logistic regression was carried out. Individual, non-identifiable data from a household budget survey conducted by the Central Statistics Office in Poland in the years 2012-2016 were used in the analysis.

Keywords: *digital divide, digital development, information society*

JEL Classification: D31, I31

DOI: 10.14659/SEMF.2018.01.08

1 Introduction

In the era of the information society, i.e., a society whose existence is largely based on the flow of information and the use of IT solutions, the lack of ability or skill to efficiently use the tools of acquiring information can become the reason of the so-called digital exclusion. According to the definition used by the OECD, the digital divide (digital exclusion) is the difference in access to modern technologies and in the use thereof between persons, households, entrepreneurs and geographic areas at different levels of socioeconomic development (OECD 2001). The problem of digital exclusion has been noted by scientists and government authorities of the USA and the EU in the mid-1990s (Hargittai, 1999).

The concept of the digital divide is linked with the concept of social exclusion. Social exclusion is a multidimensional phenomenon and exceeds the category of poverty, referring also to the non-financial constraints that do not allow the human individual to live at the level acceptable in their country (Panek, 2011; Torraco, 2018). According to the National Strategy of Social Integration for Poland, social exclusion is a situation that prevents or significantly handicaps an individual's or a group's playing a social role within the legal norms, making use of public goods and social infrastructure, gathering resources and earning income in a dignified way (Ministerstwo Gospodarki…, 2004).

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

In the literature, four areas in which an individual may be subject to social exclusion are mentioned (Burchard et al., 2002): consumption (the individual is subject to exclusion because of a low level of income), production (an individual is subject to exclusion because of unemployment and lack of opportunities to raise qualifications and find a job), political commitment (an individual is subject to exclusion due to limited active or passive electoral rights) and social integration (an individual is subject to exclusion because of no contacts with other members of the society). In accordance with this approach, digital divide could result in social exclusion in three of the four listed aspects (consumption, production and social integration). No access to the Internet does not cause true restrictions on electoral rights, but it significantly impedes access to the information needed to make an informed political choice.

The digital divide is usually considered is on two levels (Zhao et al., 2014). On the basic level the participation of citizens and businesses in the information society depends on access to information and communication technologies (ICT), i.e., the presence of electronic devices such as computers and the Internet. The digital divide in this aspect is the so-called first order effect implying inequalities in access to technology. The second order effect is the inequality in the use of technology among people who have access to the Internet (Helbig et al., 2009). It is believed that the essence of digital divide does not apply to the use of the Internet only, but rather opportunities flowing from possibilities to participate in social and cultural life and access to educational resources and the labour market (Batorski and Płoszaj, 2012).

The purpose of the article is to know the trends concerning the Polish households' access to information and communication technologies. The research issues include the analysis of access of Polish households to a computer with Internet connection (including broadband) in 2013-2016. Subsequently, an attempt to identify the factors affecting the access of the Polish households to the tools for acquiring information was made. In the study, such characteristics of the households were included as: the biological type of household, socioeconomic groups of households, the level of education of the head of household and the class of the locality where the household is found.

2 Data and research method

For the purposes of this study, individual non-identifiable data from the household budget survey carried in the years 2013-2016 by the Central Statistical Office were used. The subject of research were the households, and the object of the study included the equipment of the home with a computer with broadband Internet access (32,786 observations).

Identification of the factors that affect the use of the Internet was conducted using econometric modelling. In view of the fact that the explained variable—having access to a computer with a broadband Internet connection—is dichotomic (adopts two values: $Y = 1$ identifies the households with a computer with an Internet access or $Y = 0$ households without a computer with Internet access) a logit model was used. In the case of this model, depending on certain factors (x_j), probability can be interpreted as the value of the distribution function expressed by the formula (Maddala, 2006):

$$\mathrm{P}(\mathrm{Y_i} = 1) = \frac{\exp(\alpha_0+\alpha_1 x_{i1}+\alpha_2 x_{i2}+\cdots+\alpha_k x_{ik})}{1+\exp(\alpha_0+\alpha_1 x_{i1}+\alpha_2 x_{i2}+\cdots+\alpha_k x_{ik})}.$$

The parameters of the above model are usually estimated using the maximum likelihood estimation, maximizing the logarithm function reliability relative to the parameters of the model using the iterative numerical procedures.

As the explanatory variables, characteristics of the household and the head of the household were assumed:

- quintile group of household's income per capita (five zero-one variables; the reference group was the first quintile group),
- class of place of residence (four zero-one variables; the reference group were households in the countryside),
- education (three zero-one variables; the reference was middle school education or less),
- biological type of household (two zero-one variables; the reference group were households without children),
- socio-economic group of households (five zero-one variables; the reference was the group of households of employed people).

In order to match the model, the McFaddenR^2 formula was applied (Gruszczyński, 2012):

$$\mathrm{McFaddenR^2} = 1 - \frac{lnL_{fit}}{lnL_0}.$$

where: lnL_{fit} is a reliability function of the full model, and lnL_0 is the logarithm of the model reliability function where only a constant term occurs. If the model perfectly forecasts the variable then $lnL_0 = 0$, therefore $\mathrm{McFaddenR^2} = 1$. In practice, however, R^2 McFadden values are small, closer to 0 than 1 (Gruszczyński, 2012).

3 Internet access in Polish households

In 2017, 81.9% Polish households with at least one person aged 16-74 had access to an Internet connection (Central Statistical Office, 2017). This percentage was higher by 1.5

percentage points than in the previous year and 10 percentage points compared to 2013 (see Fig. 1). In comparison with 2016 the share of households using a broadband Internet connection increased by 1.9 percentage points. Internet access both in general and broadband was varied, depending on the type of household, the class of its location, the degree of urbanization and also the part of Poland where it is found.

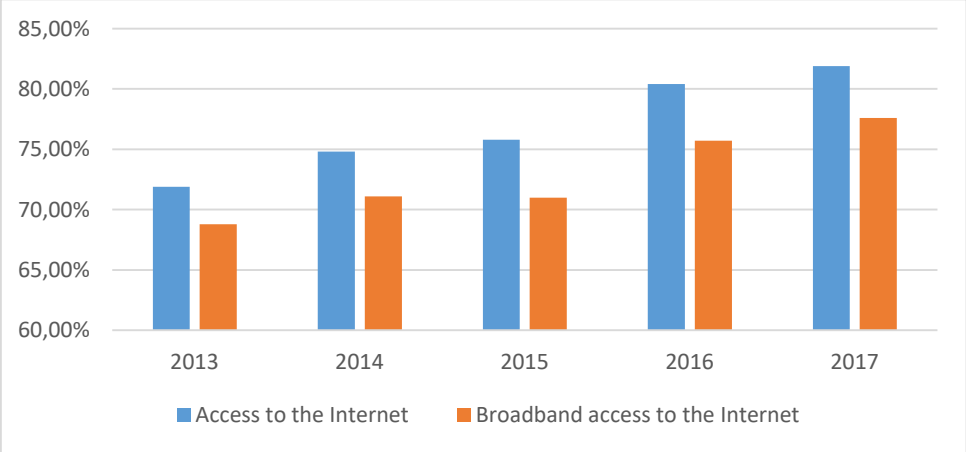


Fig. 1. Households with access to the Internet and broadband access to the Internet.

Among households with access to the Internet, households with children dominated (see Table 1). Taking into account the class of the location, more households in urban areas had Internet access than in rural ones. Due to the region, the most households with Internet access were located in the central regions of Poland. The level of urbanization implied that more households had access to the Internet in areas with higher population density. The same situation concerned access to a broadband connection.

The reasons for not having Internet most often included the lack of need to use it (70.6%). Another important reason was the lack of appropriate skills (see Table 2). Frequently mentioned reasons also pointed out that the of access and hardware costs were too high. The drop of the share of households that do not have Internet because of their technical capabilities can be regarded as a positive change. Meanwhile, the share of households that point to safety reason for not having Internet has risen.

Table 1. Households with access to the Internet at home.

Specification	**Access to the Internet**		**Broadband access to the Internet**	
	2013	**2016**	**2013**	**2016**
Total	71.9	80.4	68.8	75.7

Household type				
Households with children	93.1	97.7	89.9	92.8
Households without children	61.2	71.9	58.1	67.2
Domicile				
Large cities	76.9	82.9	75.1	79.1
Small cities	70.8	80.6	68.0	76.6
Rural areas	67.8	77.8	63.0	71.3
Degree of urbanisation				
Thinly-populated	67.4	79.0	62.7	73.3
Intermediate	71.0	79.6	68.2	74.3
Densely-populated	76.4	82.2	74.6	78.7
Regions				
Eastern Poland	70.6	78.0	65.7	74.8
Central Poland	73.2	81.2	70.3	75.9
Western Poland	70.1	81.0	67.9	76.0

Table 2. Households without access to the internet by reasons for not having access to the Internet (in % of households without access to the Internet).

Reason	2013	2016
No need	64.9%	70.6%
Lack of skills	35.8%	52.0%
Equipment costs too high	28.0%	28.0%
Access costs too high	21.9%	21.3%
Have access to the Internet elsewhere	3.8%	3.6%
Reluctance to the Internet	4.7%	9.9%
Lack of technical possibility to connect to the Internet	1.8%	0.8%
Security concerns	1.4%	3.9%

4 Determinants of Internet access of Polish households

The analysis of individual data from 2016 year showed that having a computer with Internet access is associated with the financial situation of the household, as illustrated in Fig. 2. The empirical distribution curse of the households with a computer and Internet access is visibly shifted right relative to the households with no computer with an Internet access. Average

income per capita in the households with a computer with Internet access is about 24% higher than in those that do not possess such equipment.

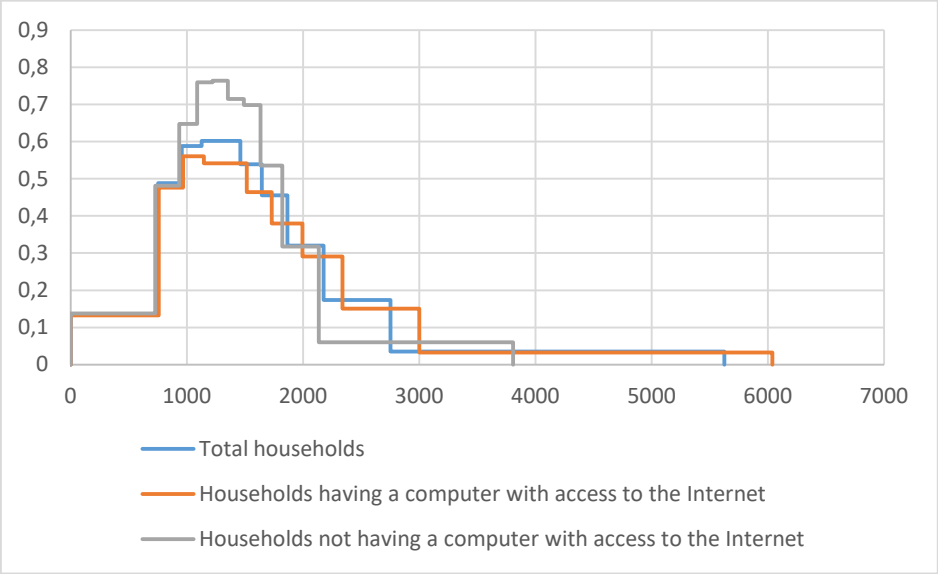


Fig. 2. Empirical distributions of household's income per capita in general, households with a computer with Internet access and those which do not have a computer with Internet access from 2016.

Interestingly, the analysis of access to the Internet in quintile groups of income per capita indicates that the percentage of households having access to the Internet is higher in the first quintile group than in the second and third (see Fig. 3). It is worth noting, however, that in the fifth quintile group the share of households having access to the Internet is higher by more than 15 percentage points than in the first group.

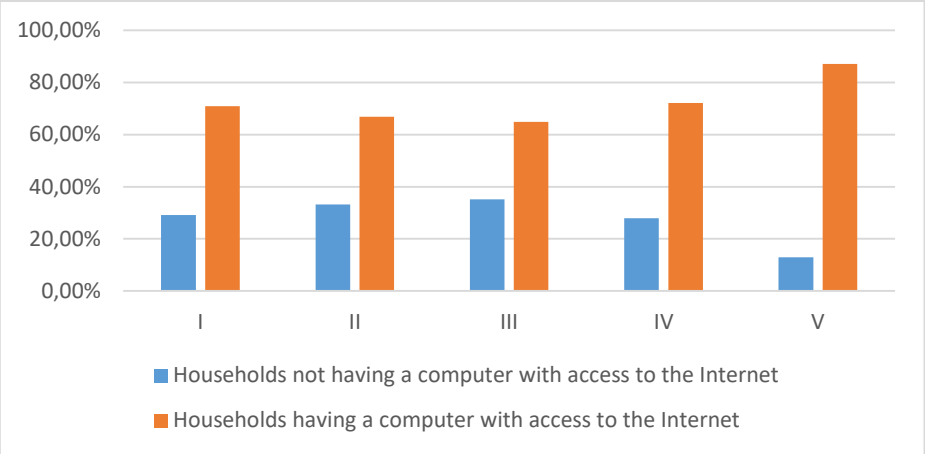


Fig. 3. Households by equipping a computer with internet access in quintile groups of household's income per capita.

In order to extract the socioeconomic qualitative factors, logistic regression was carried out. The information contained in Table 3 shows that most of the proposed variance significantly affect the likelihood of having a computer with an Internet access at home. The McFadden determination coefficient indicating the quality of the fit of the binominal model to the data should not be interpreted as R^2 for the linear model, only in accordance with its definition. Taking into account the number of observations and the nature of the data, the obtained factor level it can be considered satisfactory (Gruszczyński, 2012).

The likelihood of having the equipment necessary for the use of information and communication technologies is the most strongly influenced by family type. With the increase in class of the place of residence and income per capita, the threat of the digital divide is diminishing. Those most at risk are the households belonging to retirees and other pensioners, farmers and people living off various benefits. It is a little surprising that the education of the head of the household turned out to be statistically insignificant.

The results are partly consistent with those of other studies found in the literature. Social and economic factors were identified as the main predictors of the digital divide at work (Kiiski et al., 2005). Other studies of the determinants of social exclusion point to the role of GDP per capita, telecommunications infrastructure and the quality of regulations (Chinn and Fairlie, 2006) and gender (Dixon et al., 2014).

Table 3. Evaluation of the parameters of the logistic model of equipment of Polish households with computers with Internet access.

Specification	Parameter	Standard error	Wald statistics	p-value
Constant	0.3589	0.0532	45.48	0.0000
Households with children	2.4437	0.0522	2195.87	0.0000
Cities with population over 500 000	0.5640	0.0543	107.73	0.0000
Cities with population between 100 000 and 499 000	0.5106	0.0465	120.42	0.0000
Towns with population under 100 000	0.3177	0.0382	69.25	0.0000
Household of farmers	-0.4909	0.0824	35.50	0.0000
Household of pensioners	-1.8019	0.0358	2532.71	0.0000
Household maintained from non-earned sources	-1.0277	0.0740	192.86	0.0000
Secondary education	-0.0148	0.0217	0.47	0.4952
Higher education	0.0443	0.0422	1.11	0.2928
Second quintile group	0.2914	0.0519	31.54	0.0000
Third quintile group	0.5829	0.0520	125.83	0.0000
Fourth quintile group	0.4777	0.0448	113.60	0.0000
Fifth quintile group	1.2664	0.0516	602.63	0.0000
McFaddenR^2=0,3191, Chi^2(14)=12308, p=0.0000				

Conclusions

Full participation in the information society is not possible without access to the Internet and digital skills at the appropriate level. However, it is the Internet access that is the primary condition to acquire skills in information and communication technologies and their use. In Poland, despite improvements in recent years, ca. 20% of households still do not have access to the Internet. The most at risk are for retirees and other pensioners, as well as other individuals who live off various benefits as well as households located in rural areas.

The research conducted shows the need to increase access to the Internet in rural areas and in households of pensioners. While in some rural areas lack of Internet access can be a problem related to the lack of infrastructure, in the households of pensioners the lack of access to the Internet implies (at least partly) a lack of skills needed to use the Internet. To reduce the risk of digital divide the elderly should be enabled to participate in courses in

computer and Internet skill development. Due to the social importance of the problem of digital divide, the analysis of the issue of access to the Internet and the level of digital literacy should be continued in the coming years.

Acknowledgements

The publication was funded by appropriations of the Faculty of Management at the University of Economics in Krakow, as a grant to maintain the research capacity.

References

Batorski, D. & Płoszaj, A. (2012). *Diagnoza i rekomendacje w obszarze kompetencji cyfrowych społeczeństwa kompetencji cyfrowych społeczeństwa i przeciwdziałania wykluczeniu cyfrowemu i przeciwdziałania wykluczeniu cyfrowemu w kontekście zaprogramowania wsparcia w latach 2014-2020*. Warszawa.

Central Statistical Office, (2017). *Information society in Poland. Result of statistical surveys in the years 2013-2017*. Warszawa, Szczecin.

Chinn, M. D. & Fairlie, R.W. (2006). ICT use in the developing world: an analysis of differences in computer and internet penetration. *IZA Discussion Paper,* 2206.

Dixon, L. J., Correa, T., Straubhaar, J., Covarrubias, L., Graber, D., Spence, J. & Rojas, V. (2014). Gendered Space: The Digital Divide between Male and Female Users in Internet Public Access Sites. *Journal of Computer-Mediated Communication, 19* (2014), 991–1009.

Gruszczyński, M. (2012). *Mikroekonometria: Modele i metody analizy danych indywidualnych*. Warszawa: Wolters Kluwer Polska.

Hargittai, E. (1999). Waving the western web: explaining the differences in internet connectivity among OECD countries. *Telecommunications Policy, 23*, 701-718.

Helbig, N., Gil-García, R. J. & Ferro, E. (2009). Understanding the complexity of electronic government: Implications from the digital divide literature. *Government Information Quarterly, 26*(1), 89–97.

Kiiski, S. & Pohjola, M. (2002). Cross-country diffusion of the internet. *Information Economics and Policy*, *14*(2), 297-310.

Maddala, G. S. (2008). *Ekonometria*. Warszawa: Wydawnictwo Naukowe PWN.

Ministerstwo Gospodarki, Pracy i Polityki Społecznej, (2004). *Narodowa Strategia Integracji Społecznej dla Polski*. Warszawa.

OECD, (2001). *Understanding the digital divide*. Paris.

Panek, T. (2011). *Ubóstwo, wykluczenie społeczne i nierówności. Teoria i praktyka pomiaru*. Warszawa: Oficyna Wydawnicza Szkoła Główna Handlowa w Warszawie.

Torraco, R. (2018). Economic Inequality, Educational Inequity, and Reduced Career Opportunity: A Self-perpetuating Cycle?. *New horizons in adult education and human resource development, 30*(1), 19-29.

Zhao, F., Collier, A. & Deng, H. (2014). A multidimensional and integrative approach to study global digital divide and e-government development. *Information Technology & People, 27*(1), 38-62.

Sources of real exchange rate variability in Poland – Evidence from a Bayesian SVAR model with Markov Switching Heteroscedasticity

Marek A. Dąbrowski[1], Łukasz Kwiatkowski[2], Justyna Wróblewska[3]

Abstract

This paper investigates the sources of real exchange rate fluctuations in Poland in 2000-2016. The objective is to assess the relative importance of cost, demand and monetary shocks in driving the exchange rate in Poland. A two-country and two-good New Keynesian open economy model as developed by Engel and West (2006) is used as a theoretical framework. A Bayesian SVAR model with Markov switching heteroscedasticity is used in empirical part. The structural shocks are identified on the basis of the changes in volatility and named with reference to the sign restrictions derived from the economic model. We identify two regimes/states: one with high volatility and the other featuring low volatility. Estimated impulse response functions are in line with the theoretical model though uncertainty is rather large. The main finding is that the contribution of cost and demand shocks to exchange rate variability is about 50 percent in normal times implying that the flexible exchange rate acts as a shock absorber. In turbulent times monetary shocks dominate the exchange rate variability, but their impact on a real economy is short-lived. This undermines the claim that the exchange rate flexibility is dangerous due a shock-propagating nature of the exchange rate.

Keywords: *real exchange rate, open economy macroeconomics, Bayesian MS-VAR models*
JEL Classification: F31, F41; C11
DOI: 10.14659/SEMF.2018.01.10

1 Introduction

Poland is a small open economy with flexible exchange rate regime. It is open both to trade and financial flows: openness to trade as measured by a ratio of the sum of exports and imports to GDP was more than 100 percent in 2016 and the capital openness was above the world median in 2014 (as measured by the Chinn-Ito index; see Chinn and Ito, 2008). According to the IMF's exchange-rate-regime classification, the Polish currency freely floats since April 12, 2000.

The objective of this research is to examine the relative importance of cost, demand and monetary shocks in driving the exchange rate in Poland, and thus to provide evidence that is useful in assessing the desirability of the exchange rate flexibility. Existing evidence is far

[1] Corresponding author: Cracow University of Economics, Department of Macroeconomics, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: marek.dabrowski@uek.krakow.pl.
[2] Cracow University of Economics, Department of Econometrics and Operations Research, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: lukasz.kwiatkowski@uek.krakow.pl.
[3] Cracow University of Economics, Department of Econometrics and Operations Research, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: eowroble@cyf-kr.edu.pl.

from unambiguous. In some studies nominal shocks are found to be the main driver of exchange rate flexibility (see, e.g., Shevchuk, 2015), but in others real shocks are found to be the primary source of exchange rate variability (see, e.g., Arratibel and Michaelis, 2014; Dąbrowski and Wróblewska, 2016). Yet another view was expressed by Alexius and Post (2008) who found that 'exchange rates display some stabilizing properties but can mainly be characterized as disconnected from the rest of the economy.'[4] We use a Bayesian Markov Switching Heteroskedastic VAR model to identify shocks hitting an economy. Overidentifying sign restrictions derived from a New Keynesian small open economy model are imposed to obtain economically interpretable shocks.

2 Theoretical framework and data

A small open economy model is used as a theoretical model. It is a two-country and two-good New Keynesian model developed by Engel and West (2006). The model consists of four main equations: interest rate rule, Phillips curve, IS relation and uncovered interest rate parity condition. It is assumed that the domestic central bank cares about the exchange rate stability, so the exchange rate term enters the interest rate rule. There are three exogenous stationary AR(1) disturbances: demand, cost and monetary that are driven by respective structural shocks. Each variable is defined as a difference between a domestic variable and a foreign one. This allows us to focus on asymmetric shocks, i.e. the ones that matter when the usefulness of the exchange rate flexibility is to be assessed (when shocks are symmetric the flexible exchange rate is of limited use). As pointed out by Artis and Ehrmann (2006), the drawback of such an approach is that it 'yields no information on the comparative frequency of symmetric and asymmetric shocks' and only the latter 'necessitate exchange rate adjustments.' The approach we follow, however, is much more parsimonious, well-established in the literature and corresponds to the nature of the exchange rate which is itself a relative variable.[5]

A system of stochastic difference equations yields the solution that allows us to determine the direction of the short-term reactions of three main endogenous variables to structural shocks. The relative output gap reacts positively to a demand shock and negatively to cost and monetary shocks. All shocks have a positive impact on the real exchange rate (an increase

[4] The result, however, was obtained for advanced small open economies, Canada and Sweden.
[5] Moreover, when a shock is symmetric, a flexible exchange rate can be useful if domestic and foreign preferences with respect to its accommodation are different.

corresponds to an appreciation of domestic currency). The relative inflation gap's response to both demand and cost shocks is positive, but it is negative to a monetary shock.

We use quarterly data for Poland and the euro area (EA-12) spanning the period 1995q1-2017q2. The data for the underlying variables were retrieved from the Eurostat database and include: real GDP, implicit GDP deflator, average euro/ECU exchange rate and harmonised index of consumer prices (HICP). These data are used to construct three main endogenous variables. The relative output gap is a difference between domestic and foreign output gap and each gap is a cyclical component of the log of relevant GDP identified with the HP filter. The real exchange rate (RER) is constructed on the basis of the (index of the) nominal exchange rate and (seasonally adjusted) HICP indices for Poland and the euro area and then a cyclical component of the log of RER is extracted with the HP filter. The relative inflation gap is a difference between domestic and foreign inflation gaps and each gap is a cyclical component of relevant QoQ inflation (based on the implicit GDP deflator) identified with the HP filter.

3 Methodology

We base our empirical analysis upon Bayesian VAR(2) model with Markov switching heteroscedasticity (henceforth MSH-VAR(2)). As we also take into account the possibility of cointegration, the basic model is presented in the VEC form:

$$\Delta y_t = \alpha\beta^{'} y_{t-1} + \Gamma_1 \Delta y_{t-1} + \nu + \varepsilon_t, \quad \varepsilon_t | S_t \sim iiN\left(0, \Sigma_{S_t}\right), \quad t = 1,2,\dots,T, \tag{1}$$

where $y_t = [Output_GAP_t \quad Ex_Rate_t \quad Inf_GAP_t]'$ represents the column vector of the endogenous variables introduced in the previous section, and $\{S_t,\ t \in \mathbb{Z}\}$, $S_t \in \{1,2\}$, is a two-state homogenous and ergodic Markov chain with transition probabilities denoted by p_{ij} ($\forall_{t\in\mathbb{Z}}\, P(S_t = j | S_{t-1} = i) = p_{ij},\ p_{i1} + p_{i2} = 1,\ p_{ij} \in (0,1),\ i,j \in \{1,2\}$). Matrices α and β collect the adjustment coefficients and the cointegration vectors, respectively.

For the covariance matrices we impose the inverted Wishart priors $\Sigma_1 \sim iW(10S_\Sigma, n+11)$, $\Sigma_2 \sim iW(5S_\Sigma, n+11)$, where hyperparameter S_Σ is obtained fa training sample covering the period 1996q2-1998q3. The difference between the priors assumed for Σ_1 and Σ_2 reflects our prior belief according to which the second-state variance is lower than in the first state. The joint prior for α, Γ_1 and ν is matrix normal:

$$(\alpha, \Gamma_1, \nu)^{'} \sim mN\left(0, I_3, \begin{pmatrix} 0.01I_3 & 0_{3\times4} \\ 0_{4\times3} & 0.5I_4 \end{pmatrix}\right).$$

Finally, for the transition probabilities we impose uniform distribution $p_{ii} \sim U(0,1) = Beta(1,1), i = 1,2$. It has turned out that the analysed data may be treated as a realisation of covariance stationary three-dimensional VAR(2) process, so we assume $\beta = I_3$. The joint prior distribution is truncated by the stability condition. The states are identified by the assumption that the conditional volatility of exchange rate in the first state is higher than in the second one.

The reduced-form errors (ε_t) are linear combinations of the structural shocks (denoted as u_t), i.e. $\varepsilon_t = Bu_t$. To identify the structural shocks we employ the method proposed by Lanne et al. (2010), in which we make use of the changes in covariance matrices (Σ_1, Σ_2) of the reduced-form shocks. In the two-state case there always exist matrices B and Λ_2 such that

$$\Sigma_1 = BB', \qquad \Sigma_2 = B\Lambda_2 B', \qquad \Lambda_2 = diag(\lambda_{21}, \lambda_{22}, \lambda_{23}). \tag{2}$$

Lanne et al. (2010) show that this decomposition is locally unique, i.e. up to changing the signs and ordering of B's columns, if the diagonal elements of Λ_2 are all distinct. Using these results the structural shocks defined as $u_t = B^{-1}\varepsilon_t$ are locally identified, so to obtain global identification we additionally impose the shocks' order and their signs by assuming the descending order for $\lambda_{2i}, i = 1,2,3$, and $B_{11} < 0$, $B_{22} > 0$, $B_{33} < 0$, with B_{ii} (i = 1, 2, 3) denoting the i-th diagonal element of B. The three latter inequalities, constituting normalisation restrictions, are set according to the economic model underlying this study, and are also further employed in another MSH-VAR(2) model, which we discuss below. It is worth emphasising that structural shocks obtained in this framework feature switching (therefore, time-varying) covariance matrices (I_3 in the first state and Λ_2 in the second one), but their impacts remain unchanged. Also, since Λ_2 is diagonal, the structural shocks are orthogonal also in the second state, with $\lambda_{2i}, i = 1,2,3$, being their variances. Finally, since $Var(u_{ti}|S_t = 1, \theta) = 1$ and $Var(u_{ti}|S_t = 2, \theta) = \lambda_{2i}$, each λ_{2i} may be regarded a relative change of u_{ti}'s variance in the second state with respect to its (unit) value in the first regime.As mentioned above the structural shocks can be statistically identified if their variances in the second state are all distinct. Therefore, in Table 1 we inspect the posterior quantiles for $\lambda_{2i}, i = 1,2,3$, and their differences.

It can be noticed that the 90%-credible intervals of variances overlap and the posterior marginal densities of their differences are not well-separated from zero, so the conditions for local identification appear not fulfilled. For this reason we have decided to facilitate the identification of structural shocks in our study by introducing sign restrictions (posited by the

economic model) for some other (rather than only diagonal) elements of *B*. Such an approach, formally resulting in a different Bayesian model (further referred to as the second model), combines the purely statistical procedure delivered by the model we discussed hitherto, with one of the most common approach to shock identification by sign restrictions (in accordance with the economic model). We refer the reader to Herwartz and Lütkepohl (2014) for a broader discussion on combining identification by heteroscedasticity with typical short-run restrictions, though only in the frequentist (rather than Bayesian) setting.

Table 1. Posterior quantiles of the structural shocks' variances in the second state.

probability	**0.05**	**0.16**	**0.5**	**0.84**	**0.95**
λ_{21}	0.408	0.507	0.716	1.030	1.325
λ_{22}	0.131	0.177	0.283	0.429	0.547
λ_{23}	0.047	0.060	0.086	0.124	0.157
$\lambda_{21} - \lambda_{22}$	0.110	0.203	0.415	0.727	1.009
$\lambda_{22} - \lambda_{23}$	0.047	0.087	0.188	0.336	0.456
$\lambda_{21} - \lambda_{23}$	0.310	0.412	0.625	0.941	1.237

As regards our particular choice of sign restrictions for the elements of *B* (i.e. instantaneous effects of the shocks upon the endogenous variables), we confronted their posterior distributions in the model discussed above (results available upon request) with reactions predicted by the economic theory, and it enabled us to name the shocks and then to combine both types of identification restrictions. Ultimately, we constrained only (immediate) reactions to the cost and monetary shocks, corresponding to the first and third column of *B* (with imposing the same signs upon B_{11}, B_{22} and B_{33} as previously). Fig. 1 displays the priors and posteriors of *B*'s elements in the resulting (second) model, with sign restrictions (postulated by the theory) indicated in parentheses (in cases when the sign is not actually imposed in the model) and brackets (when the sign is imposed in the model). It is also worth mentioning that the posterior distributions of *B*'s elements in the previous model, i.e. the one without additional sign constraints, tend to feature slight bimodality, with, and most importantly, the prominent mode in each distribution being supported by the economic model (results available upon request). Introducing additional restrictions removed the bimodalities almost completely, perhaps with a sole exception of B_{12} (see Fig. 1).

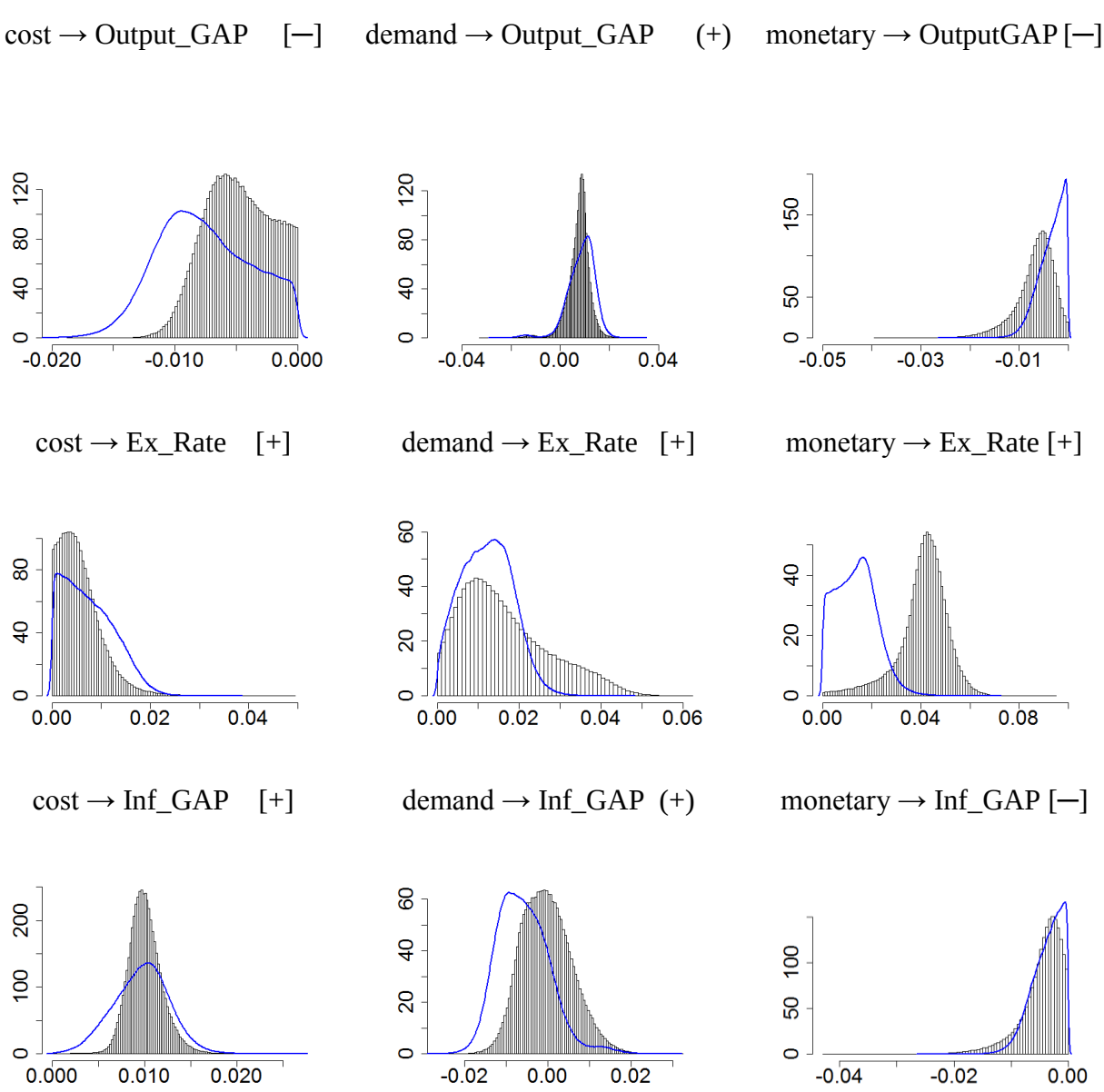


Fig. 1. Marginal posteriors (histograms) and priors (lines) of the elements of the instantaneous reactions matrix (*B*) obtained within the second model.

4 Empirical results

In this section we focus only on the results obtained within the second model, i.e. the one with additional sign restrictions imposed. Conventional tools of structural vector autoregressive analysis were used, i.e. structural impulse responses and forecast error variance decompositions. Our research question is whether fluctuations in the exchange rate in Poland were related to changes in the real economy or were driven by volatility in financial markets. In the former case the exchange rate flexibility can be considered a useful adjustment

mechanism, whereas and in the latter a factor that contributes to the transmission of undesired financial volatility into the real economy.

In Fig. 2 the impulse response functions of the real exchange rate to identified shocks over five-year period are illustrated (results for other variables are available upon request). Solid lines display the posterior mean reactions, whereas and broken lines represent the intervals of plus/minus two standard deviations. All reactions are positive and rather smooth: after a shock the output gap, real exchange rate and inflation gap gradually return towards their long-run paths. In principle, the IRFs obtained are consistent with the model of a small open economy, although the degree of uncertainty if rather large. It is worth emphasizing that sign restrictions are imposed only on the instantaneous reactions (see Fig. 1) and all other (i.e. the lagged) reactions are left unrestricted, so the trajectories of the IRFs are not determined a priori. The IRFs to demand and cost shocks are in line with the shock absorbing property of the flexible exchange rate. In the wake of a positive demand shock the aggregate demand and output increase, but their reaction is mitigated by a real appreciation of the domestic currency. Interestingly, the inflation gap remains almost unchanged, which means that the adjustment process is through the exchange rate changes. This could be an important advantage if prices are sticky. A positive cost shock raises inflation gap and thus induces a real appreciation of the domestic currency. These changes, together with a rise in the real interest rate due to the central bank's focus on price stability, contribute to a decrease in the aggregate demand and output. Since the shock is temporary, the change in the real exchange rate will fade away in the long term (see also Fig. 2). This, however, means that the domestic currency needs to depreciate in nominal terms both in the short and long run. Without the exchange rate flexibility the nominal depreciation would not be possible, and for the real exchange rate to return to its long-run path a period of low domestic inflation or even deflation would be required.[6] Interestingly, the real exchange rate is found to be back to its long-term path faster than both the output and inflation. The implication is that the *nominal* exchange rate flexibility remains important beyond the horizon of the adjustment of the *real* exchange rate. A positive monetary policy shock results in a strong real appreciation of the domestic currency which strengthens an adverse impact of the shock on the output. Moreover, the reaction of inflation is almost non-existent, so the induced appreciation does not seem to be conducive to a reduction of inflation (through the exchange rate channel of monetary

[6] More precisely, the domestic inflation would have to be lower than the foreign inflation. This, however, given that the inflation abroad were low, would indeed require a domestic *deflation*.

transmission). Thus, the IRFs to a monetary shock are consistent with the claim that the exchange rate propagates a shock to the real economy.

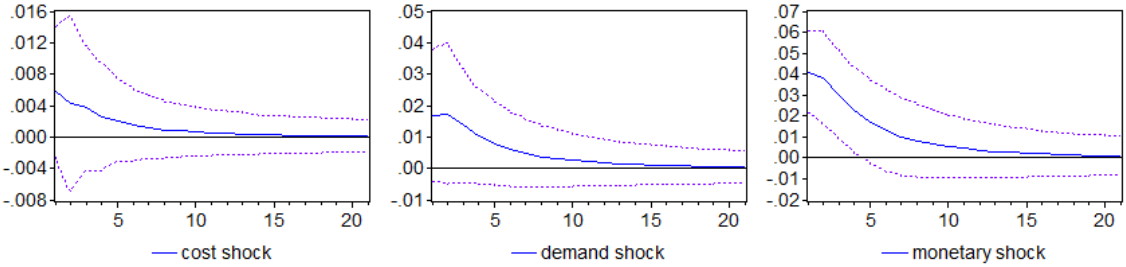


Fig. 2. Impulse response functions of the real exchange rate in Poland.

Table 2. Posterior means of forecast error variance decomposition of the relative output and real exchange rate and relative inflation in Poland.

Forecast	Shocks in turbulent times			Shocks in normal times		
horizon	cost	demand	monetary	cost	demand	monetary
			Relative output gap			
0	17.5	49.0	33.5	35.7	52.0	12.3
20	12.3	36.1	51.6	32.0	43.7	24.3
			Real exchange rate			
0	2.4	18.9	78.7	11.7	31.5	56.8
20	2.5	20.8	76.7	12.4	33.9	53.7
			Relative inflation gap			
0	69.4	16.1	14.5	87.6	9.5	2.9
20	33.8	22.6	43.6	64.7	20.0	15.3

Before the importance of shocks is assessed, the two volatility regimes need to be discussed. Our empirical approach enables us to identify two distinct states: one with high volatility and the other with low volatility. The former, which we refer to as ‘turbulent times,’ comprises (the posterior probability of the regime above 0.5): (1) the EU pre-accession period (2000q1-2005q2), (2) global financial and economic crisis (2007q4-2010q1), and (3) the debt crisis in the euro area (2011q3–2012q2). Interestingly, our findings are consistent with the evolution of the common measure of volatility, i.e. the VIX index (results available upon

request). The second regime encompasses the remaining periods and is henceforth labelled as 'normal times.'

In order to assess the contribution of shocks to the variability of output, real exchange rate and inflation, we analyse the forecast error variance decomposition. Table 2 reports the results obtained for two forecast horizons: zero quarters (the short run) and twenty quarters (the long run). The results depend on the volatility regime. During the normal times the relative output gap is driven mainly by demand shocks, the relative inflation gap by cost shocks, and the real exchange rate by monetary shocks. In turbulent times a contribution of monetary shocks is much larger than in tranquil times, whereas that of cost shocks decreases substantially. Two more important observations can be made with reference to the usefulness of exchange rate flexibility. First, even though the variability of the real exchange rate is mainly driven by monetary shocks, the contribution of the two other shocks is close to 50 percent in normal times. Bearing in mind that these shocks account for 75-90 percent of output variability, one can argue that the flexible exchange rate acts as a shock absorber. Second, in turbulent times, when monetary shocks become more prevalent, the exchange rate acts more like a shock propagator. The cost and, in particular, demand shocks remain, however, important sources of both exchange rate variability (20-25 percent) and output variability (50-65 percent). Therefore, in turbulent times, the shock-absorbing property of the flexible exchange rate is limited rather than non-existent. Moreover, the reaction of the output to a monetary shock, as measured by the IRFs, is rather short-lived (after six quarters it peters out), so the shock-propagating property of the flexible exchange rate should not be overstated.

Conclusions

Our main findings can be summarised as follows. First, monetary shocks account for about 50 percent of exchange rate variability, but the contribution of cost and demand shocks is large, especially in normal times. Together with the theory-consistent trajectories of impulse response functions, the result lends support to the hypothesis that the flexible exchange rate in Poland acted as a shock absorber. Second, in turbulent times the relative importance of monetary shocks is much higher than in tranquil times. We abstain, however, from interpreting this finding as an argument in favour of the hypothesis that the exchange rate acted as a shock propagator, because the impact of a monetary shock on the real economy was found to be rather short-lived, unlike that of either cost or demand shock. We see two things that can be done to strengthen the results. First, our approach needs to be extended to allow for identification of financial shocks that are presumably hidden behind monetary shocks.

Second, in order to establish more convincing insights into the role of the exchange rate, we need to compare countries with fixed and flexible exchange rate. We leave these issues for further research.

Acknowledgements

The authors would like to acknowledge financial support from research funds granted to Faculty of Management and Faculty of Economics and International Relations at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

Alexius, A. & Post, E. (2008). Exchange rates and asymmetric shocks in small open economies. *Empirical Economics*, *35*(3), 527-541.

Arratibel, O. & Michaelis, H. (2014). The impact of monetary policy and exchange rate shocks in Poland evidence from a time-varying VAR. *European Central Bank Working Paper Series*, No. 1636.

Artis, M. & Ehrmann, M. (2006). The exchange rate – A shock-absorber or source of shocks? A study of four open economies. *Journal of International Money and Finance, 25*(6), 874-893.

Chinn, M. D. & Ito, H. (2008). A new measure of financial openness. *Journal of Comparative Policy Analysis*, *10*(3), 309-322.

Dąbrowski. M. A. & Wróblewska, J. (2016). Exchange rate as a shock absorber in Poland and Slovakia: Evidence from Bayesian SVAR models with common serial correlation. *Economic Modelling, 58,* 249-262.

Engel, C. & West, K. D. (2006). Taylor rules and the Deutschemark-Dollar real exchange rate. *Journal of Money, Credit, and Banking*, *38*(5), 1175-1194.

Herwartz H. & Lütkepohl, H. (2014). Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks. *Journal of Econometrics, 183*(1), 104-116.

Lanne, M., Lütkepohl, H. & Maciejowska, K. (2010). Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control, 34*(2), 121-131.

Shevchuk, V. (2015). Flexible Exchange Rates as Shock Absorbers in Central and Eastern Europe. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie*, *5*(941), 17-35.

Is there a trade-off between monetary independence and exchange rate stability in Central and Eastern European economies?

Marek A. Dąbrowski[1], Monika Papież[2], Sławomir Śmiech[3]

Abstract

The paper examines the relation between the exchange rate flexibility and monetary independence in Central and Eastern European (CEE) economies. According to the conventional open economy model one of the important advantages of the floating exchange rate regime is that a country can pursue an autonomous monetary policy. Recently the empirical validity of the trilemma has been questioned: H. Rey argued that the choice is between openness to financial flows and monetary independence, 'irreconcilable duo', no matter which exchange rate regime prevails. We derive a contagion model of monetary policy to examine whether the degree of 'policy contagion' is indeed unrelated to the exchange rate flexibility in CEE economies. Our main findings are that: (1) the spillover effect from the euro area monetary policy is strong for the Czech Republic, but rather weak for Hungary, Poland and Romania; (2) the trilemma monetary policy independence indices are rather crude and their informativeness is substantially limited; (3) the CEE countries, except for the Czech Republic, are likely to maintain their monetary independence in the face of future ECB's exit from the zero-interest-rate policy.

Keywords: *monetary policy autonomy, exchange rate regime, open economy macroeconomics, quantitative policy modelling*

JEL Classification: F41, F33, E44; C54

DOI: 10.14659/SEMF.2018.01.09

1 Introduction

The macroeconomic trilemma links monetary autonomy, exchange rate stability and openness to international financial flows. It is the trilemma, because one can choose two out of three options. Using this insight Aizenman et al. (2013) constructed the indices that measure the trilemma aspects and demonstrated that emerging market economies have retained some degree of monetary autonomy. More recently, however, Rey (2016, 2015) questioned the validity of the trilemma arguing that 'whenever capital is freely mobile, the global financial cycle constrains national monetary policies regardless of the exchange rate regime.' Empirical evidence on the relevance on Rey's 'irreconcilable duo' hypothesis for emerging market economies so far have remained mixed.

[1] Corresponding author: Cracow University of Economics, Department of Macroeconomics, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: marek.dabrowski@uek.krakow.pl.
[2] Cracow University of Economics, Department of Statistics, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: monika.papiez@uek.krakow.pl.
[3] Cracow University of Economics, Department of Statistics, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: slawomir.smiech@uek.krakow.pl.

The objective of this study is to establish whether the exchange rate flexibility contributes to monetary policy autonomy in a small open economy. We examine the case of four Central and Eastern European (CEE) economies. The degree of monetary independence with respect to monetary policy pursued in the euro area is assessed in the period spanning January 2002 to July 2012.

Our study is not the first one to examine monetary independence in the CEE countries. Angeloni et al. (2007) investigated the homogeneity of monetary policy rules of the CEE countries with the euro area and found that rules are broadly similar with some differences in the speed of adjustment only. Unfortunately, they did not include the foreign rate of interest in their regressions, so it was not possible to assess the degree of monetary independence. Goczek and Mycielska (2013), using a VAR approach, found that 'monetary policy in Poland seems to be dependent on the ECB policy.' Using dynamic regressions on country-by-country basis Obstfeld (2015) examined monetary policy independence in 56 countries. Results for the CEE economies were mixed: the long-run relation between domestic and foreign interest rates was found for the Czech Republic and Poland, but not for Hungary and Romania. In a related study Dąbrowski et al. (2015) examined resilience of 41 emerging market economies to the global financial crisis and found that it was not the exchange rate regime *per se* that mattered for the resilience, but the actually adopted monetary policy option.

Table 1. Indices of macroeconomic trilemma in CEE countries, 2002-2012.

Variables	Czech Rep.	Hungary	Poland	Romania
Monetary	0.43	0.56	0.32	0.63
independence	(0.21)	(0.38)	0.61	(0.54)
Exchange rate	0.38	0.29	0.28	0.23
stability	(0.30)	(0.39)	(0.18)	(0.34)
Capital account	1.00	1.00	0.45	1.00
openness	(0.12)	(0.18)	(0.00)	(0.55)
International reserves	0.20	0.19	0.15	0.22
(minus gold, to GDP)	(0.13)	(0.20)	(0.10)	(0.13)

Note: medians; maximum minus minimum in parentheses.

Aizenman et al. (2013) developed indices of macroeconomic trilemma for a large set of countries. Their medians are tabulated for the CEE countries in Table 1 (international reserves-to-GDP ratios are based on data from the World Development Indicators). The index

of monetary independence was calculated on the basis of the correlation of the interest rates between the home country and the base country. The results are far from being unambiguous. For example, Poland had an index of exchange rate stability (0.28) close to that in Hungary (0.29), but substantially less monetary independence (0.32 vs 0.56), even though capital account in Poland was found to be less open than in Hungary (0.45 vs 1.00). More generally, the problem is that the index of monetary independence could be misleading in the face of global shocks (see, e.g. Obstfeld, 2015).

The paper is structured as follows. The following section briefly lays out theoretical issues and describes the data. Empirical methodology is discussed in Section 3, whereas empirical results are reported in Section 4. The final section concludes.

2 The 'contagion' model of monetary policy and data

Dornbusch (1976) demonstrated that the floating exchange rate reaction to foreign monetary policy changes is excessive in the short term in that sense that it overshoots its medium-term level. Since such changes – especially in the face of nominal rigidities – translate into a real economy, the central bank can intend to limit excessive exchange rate volatility. Thus, as pointed by Edwards (2015) it is quite likely that the term in the foreign policy rate will be included in the central bank policy rule.

The line of reasoning can be formalized by reference to the extended monetary policy rule, i.e. the one that includes the term in the exchange rate:

$$i_t = -\alpha(s_{t-1} - E_{t-1}s_t^e) + \gamma' x_{t-1} \quad (1)$$

where i is the domestic rate of interest, s is the actual exchange rate, $E_{t-1}s_t^e$ is the equilibrium exchange rate expected in $t-1$ to prevail in t and x includes other determinants of the interest rate (like deviation of actual inflation from target and/or output gap). Parameters α, γ are positive. Variables on the right hand side of equation (1) are indexed $t-1$ following the assumption that information from the current month is not available to the central bank when making the decision. Thus, we consider the so-called backward-looking interest rate rule. Using the uncovered interest rate parity condition one can demonstrate that

$$i_t = \alpha i_{t-1}^* - \alpha i_{t-1} + \alpha\rho_{t-1} - \alpha E_{t-1}\tilde{s}_t + \gamma' x_{t-1} \quad (2)$$

where i^* is the foreign interest rate, ρ is a foreign currency risk premium, $E_{t-1}\tilde{s}_t$ is the deviation of the actual from equilibrium exchange rate expected in $t-1$ to prevail in t.

In a freely floating exchange rate regime α is zero and the interest rate is set as implied by determinants included in x (γ is statistically different from zero). In a fully credible fixed

exchange rate regime the domestic interest rate is endogenously determined (within the uncovered interest rate parity) by the foreign interest rate and a risk premium only, so other determinants are unimportant ($\gamma = 0$).

Central banks avoid large swings in the interest rate. If we allow for the interest rate smoothing then the equation becomes

$$\Delta i_t = (1-\lambda)[\alpha i_{t-1}^* - (1+\alpha) i_{t-1} + \alpha \rho_{t-1} - \alpha E_{t-1}\tilde{s}_t + \gamma' x_{t-1}] \quad (3)$$

where the degree of interest rate smoothing is measured by the parameter $0 \leq \lambda \leq 1$.

We use monthly data for the four CEE countries: the Czech Republic, Hungary, Poland and Romania, and for the euro area that span the period January 2002 to July 2017. The main variable of interest is the rate of interest. Three-month money market interest rates from the Eurostat database are used. Inflation rates are the CPI changes on annual basis and are from the IMF IFS dataset. Industrial production indices from the Eurostat database are used to proxy for the output gap. Broad indices of the nominal effective exchange rates (NEER) from the Bank of International Settlements are used to derive the deviation of actual from the equilibrium level. The latter is proxied as a cyclically-adjusted component of the NEER (the Hordick-Prescott filter was used). The foreign risk premium is proxied with the Chicago Board Options Exchange's equity option volatility index (VIX) obtained from the Federal Reserve Economic Data.

3 Methodology

The equation we derived in Section 2 is similar to the one used in other studies. For example Edwards (2015) also assumed that the central bank adjusted its policy with a lag and did so gradually. Following Edwards (2015) and Obstfeld (2015) we estimate the dynamic regression model of the form

$$\Delta i_t = \beta_0 + \beta_1 i_{t-1}^* + \beta_2 i_{t-1} + \beta_3 \rho_{t-1} + \beta_4 \tilde{s}_{t-1} + \gamma' x_{t-1} + \\ + \delta_1 \Delta i_{t-1}^* + \delta_2 \Delta i_{t-1} + \delta_3 \Delta \rho_{t-1} + u_t \quad (4)$$

where u_t is assumed to be iid white noise process.

Foreign monetary policy spillover in the short run is given by $\beta_1 + \delta_1$. It is quite likely, however, that it takes some time for the full impact of the foreign interest rate change to be transmitted to the domestic rate. Thus, we calculate long-run policy spillover as well – it is $-\beta_1/\beta_2$. In order to assess statistical significance of both policy spillover effects, the delta method is used (Greene, 2018, p. 78-81).

In order to assess the robustness of results on policy spillover effects, the bootstrap approach is used. The main reason for utilizing the bootstrap is violation of normality

assumption in the regression (4) residuals. There are, however, another advantages of bootstrap, which include: better properties in small samples in comparison to asymptotic approach and feasible inference about smooth differentiable functions of regression parameters (see Kilian and Lütkepohl, 2017). The study use residual-based fixed-design bootstrap approach which allows to infer under stationarity condition, even when randomness of the regressors is present.[4] The only condition that has to be fulfilled is that the residuals of equation (4) are iid. Bootstrap samples are created by utilizing resampled residuals $u_t^{\#r}$ of the fitted model, holding the regressor matrix fixed in every sample. Given a sequence of data $\{\Delta i_t^{\#r}\}_{t=1}^{T}$, and original regressors $[i_{t-1}^*, i_{t-1}, \rho_{t-1}, \tilde{s}_{t-1}, x'_{t-1}, \Delta i_{t-1}^*, \Delta i_{t-1}, \Delta\rho_{t-1}]$ new estimates of $B^{\#r} = [\beta_0^{\#r}, \beta_1^{\#r}, \beta_2^{\#r}, \beta_3^{\#r}, \beta_4^{\#r}, \gamma'^{\#r}, \delta_1^{\#r}, \delta_2^{\#r}, \delta_3^{\#r}]^T$ are obtained. The procedure is repeated $N = 10000$ times.

The covariance matrix of vector parameters is estimated using the Monte Carlo approximation (Efron, 1982 , p. 36):

$$\Sigma^{\#} = (N-1)^{-1} \sum_{r=1}^{N} (B^{\#r} - B^{\#})(B^{\#r} - B^{\#})' \tag{5}$$

where $B^{\#r}$ is the *r*th bootstrap estimate for $r = 1,2,\dots,N$ and $B^{\#} = N^{-1} \sum_{r=1}^{N} B^{\#r}$.

4 Empirical results

We present the results obtained for the sample covering period from January 2002 to July 2012. That choice is motivated by three considerations. First, in the period before 2002 the CEE countries had relatively high and volatile inflation. Moreover, they liberalised their capital accounts not earlier than at the beginning of the 21st century, i.e. in the run-up to the EU membership. Both these factors, i.e. high inflation and barriers to capital flows, could hinder the identification of actual importance of the exchange rate flexibility to monetary independence. Second, the choice of July 2012 is motivated by the decrease of the euro area interest rate close to zero lower bound after the 'whatever it takes' speech by the President of the ECB Mario Draghi.[5]

An additional argument behind the choice of the sample period is that CEE countries maintained relatively fixed exchange rate regimes in 1990s when and shifted towards more

[4] Gonçalves and Kilian (2004) suggest that this algorithm is almost as accurate in finite sample as the recursive-design bootstrap for autoregressive processes.

[5] At the Global Investment Conference in London on 26 July 2012 he said 'Within our mandate, the ECB is ready to do whatever it takes to preserve the euro. And believe me, it will be enough' (Draghi, 2012).

flexible exchange rate arrangements at the turn of the centuries. According to the Reinhart-Rogoff classification Hungary widened the band for exchange rate fluctuations in 1999, Poland moved from the crawling band to managed float in 2000, Romania managed to get out from 'freely falling' regime to managed floating in 2001 and only the Czech Republic was classified as a soft pegger with de facto crawling band that was narrower than or equal to +/-2% (Ilzetzki et al., 2017). Moreover, one can observe a shift of these countries to soft peg arrangements at the end of our sample period: Hungary in 2009, Poland in 2012 and Romania in 2007 (and to a de facto peg in 2013).[6]

The main empirical results are presented in Table 2. It includes estimates of coefficients of equation (4). In general, three observations can be made with respect to the results. First, the results are well in line with the conventional interest rate rule. All central banks react positively to output gap (LIPGAP), although the coefficient is insignificant in the Czech case. A rise in inflation (INF) results in a monetary tightening in all CEE countries except for Hungary where the coefficient is negative but statistically insignificant. Second, the deviation of the actual from equilibrium exchange rate (LNERGAP) was highly statistically significant in all countries but Poland. This indicates that the National Bank of Poland was oriented at the exchange rate stability to a considerably smaller extent than the other CEE central banks. Third, a foreign currency risk premium, proxied with the VIX index (LVIXCLS), does not seem to be an important factor behind the interest rate changes as it turned out to be statistically insignificant.

The spillover effect from the euro area monetary policy can be calculated from the regression coefficients for interest rates. The short-run effect is the sum of coefficients on foreign interest rate level and difference (both lagged) and the long-run effect is the ratio of coefficient on (lagged) foreign and domestic interest rates. The results are in the bottom of Table 1. There is no clear pattern in the estimated policy spillover effects. On the one hand a short-run effect is smaller than a long-run one and both work in the expected direction. The only exception is Romania in which there is a negative coefficient on the short-run effect. On the other hand, the long-run coefficients are insignificant (except for the Czech Rep.), which suggests that the transmission of foreign monetary policy is rather short-lived.

[6] For details see Ilzetzki et al. (2017). One should, however, point out that according to the IMF classification or classification developed by Dąbrowski et al. (2017) our CEE countries were floaters for the large fraction of our sample.

Table 2. Interest rate dynamic equations.

Variables	Czech Rep.	Hungary	Poland	Romania
I3_EA(-1)	0.066***	0.065	0.017	0.085
I3_country(-1)	-0.096***	-0.077**	-0.046***	-0.084***
ΔI3_country(-1)	0.030	-0.120*	0.182**	0.483***
ΔI3_EA(-1)	0.181**	0.372	0.352***	-0.933**
INF_country(-1)	0.017**	-0.018	0.043***	0.064***
LNERGAP_country(-1)	-1.765***	-2.732**	-0.471	-4.223**
LIPGAP_country(-1)	0.332	2.849*	1.682***	3.918**
LVIXCLS(-1)	0.029	0.051	-0.031	-0.248
ΔLVIXCLS(-1)	-0.020	-0.158	0.072	0.267
C	-0.073	0.339	0.151	0.852
Obs.	127	127	127	127
Adj. R-squared	0.642	0.629	0.615	0.663
F-statistic	19.827	14.372	21.102	21.687
Durbin-Watson	1.880	2.285	1.920	2.068
Short-run coeff.	0.247***	0.437	0.369***	-0.848*
Long-run coeff.	0.687***	0.843	0.372	1.007

Such results are, to a certain extent, in line with those obtained by Obstfeld (2015). He ran dynamic equations for 56 countries, including our CEE countries, although he used the US interest rate as a foreign rate of interest. He was unable to reject the hypothesis of no long-run relation for Hungary and Romania, but rejected it for the Czech Republic and Poland. Some differences, e.g. with respect to Poland, could be due to the use of different sample period. In fact, Obstfeld (2015) used country-specific samples starting quite early in the 1990s (e.g. July 1991 – February 2014 for Poland), whereas we have used the common sample period that corresponds to the relative exchange rate flexibility.

Our results seem to be at odds with those obtained by Goczek and Mycielska (2013) for Poland. They found that the degree of monetary independence in Poland was rather low. They admitted, however, that the interest rate dependence is not one-for-one and that their approach could ‘understate the actual degree of monetary independence offered by the floating exchange rate,’ for example because both interest rates could be driven by global shocks.

Statistical significance of the results presented in Table 1 was assessed under the assumption of asymptotic normality of residuals. In order to have normally distributed residuals we used a set of dummies in each regression (either dummies or Jarque-Bera statistics not reported due to space constraints, but available upon request). It could, however, be claimed that even though such an approach removes non-normality, it also hides the true distribution of the residuals. In order to check the robustness of our results we run anew all the regressions using the bootstrap procedure to obtain coefficients and their covariance matrix. The results are reported in Table 3.

Table 3. Interest rate dynamic equations – bootstrap approach.

Variables	Czech Rep.	Hungary	Poland	Romania
I3_EA(-1)	0.059***	0.099*	0.019	0.104
I3_country(-1)	-0.093***	-0.073	-0.052***	-0.119***
ΔI3_country(-1)	0.092	-0.060	0.289***	0.313***
ΔI3_EA(-1)	0.387***	0.913*	0.305**	-1.080**
INF_country(-1)	0.018**	-0.008	0.037***	0.094***
LNERGAP_country(-1)	-1.529***	-5.735***	-0.458	-5.662*
LIPGAP_country(-1)	0.298	1.260	1.347**	6.723**
LVIXCLS(-1)	0.041	0.332*	-0.028	-0.360
ΔLVIXCLS(-1)	0.032	0.166	0.077	-0.011
C	-0.099	-0.586	0.187	1.212
Obs.	127	127	127	127
Adj. R-squared	0.512	0.118	0.581	0.264
F-statistic	15.709	2.870	20.401	6.021
Durbin-Watson	1.879	1.968	2.081	2.127
Short-run coeff.	0.446***	1.012**	0.324***	-0.976
Long-run coeff.	0.628***	1.355	0.371	0.872

The general finding is that the results remained unchanged for the Czech Rep., Poland and Romania. Moreover, the puzzling reaction of the Romanian interest rate disappears: although the short-run spillover effect remains negative, it is no longer significant. The differences can be observed for Hungary: terms in the (lagged) foreign interest rate are now weakly

significant, whereas those in the domestic interest rate are insignificant. This results in the strong and significant short-run (over)reaction to the euro area interest rate with the spillover coefficient slightly more than unity. The long-run coefficient, however, is insignificant like in the previous regression.

Conclusions

According to the macroeconomic trilemma the exchange rate flexibility can bring the interest rate independence in the face of free capital movement. The objective of this study was to examine the dependence between interest rates in four CEE countries and the euro area. The main findings can be summarised in three points. First, the relative exchange flexibility of CEE currencies insulated these economies against euro area monetary policy spillovers to a limited extent. There is evidence of strong monetary policy spillover for the Czech Republic, but rather weak or non-existence for Hungary and Romania. Poland is somewhere between these two extremes with moderate spillover in the short run, but not in the long run. Second, the results obtained do not conform to crude monetary independence indices developed by Aizenman et al. (2013), which is in line with the conjecture that the simple correlation is insufficient to describe monetary independence. Third, using historical evidence we think that the CEE countries, except for the Czech Republic, will be able to retain their monetary independence when the ECB will decide to exit its de facto zero-interest-rate policy.

Acknowledgements

The authors gratefully acknowledge financial support from the National Science Centre in Poland (grant no. DEC-2015/17/B/HS4/02681).

References

Aizenman, J., Chinn, M. D. & Ito, H. (2013). The “impossible trinity” hypothesis in an era of global imbalances: Measurement and testing. *Review of International Economics, 21*(3), 447-458.

Angeloni, I., Flad, M. & Mongelli, F. P. (2007). Monetary integration of the new EU Member States: What sets the pace of euro adoption?. *JCMS: Journal of Common Market Studies, 45*(2), 367-409.

Dąbrowski, M. A., Papież, M. & Śmiech, S. (2017). A statistically-based classification of de facto exchange rate regimes. In: Papież M. and Śmiech S. (Eds.), *The 11th Professor*

Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings. Cracow: Foundation of the Cracow University of Economics, pp. 47-56.

Dąbrowski, M. A., Śmiech, S. & Papież, M. (2015). Monetary policy options for mitigating the impact of the global financial crisis on emerging market economies. *Journal of International Money and Finance, 51*, 409-431.

Dornbusch, R. (1976). Expectations and Exchange Rate Dynamics. *The Journal of Political Economy, 84*(6), 1161-1176.

Draghi, M. (2012). Verbatim of the remarks made by Mario Draghi—Speech by Mario Draghi, President of the European Central Bank at the Global Investment Conference in London 26 July.

Edwards, S. (2015). Monetary policy independence under flexible exchange rates: an illusion?. *The World Economy, 38*(5), 773-787.

Efron, B. (1982). *The jackknife, the bootstrap, and other resampling plans* (Vol. 38). Siam.

Goczek, Ł. & Mycielska, D. (2013). Ready for euro? Empirical study of the actual monetary policy independence in Poland. *University of Warsaw Faculty of Economic Sciences Working Papers*, (13), 98.

Gonçalves, S. & Kilian, L. (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics, 123*(1), 89-120.

Greene, W. H. (2018). *Econometric analysis (8th Edition)*. Pearson

Ilzetzki, E., Reinhart, C. M. & Rogoff, K. S. (2017). *Exchange arrangements entering the 21st century: Which anchor will hold?* (No. w23134). National Bureau of Economic Research.

Kilian, L. & Lütkepohl, H. (2017). *Structural vector autoregressive analysis*. Cambridge University Press.

Obstfeld, M. (2015). *Trilemmas and trade-offs: living with financial globalisation.* BIS Working Paper No. 480. Available at SSRN: https://ssrn.com/abstract=2552572.

Rey, H. (2016). International channels of transmission of monetary policy and the Mundellian trilemma. *IMF Economic Review, 64*(1), 6-35.

Rey, H. (2015). *Dilemma not trilemma: the global financial cycle and monetary policy independence* (No. w21162). National Bureau of Economic Research.

Robust estimation of revenues of Polish small companies by NACE section and province

Grażyna Dehnel[1], Łukasz Wawrowski[2]

Abstract

Sample surveys conducted by the Central Statistical Office are currently the main source of information about revenues earned by small companies. Given the sample size, sampling scheme and the estimation method used in the survey, reliable estimates can only be produced for domains at the level of country, province or section of business classification. The market economy, however, creates a demand for local level information about businesses and economic conditions, which is provided on a regular basis at short intervals.

The article describes an empirical study designed to test a small area estimation method. The goal of the study is to apply a robust version of the Fay-Herriot model, which, unlike the classical Fay-Herriot model, makes it possible to meet the assumption of normality of random effects under the presence of outliers. These alternative models will be supplied with auxiliary variables in order to estimate revenues of small businesses (with between 10 and 49 employees). Other sources of data used in the analysis include the DG1 report, Poland's largest enterprise survey, and administrative registers. The study is expected to provide information about patterns and characteristics of the small business sector in Poland for territorial units on low level of aggregation.

Keywords: *small area estimation, indirect estimation, robust Fay-Herriot model, administrative registers, business statistics*

JEL Classification: C13, C51, M20

DOI: 10.14659/SEMF.2018.01.11

1 Introduction

The shape of Poland's modern-day economy is the result of dynamic changes during the period of economic transformation. One of the sectors that plays a significant role in the development of the economy is the sector of small companies (employing between 10 and 49 persons), which currently includes about 57,000 small businesses. Small companies are characterised by a high degree of flexibility, profitability and efficiency of economic activity. There is also a strong correlation between the development of small companies and regional development. This impact can be observed in both directions: a higher level of regional development encourages entrepreneurs to start business activity, at the same time, however,

[1] Corresponding author: Poznań University of Economics and Business, Department of Statistics, al. Niepodległości 10, 61-875 Poznań, g.dehnel@ue.poznan.pl.

[2] Poznań University of Economics and Business, Department of Statistics, al. Niepodległości 10, 61-875 Poznań, lukasz.wawrowski@ue.poznan.pl.

a growth in the number of small companies contributes to the improvement of the region's economic situation (Główny Urząd Statystyczny, 2017).

Taking into consideration the classification of economic activity, manufacturing and trade are the two most important sections. Companies conducting activities classified into these two categories account for 38% of all small businesses, and their revenues make up about 70% of all revenues in this sector. They also provide 50% of jobs that exist in the small business sector (Główny Urząd Statystyczny, 2017). An analysis of financial results of small trading and manufacturing companies is conducted and published by the Central Statistical Office only at country level. However, given the demand for more detailed information expressed by data users, the present article describes a study whose goal was to estimate certain variables at the level of province (NUTS 2). So, the target domain of estimation is province cross-classified by NACE section. Information about net revenues in the domains is not available in official publications of the Central Statistical Office (Dehnel, 2017).

The aim of the study was to estimate two variables: net revenues from the sale of goods and materials (*SH*) and net revenues from the sale of products (*SW*) for companies which employ from 10 to 49 employees. These characteristics were estimated using direct estimates from DG1 survey and auxiliary variables from administrative registers. This study is a continuation of the study described in Dehnel et al. (2017).

The article is divided into three parts. The first one provides a description of the DG1 survey. The second, theoretical part is devoted to the presentation of estimators used in the study. Estimation results are described in the third part. The article ends with conclusions and suggestions for further research.

2 DG1 survey

The study is based on data from the DG1 survey, which is the main source of information about Polish enterprises. The survey includes a 10% sample of small companies (employing more than 10 people), which are asked to complete a questionnaire about basic characteristics of the company.

By applying the Horvitz-Thompson (1952) estimator to DG1 data it is only possible to produce reliable direct estimates at province level or for NACE sections. There is, however, a growing demand for more detailed information about companies' characteristics.

3 Fay-Herriot model and its robust version

The Fay-Herriot model belongs to the class of area-level models, which means that it utilizes aggregated data instead of unit-level information. This approach was developed in 1979 as a tool for estimating income for small areas in the USA (Fay and Herriot, 1979). The construction of a Fay-Herriot model is divided into two stages. Firstly, it is assumed that the direct estimate is unbiased and can be written as the sum of the true value of the estimated parameter and random error:

$$\hat{\theta}_d = \theta_d + e_d . \tag{1}$$

Where $e_d \overset{ind}{\sim} N(0, \sigma_{ed}^2)$. In practice, variance σ_{ed}^2 is unknown and is estimated based on survey data.

In the second stage, the true value of the parameter is treated as a dependent variable in the linear model with area random effect:

$$\theta_d = x_d^T \beta + u_d \tag{2}$$

where x_d is a vector of auxiliary information for area d, β is a vector of regression parameters and u_d is area random effect with distribution $u_d \overset{iid}{\sim} N(0, \sigma_u^2)$.

By combing equations (1) and (2) we obtain the Fay-Herriot model given by:

$$\theta_d = x_d^T \beta + u_d + e_d \tag{3}$$

EBLUP (Empirical Best Linear Unbiased Predictor) is the estimator of the Fay-Herriot model and is given by the following formula:

$$\hat{\theta}_d^{FH} = x_d^T \hat{\beta} + \hat{u}_d = \hat{\gamma}_d \hat{\theta}_d + (1 - \hat{\gamma}_d) x_d^T \hat{\beta} \qquad d = 1,.....D \tag{4}$$

where $\hat{\beta} = \left(\sum_{d=1}^{D} \hat{\gamma}_d x_d x_d^T \right)^{-1} \sum_{d=1}^{D} \hat{\gamma}_d x_d \hat{\theta}_d$ and $\hat{\gamma}_d = \dfrac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_{ed}^2}$.

EBLUP is a weighted average of the direct estimate and the regression model. Weight $\hat{\gamma}_d$ measures the uncertainty of the regression model. If sample variance $\hat{\sigma}_{ed}^2$ is small, then the larger part of the final estimate will come from the direct estimate (Boonstra and Buelens, 2011). Between-area variance $\hat{\sigma}_u^2$, like sample variance, is also unknown and must be estimated. It can be done with many techniques e.g. the Fay-Herriot method, Prasad-Rao method, ML or REML (Rao, 2015).

The robust version of the Fay-Herriot model uses Huber (1981) influence function to restrict the influence of u_d and e_d. The detailed process of robustifying all equations is

described in Sinha and Rao (2009) and Warnholz (2016). Robust EBLUP is given by the formula:

$$\hat{\theta}_d^{RFH} = x_d^T \hat{\beta}^{w} + \hat{u}_d^{w} \qquad d = 1,.....D \tag{5}$$

For unsampled domains and if between-area variance is equal to zero, indirect estimation is only based on the regression model.

The mean square error (MSE) of the parameters can be estimated by the parametric bootstrap method proposed by Gonzalez-Manteiga et. al. (2008). MSE can be used to calculate the relative root mean square error (RRMSE), which is treated as a common measure of precision for all approaches.

4 Estimation of net revenue with indirect estimation

In the study we constructed models for two dependent variables. The first one was *net revenue from the sale of products – SW*. The second variable was *net revenue from the sale of goods and materials – SH*. Data for companies representing two NACE sections were used: manufacturing and trade. In the group of manufacturing companies, the average net revenue from the sale of goods and materials (SH) was much lower than the average net revenue from the sale of products (SW), while in the group of trading companies, the relation was the opposite. The number of sampled companies in both groups was similar - 3927 (manufacturing) and 4094 (trade).

The first step of the analysis involved the direct estimation of the target variables in all target domains i.e. province by section. Fig. 1 presents the distribution of obtained estimates.

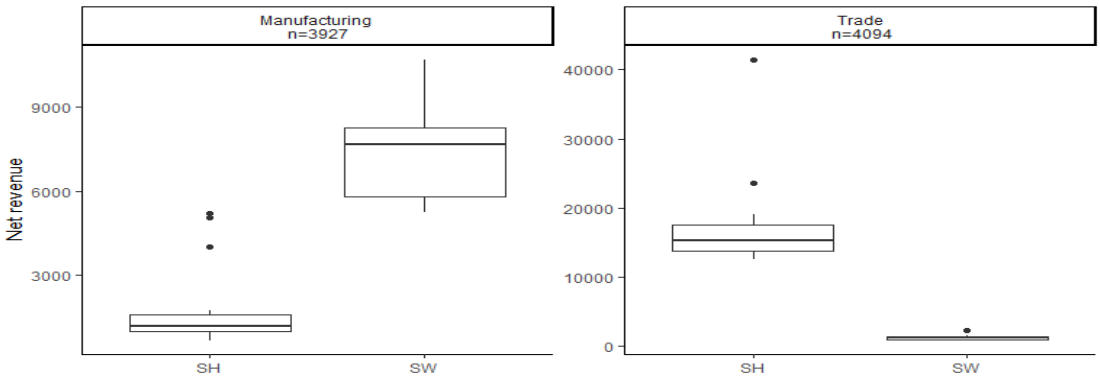


Fig. 1. Distribution of target variables by NACE sections.

The minimum value of *net revenue from the sale of goods and materials* (SH) in the manufacturing section is equal to 647,000 PLN and is observed in Zachodniopomorskie

province. Based on Fig. 1, three province outliers can be identified - Śląskie (5,200,000 PLN), Warmińsko-Mazurskie (5,039,000 PLN) and Mazowieckie (4,010,000 PLN). The maximum value of SH is close to the minimum for the net revenue from the sale of products (SW), which is equal to 5,253,000 PLN for Lubuskie province. Three provinces with the highest estimates are Warmińsko-Mazurskie (10,705,000 PLN), Wielkopolskie (9,205,000 PLN) and Zachodniopomorskie (9,133,000 PLN). The provinces vary considerably in terms of which types of business activity are identified as the main source of revenue. For example, in Warmińsko-Mazurskie province these include the food, tyre and wood industry, Śląskie province is dominated by companies mainly engaged in coal mining, steelmaking and electricity production. In Wielkopolskie province the dominant industries include the mining of salt, gypsum and lignite. Mazowieckie province is where PKN Orlen, the biggest fuel company in Poland is headquartered.

In the trade section, the sale of goods and materials (SH) is the main source of revenues. The highest estimated values are found in two provinces: Mazowieckie (41,489,000 PLN) and Małopolskie (23,585,000 PLN). The lowest estimated value is observed in the least urbanized province of Poland - Podkarpackie. The revenue from the sale of products (SW) in the trade section has a relatively marginal role - the highest value can be found in Mazowieckie province (2,377,000 PLN).

In addition to analysing the distribution of direct estimates, it is very important to consider the precision of these estimates. Table 1 contains descriptive statistics of relative root mean square errors (RRMSE) of direct estimates of net revenue.

Table 1. Descriptive statistics of RRMSE (in %) of direct estimates of net revenue.

NACE section	Variable	Minimum	Median	Mean	Maximum
Manufacturing	SH	15.8	31.4	33.9	70.9
Manufacturing	SW	9.4	11.5	14.2	37.5
Trade	SH	8.6	13.7	15.0	31.6
Trade	SW	11.7	17.1	17.7	35.0

Direct estimates of net revenue from the sale of goods and materials (SH) in the manufacturing section are characterized by very high values of RRMSE. In extreme cases, RRMSE is equal to 70.9% of the estimate (Śląskie province). For the remaining cases, the mean of RRMSE is about 15%, while the maximum exceeds 30%. The literature gives very different thresholds for precision. According to guidelines published by Eurostat for

household surveys, the precision level should depend on the survey, its purpose and the target domain. The National Institute of Statistics in Italy accepts RRMSE which does not exceed 15% for domains and 18% for small domains (Eurostat, 2013). According to the standards of the Central Statistical Office in Poland, survey results can be published if RRMSE is below 10% for target domains (Główny Urząd Statystyczny, 2013).

To obtain more precise estimates, indirect methods of estimation were applied: the Fay-Herriot model (FH) and the robust Fay-Herriot model (RFH). The models were based on information about net revenues in 2011 from the register maintained by the Ministry of Finance and the number or employees from the register of the Polish Social Insurance Institution (ZUS). Beta parameters in the models were significant and have a positive sign, which means that higher values of auxiliary variables result in higher net revenue estimates.

Fig. 2 presents the distribution of estimates obtained based on the direct Horvitz-Thompson (HT) estimator, the Fay-Herriot model (FH) and the robust Fay-Herriot model (RFH).

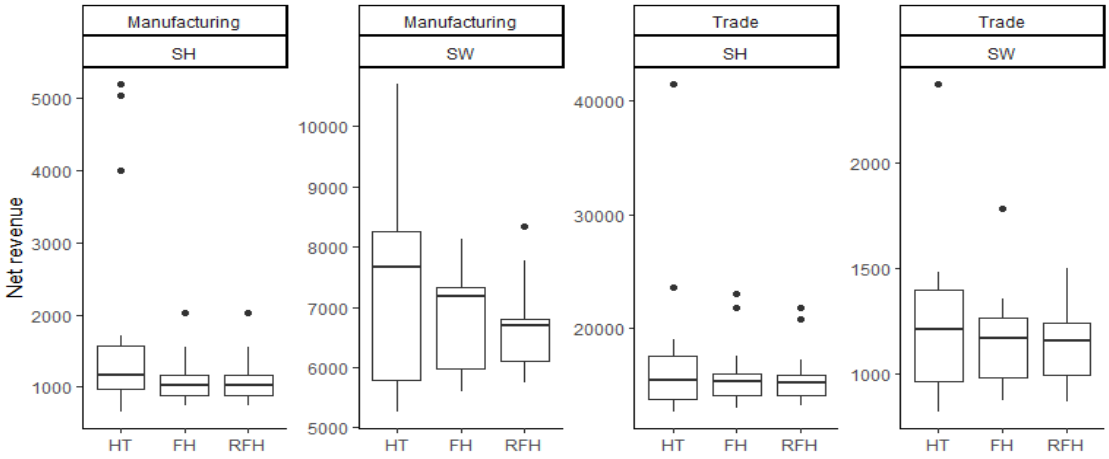


Fig. 2. Distribution of estimates by NACE section and type of net revenue.

The Fay-Herriot model belongs to the class of so-called "shrinkage" estimators, so obtained estimates have a smaller range than direct estimates. Moreover, the robust version of the Fay-Herriot model has a smaller range than the "classic" Fay-Herriot model. As regards values of net revenue from the sale of goods and materials (SH) in the manufacturing section, the maximum value is observed for Mazowieckie province and it is equal to 2,028,000 PLN for the FH model and 2,031,000 PLN for the RFH model.

The most visible change in the distribution is observed for net revenue from the sale of products (SW) in the manufacturing section. The median of direct estimates is equal to 7,664,000 PLN, 7,177,000 PLN for the Fay-Herriot model, and 6,700,000 PLN for the robust version of this model.

In the trade section estimates of both target variables obtained by applying the robust Fay-Herriot model have a smaller range in comparison to the other two approaches. With respect to net revenue from the sale of goods and materials (SH), the range of the Horvitz-Thompson estimates is equal to 28,943,000 PLN, for the Fay-Herriot model – 10,088,000 PLN, and for the robust Fay-Herriot model – 8,667,000 PLN. A very similar relation can be observed for net revenue from the sale of products (SW) in the trade section.

The next step of the study was the analysis of RRMSE. Table 2 presents values of this measure depending on section, dependent variable and estimator.

Table 2. Descriptive statistics of RRMSE (in %) of estimates by NACE section, type of net revenue and estimator.

NACE section	Variable	Estimator	Minimum	Median	Mean	Maximum
Manufacturing	SH	HT	15.8	31.4	33.9	70.9
Manufacturing	SH	FH	14.0	19.6	22.6	39.9
Manufacturing	SH	RFH	14.1	21.7	32.4	98.3
Manufacturing	SW	HT	9.4	11.5	14.2	37.5
Manufacturing	SW	FH	7.4	8.9	9.0	12.2
Manufacturing	SW	RFH	5.1	6.4	8.3	23.7
Trade	SH	HT	8.6	13.7	15.0	31.6
Trade	SH	FH	4.1	5.3	6.2	10.1
Trade	SH	RFH	4.8	7.4	8.6	15.8
Trade	SW	HT	11.7	17.1	17.7	35.0
Trade	SW	FH	10.3	13.9	14.1	19.7
Trade	SW	RFH	10.2	13.4	13.7	21.8

By applying indirect methods of estimation it was possible to reduce RRMSE of net revenue in unplanned domains, i.e. provinces. RRMSE of estimates obtained using the Fay-Herriot model is consistently lower than the precision of direct Horvitz-Thompson estimates. Estimates of net revenue from the sale of products (SW) calculated from the robust Fay-

Herriot have a better average precision than those given by the Fay-Herriot model in both sections. However, the maximum values of RRMSE are higher than in the case of the Fay-Herriot. The precision of estimating net revenue from the sale of goods and materials (SH) is better for the FH model than for the RFH model. This is associated with the result of estimating between-area variance. For this target variable and for this section the estimation algorithm of between-area variance in the case of Fay-Herriot model did not find a positive solution. As a result, the Fay-Herriot model generates synthetic estimates, which are characterized by low RRMSE. The same algorithm applied in the Robust Fay-Herriot model produces positive values of between-area variance, so the precision indicator also takes into account the uncertainty of direct estimation. RRMSE of estimates of net revenue from the sale of goods and materials (SH) in the manufacturing section obtained from the RFH model are even larger than direct estimates. In fact, there are two provinces (Warmińsko-Mazurskie and Śląskie) which are characterized by the largest RRMSE of direct estimates and large residuals in the robust Fay-Herriot model.

In addition to assessing estimation precision, estimates should also be analysed in terms of bias. Fig. 3 shows the sum of net revenue from the sale of products (SW) and net revenue from the sale of goods and materials (SH) compared to the true value of total net revenue in 2012.

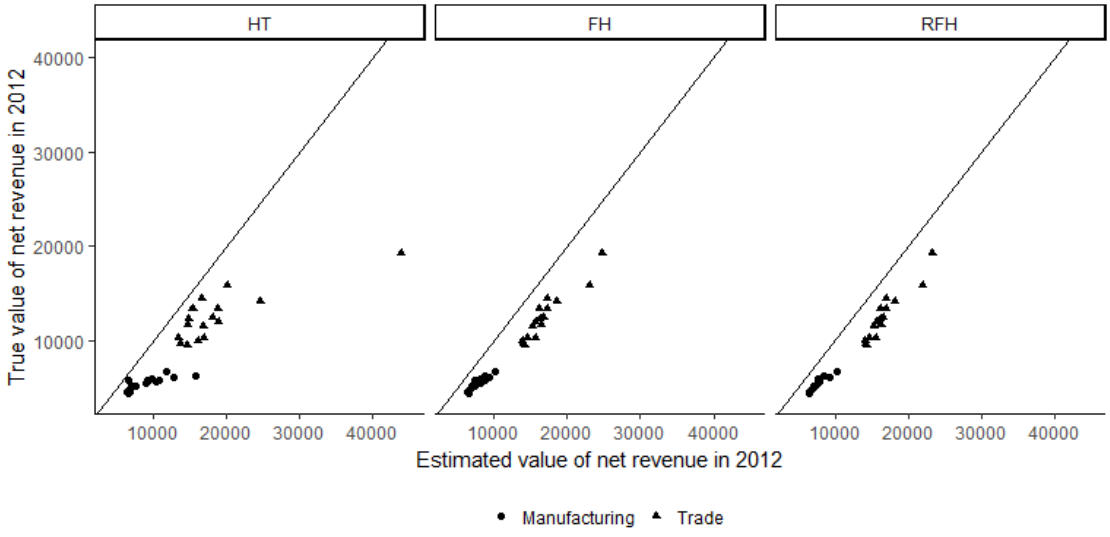


Fig. 3. Estimated and true values of net revenue in 2012.

In all cases, estimated values are overestimated in comparison with true values from the administrative register. Horvitz-Thompson estimates are characterized by the biggest bias.

Net revenue in the trade section for Mazowieckie province is overestimated by a factor of two. In the manufacturing section, Warmińsko-Mazurskie province is an outlier, with the direct estimate equal to 15,744,000 PLN, compared to the true value equal to 6,200,000 PLN. Nevertheless, the correlation between estimated and true values of net revenue is positive and strong. Spearman's correlation coefficient for the Horvitz-Thompson estimator is 0.9263, while for the indirect estimators – 0.9758 for the Fay-Herriot model and 0.9765 for its robust version. Average relative bias is equal to 55.9% for direct estimates, 40.1% for the Fay-Herriot model and 38.1% for the robust Fay-Herriot model.

Conclusions

Thanks to indirect methods of estimation, it is possible to obtain estimates of net revenue from the sale of goods and materials (SH) and net revenue from the sale of products (SW) for two NACE sections at a previously unpublished level of aggregation. Results obtained using the Fay-Herriot model and its robust version in most cases are more precise in terms of RRMSE than direct estimates. Moreover, robust estimation affects outlier values of net revenue and decreases the range of estimates. It is also worth noting that average relative bias is the smallest for estimates obtained by means of the robust Fay-Herriot model.

Further work will focus on estimating net revenue for small companies in Poland in other NACE sections. Because the level of precision of estimates generated by the Robust Fay-Herriot model is still unsatisfactory, we are considering changing the tuning factor or testing other influence functions (e.g. Tukey's, Cauchy's, Fair's, Talworth's or Welsch's) in the robust F-H model. Also, given the strategic role of the district (NUTS 4 unit), it would be interesting to apply the proposed approach could to estimate characteristics of small companies at district level

Acknowledgements

The project is financed by the National Science Centre in Poland on the basis of the decision number DEC-2015/17/B/HS4/00905.

References

Boonstra, H. J. & Buelens, B. (2011). Model-based estimation. *Statistics Netherlands.* Hague.

Dehnel, G. (2017). GREG estimation with reciprocal transformation for a Polish business survey In: Papież M. and Śmiech S. (eds) *The 11th Professor Aleksander Zelias Internetional Conference on Modelling and Forecasting of Socio-Economic Phenomena.*

Conference Proceedings (pp. 67-75). Cracow: Foundation of the Cracow University of Economics.

Dehnel, G., Pietrzak, M. & Wawrowski, Ł. (2017). Estymacja przychodu przedsiębiorstw na podstawie modelu Faya-Herriota. *Przegląd Statystyczny*, *64*(1), 79-94.

González-Manteiga, W., Lombardia, M. J., Molina, I., Morales, D. & Santamaría, L. (2008). Bootstrap mean squared error of a small-area EBLUP. *Journal of Statistical Computation and Simulation*, *78*(5), 443-462.

Główny Urząd Statystyczny (2013). Ludność. Stan i struktura demograficzno-społeczna. Narodowy Spis Powszechny Ludności i Mieszkań 2011. *Zakład Wydawnictw Statystycznych*. Warszawa: GUS.

Główny Urząd Statystyczny (2017). *Działalność przedsiębiorstw niefinansowych w 2015 roku.* Warszawa: GUS.

Eurostat (2013). Handbook on precision requirements and variance estimation for ESS households surveys. *European Union*.

Fay III, R. E. & Herriot, R. A. (1979). Estimates of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association*, *74*(366a), 269-277.

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.

Huber, P. J. (1981). *Robust Statistics*. New York: John Wiley.

Rao, J. N. K. (2015). *Small-Area Estimation*. John Wiley & Sons, Ltd.

Sinha, S. K. & Rao, J. N. K. (2009). Robust small area estimation. *Canadian Journal of Statistics*, *37*(3), 381-399.

Warnholz, S. (2016). *Small Area Estimation Using Robust Extensions to Area Level Models* (Doctoral dissertation, Freie Universität Berlin).

Application of differential evolution algorithm to group bank's individual clients

Czesław Domański[1], Robert Kubacki[2]

Abstract

Grouping methods are one of the most commonly used data mining methods in banking. Their goal is to describe population of clients. They usually are a starting point for subsequent analyzes. The aim of the article is to present the results of grouping individual clients of the bank with the differential evolution algorithm. Differential evolution algorithm is an alternative to the commonly used k-means algorithm. Algorithm is generating several competing solutions in one iteration. It allows to become independent of starting vectors and to be more effective in searching for an optimal solution. Clustering was run with preselected continuous variables characterizing all individual clients (deposit, credit and investment). The calculations were run using computer program written in SAS (4GL/SQL). The differential evolution algorithm itself has been enriched with a variable that allows the selection of the optimal number of clusters. Each iteration contained proposed solutions (chromosomes) which were evaluated by the target function built on the CS measure proposed by Chou (Das et. al., 2009) . Conducted analysis showed that the algorithm correctly grouped the bank's clients.

Keywords:*Clustering methods, Differential Evolution Algorithm, CS measure*
JEL Classification:C38, M31, G21
DOI: 10.14659/SEMF.2018.01.12

1 Introduction

In today's world more and more companies have problems with effective management of available data. The gap between the amount of data that is generated, stored and the degree of their understanding is constantly growing. According to a survey conducted by IBM among the representatives of the largest banks, over 40% of them have problems with the excess of information and the lack of appropriate tools for analyzing them (Giridhar et al., 2011).

Grouping methods are effective in describing populations. Many authors have studied these methods (Everitt et al., 2011; Jain and Dubes, 1988; Gan et al., 2007; Kaufman and Rousseeuw, 2005).

Most classic grouping algorithms have two major disadvantages:

1. Easily fall into local optima in multidimensional spaces that have multimodal objective functions.

[1] University of Lodz, Department of Statistical Methods, 90-255 Łódź, 3/5 POW Street, e-mail:czedoman@uni.lodz.pl
[2] Corresponding author: University of Lodz, Department of Statistical Methods, 90-255 Łódź, 3/5 POW Street, e-mail: robertkubacki@o2.pl

2. The efficiency of searching for a solution depends very much on the start vectors.

In literature, there is a description of grouping methods as a method without a supervisor, while most traditional algorithms require a priori knowledge of the number of clusters, which means that this is not a method without the interference of an outsider. On the other hand, in many practical applications it is impossible to provide even an approximate number of groups for an unknown data set.

The limitations of classical grouping methods, including the k-means algorithm, led the researchers to search for new, more effective grouping methods. One of the directions for the development of grouping algorithms was to treat them as an optimization problem. Over the time, the paradigm of evolutionary computation, the relationship between optimization and biological evolution, has evolved. Evolutionary calculations use the power of natural selection and allow to use the computing power of computers for automatic optimization (Das et al., 2009).

2 Differential evolution algorithm – selected issues

Differential evolution algorithm is part of heuristic methods, because the goal of optimization is not to find the exact equation describing the studied phenomenon, but to search the available space for solutions. These solutions are constructed using random elements. What is more, in one iteration of the algorithm several competing solutions are created. Subsequent solutions are created using similarities to the evolutionary mechanisms occurring in nature. These are the ones that, according to the defined objective function, are the best. The characteristic feature of the differential evolution algorithm is that solutions are created on the basis of real variable vectors, not vectors coded to zero-one sequences

Since 1995 differential evolution algorithm (Storn, 1995; Storn and Price, 1997) drew practitioners' attention in optimization due to the degree of resistance, the speed of convergence and the accuracy of solutions for real optimization problems. The differential evolution algorithm has defeated many algorithms, such as genetic algorithms, evolutionary strategies and memetic algorithms (Das et al., 2016).

Suppose we have a set of objects *Np* vectors, each has *D* dimensions. In addition, we mark $\boldsymbol{P_X}$ as the current population of solutions to the optimization problem, which was created as an initial solution or at any subsequent stage of the algorithm's operation.

$$\boldsymbol{P}_{\boldsymbol{X},g} = (\boldsymbol{X}_{i,g}), \qquad i = 0,1,\dots,Np-1, \qquad g = 0,1,\dots,g_{max} \tag{1}$$

$$\boldsymbol{X}_{i,g} = (x_{j,i,g}),\ j = 0,1,\dots,D-1. \tag{2}$$

Index$g = 0,1,\dots,g_{max}$ denotes the generation to which the vector belongs. Each vector is assigned to the corresponding population index $i = 0,1,\dots,Np-1$. The dimensions of the vector are marked by $j = 0,1,\dots,D-1$.

The differential evolution algorithm generates mutant vectors in the next step, which will be marked as follows:

$$\boldsymbol{P}_{\boldsymbol{V},g} = (\boldsymbol{V}_{i,g}), \qquad i = 0,1,\dots,Np-1, \qquad g = 0,1,\dots,g_{max} \tag{3}$$

$$\boldsymbol{V}_{i,g} = (v_{j,i,g}), \qquad j = 0,1,\dots,D-1. \tag{4}$$

However, the vectors after crossover will be marked as follows:

$$\boldsymbol{P}_{\boldsymbol{U},g} = (\boldsymbol{U}_{i,g}), \qquad i = 0,1,\dots,Np-1, \qquad g = 0,1,\dots,g_{max} \tag{5}$$

$$\boldsymbol{U}_{i,g} = (u_{j,i,g}),\ j = 0,1,\dots,D-1. \tag{6}$$

The first stage, i.e. setting the initial vectors, consists in generating starting vectors. Initial parameters (for *g*=0) are set within limits that correspond to a range that is acceptable for the intended solution. Therefore, if *j*-th the search task parameter has ranges marked as $x_{min,j}$ and $x_{max,j}$ and $rand_{i,j}(0,1)$ means *j*-th realizations of a uniform distribution from the range from 0 to 1 for *i*-th vector then can be determined *j*-th component *i*-th population element, as:

$$x_{i,j}(0) = x_{min,j} + rand_{i,j}(0,1) * \left(x_{max,j} - x_{min,j}\right). \tag{7}$$

The differential evolution algorithm searches for the global optimum in *D*-dimentional continuous hyperspace. It starts with a randomly selectedpopulation *NpD*-dimensional values of parameter vectors. Each vector, also known as genome / chromosome, is a proposed solution in a multidimensional optimization issue. The next generations of solutions in the differential evolution are marked as $g = 0,1,2,\dots,g,g+1$.

The vector parameters may change with the appearance of new generations, therefore the notation for which it will be accepted, for which*i*-th population vector for the current generation over time (*g*=*g*) as:

$$\vec{X}_i(g) = [x_{1,1}(g), x_{i,2}(g), \dots, x_{i,D}(g)]^T \tag{8}$$

where *i*=1,2,…,*Np*.

Mutation means a sudden change in the characteristics of the chromosome gene. In the context of evolutionary computation, a mutation means a change or disorder of a random component. Most evolutionary algorithms simulate the effect of mutations through the additivity of the component generated with a given probability distribution. In the differential evolution algorithm, a uniform distribution of the vector of the form differences was used:

$$\Delta\vec{X}_{r2,r3} = (\vec{X}_{r2} - \vec{X}_{r3}). \tag{9}$$

In the differential evolution algorithm, the mutation creates a successor vector $\vec{\mathrm{V}}_{\mathrm{i}}(\mathrm{g})$ for changing the population element$\vec{\mathrm{X}}_{\mathrm{i}}(\mathrm{g})$ in every generation or iteration of the algorithm.

To create a vector$\vec{\mathrm{V}}_{\mathrm{i}}(\mathrm{t})$for each *i*-th element of the current population, the other three disjoint vectors $\vec{\mathrm{X}}_{\mathrm{r}_1^{\mathrm{i}}}(\mathrm{g}), \vec{\mathrm{X}}_{\mathrm{r}_2^{\mathrm{i}}}(\mathrm{g}), \vec{\mathrm{X}}_{\mathrm{r}_3^{\mathrm{i}}}(\mathrm{g})$ are randomly selected from the current population. Indexes $\mathrm{r}_1^{\mathrm{i}}, \mathrm{r}_2^{\mathrm{i}}, \mathrm{r}_3^{\mathrm{i}}$ are mutually exclusive integers selected from a range [1,*NP*], which are also different from the index and the base vector. Indexes are generated randomly for each mutated vector. Then, the difference of any two of the three vectors is scaled by the number *F* and added to the third vector. In this way, we get a vector $\vec{\mathrm{V}}_{\mathrm{i}}(\mathrm{g})$ expressed as:

$$\vec{V}_i(g) = \vec{X}_{r_1^i}(g) + F.(\vec{X}_{r_2^i}(g) - \vec{X}_{r_3^i}(g)). \tag{10}$$

The mutation scheme shows different ways of differentiating the proposed solutions.

The crossover operation is used to increase the diversity of the population of solutions. Crossing takes place after generating a donor vector through a mutation. The algorithms of the differential evolution family use two intersection schemes - exponential and binomial (zero-one). The donor vector lists the components with the target vector$\vec{X}_i(g)$to create a trial vector

$$\vec{U}_i(g) = [u_{1,1}(g), u_{i,2}(g), \dots, u_{i,D}(g)]^T. \tag{11}$$

In exponential crossover, we first select a random integer *n* from range [0,*D*-1]. The drawn number is the starting point for the target vector from which the components are crossed with the donor vector. An integer *L* is also selected from range[1,*D*]. *L* indicates the number of components in which the donor vector is involved. After selection *n* and *L* trial vector takes the form:

$$u_{i,j}(g) = \begin{cases} v_{i,j}(g) dla j = \langle n\rangle_D, \langle n+1\rangle_D, \dots, \langle n+L-1\rangle_D \\ x_{i,j}(g), \qquad for\ other\ j \in [0, D-1] \end{cases} \tag{12}$$

where the intervals denote the module modulo function *D*. Integer *L* is drawn from the sequence [1,2,…,*D*] according to the following pseudocode:

L=0;

Do

{

L=L+1;

} while (rand(0,1)<CR) AND (L<D));

As a result, the probability$(L \geq v) = (CR)^{v-1}$ for any $v > 0$. Crossover rate (*CR*) is a parameter the same as *F*. For each donor vector, a new set *n* and *L* must be drawn as described above.

On the other hand, binomial crossover is carried out for each *D* variables each time, when the number selected is from 0 to 1 is less than or equal to the value *CR*. In this case, the number of parameters inherited from the donor has a very similar distribution to the binomial one. This scheme can be represented in the following way:

$$u_{i,j,g} = \begin{cases} v_{i,j,\mathrm{g}}, jeśli\ (rand_{i,j}(0,1) \leq \mathrm{C}Rlubj = j_{rand}) \\ x_{i,j,\mathrm{g}}, otherwise \end{cases} \quad (13)$$

where $rand_{i,j}(0,1) \in [0,1]$is a randomly drawn number that is generated for every j-th of the *i*-th parameter of the vector. $j_{rand} \in [1,2,\ldots,D]$ is a randomly selected index that ensures that $\vec{\mathrm{U}}_{\mathrm{i,g}}$ contains at least one component from the vector$\vec{\mathrm{V}}_{\mathrm{i,g}}$.

This is determined once for each vector in a given generation. *CR* is an estimate of true probability $\mathrm{p_{Cr}}$ the event that the component of the sample vector will be inherited from the parent. It may also happen that in the two-dimensional search space, three possible test vectors can be the result of one-dimensional mating of the mutant / donor vector $\vec{V}_i(g)$ with the target vector$\vec{X}_i(g)$. Trial vectors:

a) $\vec{U}_i(g) = \vec{\mathrm{V}}_{\mathrm{i}}(\mathrm{g})$ both components$\vec{U}_i(g)$ inherited from the vector $\vec{\mathrm{V}}_{\mathrm{i}}(\mathrm{g})$

b) $\vec{U}_i{}'(g) = \vec{V}_i(g)$ one component (*j*=1) comes from vector $\vec{V}_i(g)$, second (*j*=2) from vector $X_i(t)$

c) $\vec{U}_i{}''(g) = \vec{\mathrm{V}}_{\mathrm{i}}(\mathrm{g})$ one component (*j*=1) comes from vector $X_i(g)$, second (*j*=2) from vector $\vec{\mathrm{V}}_{\mathrm{i}}(\mathrm{g})$

The last stage of the differential evolution algorithm is selection, i.e. the choice between the vector$\vec{X}_i(g)$ and a newly designated test vector$\vec{U}_i(g)$. The decision which of the two vectors will survive in the next generation g+1 depends on the value of the matching function.

If the values of the matching function for the sample vector is better than the value of the target vector, the existing vector is replaced with the new vector.

$$\vec{X}_i(g+1) = \begin{cases} \vec{U}_i(g) dlaf\left(\vec{U}_i(g)\right) \leq f\left(\vec{X}_i(g)\right) \\ \vec{X}_i(g) dlaf\left(\vec{U}_i(g)\right) > f\left(\vec{X}_i(g)\right) \end{cases} \tag{14}$$

where $f(\vec{X})$ is a minimized function. The selection process consists in selecting one of two variants. The adjustment of population members improves in subsequent generations or remains unchanged, but never deteriorates.

CS (Candidate Solution) Measureproposed by Chou (Das et al., 2009) is an objective function in this study. Group centroids are determined as the average vectors belonging to a given cluster

$$\bar{m}_i = \frac{1}{N_i} \sum_{\bar{Z}_j \in C_i} \bar{Z}_j \tag{15}$$

The distance between two points $\bar{Z}_p$and $\bar{Z}_y$is marked as$d(\bar{Z}_p, \bar{Z}_y)$. Then the *CS* measure can be defined as:

$$CS(k) = \frac{\frac{1}{k}\sum_{i=1}^{k}\left[\frac{1}{|C_i|}\sum_{\bar{Z}_y \in C_i} \max_{\bar{Z}_y \in C_i}\{d(\bar{Z}_p, \bar{Z}_y)\}\right]}{\frac{1}{k}\sum_{i=1}^{k}\left[\min_{j \in k, j \neq i} d(\bar{m}_i, \bar{m}_j)\right]} = \frac{\sum_{i=1}^{k}\left[\frac{1}{|C_i|}\sum_{\bar{Z}_y \in C_i} \max_{\bar{Z}_y \in C_i}\{d(\bar{Z}_p, \bar{Z}_y)\}\right]}{\sum_{i=1}^{k}\left[\min_{j \in k, j \neq i} d(\bar{m}_i, \bar{m}_j)\right]} \tag{16}$$

The measure is a function of the ratio of the amount of intra-group dispersion and the separation between groups. The *CS* measure is more effective at clusters with different density and / or different sizes than other measures.

3 Design of the study

The database of commercial bank clients was used for the study. It has been limited to the part of the population for which the actions taken will translate in the maximum way into business benefits. In particular, clients meet the following criteria: individual clients with active products, aged from 18 to 75 years, not being bank employees, with positive marketing consent, without delays in repayment of loan products.

As for the variables used for the study, the choice was not accidental. Variables selected for this study can be evaluated for each customer regardless of whether they have deposit, credit or investment products. Pre-processing of data allowed to eliminate outliers from the studied population. Due to the strong right-side skewness of the variables, a transformation

was made by adding a constant 0.001, and then their logarithmisation. As a result, the resulting distributions of variables are more symmetrical.

The final set of variables that took part in the study is presented below:

- ZM1 (DEPOZYTY) - Total funds on accounts and deposits in thousands of PLN,
- ZM2 (INWESTYCJE) - Total funds in investment products in thousands of PLN,
- ZM3 (LUDNOSC) - number of inhabitants, based on the city from the correspondence address and data published by the Statistics Poland,
- ZM4 (KREDYTY) - amount of bank loans taken in thousands of PLN,
- ZM5 (SALDO_BIK) - balance for repayment on credit products outside the bank, based on inquiries from BIK in thousands of PLN,
- ZM6 (AVG_TRN_INCOMING_ALL_3M) - average monthly income on customer's accounts in the last 3 months in thousands of PLN,
- ZM7 (AVG_TRN_INCOMING_CLEAN_3M) –cleaned average monthly income on customer's accounts in the last 3 months in thousands of PLN. Transactions between accounts belonging to the customer are not taken into account,
- ZM8 (AVG_TRN_OUTGOING_ALL_3M) - average monthly outflows from customer accounts in the last 3 months in thousands of PLN,
- ZM9 (AVG_TRN_OUTGOING_CLEAN_3M) – cleaned monthly average outflows from customer accounts in the last 3 months in thousands of PLN. Transactions between accounts belonging to the customer are not taken into account,
- ZM10 (AVG_TRN_OUT_DEBIT_3M) - average monthly transaction amount on the debit card from the last 3 months (cash and non-cash transactions) in thousands of PLN,
- ZM11 (AVG_TRN_OUT_CREDIT_3M) - monthly average amount of credit card transactions from the last 3 months (cash and non-cash transactions) in thousands of PLN,
- ZM12 (WIEK_LATA) - customer's age in years,
- ZM13 (STAZ_LATA) - customer experience in years.

Table 1 outlines constants used in the algorithm.

For the purpose of optimizing number of centroids dimensional matrix is created $MR_{c,k,z}$, where c means the number of chromosomes, k means the number of clusters, z means the number of variables. Number of variables is increased by 1. An additional variable is used to store information on whether the cluster is active or inactive in the given iteration. Values for individual matrix elements are generated according to the formula (7). An additional variable

indicating focus activation is determined based on the rule: If the randomly generated number from the range 0 to 1 is smaller than the activation constant (SA) then the variable takes the value 0, otherwise it takes the value 1.

Table 1. Constants used in the study.

Constant	value	Description of the constant
LZ	13	Number of variables describing the client
LC	13	Number of chromosomes
LK	15	Maximum number of clusters
SA	0.2	Constant activation of the vector
F	0.7	Mutation operator
Iterations	15	Number of iterations
CR	1	Crossover rate

4 Results of empirical analyses

The smallest value of the CS function in the fifteenth iteration was obtained for chromosome number 3. This solution was chosen as the optimal solution.

Table 2 contains the characteristics of chromosome 3, which divided the surveyed population of the bank's clients into 9 groups (the maximum number of groups on which the population could be divided into 13).

The results of grouping in Table 2 indicate that the distinguished groups are characterized by nonequal distribution of the number of clients in groups. Group 8 is more selective and gathers 45.71% of clients, group 4 contains 22.11% of clients, and the third group 6 includes 14.41% of clients. The three mentioned groups gather over 80% of the surveyed population.

More detailed characteristics of the distinguished groups of clients are presented in the Table 3, which contains average values of features in individual groups. The data presented in Table 3 indicate that individual groups differ from each other. Thanks to the knowledge of average values for particular groups, it is possible to indicate groups of transactionally active customers (groups 14,5,6) and customers who use accounts less frequently (group 3,8,4,1). The most-affluent group of customers with very high means is without a doubt group number 14.

Table 2. Numbers and share of groups for chromosome 3.

Group	No of Clients	% of total
8	92 109	45.71%
4	44 545	22.11%
6	29 047	14.41%
3	20 003	9.93%
5	5 476	2.72%
14	3 582	1.78%
1	2 839	1.41%
15	2 075	1.03%
12	1 832	0.91%
SUM	201508	100.00%

Table 3. Average values of variables ZM1-ZM13 for clusters obtained by the differential evolution algorithm.

Cluster	ZM1	ZM2	ZM3	ZM4	ZM5	ZM6	ZM7	ZM8	ZM9	ZM10	ZM11	ZM12	ZM13
8	8	0	452	137	88	5	4	5	3	0	0	43	5
4	2	5	273	5	1	5	4	5	4	0	0	43	6
6	20	10	500	182	146	22	17	22	17	1	0	42	6
3	28	4	453	7	45	1	1	1	0	0	0	47	6
5	60	61	627	325	0	26	20	27	19	0	1	41	6
14	113	97	758	414	144	113	84	106	73	2	1	42	6
1	20	67	473	223	114	4	3	4	2	0	0	43	6
15	24	48	321	264	10	10	8	7	5	0	0	41	5
12	0	2	131	7	117	15	11	18	15	0	0	41	3

Thanks to the use of the differential evolution algorithm to group the bank's clients, we can get information on how many natural groups exists in a short time. Moreover, the number of groups has been calculated, not imposed in advance. The algorithm evaluated and compared obtained results for other candidate solutions in subsequent iterations, recognizing according to the values of the objective function that the optimal division of this group of customers contains 9 clusters.

Conclusions

The differential evolution algorithm is a promising approach to optimization, because it generates a whole set of solutions that can be easily adapted to carry out the optimization again. The fact of keeping a set of solutions, not only the best solution, allows faster adaptation to new conditions using the previously made calculations. It is resistant in terms of the choice of parameters as well as the regularity in which it finds the global optimum. Algorithm is a direct search solution method, versatile enough to solve problems whose objective function lacks the analytical description needed to determine the gradient. The algorithm is also very simple to use and modify.

Evolutionary algorithms, in particular the differential evolution algorithm do well with continuous variables when grouping clients. Customers from particular groups can be synthetically described by the mean vector for variables used in clustering. They allow to effectively separate customers with the same basket of products, but differing in the level of individual variables.

References

Das, S., Abraham, A. & Konar, A. (2009). *Metaheuristic clustering* (Vol. 178). Springer.

Das, S., Mullick, S. S. & Suganthan, P. N. (2016). Recent advances in differential evolution–an updated survey. *Swarm and Evolutionary Computation*, *27*, 1-30.

Everitt, B. S., Landau, S., Leese, M. & Stahl, D. (2011). *Cluster analysis.* Wiley series in probability and statistics.

Gan, G., Ma, C. & Wu, J. (2007). *Data clustering: theory, algorithms, and applications* (Vol. 20). Siam.

Giridhar, S., Notestein, D., Ramamurthy, S. & Wagle, L. (2011). Od złożoności do orientacji na klienta. *Executive report, IBM Global Business Services*.

Jain, A. K. & Dubes, R. C. (1988). *Algorithms for clustering data*.

Kaufman, L. & Rousseeuw, P. J. (2005). *Finding groups in data*. Hoboken.

Storn, R. (1995). *Differrential evolution-a simple and efficient adaptive scheme for global optimization over continuous spaces*. Technical Report, International Computer Science Institute, 11.

Storn, R. & Price, K. (1997). Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, *11*(4), 341-359.

Analysis of transaction prices on one of the housing estates in Szczecin: Do the buyers differentiate prices on local markets with respect to the kind of the right to own?

Iwona Foryś[1], Barbara Batóg [2]

Abstract

The results of researches on housing market dealing with dependencies between prices and attributes and the results of classifications of purchased apartments according to their attributes are very useful tool supporting decisions of housing market participants. One of the most essential attributes of real estate is the right to own.

In case of apartments it could be property, limited rights in rem and law of obligations. The widest ones are property, co-operative title to premises, tenant law and rent. The aim of the paper is comparison of transaction prices in case of two strong laws: property and co-operative title to premises in different stages of business cycle on homogeneous housing estate. The key question is: do the buyers differentiate prices with respect to the kind of the right to own? The tendencies of average quarterly prices of these two kinds of rights will be analyzed. The distributions of transaction prices in consecutive years will be assigned.

The research is based on information concerning all transactions on local housing market in Szczecin in 2006-2017 found in notary deeds collected by Authors. The transactions were conducted on one of housing estate in Szczecin named "Zawadzkiego-Klonowica". The choice of this housing estate was caused by its characteristics such as constant number of apartments and the same type and technology of buildings.

Keywords: *housing market, transaction prices, cointegration*

JEL Classification: R31, C38

DOI: 10.14659/SEMF.2018.01.13

1 Introduction

Real estate market is very dynamic market and is strongly dependent on changes on social and economic environment. Especially housing market is very sensitive to economic situation of households, changes of demand from households and their preferences concerning attributes of purchased apartments. Housing market participants are interested in dependency between price and attributes of apartments during every stage of business cycle.

This knowledge enables each transaction party to estimate the value of apartment in case of information asymmetry and make a reasonable decision to buy or sell it (Springer, 1996).

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

[2] University of Szczecin, Faculty of Economics and Management, Institute of Econometrics and Statistics, 71-101 Szczecin, Mickiewicza 64, Barbara.batog@usz.edu.pl.

The fore mentioned dependency is also important for property appraisers in case of evaluation of apartments by means of comparative approach.

The main problem in the analysis of the real estate market is its low information effectiveness (Case and Shiller, 1989) and the change of preferences of purchasers over time, especially disturbed in periods of economic downturn (Nicholas and Scherbina, 2013). An additional problem in analyses using econometric methods (Francke, 2010) to study sales contracts is the heterogeneity of data (Foryś, 2011) and features measured on weak scales (Foryś and Gaca, 2016). Real estate properties are by definition not identical but similar. Similarity may be interpreted as the absence of distinguishing features, which significantly affect value or comparison by virtue of the list of legally imposed features.

In this approach the first step is choice of similar apartments in respect of attributes strongly influencing market value of apartment. Therefore the results of researches on housing market dealing with dependencies between prices and attributes and the results of classifications of purchased apartments according to their attributes are very useful tool supporting decisions of housing market participants (Foryś and Nowak, 2012). One of the most essential attributes of real estate is the right to own. In case of apartments it could be property, limited rights in rem and law of obligations. The widest ones are property, co-operative title to premises, tenant law and rent.

The aim of the paper is comparison of transaction prices in case of two strong laws: property and co-operative title to premises in different stages of business cycle on homogeneous housing estate. The key question is: do the buyers differentiate prices with respect to the kind of the right to own? The tendencies of average quarterly prices of these two kinds of rights will be analyzed. The distributions of transaction prices in consecutive years will be assigned.

The research is based on information concerning all transactions on local housing market in Szczecin in 2006-2017 found in notary deeds collected by Authors. All transactions were conducted on one of housing estate in Szczecin named “Zawadzkiego-Klonowica”. The choice of this housing estate was caused by its characteristics such as constant number of apartments and the same type and technology of buildings during the research period. Sales contracts were concluded there with each property right tested.

2 Literature revive

The problem of property ownership in the real estate market is widely discussed primarily in legal literature, where the issue of ownership rights and the extent to which a given right is exercised are discussed.

Property valuation assumes that the same type of rights trading is used for comparison, which is also indicated by Polish law. On the other hand, due to the decisions of the real estate market participants, it turns out to be important whether if weaker and stronger rights are not fully recognizable by them, they differentiate them in the purchase price of real estate. In addition, is trade in these rights balanced on the local market, or does the preference of purchasers change over time due to the lower price of one of them?

Harding et al. in their research note that ownership is associated with a greater tendency to care for your property, and thus better maintained properties have a higher value on the market (Harding et al., 2000; Zeithaml, 1988). Germans prefer to rent apartments rather than buy them, unlike the British who value having their own flat or house (Sivesand, 2005). This applies especially to the young generation entering the labor market and thus the real estate market. In the countries of Central and Eastern Europe, the communist system deprived many citizens of the ownership of real estate, limited the acquisition of real estate, hence the desire to have unlimited ownership of real estate, especially a flat, is very strong.

Following the political changes, the right of ownership of a dwelling in Poland is a disposable, hereditary and one on which a mortgage can be established. Similarly, the limited right to a dwelling in a housing cooperative (cooperative law) is a transferable, hereditary and one that can serve as collateral for a mortgage claim (Foryś and Nowak, 2012). So what is the difference between the two rights, which could make it possible to differentiate their prices on the market?

The ownership right of a dwelling in multi-family buildings also includes the right to participate in common parts of a property not used exclusively by the owner of a dwelling. These include, for example, shares in the structure of a building, traffic areas, often in cellars and other utility rooms, and above all a share in a land property. The possession of such rights entails obligations but also permissions to, e.g. direct co-decision on real estate in accordance with the shares held. Therefore, the sense of ownership extends to the whole property and not only to the dwelling.

In the case of a cooperative right to a dwelling, there is a right to use the dwelling for its intended purpose and the joint ownership relates to all the property of a housing cooperative. So it is not assigned to a particular building, and such a loose relationship results in less care

for common property. It should not be forgotten that a potential buyer sets the price on the basis of ideas, opinions shaped by the media and, finally, his own feelings while viewing the property. Therefore, the mechanisms of the so-called first impression which is strongly connected with the visual sensations of the whole property and its surroundings, are immensely important. Expert knowledge appears at a later stage of evaluation of the subject matter of an agreement, also during valuation by a property appraiser (Trojanek, 2012).

Thus, the prevailing views on both rights, ignorance of ownership rights or impulse induced action mean that the market does not always differentiate a weaker right from a better one in terms of price. The above mentioned view that a weaker right to a dwelling is priced lower on the market than a stronger law will be verified in the study.

In the first part of the study the descriptive statistics of unit prices were analyzed in time. In the second part the analysis of cointegration was applied. Some examples of application of the analysis of cointegration concerning real estate market can be found in the literature. For example Lin and Lin have been analyzing the cointegration relationship between stock markets and real estate markets in Asia (Lin and Lin, 2011) and Baltagi and Li have been looking for cointegration relationship between the home purchase price and rental price based on nationally estimated indexes and also between area-specific home purchases and rental price indexes (Baltagi and Li, 2015).

The analysis is conducted both in the period of the real estate market boom (2006-2008), as well as in the period of economic downturn (2009-2015) and exit from it (2016-2017).

3 Data and research results

The presented study was conducted on the basis of source data from notarial deeds collected by the authors on all transactions concluded in 2006–2016 and in 10 months of 2017 on the local housing market in Szczecin.

However, one of Szczecin's housing estates was selected for the analysis for which a full set of agreements on the purchase and sale of secondary trade apartments was available, as well as those characterized by a stable stock of apartments in the analyzed period. Between 2006 and 2017, no new apartments were built there and no alterations to the buildings or their functions were made. Therefore the objects investigated (housing units) were in this respect a uniform sample. Additionally, the buildings were built in a similar industrialized technology, in the years 1960-1990. Due to the purpose of the study, information concerning the date of conclusion of the contract, type of the right to premises, usable area and transaction price were used which allowed to determine unit prices of usable area per m^2. In total, there were

1,077 transactions identified for the investigated area. Incomplete data for 2017 are related to at least a quarterly delay in recording notarial deeds in real estate cadastral units.

In the table below (Table 1) basic descriptive statistics were compiled concerning the number of concluded agreements on the sale of apartments in the investigated area and unit price[3]. Between 2006 and 2008, the number of transactions oscillated around one hundred per year, while the economic downturn in 2009 and 2010 was followed by a downward trend in the number of agreements concluded until 2015, followed by an improvement between 2016 and 2017.

Table 1. Descriptive parameters of the unit price in the years 2006–2017*.

Year	Number of transactions	Average	Median	Lower Quartile	Upper Quartile	Standard Deviation	Coefficient of Variation
2006	104	2 820	2 771	2 490	3 189	532	18.85
2007	103	4 476	4 561	4 125	4 953	813	18.16
2008	92	4 642	4 670	4 307	5 006	607	13.08
2009	125	4 934	5 039	4 402	5 514	705	14.28
2010	63	4 361	4 456	4 049	4 810	607	13.91
2011	103	4 203	4 286	3 850	4 560	756	17.99
2012	89	3 916	3 939	3 713	4 237	566	14.45
2013	74	3 810	3 873	3 402	4 129	627	16.45
2014	76	3 888	3 868	3 493	4 236	611	15.70
2015	73	3 987	4 097	3 508	4 442	837	20.99
2016	94	4 303	4 286	3 823	4 668	752	17.48
2017*	81	4 478	4 202	3 833	5 036	968	21.61

* – for 10 months

In the years 2007–2015 (except for 2014), average prices were lower than the median of unit prices, which indicates left-hand asymmetric distributions and advantage of transactions with higher prices. The strongest left-hand asymmetry occurred in 2009 and 2015. The difference in unit price in around 50% of the transactions (range of the lower and upper

[3] The nominal prices are used in the research because consumer price index is not appropriate for housing market and real estate price index is not published by the Central Statistical Office of Poland.

quartile) in the analyzed years was between PLN 524-935 per sqm with the exception of 2009 (PLN 1 112 per sqm) and 2017 (PLN 1 203 per sqm). The determined coefficients of variation indicate a differentiation of unit prices in the analyzed years. However, the highest coefficients of variation can be observed in the years 2015 and 2017, while in the remaining years the coefficient fluctuated between 13-19%.

In the years 2006-2014, the maximum unit prices did not exceed PLN 6 500 per sqm and were higher than PLN 1 000 per sqm in the whole investigated period. There is a clear leap in peak prices in 2009. The average unit prices in the surveyed years oscillated around PLN 4 000 per sqm (Fig. 2).

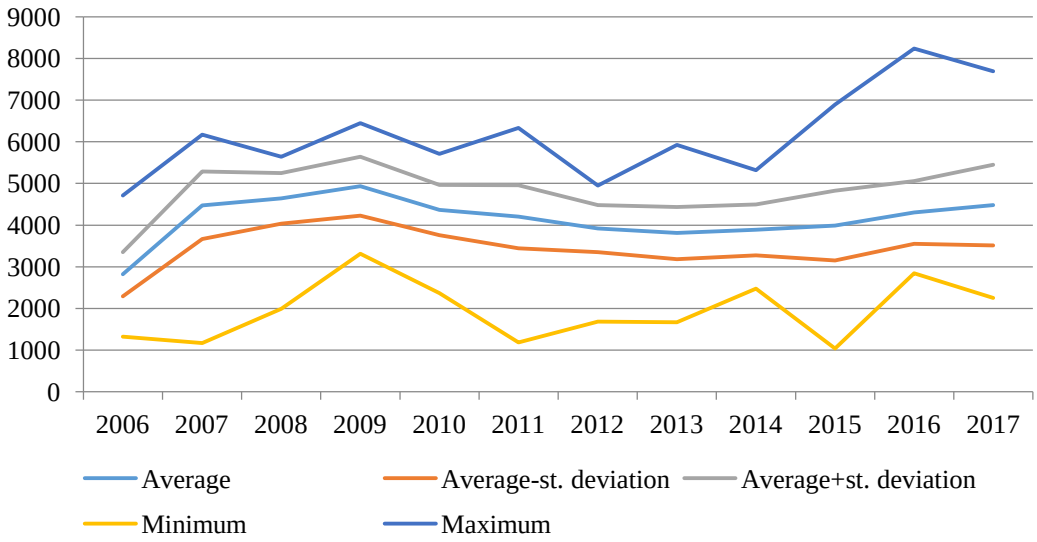


Fig. 1. Basic descriptive statistics of unit prices of apartments in 2006-2017.

In the second part of the paper the tests for cointegration have been run. The tests concern three quarterly average unit prices (see Fig. 3):

- average unit price of all transactions in given quarter (average unit price),
- average unit price of all transactions on property in given quarter (average unit price 1),
- average unit price of all transactions on co-operative title to premises in given quarter (average unit price 0).

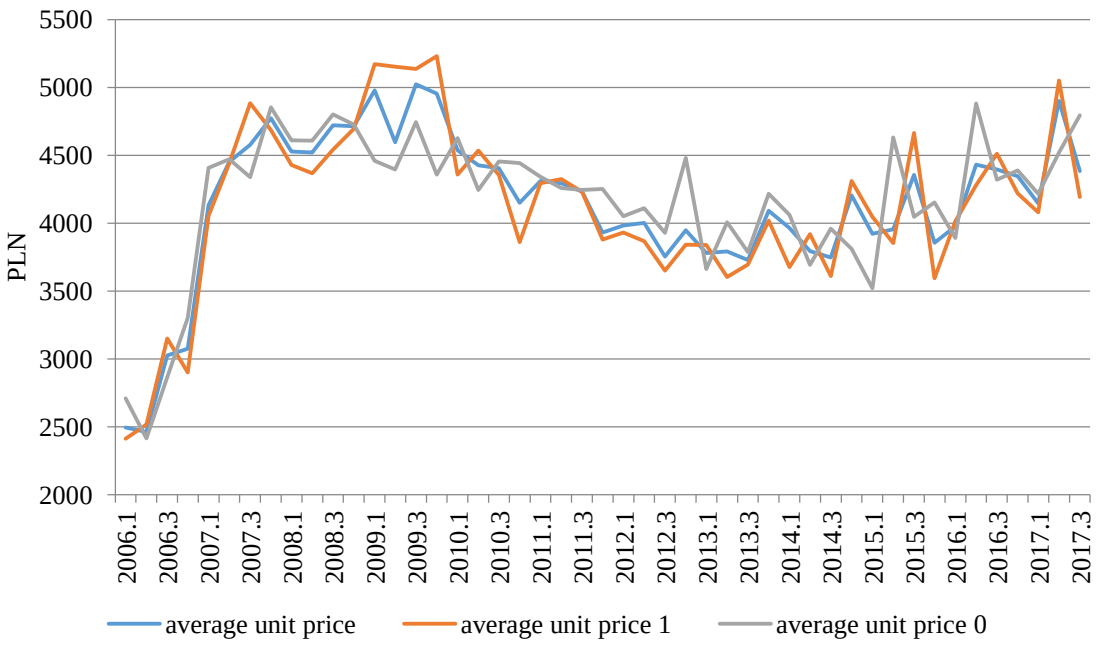


Fig. 2. Quarterly average unit prices in 2006–2017.

In order to examine cointegration the Engle–Granger test was applied (Engle and Granger, 1987; Dolado et al., 2003; Osińska, 2007). In the first step each variable was tested for a unit root using Dickey-Fuller test. Then cointegration regressions were run for every two of three examined variables. In the final step Dickey-Fuller test was run for the residuals from the cointegration regressions. In Dickey–Fuller test the null hypothesis is: variable has unit root (a = 1; process is integrated of order one, I(1)).

Table 2 presents the results of Dickey-Fuller test for analysed variables.

Table 2. Dickey-Fuller test for variables (without constant).

Variable	Average unit price	Average unit price 1	Average unit price 0
Estimate of (a-1)	0.0052	0.0017	0.0050
Test statistic	0.4661	0.1106	0.3446
p value	0.8115	0.7128	0.7805
Autocorrelation of first order residuals	-0.2160	-0.3710	-0.4110

It turned out that the null hypothesis cannot be rejected for three kinds of average unit price – all p values are very high and variables are integrated of order one. Therefore the second step was performed. Table 3 presents results of cointegration regressions.

Table 3. Cointegration regressions.

Dependent Variable	Average unit price	Average unit price	Average unit price 1
Independent Variable	Average unit price 1	Average unit price 0	Average unit price 0
Parameter	1.0006	0.9923	0.9882
Standard error	0.0059	0.0091	0.0140
t statistic	168.8100	108.8100	70.3956
p value	0.0000	0.0000	0.0000
R^2	0.9984	0.9961	0.9908

On the base of results presented in Table 3 the residuals were calculated and Dickey-Fuller test was run on them. The results of this test are presented in Table 4.

Table 4. Dickey-Fuller test for residuals.

Dependent Variable	Average unit price	Average unit price	Average unit price 1
Independent Variable	Average unit price 1	Average unit price 0	Average unit price 0
Estimate of (a-1)	-1.0401	-1.1386	-1.0994
Test statistic	-6.9070	-7.5800	-7.2969
p value	0.0000	0.0000	0.0000
Autocorrelation of first order residuals	0.0120	0.0240	0.0270

It turned out that the null hypothesis should be rejected for three kinds of average unit price – all p values are very small and every pair of variables is cointegrated. These results means that every pair of variables is in stable relation in examined period – the values change in a similar way. So the long run relationship exists for every pair of variables. It also means that unit transaction prices on property and on co-operative title to premises on the examined housing market are characterized by regularities that are similar in analysed period.

Conclusions

The research conducted on a uniform, in the sense of technical values, stock of flats did not confirm the relation between the right to the sold flat and the price. The buyers on the analyzed market did not value the ownership right higher than the limited right, i.e. the cooperative right to premises.

In the case of similar other features, it would seem that the right to more freely dispose of the premises (ownership right) has a higher value in the examined market due to the higher quality of the premises, which is also a result of taking care of ownership not only of the premises but also of the common parts of a building (Harding et al., 2000). However, the analysis for quarterly data did not confirm this correlation. There is no difference in the price levels between the rights. For average quarterly prices, it cannot be said that the price of ownership is higher, as shown in Fig. 2. There is co-integration, i.e. the evolution of these prices is similar during the period considered. For the annual data, significant differences were only in two years and their marks were different. This implies that in reality, however, a right type has no effect on prices.

References

Baltagi, B. H. & Li, J. (2015). Cointegration of matched home purchases and rental price indexes – Evidence from Singapore. *Regional Science and Urban Economics, 55,* 80–88.

Case, K. & Shiller, R. (1989). The Efficiency of the Market for Single-family Homes, *American Economic Review*, *79,* 125-137.

Dolado, J. J., Gonzalo, J. & Marmol, F. (2003). Cointergration. In: Baltagi B. H. (ed.), *A Com-panion to Theoretical Econometrics*, Blackwell Publishing, USA.

Engle, R. F. & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica 55*, 251–276.

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

Foryś, I. & Nowak, M. (2012). *Spółdzielnia czy wspólnota? Zarządzanie zasobem mieszkaniowym*. Wydawnictwo Poltex, Warszawa.

Foryś, I. & Gaca, R. (2016). Theoretical and Practical Aspects of Qualitative Variable Descriptions of Residential Property Valuation Multiple Regression Models. In: Papież M. & Śmiech S. (eds) *Proceedings,* Publisher: Foundation of the Cracow University of Economics.

Francke, M. K. (2010). Repeat Sales Index for Thin Markets. A Structural Time Series Approach. *The Journal of Real Estate Finance and Economics, 41,* 24-52.

Harding, J. P., Miceli, T. J. & Sirmans, C.F. (2000). Do owners Take Better Care of Their Housing than Renters? *Real estate Economics, 4*(28), 663-681.

Lin, T. C., & Lin, Z-H. (2011). Are stock and real estate markets integrated? An empirical study of six Asian economies. *Pacific-Basin Finance Journal, 19,* 571–585.

Nicholas, T. & Scherbina, A., (2013). Real Estate Prices During the Roaring Twenties and the Great Depression. *Real Estate Economics*, *41* (2), 278-309.

Osińska, M. (ed.) (2007). *Ekonometria współczesna*. Towarzystwo Naukowe Organizacji i Kierownictwa, Stowarzyszenie Wyższej Użyteczności „DOM ORGANIZATORA", Toruń.

Sivesand, H. (2005). *The Buyer's Remedies for Non-conforming Goods: Should There be Free Choice Or are Restrictions Necessary?* Sellier, European Law Publisher.

Springer, T. M. (1996). Single-Family Housing Transaction: Seller Motivation, Price, and Marketing Time. *Journal of Real Estate Finance and Economics*, *3*(13), 237-254.

Trojanek, R. (2012). An analysis of changes in dwelling prices in the biggest cities of Poland in 2008-2012 conducted with the application of the hedonic method. *Actual Problems of Economics*, *7*, 5-14.

Zeithaml, V. A. (1988). Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence. *Journal of Marketing*, *52*, 2-22.

Efficiency of investments in solar power in the EU countries

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

Abstract

The aim of the study is to assess efficiency of investments in solar power in the EU countries in 2015. The Data Envelopment Analysis (DEA) method is implemented to evaluate relative efficiency of their solar power performance for electricity generation. The installed solar capacity is used as the input variable. Four variables: solar electricity generation (the baseline model), the environmental output, the economic output and the energy dependence output are used to measure efficiency of investments in solar power.

The study reveals that three countries: Germany, Spain and Ireland are efficient when the relation between the investments in solar power and the volume of energy generated by solar farms is considered. However, when additional aspects of solar power efficiency are included different countries benefit. Inclusion of the environmental output reveals that efficiency increases the most in Poland, Malta, Bulgaria, Cyprus and the Czech Republic, where coal is the main energy source. Taking into account the economic aspect, i.e. the costs of generating energy from non-renewable energy sources, an increase in the efficiency scores is noted in countries in which oil (Malta and Cyprus), or natural gas (Lithuania, Belgium, Luxembourg) are used as the main energy source.

Keywords: *Data Envelopment Analysis (DEA), solar power generation*

JEL: *C59, C61, Q2, Q01*

DOI: *10.14659/SEMF.2018.01.14*

1 Introduction

Over the last 20 years the EU countries considerably increased their dependence on renewable energy sources, especially in the area of electricity generation. The share of renewable energy sources in total electricity generation grew more than twofold - from 13.9% in 1995 to 29.9% in 2015. As a result, the share of solar power generation in renewable energy generation increased from 5.7% in 2010 to 29.1% in 2015 (w 1995 this share was only 0.01.%). Also the cumulative installed solar power capacity increased by about 2000 times between 1995 and 2015 (from 49 MW in 1995 to 94864 MW in 2015).

Such rapid development of renewable energy sources in the EU countries results from the common energy policy aimed at improving energy security and reducing greenhouse gas

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

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

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

emissions. In accordance with the Directive 2009/28/EC, the EU member states should increase the share of energy obtained from renewable sources in their overall energy consumption. The European climate and energy package specifies targets to be achieved by 2020.

Since the EU countries differ a lot when solar radiation is taken into account, the issue of the extent to which investments in solar power translate into the volume of electricity generation in particular EU countries seems worth investigating. Since solar electricity generation, which can replace non-renewable energy sources, influences their economic, social and environmental aspects, it requires in-depth analysis.

The aim of the study is to assess efficiency of investments in solar power in the EU countries (excluding Estonia and Latvia, which did not generate any solar power in the analysed period). The analysis uses cross-sectional data set from 2015. The DEA method is implemented to evaluate relative efficiency of particular countries' solar power performance in electricity generation. The installed solar capacity is used as the input variable, and it approximates solar power investment or capital. The output variables refer to the volume of solar electricity generation, which allows for measuring solar power performance in electricity generation, and selected factors from the areas of energy security (energy dependence), environmental protection, and economy, which allows for identifying the benefits of investments in solar power in the EU countries.

Efficiency of renewable energy is a frequently studied issue, and most papers investigating to use Data Envelopment Analysis (DEA) as the empirical framework. Mardani et al. (2017) offer a comprehensive review of DEA applications in energy efficiency. A comparison of different sources of renewable energy is proposed in Cristóbal (2011), Kim et al. (2015), Karpińska (2016) and the assessment of efficiency of renewable energy sources referring to wind can be found in, among others, Wu et al. (2016), Sağlam (2017), Frodyma et al. (2018), to biogas – in Lijó et al. (2017), and to hydroelectric power – in Barros et al. (2017). Only several papers deal with solar power efficiency (Sueyoshi and Goto, 2014; Sueyoshi and Goto, 2017; Imteaz and Ahsan, 2018).

The paper contributes to the existing literature in two aspects. Firstly, its novelty lies in the selection of objects for the analysis, that is the EU countries, which have not been compared in this context so far. There are two reasons why the EU countries are an interesting object of study. First, all EU countries are obliged to meet the targets set in the climate and energy package. Some of them offer certain incentives (e.g. feed-in tariffs, green certificates), which are supposed to increase investments in solar power but might result in inefficiency of solar

power installed. Second, the EU countries are highly diversified with reference to their solar resources. Southern Europe countries have a great advantage in the solar potential over the northern ones. Secondly, the DEA models include variables describing environmental, economic and energy security aspects, which have not been analysed in previous studies. The inclusion of these factors offer a much broader view on solar efficiency in EU the countries.

2 Methodology and data

The CCR model, proposed by Charnes, Cooper, Rhodes in 1978 (Charnes et al., 1978), assumes constant returns to scale, and the modification of the model presented in 1984 by Banker, Charnes and Cooper (Banker et al., 1984) (the BCC model) allows for variable returns to scale. The BCC model allows for assessing countries not only from the perspective of pure technical efficiency (the best use of input/investments) but also from the perspective of economies of scale (operating within the area of optimal benefits), thanks to which it is possible to discover whether inefficiency of a given country in the area of solar power results from wasting investments or from operating within non-optimal economies of scale. The DEA method is based on a comparison of a group of decision making units (DMUs), in which each unit has a certain degree of freedom of decision. In other words, this method allows for identifying efficient units which later set the efficiency level desired for and possible to obtain by other units. In our study one European country is treated as one decision making unit. Following the assumptions of the DEA method, the inputs and the outputs are greater than or equal zero, and for each decision making unit there exists at least one input and one output greater than zero. The task is to find the minimum value of parameter θ, which makes it possible to decrease inputs in such a way that the efficiency level is not changed. A decision making unit is efficient if $\theta = 1$; a unit is inefficient if $\theta < 1$.

The data on which the empirical analysis is based describe different aspects of solar power efficiency in the EU countries. The sample of 26 European Union member states includes: Austria, Belgium, Bulgaria, Croatia, Cyprus, the Czech Republic, Denmark, Finland, France, Hungary, Greece, Germany, Ireland, Italy, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom. All data describe the EU countries in 2015 and are obtained from the European Commission webpage[4].

The DEA models include one input variable and a subset of four output variables.

[4]Energy datasheets: EU-28 countries (https://ec.europa.eu/energy/en/data-analysis/country) accessed on 10.01.2018.

The installed solar capacity (MW) per capita in 2015 (**CAP**) serves as the input variable in the DEA models. The variable is a proxy for total investment in solar power in particular countries, which cannot be measured directly. The data describing investments in renewable energy (including solar power) are available since 2010 for the EU counties and do not include previous investments (*EurObserv'ER*).

The most significant variable used as the output in the DEA models in our study is gross electricity generation from solar (TWh per capita) in 2015 (**GEN**), which is the most common output variable used in previous studies assessing efficiency of solar power.

The environmental output of solar power (**ENV**) is the second output variable in the DEA models, and it measures carbon dioxide emissions avoided thanks to solar power. In order to calculate the CO_2 avoided, the following formula is used:

$$ENV = \frac{GHE}{TPES} \cdot GEN,$$

where: GHE is total CO_2 emission from power station, TPES is total primary energy supply, and GEN is gross electricity generation from solar power per capita in 2015. For countries with the highest CO_2 emissions from power stations in total primary energy supply, the change into solar power generation is the most significant, as vast amounts of CO_2 are avoided.

The economic output of solar power (**ECON**) is the third output variable in our study. To calculate **ECON**, the prices of energy sources and the contribution of each energy source in TPSE in a particular country are needed. The following formula is used:

$$ECON = CFF \cdot GEN,$$

where: $CFF = \sum_{i=1}^{3} TPES_i \cdot price_i;$ where i indicates: coal, oil, natural gas, $TPES_i$ - total primary energy supply generated from i energy sources in 2015, *GEN* - gross electricity generation from solar per capita in 2015. The more expensive energy sources are used in a given country, the more efficient solar power in terms of economy is in it.

The energy dependence output of solar power (**DEP**) is the fourth output variable. This variable is calculated as:

$$DEP = \frac{IMPORT}{TPES} \cdot GEN,$$

where: *IMPORT* is total import of energy in 2015, *TPES* – total primary energy supply in 2015, *GEN* – gross electricity generation from solar per capita in 2015.

The more energy dependent a country is, the more beneficial - in terms of energy security - solar power generation is. On the other hand, if countries do not import any energy (i.e. they are energy self-sufficient), there is no benefit of solar power in this aspect.

The variables described above are used in the DEA model to analyse efficiency. The comparison of the analysed countries reveals that Germany has the highest installed solar capacity (490 MW per capita (**CAP**)), and the countries above the third quartile are: Italy, Belgium, Greece, Luxemburg, the Czech Republic and Malta. The countries below the first quartile are, apart from Ireland, which is the country with the lowest capacity, Finland, Poland, Sweden, Croatia, Hungary and Lithuania. When gross electricity generation from solar power (**GEN)** is considered, Germany is the leader (477 TWh per capita). Other countries with high values are: Italy, Greece, Spain, Belgium, Malta and the Czech Republic. The highest values of the environmental output (**ENV**) per capita are noted in Greece, Germany and Bulgaria. The greatest reduction of costs of energy sources replaced by energy generated by solar per capita (**ECON**) is noted in Italy, Malta and Belgium. Countries with the greatest volatility in the area of energy dependency per capita (**DEP**) are Malta, Greece and Belgium. Energy dependency indicates to what extent a country relies on import to meet its energy requirements. Summing up, the distribution of particular variables is skewed, which means that most EU countries have all variables below the average for the whole European Union. The lowest values of the variables are found in Ireland, Poland, Finland and Sweden.

3 Results

In order to analyse efficiency of investments in solar power in the EU countries, we consider various models with the capacity as the input variable and selected combinations of variables described in section 2 as the output variables. All models are presented in Table 1.

Table 1. Input-output variables of eight models.

	GEN	***GEN_ENV***	***GEN_ECON***	***GEN_DEP***	***ALL***
CAP	X	X	X	X	X
GEN	X	X	X	X	X
ENV		X			X
ECON			X		X
DEP				X	X

Note: in bold: variable name, in italics: the name of the model.

The relative efficiency scores (θ) of investments in solar power in 26 EU countries in 2015 for input-oriented BCC models are presented in Table 2. The position of a given country in the rank is given in the parentheses for each model. The efficiency levels of investments in solar power in the first model (*GEN*) are not very high. The overall efficiency (BCC efficiency) scores range from 0.26 to 1.00, and the average BCC efficiency score is 0.50. The relative efficiency scores in models with two output variables are higher than in models with one output variable. The average BCC efficiency scores are about 0.54-0.56 in all these models (*GEN_ECON*, *GEN_ENV*, *GEN_DEP*), and the relative efficiency scores range from about 0.27 to 1.00. The highest average BCC efficiency score equals 0.61 in the model with all outputs (*ALL*), and the relative efficiency scores range from 0.27 to 1.00. The results reveal that the efficiency scores of the last model (ALL), which is the most comprehensive one regarding the number of input and output variables, are greater than the efficiency scores of the remaining seven models.

Table 2 reveals that Germany, Spain and Ireland are the countries with the maximum efficiency score of solar power generation (*GEN*) (equal 1.00), while the least efficient country is the Netherlands (with the relative efficiency score 0.26). When the models account for additional aspects, such as the environmental output, the economic output and the energy dependence output, it is possible to obtain a broader picture of benefits connected with investing in solar power in the EU countries. Taking into consideration both solar power generation and the environmental output (*GEN_ENV*) (Table 2), increased efficiency is noted in Greece, in which, as in Germany, Spain and Ireland, the efficiency scores of solar power generation equal 1.00. Countries with coal as the main source used in electricity generation (such as Poland (in 2015 79% of its total electricity generation comes from plants using coal), the Czech Republic (49%), Bulgaria (46%), Greece (43%), and Germany (44%), together with Cyprus and Malta, in which 91% of electricity is generated in plants using heating oil) gain the most: the efficiency score in Bulgaria increases by 87%, in Poland by 63%, in Cyprus by 59%, the Czech Republic by 52% and in Malta by 44%. When both solar power generation and the economic output (*GEN_ECON*) (Table 2) are taken into account, the most efficient countries with respect to investing in solar power include, apart from Germany, Spain and Ireland, also Cyprus, Italy and Malta. Taking into account the economic aspect, i.e. the costs of generating energy from non-renewable energy sources, an increase in the efficiency scores (a greater value of parameter θ by 127% in Malta, in Cyprus by 72%, in Belgium by 48%, in Lithuania by 33%, in Luxembourg by 28%) and a higher position in the rank is noted in countries in which the main energy source is oil (i.e. Malta, in which in 2015 as much as 92% of total electricity generation

comes from plants using heating oil and Cyprus (91%)), or natural gas (i.e. Lithuania (40% of its energy comes from plants using natural gas), Belgium (35%), Luxembourg (30%)). Apart from the improvement of their position in the rank, the value of parameter θ also increases.

Table 2. The efficiency scores of the input- and output oriented BCC models.

Country	***GEN***	***GEN_ENV***	***GEN_ECON***	***GEN_DEP***	***ALL***
AT	0.35 (17)	0.35 (18)	0.39 (17)	0.35 (19)	0.39 (20)
BE	0.34 (18)	0.34 (19)	0.50 (10)	0.42 (14)	0.50 (13)
BG	0.47 (9)	0.88 (7)	0.47 (12)	0.47 (12)	0.88 (8)
CY	0.58 (7)	0.92 (5)	**1.00 (1)**	0.60 (8)	**1.00 (1)**
CZ	0.38 (13)	0.58 (11)	0.38 (19)	0.38 (16)	0.58 (12)
DE	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**
DK	0.27 (25)	0.28 (25)	0.27 (25)	0.28 (25)	0.29 (24)
ES	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**
FI	0.33 (19)	0.33 (20)	0.33 (20)	0.33 (20)	0.33 (22)
FR	0.37 (15)	0.37 (16)	0.40 (14)	0.37 (17)	0.40 (17)
GR	0.98 (4)	**1.00 (1)**	0.98 (7)	**1.00 (1)**	**1.00 (1)**
HR	0.43 (11)	0.43 (13)	0.43 (13)	0.43 (13)	0.43 (16)
HU	0.27 (24)	0.27 (26)	0.27 (24)	0.27 (26)	0.27 (26)
IE	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**
IT	0.88 (5)	0.88 (6)	1.00 (1)	0.88 (6)	1.00 (1)
LT	0.38 (14)	0.38 (15)	0.50 (11)	0.57 (9)	0.58 (11)
LU	0.31 (21)	0.31 (22)	0.40 (16)	0.31 (22)	0.40 (19)
MT	0.44 (10)	0.63 (9)	**1.00 (1)**	**1.00 (1)**	**1.00 (1)**
NL	0.26 (26)	0.30 (23)	0.26 (26)	0.48 (11)	0.48 (14)
PL	0.28 (23)	0.46 (12)	0.28 (23)	0.28 (24)	0.46 (15)
PT	0.62 (6)	0.68 (8)	0.62 (8)	0.66 (7)	0.69 (9)
RO	0.52 (8)	0.62 (10)	0.52 (9)	0.52 (10)	0.62 (10)
SE	0.35 (16)	0.35 (17)	0.38 (18)	0.35 (18)	0.38 (21)
SI	0.40 (12)	0.40 (14)	0.40 (15)	0.40 (15)	0.40 (18)
SK	0.33 (20)	0.33 (21)	0.33 (21)	0.33 (21)	0.33 (23)
UK	0.28 (22)	0.28 (24)	0.28 (22)	0.28 (23)	0.28 (25)

Note: In bold: the maximum efficiency score of the solar power (equal 1.00).

After adding the energy security aspect to solar power generation (*GEN_DEP*) (Table 2), Malta joins the list of the most efficient countries (its efficiency score increases by 127% in comparison with the basic model) and Greece (2%). Considerable improvement is also noted in Netherlands (the efficiency score grows by 87%), Lithuania (49%) and Belgium (by 23%).

Including all variables in the analysis of efficiency yields the greatest number of solar efficient countries. Countries already mentioned above (efficient with respect to two various input variables), remain efficient also in this stage: i.e. Germany, Spain, Ireland, Greece, Malta, Italy and Cyprus. Hungary turns out to be the least efficient country in this extended model. In other words, if efficiency of investments in solar power is considered not only from the perspective of electricity generation but also from the perspective of social, economic and environmental benefits resulting from replacing non-renewable energy sources with solar power, the countries which benefit the most are: Malta (the efficiency score increases by 127%), Bulgaria (87%), Netherlands (87%), Cyprus (72%), and Poland (63%).

Conclusions and discussion

The objective of the study is to assess efficiency of investments in solar power in the EU countries. The DEA method is used as an empirical framework. The installed solar capacity is used as the input variable. Four variables: solar electricity generation, the environmental output, the economic output and the energy dependence output are used to measure efficiency of investments in solar power, which also allows for identifying the benefits of investing in solar power in the EU countries.

In particular, our findings can be summarised in two main points.

First, the study reveals that Germany, Spain and Ireland are efficient both when only solar power generation is considered and when the relation between the investments in solar power and the volume of energy generated by solar farms is considered. The results are quite surprising, since only one of these countries is located in the south of Europe, however, they are in line with the results obtained by Sueyoshi and Goto (2014).

Second, the inclusion of environmental, economic or energy dependence outputs, demonstrates that different countries benefit from solar power investments. If efficiency of solar power takes into account socio-economic and environmental benefits resulting from replacing non-renewable energy sources with solar power, Malta, Cyprus, Greece, and Italy gain the most.

Poland, Malta, Bulgaria, Cyprus and the Czech Republic gain the most as far as investments in solar powers and the improvement of the air quality and the reduction of greenhouse gases

emissions are concerned. Including the environmental output in the analysis reveals that efficiency increases the most in countries where the main energy source is coal. This result is connected with a high level of carbon dioxide emissions resulting from burning coal, which is considerably reduced when coal is replaced with solar power. Malta, Cyprus, Belgium, Lithuania and Luxembourg gain the most in the economic aspect, as power generated from solar is much cheaper than power generated from fossil fuels. That is why including the economic output indicates the greatest improvement in efficiency in countries which obtain electricity from natural gas or heating oil instead of coal. This effect is the consequence of high prices of these fuels.

Malta, the Netherlands, Lithuania and Belgium gain the most in the aspect of energy security taking into account the dependence output in the assessment of efficiency of investments in solar power.

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

Banker, R. D., Charnes, A. & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, *30*(9), 1078-1092.

Barros, C. P., Wanke, P., Dumbo, S. & Manso, J. P. (2017). Efficiency in angolan hydro-electric power station: A two-stage virtual frontier dynamic DEA and simplex regression approach. *Renewable and Sustainable Energy Reviews*, *78*, 588-596.

Charnes A., Cooper W. W. & Rhodes E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research, 2*, 429 – 444.

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). Efektywność inwestycji w energię wiatrową w krajach UE. *Research Papers of the Wroclaw University of Economics. PN 507 Taksonomia, 30,* 158-162.

Imteaz, M. A. & Ahsan, A. (2018). Solar panels: Real efficiencies, potential productions and payback periods for major Australian cities. *Sustainable Energy Technologies and Assessments*, *25*, 119-125.

Karpińska, L. (2016). Development of EU energy market – an insight to its efficiency. *Folia Oeconomica Cracoviesia*, 57, 69-82.

Kim, K. T., Lee, D. J., Park, S. J., Zhang, Y. & Sultanov, A. (2015). Measuring the efficiency of the investment for renewable energy in Korea using data envelopment analysis. *Renewable and Sustainable Energy Reviews*, *47*, 694-702.

Lijó, L., Lorenzo-Toja, Y., González-García, S., Bacenetti, J., Negri, M. & Moreira, M. T. (2017). Eco-efficiency assessment of farm-scaled biogas plants. *Bioresource technology*, *237*, 146-155.

Mardani, A., Zavadskas, E. K., Streimikiene, D., Jusoh, A. & Khoshnoudi, M. (2017). A comprehensive review of data envelopment analysis (DEA) approach in energy efficiency. *Renewable and Sustainable Energy Reviews*, *70*, 1298-1322.

Sağlam, Ü. (2017). A two-stage data envelopment analysis model for efficiency assessments of 39 state's wind power in the United States. *Energy Conversion and Management*, *146*, 52-67.

Sueyoshi, T. & Goto, M. (2014). Photovoltaic power stations in Germany and the United States: A comparative study by data envelopment analysis. *Energy Economics*, 42, 271-288.

Sueyoshi, T. & Goto, M., (2017). Measurement of returns to scale on large photovoltaic power stations in the United States and Germany. *Energy Economics*, *64*, 306-320.

Wu, Y., Hu, Y., Xiao, X. & Mao, C. (2016). Efficiency assessment of wind farms in China using two-stage data envelopment analysis. *Energy Conversion and Management*, *123*, 46-55.

Spatial differentiation of the remuneration and labour productivity in Polish agriculture

Justyna Góral [1], Włodzimierz Rembisz [2]

Abstract

The objective of the study was to show the strength of the dependency of the distribution of the labour factor productivity variable in Polish agriculture and remuneration variable in regional terms. The hypothesis adopted is that the remuneration differentiation is correlated with the labour productivity differentiation in spatial terms. Analysis was conducted on a basis of the analytical aspect and empirical studies based on the information from the CSO Local Data Bank. To examine these dependencies, the authors used the Gini coefficient and regression and correlation analysis. We observed the greater differentiation and variation in the labour factor productivity rather than in its remuneration. The distributions of those variables in spatial terms did not match each other in terms of their equality. The labour productivity differentiation resulted mainly from the differentiation in the capital-labour ratio and land-labour ratio.

Keywords: *labour productivity, remunerations, Gini coefficient, agriculture*
JEL Classification: *E24, O11, R11, Q12*
DOI: 10.14659/SEMF.2018.01.15

1 Introduction

In microeconomics, it is assumed, according to the Neo-Classical trend, that the relationships between the labour factor productivity and its remuneration are important. Taking, above all, these variables into account, it is assumed that the level and changes in the labour factor remuneration should result from, or be shaped by, the level of and changes in its productivity. We can also identify other determinants undermining the relationship between those two variables, but the literature enables a sufficient justification of the importance of the analysed relationship. It is also undertaken in economics of agriculture.

The objective of the study was to show the relationship and strength of the dependency of the distribution of the labour factor productivity variable in Polish agriculture and remuneration variable in spatial (regional) terms. The main question was: to what extent the unequal distribution of the labour factor remuneration is due to the unequal distribution of its productivity? Therefore, the first hypothesis was adopted that the differentiated remuneration is correlated with the differentiated labour productivity in regional terms. Moreover, the secondary study objective was to show the

[1] Corresponding author: Institute of Agricultural and Food Economics - National Research Institute in Warsaw, Mathematics Application in Agricultural Economics Department, justyna.goral@ierigz.waw.pl.
[2] Institute of Agricultural and Food Economics - National Research Institute in Warsaw, Mathematics Application in Agricultural Economics Department, wrembisz@gmail.com.

main reasons for the differentiation of the labour factor remuneration in the regional system. These reasons, according to the theory of economics of agriculture and microeconomics, include the capital-labour ratio and land-labour ratio. They are, respectively: relationship between the employed capital factor and the number of the employed and the relationship between the land factor and one employed person. In fact, they are characteristic of the agrarian structure, i.e. the factor structure of agriculture. This can be considered as normal. It is assumed that structural changes in agriculture are a major source of growth in income of agricultural producers. Consequently, the second hypothesis appeared stating that the unequal distribution of the labour productivity corresponds to the unequal distribution of the capital-labour ratio and land-labour ratio analysed in regional terms. The authors proved this on a basis of theoretical assumptions, as well as analytically.

2 Analytical assumptions

The issue of the relationship between the labour remuneration and productivity results from the theory of the producer's equilibrium. The producer maximising his objective function should balance the remuneration level with the marginal productivity of each production factor. This equilibrium is reached when the production factor remunerations are equal to their marginal productivities and when the marginal productivities are equal to average productivities (production function theory). This latter point also defines the sphere of rational management in terms of technical efficiency (Rembisz and Sielska, 2015). We can adopt the following relationship for the given product prices p_y as constants (which meets the conditions of competitive equilibrium):

$$\frac{\partial_y}{\partial_L} = \frac{y}{L} = p_L \quad \text{and} \quad \frac{y}{L} = w_L \quad \text{we have:} \quad w_L \approx p_L.$$

In fact, when we cancel the assumption that the product prices are determined, the remuneration is defined both by the labour productivity and by the level of product prices. However, in our analysis the prices are determined, therefore, we disregard their impact $w_L \cdot p_Y \approx p_L$.

Therefore, the labour factor productivity should determine its remuneration. The inequality in terms of concentration of both these values, i.e. w_L and p_L for the agricultural sector in regional terms has been analysed (Tsoku and Matarise, 2014; Guiteras and Jack, 2018).

It has been assumed that the labour factor productivity is determined in the sense of identity by the capital-labour ratio (relationship between the capital factor K and the labour factor L) and by the land-labour ratio (relationship between the land factor Z and the labour factor L). For the given capital factor productivity:

$\frac{y}{K} = a$ and the land factor productivity $\frac{y}{Z} = b$, we have: $w_L \approx \frac{K}{L} \cdot a$ and $w_L \approx \frac{Z}{L} \cdot b$.

3 Methodological assumptions

These dependencies were verified against empirical data based on the statistics i.e. the Gini coefficient (Barabesi et al., 2015; Ghosh, 2015; Rad et al., 2016; Tyrowicz et al., 2017). This is a commonly used indicator of the unequal distribution of variables (values), mostly income. It is within the range (0-1). The higher is the value of the indicator (1), the greater is the degree of concentration of the variable and thus the greater are the inequalities of, e.g. income (Ghosh, 2015; Rad et al., 2016; Prendergast and Staudte, 2016). The Gini coefficient is a normalised value, which makes it easier to make comparisons. On a basis of this statistics, the differentiation of the productivity and remuneration has been compared, by voivodeships. They were shown in the order adequate to the presented analytical aspect[3].

4 Differentiation of the labour productivity and remuneration – study results

The Gini coefficient and the coefficient of variation for section A (Polish Classification of Activity 2007, where section A is agriculture, forestry, fishing and fisheries) have been estimated. It is assumed conventionally that section A illustrates agriculture. Data from the CSO Local Data Bank for the years 2005-2014 has been analysed (Fig. 1 in appendix). The results for the labour factor productivity (gross value added per employee – GVA per employee in constant prices) and the average monthly gross remunerations in section A are presented in Table 1.

The inequality and variation of the labour productivity were relatively high. These results proved to be relatively stable over time (Harasim, 2006; Nowak, 2010; Kuźmar, 2017). The inequality of remunerations was significantly lower. There was an important condition that the differentiation and variation of remunerations proved to be lower than those of the labour productivity in spatial terms. In addition, small spatial inequalities of remunerations decreased over time, despite the fact that the differentiation of the labour productivity was invariable. This indicates a weak relationship between (in)equalities of the remuneration and productivity. This phenomenon may result from public aid transfers (Alexandri, 2017). The differentiation of remunerations is much lower and characterised by the more equal distribution. This confirms the above finding negatively verifying the first hypothesis. A discrepancy between the distributions of these two indicators is visible. From the point of view of social objectives, this is a positive phenomenon. From the purely economic point of view – it is not.

[3] The method of presenting the results refers to the approach applied by S. Kuźmar entitled *Differentiation of the regional labour productivity in Poland* (Contemporary Economic Issues, No. 11, 2015, pp. 137-147), which the authors found very accurate and comprehensible.

Table 1. Differentiation of the labour productivity per employee and average monthly gross remunerations in section A, by voivodeships in Poland in the years 2005-2014, 2005 = 100.

Years	Gini coefficient (labour productivity)	Coefficient of variation of labour productivity [%]	Gini coefficient (remunerations)	Coefficient of variation of remunerations [%]
2005	0.264	48.21	0.048	8.71
2006	0.265	48.27	0.042	7.84
2007	0.227	41.26	0.038	7.06
2008	0.218	39.45	0.046	9.42
2009	0.254	46.10	0.031	5.72
2010	0.251	45.47	0.034	6.37
2011	0.239	43.34	0.037	7.05
2012	0.252	45.53	0.042	8.26
2013	0.254	45.97	0.037	6.96
2014	0.260	46.85	0.035	6.84

The above finding has been confirmed by panel data analysis[4] (Table 2). The analyzed data was of cross-sectional nature, therefore a panel of 16 voivodeships in the years 2005-2014 was created. These dependencies were statistically significant in the case of the correlation coefficient (0.883). Panel data analysis showed slight impact of labour productivity on the remunerations. These findings, however, did not deny the logic of a theoretical relationship between the productivity and remuneration of the labour factor. They only showed that the inequality of the labour productivity was different than that of remunerations. It is obvious to restore this relationship and treat it as a standard. However, referring to the reasons (according to the analytical aspect[5]), the differentiation and inequality of the distribution of the relationship between the capital and land

[4] Panel data models are special models built on the basis of time-space data that describe a fixed group of objects in more than one period. The models can be in the form of (1) models with free expression decomposition (FEM) or (2) models with decomposition of a random component (Random Effects Model - REM). The model's assessment is based on chi-square statistics, which in turn is based on the reliability function (statistic Likelihood Ratio Test) and F statistics. The choice between the FEM and REM model is made using the Hausman test (Wooldrige, 2002).

[5] $w_L \approx \frac{K}{L} \cdot a$ and $w_L \approx \frac{Z}{L} \cdot b$ for the given voivodeships and given values ij.

factors to the labour factor in regional terms have been shown. The values of the Gini coefficients and the coefficients of variation have been estimated, which was shown in Table 3.

Table 2. Panel data model (REM) illustrating the dependency of remunerations (Y) on the labour productivity (X – GVA per employee), 2005 = 100.

Specification	***Coefficients***	***Standard error***	***t Stat***	***p-value***	
Constans	17.919	4.967	3.607	0.0003	***
GVA per emploee	0.069	0.015	4.507	<0.0001	***
GVA per emploee in previous year	0.013	0.019	0.679	0.4970	
Remunerations in previous year	0.948	0.031	3.782	<0.0001	***
Number of observations				*160 (16 voivodeships and 10 years)*	
Log Likelihood				*648.602*	
Hausmann test:		*1681.241 (p-value = 0.006)*			
Pesaran CD test for cross-sectional dependence:		*-1.972 (p-value = 0.048)*			
Test for normality of residual:		*6.636 (p-value = 0.036)*			

Table 3. Differentiation of the relationship of the land-labour ratio and capital-labour ratio (2005 = 100), by voivodeships, in the years 2005-2014.

Years	**Gini coefficient (land-labour ratio)**	**Coefficient of variation of land-labour ratio [%]**	**Gini coefficient (capital-labour ratio)**	**Coefficient of variation of capital--labour ratio [%]**
2005	0.188	37.55	0.209	38.59
2006	0.194	38.33	0.210	39.00
2007	0.201	38.86	0.213	38.95
2008	0.195	39.00	0.206	39.15
2009	0.207	41.87	0.217	41.61
2010	0.210	44.85	0.225	44.44
2011	0.205	43.52	0.237	46.19
2012	0.211	46.09	0.228	45.08
2013	0.218	47.03	0.248	46.55
2014	0.219	48.23	0.249	47.00

As we can see, the land-labour ratio (indicator roughly illustrating the agrarian structure) was highly differentiated with the unequal distribution. However, when comparing the data from Tables 1 and 3, it is evident that the inequality and differentiation of the land-labour ratio did not deviate from that for the labour productivity. This indicates the validity of adopting the second hypothesis and analytical assumption regarding the labour productivity conditions. The similar finding was outlined from comparing the differentiation and inequality of the capital-labour ratio and the labour productivity (Table 3). As we can see, the capital-labour ratio was characterised by the highly unequal distribution, for which the values of the Gini coefficient oscillated within the limits of 0,209-0,249. Even higher values were reached by the coefficient of variation (38.59-47.00%). In addition, the analysed differentiation of this coefficient increased in the analysed period of time more than in the case of the labour productivity.

From this analysis, it can be concluded that the capital-labour ratio and land-labour ratio (to some extent, the approximation of the indicator of the structure and concentration in agriculture) highly explain the differentiation in the labour factor productivity.

Analysis of the unequal distribution curves of the land-labour ratio and capital-labour ratio in the above charts confirms previous observations. The inequality and differentiation of the labour productivity and the land-labour ratio and capital-labour ratio can be linked, which verifies the second hypothesis positively. The same cannot be said regarding the differentiation of remunerations in relation to the labour productivity.

In order to verify the above observations, the dependency between the analysed values has been analysed statistically (correlation and panel data analysis). The results are shown in Table 4. These dependencies were statistically significant. Correlation between the labour productivity level and the land-labour ratio and capital-labour ratio amount to 0.919. and 0.868. This was a basis for positive verification of the second hypothesis (of the relationship between the differentiation of the labour productivity and its land-labour ratio and capital-labour ratio).

In analytical terms, in accordance with the last formulae, the regional differences in the productivity levels of production factors, i.e. capital and land, were analysed[6] (Table 5). The differentiation and inequality of the distribution in spatial terms in the case of the land productivity was milder than in the case of the productivity and land-labour ratio. This factor was similarly productive in various parts of the country. A similar conclusion resulted from analysis of the

[6] $\frac{y_{ij}}{K_{ij}} = a_{ij}$, $\frac{y_{ij}}{Z_{ij}} = b_{ij}$ for given value (production, capital, labour) and for the given voivodeship.

differentiation of the capital productivity. Here, the estimated indicators demonstrated the smallest differentiation and discrepancies.

Table 4. Panel data model (REM) illustrating the dependency between the differentiation of the labour productivity and its land-labour ratio and capital-labour ratio (2005 = 100).

Specification	*Coefficients*	*Standard error*	*t Stat*	*p-value*	
Constans	91.624	9.728	9.418	<0,0001	***
Land-labour ratio	0.145	0.022	6.547	<0,0001	***
Capital-labour ratio	0.192	0.073	2.629	0.008	***
Number of observations				*160 (16 voivodeships and 10 years)*	
Log Likelihood				*834.739*	
Hausmann test:	*51.394 (p-value = 0.000)*				
Pesaran CD test for cross-sectional dependence:	*13.829 (p-value = 0.000)*				
Test for normality of residual:	*21.289 (p-value = 0.000)*				

Table 5. Inequality and differentiation of the land productivity (agricultural production value per ha) and capital productivity (agr. production value in relation to the intermediate consumption and depreciation value) in section A, by voivodeships, in the years 2005-2014, 2005 = 100.

Years	**Gini coefficient for land productivity**	**Coefficient of variation of land productivity [%]**	**Gini coefficient for capital productivity**	**Coefficient of variation of capital productivity [%]**
2005	0.125	23.40	0.053	10.09
2006	0.123	22.51	0.048	8.98
2007	0.105	19.48	0.045	8.69
2008	0.112	20.23	0.050	9.17
2009	0.105	19.42	0.044	8.48
2010	0.108	19.57	0.054	10.17
2011	0.109	19.64	0.072	13.41
2012	0.117	21.36	0.063	12.22
2013	0.121	22.35	0.060	10.91
2014	0.123	22.45	0.061	11.38

The productivity distributions are more equal when compared to the labour productivity distribution, land-labour ratio and capital-labour ratio. This may point to the fact that the land-labour ratio and capital-labour ratio actually affect the differentiation of the labour productivity. The distributions of these values were similarly unequal. This confirmed the second hypothesis and the adopted analytical assumptions. Indeed, in the sense of the rules of agricultural economics, this means that the agrarian structure and capital-intensive production techniques are of paramount importance. This is confirmed indirectly by analysis of the statistics for the last analysed values i.e. productivity indicators in Table 6 (data base from Eurostat and CSO).

Table 6. Panel data model (REM) illustrating the dependency between the labour productivity and the capital and land factor productivity in the years 2005-2014, 2005 = 100.

Specification	***Coefficients***	***Standard error***	***t Stat***	***p-value***	
Constans	30.947	11.644	2.658	0.008	***
Land factor productivity	0.826	0.063	13.382	<0,0001	***
Capital factor productivity	4.678	6.589	0.710	0.478	
Number of observations				*160 (16 voivodeships and 10 years)*	
Log Likelihood				*788.099*	
Hausmann test:		*16.081 (p-value = 0.000)*			
Pesaran CD test for cross-sectional dependence:		*16.487 (p-value = 0.000)*			
Test for normality of residual:		*5.247 (p-value = 0.072)*			

* Correlation between the labour productivity and the productivity of land amount to 0.638. Correlation between the labour productivity and capital productivity amount to 0.096.

Conclusions

The article deals with the issue of the relationship between the labour remuneration and productivity and the factors shaping this relationship, by voivodeships. We observed the greater differentiation and variation in the labour factor productivity rather than in its remuneration. The distributions of those variables in spatial terms did not match each other in terms of their equality. The labour productivity differentiation resulted mainly from the differentiation in the capital-labour ratio and land-labour ratio. Here, the inequalities in the distribution of these variables matched each other. This had a negligible impact on the spatial differentiation of the capital and land productivity.

Analysis was conducted on a basis of the analytical aspect, which was therefore positively verified. The overall conclusion, the alignment of the differentiation in the labour factor

remuneration (agricultural income) in spatial terms did not result from the decreasing labour productivity differentiation. The alignment of the distribution of this latter indicator was more due to the decrease in the differentiation of the capital-labour ratio and land-labour ratio, and less due to the productivity of these factors. As a result, the alignment of differences in the amount of the analysed indicators in spatial terms may not be an important source of the agricultural development, as it is most often assumed. This can, however, be relevant for determining the agricultural policy, in particular in terms of cohesion. Authors disregarded the impact of subsidies and found them to be a less important source of agricultural development than labor productivity (the main source of growth and development). However, it is necessary to recognize these relationships in further research due to the high values of constant parameters and standard errors in all panel data models.

References

Alexandri, C. (2017). Labour productivity in Romania's agriculture. A regional analysis by farm types. *Agricultural Economics and Rural Development, New Series, Year XIV, 1,* 3-13.

Benito, J. & Ezcurra, R. (2007). Spatial Disparities in Productivity and Industry Mix: The Case of the European Regions. *Urban Studies 44*, 177-194.

Ghosh, S. (2015). Computation of Spatial Gini Coefficients. *Communications in statistics – theory and methods, 44 (22),* 4709-4720.

Gołaś, Z. (2010). Wydajność i dochodowość pracy w rolnictwie w świetle rachunków ekonomicznych dla rolnictwa. *Zagadnienia Ekonomiki Rolnej, 3,* 19-42.

Guiteras, R. P. & Jack, B. K. (2018). Productivity in piece-rate labor markets: Evidence from rural Malawi. *Journal of Development Economics, 31,* 42-61.

Harasim, A. (2006). Dobór wskaźników do oceny regionalnego zróżnicowania rolnictwa. *Regionalne zróżnicowanie produkcji rolniczej w Polsce, Monografia 3,* IUNG-PIB, 53-71.

Kuźmar, S. (2015). Zróżnicowanie regionalnej wydajności pracy w Polsce. *Współczesne Pro-blemy Ekonomiczne, 11,* 137-147.

Nowak, A. (2010). Zróżnicowanie regionalne wydajności pracy w rolnictwie. *Zagadnienia Doradztwa Rolniczego, 1-2,* 103-112.

Prendergast, L. A. & Staudte, R. G. (2016). Quantile versions of the Lorenz curve. *Electronic Journal of Statistics, 10*(2), 1896-1926.

Rad, N. N., Borzadaran, G. R. & Yari, G.H. (2016). Maximum entropy estimation of income share function from generalized Gini index. *Journal of Applied Statistics, 43(16),* 2910-2921.

Rembisz, W. & Sielska, A. (2015). Wydajność pracy a dochody producentów rolnych. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach, 235,* 208-221.

Tsoku, J. T. & Matarise, F. (2014). An Analysis of the Relationship between Remuneration (Real Wage) and Labour Productivity in South Africa. *Journal of Educational and Social Research* MCSER Publishing, Rome-Italy, 4(6), 59-66.

Tyrowicz, J., van der Velde, L. & Svejnar, J. (2017). Effects of labour reallocation on productivity and inequality insights from studies on transition. *Journal of Economic Surveys*, *31*(3), 712-732.

Wooldrige, J. (2002). *Econometric analysis of cross section and panel data*. The MIT Press, Londyn, 19-46.

APPENDIX



Fig. 1. Labour productivity (blue line) and remunerations (red) in 2005-2014 (2005 = 100).

Application of Generalized Data Envelopment Analysis on Warsaw Stock Exchange

Sergiusz Herman[1]

Abstract

Fundamental analysis is one of the techniques used by stock investors. Its main aim is the estimation of financial situation of companies. Companies' financial statements are the source of information that allow for estimation of financial ratios for various operations of these companies. Elements of financial statements are also used to analyse analyse operating efficiency of companies with the use of parametric and non-parametric methods. Efficiency estimated with the use of these methods is rarely used for the assessment of joint-stock companies. There are three main goals of the research. Firstly, it aims to verify whether there is a statistical valid dependence between operating efficiency of joint-stock companies in Poland and their return rates on the Warsaw Stock Exchange. The second aim of the research is to point variables which should be used to estimate operating efficiency of companies in chosen industries. The other aim is to verify if an investor may benefit from information about operating efficiency of joint-stock companies when building an investment portfolio. Empirical studies were conducted on 72 joint-stock companies in Poland. These companies represent following industries: construction, clothes and cosmetics industry, food and drinks industry. The calculations were performed using the Generalized Data Envelopment Analysis.

Keywords: *operating efficiency, stock returns, Data Envelopment Analysis, stock exchange*
JEL Classification: C610, G300
DOI: 10.14659/SEMF.2018.01.16

1 Introduction

A good choice of securities for an investment portfolio is key for every investor on the Warsaw Stock Exchange. A fundamental analysis makes it possible to differentiate between companies in poor and good financial condition. One of company functional areas assessed by an investor is its performance (effectiveness). The inventory turnover, accounts receivable and payable ratios are the most popular ratios used for that purpose.

Numerous parametric (e.g. Stochastic Frontier Approach, Distribution-Free Approach) and non-parametric (e.g. Data Envelopment Analysis, Free Disposal Hull) methods for measuring efficiency are listed in the literature. Due to its computation complexity, results of these analysis are rarely used on the Warsaw Stock Exchange. G. Szafrański (2004) is a Polish author who conducted research on companies' operating efficiency. The study aimed to show whether DEA allows for the optimal choice of ratios which are widely used in

[1] Corresponding author: Poznań University of Economics and Business, Department of Econometrics, al. Niepodległości 10, 61-875 Poznań, Poland, e-mail: sergiusz.herman@ue.poznan.pl.

fundamental analysis. The study confirmed that there is a relation between a efficiency ratio and a change of stock price. It is worth mentioning that the author chose a priori a set of inputs and outputs of companies' operations. Moreover, the sample consisted of companies from various industries.

The Generalized DEA (GDEA) solved the problem of the a priori choice of variables for inputs and outputs. The method was used on a stock exchange by Edirisinghe and Zhang (2007) as well as Avkiran and Morita (2010). However, authors of both publications did not estimate companies' efficiency ratio. Financial indicators used in a standard fundamental analysis served as inputs and outputs and their results showed a financial condition of studied companies. The study presented in this article aimed to verify whether there is a valid statistical dependence between operating efficiency of joint-stock companies listed on the Warsaw Stock Exchange and their return rates. Next, the author investigated what variables should be used to estimate operating efficiency of companies from diverse industries of the Polish economy. The Generalized Data Envelopment Analysis (GDEA) was used for that purpose. The results revealed whether, when creating a securities portfolio, information on companies' operating efficiency may benefit an investor on the stock exchange.

2 Methodology

To conduct the research, a relevant sample of joint-stock companies listed on the Warsaw Stock Exchange had to be collected. Taking into account the methodology and analysis goal, the author chose companies representing three industry indexes: WIG-construction, WIG-food and WIG-clothes. Companies were chosen based on the following criteria: availability of their financial data and having been listed on the Warsaw Stock Exchange since 2011 or earlier. Information from the Notoria Serwis database was used to choose the sample. As a result, financial information of the following companies was collected:

- 37 joint-stock companies from the WIG-construction index (financial data for 2010-2016),
- 15 joint-stock companies from the WIG-clothes index (data for 2010-2016),
- 20 joint-stock companies from the WIG-food index (data for 2011-2016).

Data from financial statements was used as inputs and outputs of studied companies` operations. In total, 23 items from balance sheets and profit and loss accounts were used in the study (Table 1). The Data Envelopment Analysis (DEA) was used to measure operating efficiency of joint-stock companies. It is one of the most popular non-parametric methods of measuring companies' operating efficiency which allows for assessing efficiency of Decision

Making Units (DMU) defined by various inputs and/or outputs by solving a relevant optimization problem.

Table 1. Inputs and outputs underlying the research.

Potential inputs	**Potential outputs**
total assets	*net revenues from sales of products, goods and materials*
fixed assets	*gross profit (loss) on sales*
tangible fixed assets	profit (loss) on sales
intangible assets	*other operating revenues*
current assets	profit (loss) on operating activities
Inventory	*financial revenues*
short-term receivables	gross profit (loss)
cash and other pecuniary assets	*net profit (loss)*
equity	
share capital	
supplementary capital	
long-term liabilities	
short-term liabilities	
other operating expenses	
financial expenses	

The research used the BCC model named after its authors R.D. Banker, A. Charnes and W.W. Cooper (1984) who created it in 1984. The input-oriented BCC model (BCC-I), named after first letters of authors' names, may be expressed by:

$$min\theta_o \tag{1}$$

subject to:

$$\sum_{j=1}^{n} x_{ij}\lambda_j \leq \theta_o x_{io} \quad i = 1, \dots, m, \tag{2}$$

$$\sum_{j=1}^{n} y_{rj}\lambda_j \geq y_{ro} \quad r = 1, \dots, s, \tag{3}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \;\; \lambda_j \geq 0 \quad j = 1, \dots, n. \tag{4}$$

where:

θ_0 – represents the efficiency score of unit *o*,

x_{ij} – is the amount of input *i* used by unit *j* (*i*=1, … , *m*),

y_{rj} – is the amount of output *r* used by unit *j* (*r*=1,2, …, *s)*.

λ_j – represent the variables that identify the benchmarks for inefficient units.

The reason for choosing the model was its result invariance which means that adding a constant to any result does not alter the optimum result of the problem.

Global analysis was carried out with the use of relevant financial data as inputs and outputs. The same joint-stock company was analyzed for various years.

3 Study of the relationship between operating efficiency and return ratios of joint-stock companies listed on the Warsaw Stock Exchange

First, the author verified whether the relationship between operating efficiency and return ration of joint-stock companies listed on the Warsaw Stock Exchange is statistically valid. In order to estimate efficiency ratios with the use of the BCC model, financial data serving as inputs and outputs of companies' operations needs to be defined. In diverse studies around the world, DEA was often used for companies' operating efficiency. Among authors who studied companies from the processing industry are: Ma et al. (2002), Fang et al. (2009), Hassan et al. (2010). Authors who studied companies from the construction industry are among others: Zheng et al. (2011), Horta et al. (2012), Wong et al. (2012) and Kapelko et al. (2014). Each author used a different set of financial variables in their research. However, there were some common features. Assets and a number of employees were often used as inputs. Whereas, items from profit and loss accounts were used as outputs. Taking into account literature and availability of financial data, the author decided to assess operating efficiency of joint-stock companies from relevant industries based on the following variables:

- version 1 – inputs: tangible assets, intangible assets, output: net profit/loss,
- version 2 – inputs: tangible assets, equity, output: net profit/loss.

Efficiency ratios were estimated for the years 2010-2016 (in the case of construction as well as clothing and cosmetic industries) and 2011-2016 (food industry). It was verified what is the relation between the results for consecutive years and companies' simple rate of return on the Warsaw Stock Exchange for the same period. Pearson correlation coefficient ($\gamma_{j,h}$ where j-th company is from the h industry) was used for that purpose. An average coefficient value was estimated for each studied industry:

$$\bar{\gamma}_h = \frac{1}{J}\sum_{j=1}^{J}\gamma_{j,h} \tag{5}$$

The following hypotheses were formulated to verify if there is a valid statistical dependence (where ρ_0 is an assumed positive correlation coefficient value).

$$H_0\colon\ \bar{\gamma}_h \leq \rho_0 \tag{6}$$

$$H_1\colon\ \bar{\gamma}_h > \rho_0 \tag{7}$$

The inference included the following test statistic (Edirisinghe and Zhang, 2007):

$$\bar{\Psi}_h = \frac{1}{J_h}\sum_{j=1}^{J_h}\frac{1}{2}log_e\frac{1+\gamma_{j,h}}{1-\gamma_{j,h}} \tag{8}$$

where: J_h – number of firms in each industry h.

Table 2 presents results for the studied industries.

Table 2. Dependence between performance and return rates.

	Version I			**Version II**		
	Construction	Food and drinks	Clothes and cosmetics	Construction	Food and drinks	Clothes and cosmetics
Correlation metric	0.218	0.248	0.323	0.220	0.297	0.338
Test statistic	0.261	0.326	0.367	0.263	0.389	0.387
Critical value	0.236	0.313	0.313	0.236	0.313	0.313

The results show that the strongest correlation can be observed for efficiency ratios and return rates of joint-stock companies from the clothing and cosmetic industry. Companies from the construction industry had the lowest value of correlation coefficient. The value was positive for all industries – the higher operating efficiency, the higher rates of return on the Warsaw Stock Exchange. According to test statistics, the null hypothesis stating that an average correlation coefficient is lower than 0.1 can be rejected for all cases (at significance level of 0.05).

4 Optimum variables for estimating operating efficiency of joint-stock companies from diverse industries

The second goal of the research is to point variables which should be used to estimate operating efficiency of joint-stock companies from diverse industries. The answer will

provide us with financial variables that should be taken into account when assessing operating efficiency of companies from various industries. The generalized DEA was used to solve that task. Its main purpose is to find a set of variables to be used as inputs and outputs based on a defined criteria - a relevant objective function. Owing to GDEA, variables are not attributed to inputs/outputs a priori. To measure efficiency ratios, the author introduced scale variables and variables which values determine how a variable (i-th variable describing j-th element) is attributed to inputs and outputs in the study. Due to the fact that the study concerns a defined number of variables *I*, variables create vectors scaling inputs and outputs. The GDEA approach makes it possible to find a pair which allows for maximizing (minimizing) the objective function of the optimization problem. According to Edirisinghe and Zhang (2007), the pair needs to belong to the following Binary Complementary Domain (BCD):

$$\Omega := \{(y,z): \sum_{i=1}^{I} y_i \geq 1, \quad y_i + z_i \leq 1, \quad y_i, z_i \epsilon \{0,1\}, \ i = 1, \dots, I\}. \tag{9}$$

Only then, none of the studied variables can simultaneously be an input and output while the study includes at least one input and output. Furthermore, in the study it was assumed that only 15 variables can be inputs and the remaining 8 can be outputs (Table 1). As a result, a pair $(\boldsymbol{y}, \boldsymbol{z})$ must belong to the Restricted BCD:

$$\Omega^* := \left\{(y,z) \in \Omega: \sum_{i=16}^{23} y_i = 0, \ \sum_{i=1}^{15} z_i = 0, \quad \right\}. \tag{10}$$

The above results confirmed that there is a valid statistical dependence between companies' operating efficiency and their rates of return on the Warsaw Stock Exchange. Therefore, a further research aims to find a set of inputs and outputs with the strongest dependence for each industry – the highest average value of a correlation coefficient for a given industry. Thus, a separate optimization problem needs to be solved for every industry:

$$\max_{y,z} \bar{\gamma}_h(y,z) \tag{11}$$

$$(y,z) \in \Omega^* \tag{12}$$

where:

$\bar{\gamma}_h$ – industry-correlation metric,

(y,z) – input/output scaling vector pair.

It is difficult to solve such formulated optimization problem as its objective function includes efficiency ratios for estimation of which it is required to solve optimization problems for each studied company. The simulated annealing algorithm was used to find an optimal solution. In order to define an initial solution of the algorithm, the author searched for all

possible solutions to a set of 10 potential inputs and 5 potential outputs (variables in italics in the Table 1). Additionally, it was assumed that variables strongly correlated with other variables are removed from the analysis (correlation coefficient higher than 0.90)[2]. Table 3 presents results achieved with the use of the simulated annealing algorithm for each studied industry.

Table 3. Dependence between operating efficiency and rates of return for GDEA.

	Construction	**Food and drinks**	**Clothes and Cosmetics**
Industry-correlation metric	0.319	0.411	0.429
Test statistic	0.396	0.490	0.505
Rejected null hypothesis	$\bar{\gamma}_h \leq 0.25$	$\bar{\gamma}_h \leq 0.25$	$\bar{\gamma}_h \leq 0.25$

The analysis of results indicated that again the strongest dependence between companies' efficiency ratios and their rates of return on the stock exchange was observed for clothes and cosmetic industry. The construction industry had the lowest correlation coefficient. After comparing the results with outcome of the prior study, it turned out that an average correlation coefficient for every industry increased due to the choice of variables with the use of GDEA. In the case of all analyzed industries, the null hypothesis can be rejected in favor of an alternative hypothesis stating that an average correlation coefficient is higher than 0.25. According to the results, it is advisable to use other methods than the expert method to choose input and output variables. It makes it possible to estimate efficiency ratios in such a way that better represents a market situation of companies on the Warsaw Stock Exchange.

The second aspect of the study is focused on an assessment of dependence between the estimated efficiency ratios and rates of return as well as an analysis of financial data that were considered owing to GDEA. They are presented in Table 4. A different set of variables serving as inputs and outputs was used for each studied industry. Similar variables were used for the food as well as clothes and cosmetic industries. In both cases, inputs included variables from liabilities which show sources of companies' equity. Similarities might stem from a comparable activity of studied companies. A completely different set of variables was used as inputs in the case of the construction industry. The study included only tangible and intangible assets. Results shows that variables used to estimate operating efficiency of joint-

[2] If two variables are strongly correlated, a variable with a higher average of absolute values of correlation coefficients is removed from the analysis.

stock companies differ from one industry to another. It should be taken into consideration in studies on their activity.

Table 4. Optimum inputs/outputs for relevant industries.

	Construction	Food and drinks	Clothes and cosmetics
tangible fixed assets	input		
inventory	input		
short-term receivables	input		
equity		input	
short-term liabilities		input	input
long-term liabilities		input	input
profit (loss) on sales			output
gross profit (loss) on sales	output	output	
other operating expenses			output
net profit (loss)	output		

5 Use of information on operating efficiency of joint-stock companies while creating investment portfolio

The results from previous stages of the study confirmed that there is a valid statistical dependence between operating efficiency of companies listed on the stock exchange and their rates of return. Companies performing effectively had higher rates of return on the Warsaw Stock Exchange. Finally, this dependence was used to create an investment portfolio. For this purpose, the author had to estimate efficiency ratios for studied companies in 2017. The author used values of efficiency ratios for periods 2010-2016 (construction as well as clothes and cosmetic industries) and 2011-2016 (food industry) with the use of the BCC model and variables from the previous stages of the study. Using the data, a comparative analysis of the following prediction methods was carried out (based on an ex-post forecast error): naive methods, methods based on simple average, methods based on weighted average and trend. On that basis, it was decided that performance of companies in 2017 will be predicted using a simple 4-element average.

A model developed by H. M. Markowitz was used to create an investment portfolio. Two portfolios were created for the period April 3 – October 2, 2017. The first portfolio was created with the use of share prices of all 72 studied companies (the “Markowitz” portfolio). The Markowitz model was also used to build a second portfolio. However, only 20 companies

considered as effective in 2017 were included (correlation coefficient higher than 0.7). Fig. 1 presents the rate of return of such portfolios in the studied time frame.

Results show that the compound return rate of the portfolio built based on effective companies is substantially higher (10.02%) than for the one based on all studied companies (-4.26%). Most importantly, the portfolio based on effective companies allows an investor to achieve higher return rates uninterruptedly for 70% of the time.

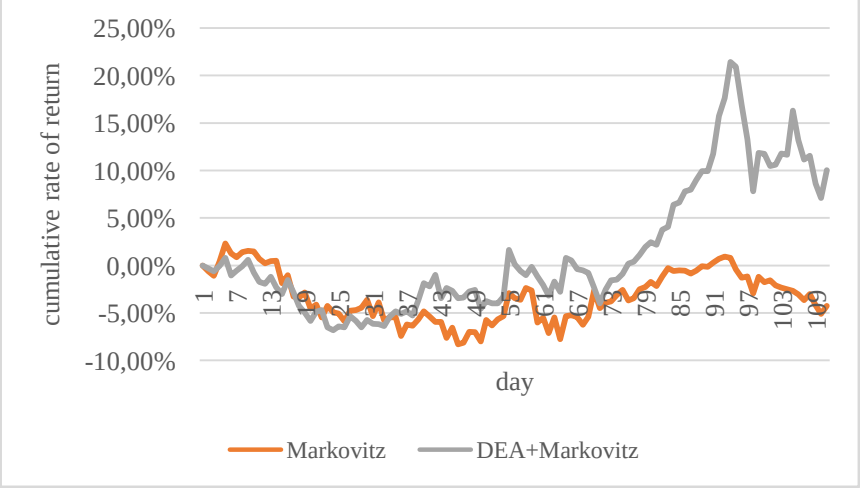


Fig. 1. Cumulative rate of return of two investment portfolios.

Conclusions

The results demonstrate that there is a valid statistical dependence between operating efficiency of Polish joint-stock companies and their return rates. It is a positive dependence which means that effective companies have higher return rates than those performing poorly. Variables used to estimate operating efficiency of joint-stock companies from various industries were discovered owing to the GDEA approach. Results demonstrated that a different set of variables serving as inputs and outputs should be used for operating efficiency estimation of joint-stock companies from various industries.

What's more, results showed that information on operating efficiency of joint-stock companies might be crucial for investors on the Warsaw Stock Exchange. If they take that into consideration when creating an investment portfolio, investments in companies performing well will give them higher return rates.

Further research should include the analysis of more industries in Poland. Moreover, the use of a different method (other DEA-like models) should give interesting results.

References

Avkiran, N. K. & Morita, H. (2010). Predicting Japanese bank stock performance with a composite relative efficiency metric: A new investment tool. *Pacific-Basin Finance Journal, 18*(3), 254-271.

Banker, R. D., Charnes, A. & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science, 30(9),* 1078-1092.

Edirisinghe, N. & Zhang, X. (2007). Generalized DEA model of fundamental analysis and its application to portfolio optimization. *Journal of Banking & Finance, 31*(11), 3311-3335.

Fang, H., Wu, J. & Zeng, C. (2009). Comparative study on efficiency performance of listed coal mining companies in China and the US. *Energy Policy, 37*(12), 5140-5148.

Hassan, M. K., Isik, I. & Mamun, A. (2010). Trade liberalization and industry performance in Bangladesh. *Journal of Policy Modeling, 32*(3), 399-417.

Horta, I. M., Camanho, A. S., Johnes, J. & Johnes, G. (2012). Performance trends in the construction industry worldwide: an overview of the turn of the century. *Journal of Productivity Analysis, 39*(1), 89-99.

Kapelko, M., Lansink, A. O. & Stefanou, S. E. (2014). Assessing dynamic inefficiency of the Spanish construction sector pre- and post-financial crisis. *European Journal of Operational Research, 237*(1), 349-357.

Ma, J., Evans, D. G., Fuller, R. J. & Stewart, D. F. (2002). Technical efficiency and productivity change of Chinas iron and steel industry. *International Journal of Production Economics, 76*(3), 293-312.

Szafrański, G. (2004). Metoda DEA w analizie fundamentalnej polskiego rynku akcji. *Acta Universitatis Lodziensis. Folia Oeconomica, 177/2004*, 347-363.

Wong, W. P., Gholipour, H. F. & Bazrafshan, E. (2012). How efficient are real estate and construction companies in Irans close economy? *International Journal of Strategic Property Management, 16*(4), 392-413.

Zheng, X., Chau, K. & Hui, E. C. (2011). Efficiency assessment of listed real estate companies: an empirical study of China. *International Journal of Strategic Property Management, 15(2),* 91-104.

Changes of the consumption structure in European countries considering its modernization process

Mateusz Jankiewicz[1]

Abstract

The aim of the article is to assess the diversity and changes in the structure of household consumption of European countries, with particular emphasis on the division into consumption of necessity goods and luxury goods in the years 1995-2015. The progressive process of modernizing consumption in Europe causes a change in the consumption pattern. According to this process, the consumption of luxury goods plays a more and more significant role in the consumption structure. The conducted study is the basis for checking whether this process is reflected in all European countries. In particular, this analysis is to check whether and to what extent there was an increase in the share of consumption of luxury goods in total consumption in each of the countries in the analyzed period. The analysis of the diversification of the consumption structure in the countries in the entire period is carried out using the cluster analysis method that allows the identification of territorial units with similar characteristics in this matter. The study is supplemented with the analysis of the consumption structure in the European countries over the period 1995-2015, with the particular emphasis on the location factor.

Keywords: *consumption structure, necessity goods, luxury goods, cluster analysis, consumption convergence*
JEL Classification: D13, E21, I0
DOI: 10.14659/SEMF.2018.01.17

1 Introduction

Consumption is a very important element of generating GDP in every country. It is determined by economic, demographic, cultural or biological factors. However, the most important consumption factor is the disposable income of households. Based on this income, households decide what goods and what quantities they are able to acquire. Richer households can afford to buy more luxury goods than poorer households. This affects the structure of consumption in a country. Increasing the share of expenditures on the consumption of luxury goods is defined in the literature as modernization of consumption.

The economies of European countries have been converging for a long time (Carnicky et al., 2016; Corrado et al., 2005; Dall'Erba and Le Gallo, 2008; López-Bazo et al., 1999; Quah, 1996; von Lyncker and Thoennessen, 2017). The progressing process of economic convergence in the European Union (EU) countries seems to be the reason for the leveling off of consumption and assimilating to its structure. The consumption structure points towards a pattern in which the share of consumption of luxury goods in total consumption increases.

[1] Corresponding author: Nicolaus Copernicus University, Faculty of Economic Sciences and Management, Gagarina 13A, 87-100 Toruń, Poland, e-mail: mateuszj@doktorant.umk.pl

This means that poorer countries should strive to achieve a consumption structure that characterizes wealthier countries.

In the current research, the issues of diversification of consumption structure (Kuśmierczyk and Piskiewicz, 2012; Grzega, 2015) and modernization of consumption (Kusińska, 2009) in the European countries have already been discussed. There were also many studies conducted on the expenditure of the inhabitants of European countries on various categories of goods (Berbeka, 2014; Hoszman, 2013; Piekut, 2014). The convergence process of consumption patterns in Europe was also analyzed (Nowak and Kochkova, 2011).

The main objective of this study is to assess the variation and changes in the structure of household consumption in EU countries (excluding Croatia) in the years 1995-2015. An additional aspect raised in the study is the assimilation of the consumption structure on the European continent. Previous studies show structure of the consumption in European countries from 2004 to 2012. Moreover convergence of the consumption was analyzed only to 2007. In this paper time range in both investigations is expanded to 2015. Moreover another method to choose number of clusters is used. This study is the preview to the wider look at the consumption problem in Europe.

2 Subject and range of the study

The study focuses on the consumption structure in the European countries in the years 1995-2015. The consumption of groups of goods in total consumption was analyzed according to the Classification of Individual Consumption According to Purpose (COICOP). According to the above classification, consumer goods are divided into the following groups: food and non-alcoholic beverages; alcoholic beverages, tobacco and narcotics; clothing and footwear; housing, water, electricity, gas and other fuels; furnishing, household equipment and routine household maintenance; health; transport; communication; recreation and culture; education; restaurants and hotels; miscellaneous goods and services. In the case of longer group names, only the first word describing a given group of consumer goods was used in the study.

In order to examine the similarity of the consumption structure between countries, methods of cluster analysis were used. The agglomeration method of hierarchical clustering was used - the Ward's method using the Euclidean distance as a measure of similarity of objects. The objects were divided into three groups with similar properties regarding the consumption structure. Then, the consumption structure was modeled using the β convergence model in the classical approach.

The study verifies the hypothesis of a modern consumption structure in the countries of Western and Northern Europe as well as the hypothesis about the blurring of differences in the consumption structure between the EU countries.

3 Data

The data applied in this study come from the European Statistical Office database. Variables regarding the consumption of individual categories of goods were taken directly from the database. The values of disposable income per capita were calculated – the values of disposable income in each countries were divided by the number of population.

Fig. 1 shows the development of disposable income per capita value in the EU countries (except for Croatia) in 2000, 2005, 2010 and 2015 respectively. The highest values of that variable were observed in the countries of Western and Northern Europe. In contrast, disposable income per capita in the Central-Eastern part of the continent was below the median. This trend has continued throughout the considered period. This means that the inhabitants of Northern and Western Europe have the opportunity to consume not only goods that allow them to meet basic needs but also they can consume more luxury goods than the others in Europe.

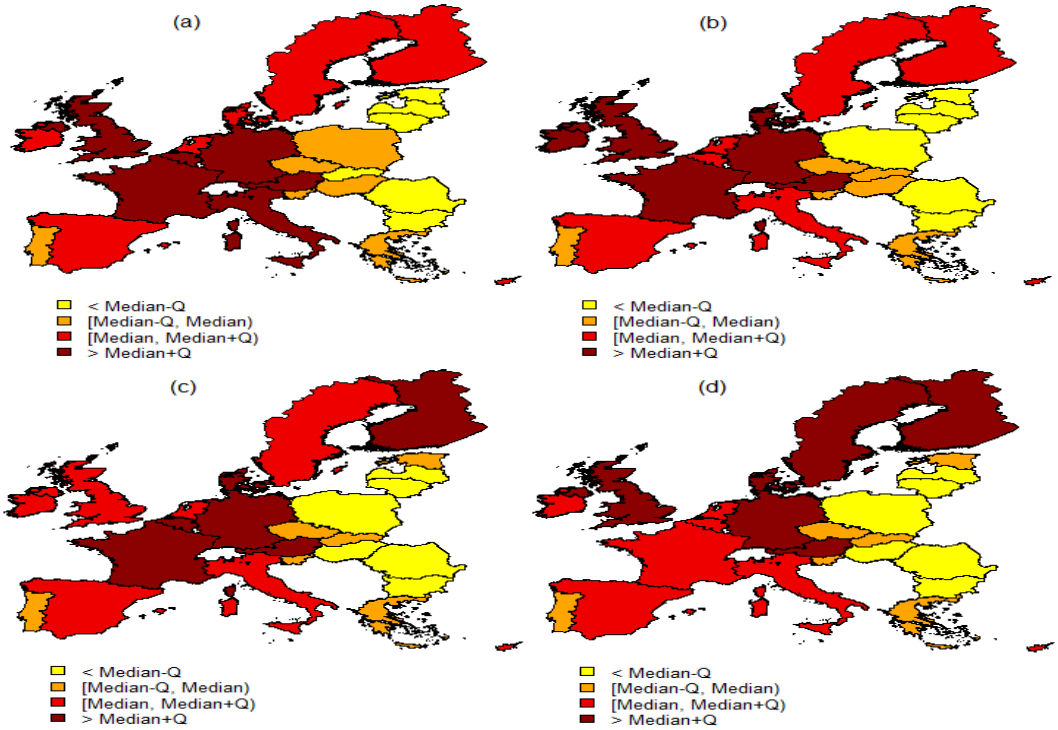


Fig. 1. Spatial distribution of disposable income per capita in Europe in 2000 (a), 2005 (b), 2010 (c) and 2015 (d).

Fig. 2 shows the average share of individual groups of consumer goods in the countries belonging to the EU-27 group in the analyzed period. The highest share of consumption was noted in goods related to housing, water, electricity, gas and other fuels has been recorded (more than 20% in each year of the study). The share of consumption between 10 and 15 percent was the consumption of goods related to food and non-alcoholic beverages, recreation and culture, transport and miscellaneous goods and services. The consumption of other groups of goods accounted for less than 10% of total consumption. The largest fluctuations in the consumption structure were noted in the consumption of non-alcoholic beverages.

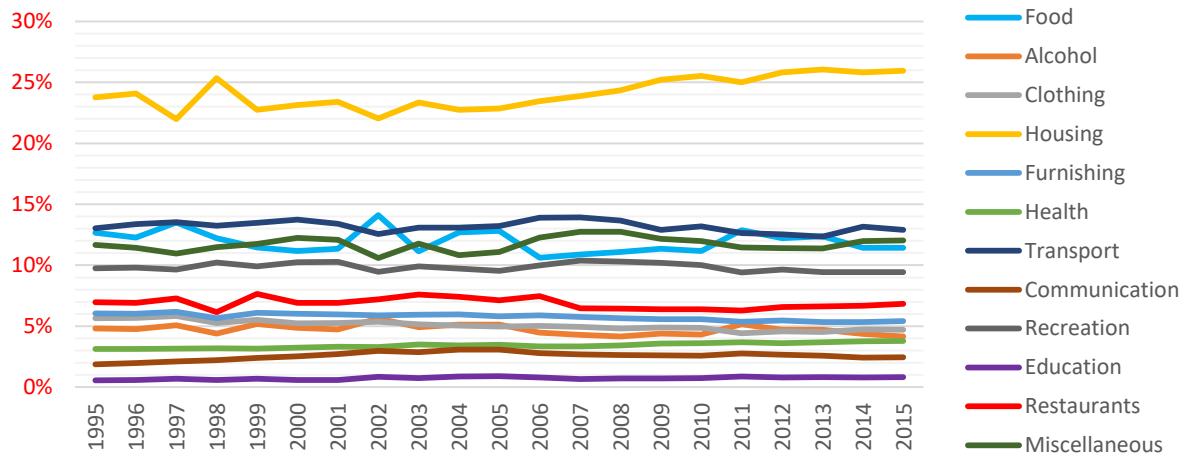


Fig. 2. Average share of consumption of particular groups of goods in total consumption.

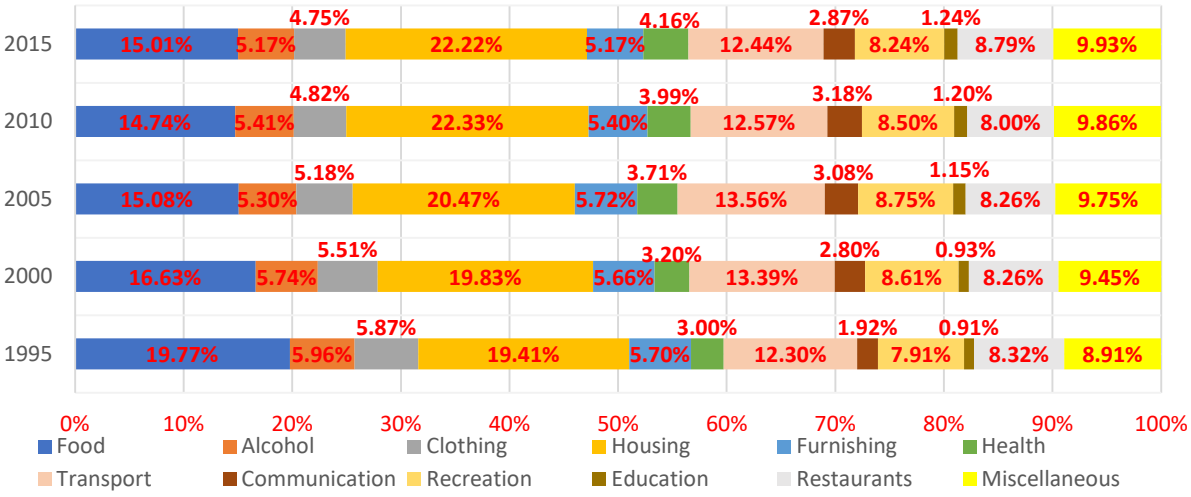


Fig. 3. The structure of consumption in Europe in selected years of the study.

Fig. 3 shows changes in the structure of consumption in 1995, 2000, 2005, 2010 and 2015. There is a noticeable drop in the share of consumption of food and non-alcoholic beverages

(from 19.77% to 15%) and clothing and footwear (from 5.87% to 4.75%), which may result from the enrichment process of European countries over the period 1995-2015. In addition, there was an increase in the average share of consumption related to housing, water, electricity, gas and other fuels (from 19.41% to 22.22%), recreation and culture (from 7.91% to 8.24%) and restaurants and hotels (from 8.32% to 8.79%) .

4 Methodology

Clustering of objects with a similar consumption structure was made using the hierarchical agglomeration method – the Ward's method. It's motivated by uncertainty as to the number of clusters to choose. It is helpful to establish number of clusters. In the future studies clustering can be done with non-hierarchical agglomeration methods. The purpose of using this method is to minimize the sum of squared deviations of two arbitrarily selected clusters (Ward, 1963). The Euclidean distance was used as a measure of the distance between the objects. The number of clusters has been specified using the average silhouette method for hierarchical clustering.

In order to verify the hypothesis of inequalities reduction in the consumption structure, the traditional model of β-convergence for cross-sectional data given by formula (Arbia, 2006) was used:

$$ln\left(\frac{y_{T,i}}{y_{0,i}}\right) = \alpha + \beta \ln(y_{0,i}) + \varepsilon_i, \quad (1)$$

where $y_{0,i}$ and $y_{T,i}$ are the values of the share of consumption of the chosen groups of goods in total consumption the i^{th} region in the first and last survey period respectively.

The β parameter is used to calculate the $t_{half-life}$ value, which presents the time needed to reduce the difference by half. It is expressed as follows:

$$t_{half-life} = \frac{\ln(2)}{b}, \quad (2)$$

where *b* expresses the convergence rate and is evaluate with the use of the following formula:

$$b = -\frac{\ln(1+\beta)}{T}. \quad (3)$$

5 Empirical results of research

5.1 Grouping countries with a similar consumption structure

The analysis of the similarity of the structure of consumption in the European Union countries began with their grouping into three clusters using the Ward's method of objects agglomeration. Table 1 presents the results obtained on the basis of that grouping.

Table 1. Concentration of countries obtained by the Ward's method in selected years of the study.

Year	Cluster 1	Cluster 2	Cluster 3
1995	Austria, Belgium, Denmark, Finland, France, Germany, Luxembourg, Netherlands, Sweden, United Kingdom	Cyprus, the Czech Republic, Greece, Hungary, Ireland, Italy, Malta, Portugal, Slovenia, Spain	Bulgaria, Estonia, Latvia, Lithuania, Poland, Romania, Slovakia
2000	Austria, Belgium, Denmark, Finland, France, Germany, Luxembourg, Netherlands, Sweden, United Kingdom	Cyprus, Greece, Ireland, Malta, Portugal, Spain	Bulgaria, the Czech Republic, Estonia, Hungary, Italy, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia
2005	Austria, Belgium, the Czech Republic, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Luxembourg, Netherlands, Poland, Slovakia, Slovenia, Sweden, United Kingdom	Cyprus, Greece, Malta, Portugal, Spain	Bulgaria, Estonia, Latvia, Lithuania, Romania
2010	Belgium, Denmark, Finland, France, Germany, Luxembourg, Netherlands, Sweden, United Kingdom	Austria, Cyprus, Greece, Ireland, Italy, Malta, Portugal, Slovenia, Spain	Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia
2015	Belgium, Denmark, Finland, France, Germany, Luxembourg, Netherlands, Sweden, United Kingdom	Austria, Cyprus, Greece, Ireland, Malta, Spain	Bulgaria, the Czech Republic, Estonia, Hungary, Italy, Latvia, Lithuania, Poland, Portugal, Romania, Slovakia, Slovenia

The countries that remained in the first group throughout the entire period considered were Belgium, Denmark, Finland, France, Germany, Luxembourg Netherlands, Sweden and United Kingdom. The core of the second cluster that remained in each year of the study were: Cyprus, Greece, Malta and Spain. The third cluster always included the following countries: Bulgaria, Estonia, Latvia, Lithuania and Romania. In the majority of the analyzed period, Poland, Slovakia and Hungary also remained in the third cluster.

Fig. 4 shows the characteristics of individual clusters. It provides a relation of the average group consumption structure of particular groups of goods to the average consumption structure of all EU-27 countries.

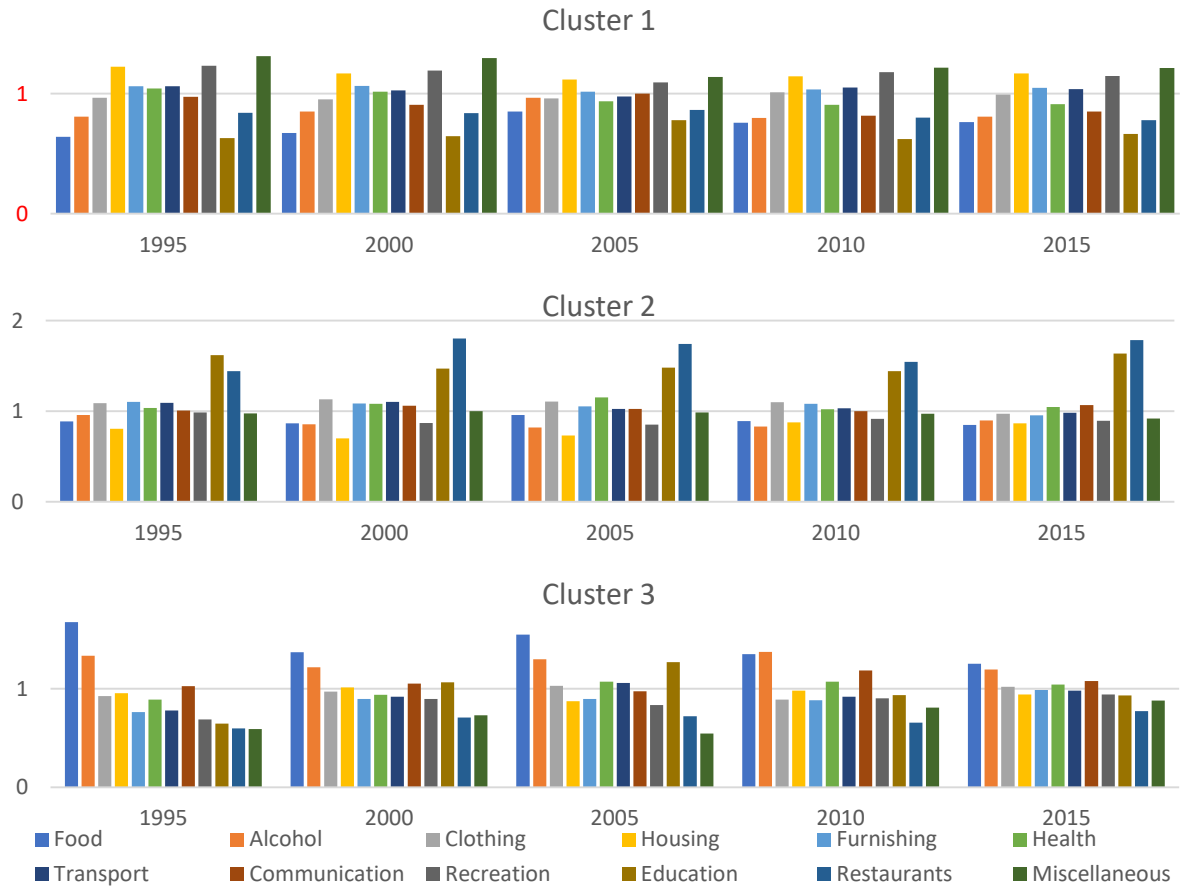


Fig. 4. The relation of the average share of individual categories of consumer goods to the European average.

The first cluster is characterized by a low share of consumption of food and non-alcoholic beverages in total consumption compared to the European average. This category of goods has a very large impact on the composition of groups. The second cluster is characterized by a higher than average European share of the consumption of goods related to restaurants,

hotels and education. In other groups, the calculated ratios are at a level below 1. In addition, the third group is characterized by a lower than average share of consumption of goods related to recreation, culture, transport and miscellaneous goods. In turn, the residents of third group of countries consume relatively more alcoholic beverages compared to the inhabitants of the countries included in the other two groups.

5.2 Convergence of consumption structure in the European countries

In order to analyze the consumption structure, the β convergence model for cross-sectional data (classical approach) was used. Table 2 presents the results of its estimation and verification for the consumption of particular groups of goods.

Table 2. Results of the estimation and verification of the convergence model for the groups of consumer goods.

Group of consumption goods	α	Speed of convergence	Half-life	p-value
Food	-0.3335	0.0203	34.1642	0.0000
Alcohol	-0.2468	0.0142	48.9082	0.0292
Clothing	-0.5521	0.0402	17.2586	0.0017
Housing	-0.4765	0.0324	21.4181	0.0006
Furnishing	-0.7140	0.0626	11.0760	0.0000
Health	-0.2606	0.0151	45.9104	0.1061
Transport	-0.6547	0.0532	13.0364	0.0000
Communication	-0.7332	0.0661	10.4916	0.0005
Recreation	-0.6456	0.0519	13.3642	0.0000
Education	-0.4110	0.0265	26.1878	0.0117
Restaurants	-0.0448	0.0023	302.7661	0.6651
Miscellaneous	-0.4690	0.0317	21.8994	0.0000

The results of the assimilation of the consumption structure analysis indicate a blurring of differences in the share of consumption in most categories of goods in the total consumption of countries. The only categories of consumption in which convergence does not occur are: health, restaurants and hotels. In other cases, the α parameter of the β-convergence model was statistically significant. The fastest rate of convergence based on the classical convergence analysis was observed in communication; furnishing, household equipment and routine household maintenance, which translates into the least amount of time needed to reduce by

half the differences in the consumption of goods belonging to this category. However, the most time is needed to align the consumption of food and non-alcoholic beverages and also alcohol beverages, tobacco and narcotics.

Conclusion

The consumption structure in the EU-27 countries is very diverse. The greatest disproportions are observed in the consumption of food and non-alcoholic beverages, as well as consumption related to restaurants and hotels. The countries in which the share of food consumption in total consumption is the largest are the countries with lower disposable income, located in the Central-Eastern part of the continent. These are also countries where alcohol consumption is higher than in other European countries. Throughout the studied period, they belonged to the same focus, characterized by the largest share of consumption of goods satisfying basic life needs. The second cluster was dominated by countries that are a destination for many tourists - this is why this concentration was characterized by the largest share of goods consumption from the restaurants and hotels category. The richest countries (with the highest level of disposable income), located in the Western and Northern parts of Europe belonged to a group with a relatively high level of consumption of luxury goods (voluntary, which are not necessary to meet basic life needs). Nevertheless, the European countries are striving to unify the structure of consumption in the direction of its modernization. The consumption of luxury goods is increasingly contributing to the total consumption of poorer countries. A certain spatial tendency has emerged in the consumption structure of European countries, the identification of which will be the subject for further research. The study should also be enriched with a more in-depth analysis of the convergence of the expenditure structure - using the conditional convergence approach or panel data models. The model should be extended with influence of additional determinants of consumption such as consumer expectations and credits. Both of them have a significant impact on volume and structure of households expenditures. Moreover in further investigation some of the non-hierarchical clustering methods can be used.

References

Arbia, G. (2006). *Spatial Econometrics. Statistical Foundations and Applications to Regional Convergence*. Springer-Verlag, Berlin–Heidelberg.

Berbeka, J. (2014). Konsumpcja usług turystyki społecznej – obecnie i w przyszłości. *Konsumpcja i rozwój*, *2*(7), 39-49.

Carnicky, S., Megyesiova, S., Conkova, M. & Zavadsky, C. (2016). Productivity development and convergence across the EU Member States. *Економічний часопис-XXI, 162*(11-12), 13-17.

Corrado, L., Martin, R. & Weeks, M. (2005). Identifying and interpreting regional convergence clusters across Europe. *The Economic Journal, 115*(502).

Dall'Erba, S. & Le Gallo, J. (2008). Regional convergence and the impact of European structural funds over 1989–1999: A spatial econometric analysis. *Papers in Regional Science*, 87(2), 219-244.

Grzega, U. (2015). Dochód i konsumpcja gospodarstw domowych w Polsce i Unii Europejskiej w latach 2004-2012. *Handel wewnętrzny*, *4*(357), 124-134.

Hoszman, A. (2013). Konsumpcja usług transportowych na tle przemian gospodarczych w Polsce w latach 2003-2012. *Konsumpcja i rozwój*, *2*(5), 63-77.

Kusińska, A. (2009). Zmiany w sferze konsumpcji i ich uwarunkowania w latach 1989-2009. *Handel wewnętrzny*, (4-5), 33-43.

Kuśmierczyk, K. & Piskiewicz, L. (2012). Konsumpcja w Polsce na tle pozostałych krajów Unii Europejskiej. *Konsumpcja i rozwój*, *2*(3), 78-93.

López-Bazo, E., Vayá, E., Mora, A. J. & Suriñach, J. (1999). Regional economic dynamics and convergence in the European Union. *The Annals of Regional Science*, *33*(3), 343-370.

Nowak, J. & Kochkova, O. (2011). Income, Culture, and Household Consumption Expenditure Patterns in the European Union: Convergence or Divergence? *Journal of International Consumer Marketing*, *23*(3-4), 260-275.

Quah, D. T. (1996). Regional convergence clusters across Europe. *European economic review*, *40*(3), 951-958.

Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, *58*(301), 236-244.

von Lyncker, K. & Thoennessen, R. (2017). Regional club convergence in the EU: evidence from a panel data analysis. *Empirical Economics*, *52*(2), 525-553.

Comparison of parameter estimators for the Dagum distribution

Alina Jędrzejczak[1], Dorota Pekasiewicz[2], Wojciech Zieliński[3]

Abstract

The Dagum model is frequently applied to the analysis of income and wage distributions all over the world. It has many desirable statistical properties and turned out to be well fitted to empirical distributions in different divisions. The estimates of its parameters can be applied to the evaluation of numerous income distribution characteristics, including inequality, poverty and wealth indices, dispersion measures based on quantiles and concentration curves. They can also be used to compare income distributions in space and over time. The estimation of these characteristics needs reliable Dagum distribution estimates. The paper is devoted to the analysis of statistical properties of various estimators of the Dagum distribution parameters.

Keywords: *size distributions, Dagum distribution, parameter estimation,*

JEL Classification: C13, C15

DOI: 10.14659/SEMF.2018.01.18

1 Introduction

The Dagum distribution was originally derived by Camilo Dagum (1977), when he investigated the income elasticity of the cumulative distribution function (CDF) in several developed and developing countries. The elasticity turned out to be a monotonic decreasing and bounded function of CDF. The resulting distribution turned out to present many advantages over its competitors and thus was frequently applied to the analyses of income distributions in many countries and by different divisions (Jędrzejczak, 2014, 2015a, 2015b). Its CDF depends on three parameters: the scale parameter and the two shape parameters, and it can be considered as a special case of the generalized beta distribution of the second kind (GB2) and of the Burr type III (inverse Burr XII) distribution (see: Kleiber and Kotz, 2003). Unlike the gamma or the lognormal distribution the Dagum distribution has an explicit mathematical formula for its CDF and its p^{th} quantile. It generally presents very good consistency with empirical income and wage distributions (that are known to be unimodal and positively skewed) as well as with the non-modal wealth distributions, depending on the

[1] University of Łódź/Institute of Statistics and Demography, Department of Statistical Methods, 41/43 Rewolucji 1905 St., 90-214 Łódź, Poland, jedrzej@uni.lodz.pl.

[2] Corresponding author: University of Łódź/Institute of Statistics and Demography, Department of Statistical Methods, 41/43 Rewolucji 1905 St., 90-214 Łódź, Poland, pekasiewicz@uni.lodz.pl.

[3] Warsaw University of Life Science /Department of Econometrics and Statistics, Nowoursynowska 159 St., 02-776 Warsaw, Poland, wojciech_zielinski@sggw.pl

product of its shape parameters. For some values of shape parameters the distribution is unimodal, otherwise it is non-modal. Also, the number of finite moments of the distribution (in practice 2 or 3) depends on the value of the shape parameter. Moreover, the model possesses the important property of weak Pareto law as its cumulative distribution function converges to the Pareto model, which is considered ideal for high income groups. Recently, the Dagum distribution has been studied from a reliability point of view and used to analyse survival data (e.g.: Domma, et al., 2011) due to its hazard function which can be monotonically decreasing, an upside-down bathtub, or bathtub and then upside-down bathtub shaped. Several estimation methods for the Dagum model parameters have been discussed in the literature. Amongst them, the maximum likelihood, pseudo maximum likelihood, the method of percentiles, the method of moments and a nonlinear least squares on the quantile function proposed by Dagum himself (Dagum, 1977) are the most frequently applied. The review of these methods can be found in: Dey et al. (2017), Kleiber and Kotz (2003). The method of maximum likelihood, the most popular one, presents many desirable properties including consistency, invariance, asymptotic efficiency and normality. A simulation study performed by Domański and Jędrzejczak (1998) revealed that the ML estimates of the Dagum model parameters are normally distributed and efficient for very large samples (greater than 7000). One needs also remember that the ML estimators from parametric distributions have robustness problems and are sensitive to extremes (see e.g. Victoria-Feser and Ronchetti, 1994; Victoria-Feser, 2000). Moreover, as it has been emphasized in Kleiber and Kotz (2003), there is a path in the Dagum distribution parameter space along which the likelihood becomes unbounded. A simulation study concerned with other estimation methods (except for the maximum likelihood) was reported in Dey et al. (2017), but it did not lead to practical conclusions due the choice of the parameters and sample sizes, unrealistic from the point of view of income distribution analyses. In the light of these finding it becomes desirable to examine the estimation methods for the Dagum model in order to assess their estimation errors for sample sizes used in practical applications.

The main objective of this paper was to recognise the methods that can be applied for moderate sample sizes which are often considered in practice (socio-economic groups, family types or NUTS2 regions in Poland). In the second section we briefly present the characteristics of the Dagum distribution. The next section comprises the review of selected estimation methods that can be applied to the case of three-parameter Dagum model. In section 4 we present the results of Monte Carlo experiments designed to assess the accuracy of estimators.

2 Dagum distribution

The probability density function of the Dagum distribution $D(a,v,\lambda)$ is given by:

$$f_{a,v,\lambda}(x)=\frac{av}{\lambda}\left(\frac{x}{\lambda}\right)^{av-1}\left(1+\left(\frac{x}{\lambda}\right)^{v}\right)^{-a-1} \qquad \text{for } x>0 \tag{1}$$

and the cumulative distribution function takes on the form:

$$F_{a,v,\lambda}(x)=\left(1+\left(\frac{x}{\lambda}\right)^{-v}\right)^{-a} \qquad \text{for } x>0, \tag{2}$$

where $\lambda>0$ is scale parameter, $v>0$ and $a>0$ are shape parameters determining the Lorenz curve and inequality measures.

The quantile function of the Dagum distribution is the following:

$$Q_{a,v,\lambda}(q)=\lambda\left(q^{-\frac{1}{a}}-1\right)^{-\frac{1}{v}} \quad \text{for } 0<q<1, \tag{3}$$

while the moments about the origin of the random variable X are:

$$E_{a,v,\lambda}X^{m}=\lambda^{m}\frac{\Gamma\left(1-\frac{m}{v}\right)\Gamma\left(a+\frac{m}{v}\right)}{\Gamma(a)} \text{ for } m<v, \tag{4}$$

where Γ is the gamma function.

From (4) it is seen that the number of finite moments of the distribution (1) does not exceed the value of the parameter *v*. Thus this parameter determines the tail of the Dagum distribution.

As the Dagum distribution is dedicated to the analysis of income and wages, it is convenient to have the explicit formulas for Gini and Zenga inequality measures based on the distribution parameters. The popular Gini index for the distribution (1) takes on the form

$$G=\frac{\Gamma(a)\Gamma\left(2a+\frac{1}{v}\right)}{\Gamma(2a)\Gamma\left(a+\frac{1}{v}\right)}-1 \quad \text{for } v>1, \tag{5}$$

while the Zenga (1984) index, defined on the basis of distribution and income quantiles, becomes:

$$Z=1-\exp\left[\frac{1}{v}\left[\Psi(a)+\Psi\left(1-\frac{1}{v}\right)-\Psi\left(a+\frac{1}{v}\right)-\Psi(1)\right]\right], \tag{6}$$

where Ψ is the digamma function.

It is worth noting that both formulae (5) and (6) depend only on the shape parameters of the Dagum distribution: *a* and *v*, which can be considered as the distribution "inequality" and "equality" parameters, respectively.

3 Estimation methods

The section is devoted to the brief description of the estimation methods which can be especially useful in the case of the three-parameter Dagum distribution applied to income data. The comprehensive review of this methods can be found in Dey et al. (2017). After the theoretical consideration of their properties, we choose the following methods of parameter estimation: Method Maximum Likelihood Method (ML), Methods of *L*-Moments (LM), Method of Ordinary Least-Squares (OLS) and Method of Weighted Least-Squares (WLS)

Let X be a random variable, $X_1, X_2, ..., X_n$ be a sample and $X_{1:n} \le X_{2:n} \le ... \le X_{n:n}$ be order statistics.

One of the estimation methods frequently applied for the Dagum model parameters is the maximum likelihood. The likelihood function of the Dagum distribution (1) takes the form:

$$L(a, v, \lambda; X_1, ..., X_n) = \left(\frac{av}{\lambda}\right)^n \prod_{i=1}^{n} \left(\frac{X_i}{\lambda}\right)^{av-1} \left(1 + \left(\frac{X_i}{\lambda}\right)^v\right)^{-a-1}. \quad (7)$$

To find maximum likelihood estimators we have to solve (numerically) the following system of equations: $\frac{\partial L}{\partial a} = 0$, $\frac{\partial L}{\partial v} = 0$, $\frac{\partial L}{\partial \lambda} = 0.$

Another estimation method that can be applied in case of the three-parameter Dagum model, is the method of *L*-moments (LM). The *L*- moments are defined as (Hosking, 1990):

$$v_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} EX_{r-k:r} \quad \text{for } r \ge 1. \quad (8)$$

For the Dagum distribution the first three *L*-moments take the following forms:

$$v_1 = \lambda \frac{\Gamma\left(a + \frac{1}{v}\right)}{\Gamma(a)} \Gamma\left(1 - \frac{1}{v}\right), \quad (9)$$

$$v_2 = \lambda \left(\frac{\Gamma\left(2a + \frac{1}{v}\right)}{\Gamma(2a)} - \frac{\Gamma\left(a + \frac{1}{v}\right)}{\Gamma(a)} \right) \Gamma\left(1 - \frac{1}{v}\right), \quad (10)$$

$$\nu_3 = \lambda \left(\frac{2\Gamma\left(3a+\frac{1}{\nu}\right)}{\Gamma(3a)} - \frac{3\Gamma\left(2a+\frac{1}{\nu}\right)}{\Gamma(2a)} + \frac{\Gamma\left(a+\frac{1}{\nu}\right)}{\Gamma(a)} \right) \Gamma\left(1-\frac{1}{\nu}\right). \tag{11}$$

Unbiased estimators of those moments are equal to (Dey et al., 2017):

$$l_1 = \frac{1}{n}\sum_{i=1}^{n} X_{i:n}, \tag{12}$$

$$l_2 = \frac{2}{n}\sum_{i=1}^{n} \frac{(i-1)}{n-1} X_{i:n} - l_1, \tag{13}$$

$$l_3 = \frac{6}{n}\sum_{i=1}^{n} \frac{(i-1)(i-2)}{(n-1)(n-2)} X_{i:n} - \frac{6}{n}\sum_{i=1}^{n} \frac{(i-1)}{(n-1)} X_{i:n} + l_1 . \tag{14}$$

Estimators obtained by the method of *L*-moments are the solution of the following system of equations: $l_m = \nu_m$ for $m = 1, 2, 3$.

The ordinary least square estimators (OLS) and weighted least square estimators (WLS) were proposed by Swain et al. (1988). It is well known that

$$E\left[F_{a,\nu,\lambda}\left(X_{i:n}\right)\right] = \frac{i}{n+1}, \quad D^2\left[F_{a,\nu,\lambda}\left(X_{i:n}\right)\right] = \frac{i(n-i+1)}{(n+1)^2(n+2)}. \tag{15}$$

The ordinary least square estimators are obtained by minimizing with respect to a, ν, λ the function:

$$\min_{a,\nu,\lambda}\left(\sum_{i=1}^{n} \left[F_{a,\nu,\lambda}\left(X_{i:n}\right) - \frac{i}{n+1} \right]^2 \right). \tag{16}$$

The weighted least square estimators are obtained by minimizing with respect to a, ν, λ the function:

$$\min_{a,\nu,\lambda}\left(\sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[F_{a,\nu,\lambda}\left(X_{i:n}\right) - \frac{i}{n+1} \right]^2 \right). \tag{17}$$

4 Simulation study

Several Monte Carlo experiments have been carried out to examine basic statistical properties of the estimators $\hat{\lambda}, \hat{\nu}$ and $\hat{a}$ of the Dagum model parameters λ, ν and a. The experiments involved four estimation methods described in section 3, namely: ML, LM, OLS and WLS. For each estimation method and each sample size (n=1000, 2000) the following steps were performed:

- drawing N=1000 independent random samples from the Dagum distribution,

- estimation of the Dagum model parameters for each sample,
- the estimators $\hat{\lambda}, \hat{v}, \hat{a}$ and their empirical relative bias and empirical relative root means squared error were calculated:

$$B(\hat{\theta}) = \frac{\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta)}{\theta} \cdot 100\%, \qquad \mathrm{RMSE}(\hat{\theta}) = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta)^2}}{\theta} \cdot 100\%.$$

Throughout the experiments we assumed a constant value of the scale parameter ($\lambda = 1$), while the values of the inequality parameter *v*, responsible for the right tail of the Dagum distribution, changed from 2 to 4 in order to comprise the wide variety of distributions, presenting light as well as heavy tails. The light-tailed distributions (*v*=3.5 and *v*=4.0) had three finite moments; among the heavy-tailed ones we considered the distributions with two moments (*v*=2.5 and *v*=3.0) or with only first moment (*v*=2). The parameter *a* fluctuated between 0.2 and 1.2 what made it possible to control the distribution inequality. Selected this way the sets of the Dagum model parameters embraced the distributions with different levels of the Gini ratio including the values typical for income analyses (Table 1).

Table 1. Gini index.

v	***a***					
	0.2	**0.4**	**0.6**	**0.8**	**1.0**	**1.2**
2.0	0.704	0.600	0.549	0.519	0.500	0.487
2.5	0.618	0.503	0.449	0.419	0.400	0.387
3.0	0.552	0.433	0.380	0.351	0.333	0.321
3.5	0.499	0.380	0.330	0.303	0.286	0.274
4.0	0.456	0.339	0.291	0.266	0.250	0.240

Fig. 1 shows the examples of Dagum densities for different sets of parameters. They confirm outstanding flexibility of this distribution discussed in section 1. Fig. 2-3 depict the values of the empirical bias and RMSE of $\hat{a}$ obtained by the above-mentioned estimation methods, while the corresponding characteristics obtained for $\hat{v}$ are presented in Fig. 4-5.

In Table 2 we present the relative root means squared errors (RMSEs) for the estimators of the Dagum model parameters obtained by four estimation methods : ML, LM, OLS and WLS and for sample sizes *n*=1000 and *n*=2000.

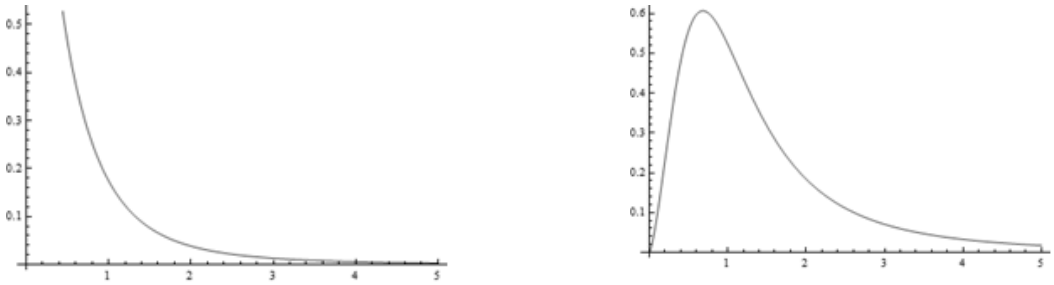


Fig. 1. Density functions of Dagum distribution: $D(0.2;2,1)$-left; $D(1.2;2,1)$-right.

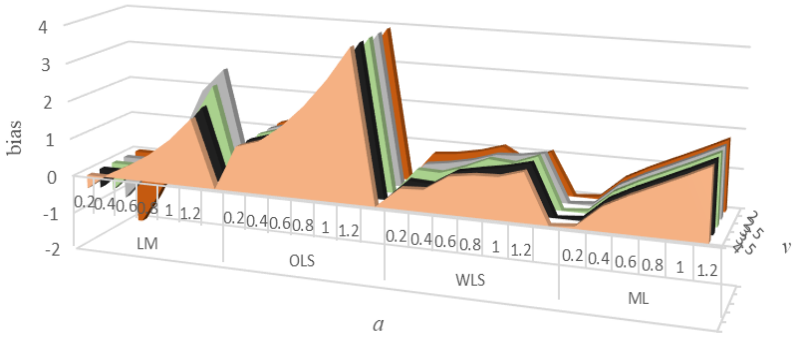


Fig. 2. Empirical relative bias of various estimators of *a*.

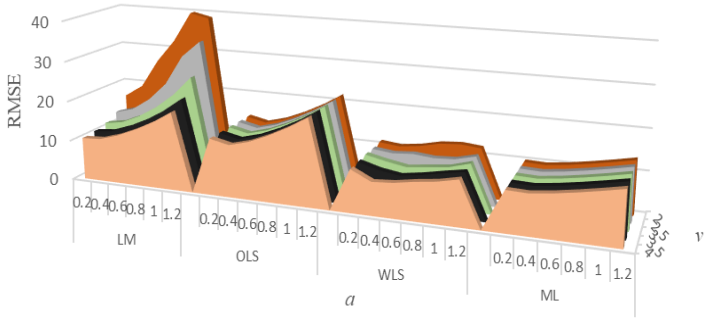


Fig. 3. Empirical relative RMSE of various estimators of *a*.

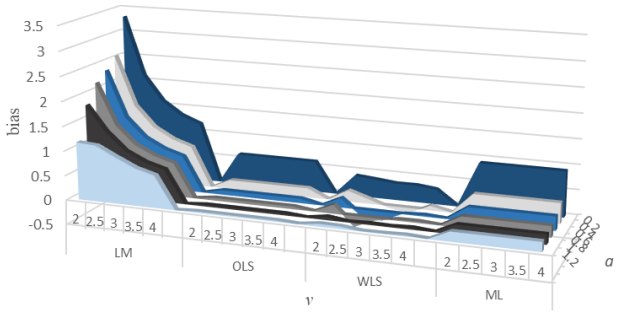


Fig. 4. Empirical relative bias of various estimators of *v*.

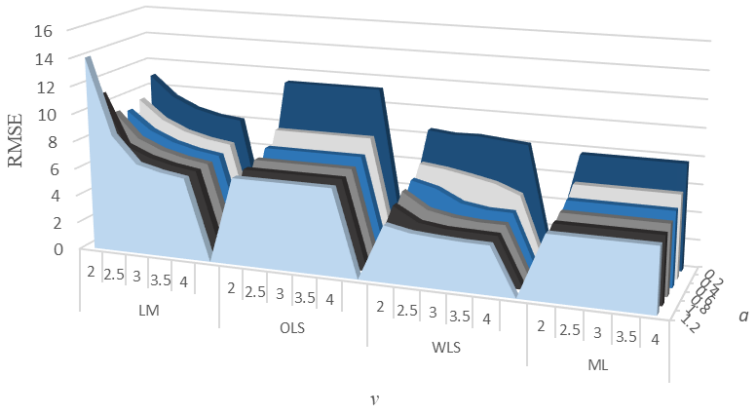


Fig. 5.. Empirical relative RMSE of various estimators of *v*.

Table 2. Characteristics of estimators of the Dagum model parameters.

Distribution	**Estimation method**	**RMSE$(\hat{a})$**		**RMSE$(\hat{v})$**		**RMSE$(\hat{\lambda})$**	
		***n*=1000**	***n*=2000**	***n*=1000**	***n*=2000**	***n*=1000**	***n*=2000**
D(0.2,2,1)	LM	16.510	12.574	10.944	8.097	11.971	9.037
	OLS	14.054	10.022	11.210	7.860	9.959	7.390
	WLS	11.088	7.883	8.522	5.976	8.848	6.544
	ML	9.993	6.982	7.703	5.240	8.354	6.065
D(0.6,2.5,1)	LM	17.572	12.359	7.659	5.572	9.727	7.204
	OLS	14.306	9.892	7.110	4.920	8.314	5.990
	WLS	10.684	7.279	5.312	3.804	6.462	4.541
	ML	10.438	7.243	5.412	3.743	6.403	4.590
D(0.8,3,1)	LM	17.261	11.820	6.742	4.825	8.053	5.840
	OLS	16.328	11.120	6.689	4.623	7.717	5.487
	WLS	9.601	6.400	3.942	2.800	4.635	3.239
	ML	11.288	7.800	5.066	3.517	5.643	4.013
D(1.2,4,1)	LM	19.513	12.884	5.195	4.180	6.552	4.659
	OLS	21.702	14.244	6.302	4.356	7.109	4.966
	WLS	11.131	7.433	3.475	2.497	3.797	2.680
	ML	13.313	9.119	4.691	3.277	4.786	3.373

The estimators of the parameter *a* obtained by the WLS or ML method are characterized by the smallest RMSE. Among considered estimators of *v* the smallest biases are observed for the WLS method. The smallest root mean squared errors of the estimator of *v* have been

noticed for the estimators obtained by ML or WLS method. LM estimators were found to present the smallest precision (highest RMSE of a and highest RMSE of v<2.5) amongst all the statistics considered in the study.

It is worth noting that the ML method tends to overwhelm the remaining ones when the sample size is increasing and when the inequality level is high what can be observed for the first distribution D(0.2,2,1), corresponding to G=0.704, shown in the Table 2. For the distributions presenting smaller inequality (e.g. D(0.8,3,1) with G=0.351) the WLS method seems more appropriate. It can also be noticed (Fig. 2-5) that the precision of the estimators of the Dagum model parameters v and a is getting better (smaller RMSE) together with the increasing values of v and decreasing values of a. It is what could have been expected as the v parameter is responsible for the right tail of the Dagum density, determining the number of finite moments of the distribution (see: eq. 4). Contrary to this, the a parameter is positively correlated with the distribution inequality (see: eq. 5 and 6) so its high values denote highly dispersed distributions. In general, the sample sizes n=1000 are still too small to confirm satisfactory level of RMSE, what is especially evident for a (Fig. 3). Just for n=2000 and for moderate inequality the RMSE values do not exceed 5% of the estimated parameters values.

5 Conclusions

The Dagum distribution is widely assumed as a theoretical model for income distributions in empirical analyses. Its parameters are to be estimated from sample data for whole countries and for different subpopulations. The main objective of this study was to recognise the estimation methods that can be applied for moderate sample sizes which are often considered in practice (socio-economic groups, family types or NUTS2 regions), The Monte Carlo experiments revealed that the smallest root mean squared errors of the estimators have been obtained when the ML or WLS method were applied. LM estimators were found to present the smallest precision (highest RMSE of a and highest RMSE of v<2.5) amongst all the statistics considered in the study. The ML method tends to overwhelm the remaining ones when the sample size is increasing and when the inequality level is higher. The precision of the estimators of the Dagum model shape parameters v and a is getting better (smaller RMSE) together with the increasing values of v and decreasing values of a. To obtain satisfactory results the sample size must be at least n=2000.

References

Dagum, C., (1977). A New Model of Personal Income Distribution: Specification and Estimation. *Economie Appliquee*, *30*, 413-437.

Dey, S., Al-Zahrani, B. & Basloom, S. (2017). Dagum Distribution: Properties and Different Methods of Estimation. *International Journal of Statistics and Probability, 6,* 74-92.

Domma, F., Giordano., S. & Zenga., M. M. (2011). Maximum Likelihood Estimation in Dagum Distribution with Censored Sample. *Journal of Applied Statistics*, *38,* 2971-2985.

Domański, C. & Jędrzejczak, A. (1998). Maximum Likelihood Estimation of the Dagum Model Parameters. *International Advances in Economic Research, 4,* 243-252.

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

Jędrzejczak, A. (2014). Income Inequality and Income Stratification in Poland. *Statistics in Transition, 15,* 269-283.

Jędrzejczak, A. (2015a). Regional Income Inequalities in Poland and Italy. *Comparative Economic Research, 18,* 27-45.

Jędrzejczak, A. (2015b). Asymptotic Properties of Some Estimators for Gini and Zenga Inequality Measures: a Simulation Study. *Statistica & Applicazioni, 13,* 143-162.

Kleiber, C. & Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*. Wiley, Hoboken.

Swain, J., Venkatraman, S. & Wilson, J. (1988). Least squares estimation of distribution function in Johnson's translation system. *Journal of Statistical Computation and Simulation, 29,* 271-297.

Victoria-Feser, M. P. (2000). Robust Methods for the Analysis of Income Distribution, Inequality and Poverty. *International Statistical Review, 68,* 277-293.

Victoria-Feser, M. P. & Ronchetti, H. (1994). Robust Methods for Personal Income Distribution Models. *The Canadian Journal of Statistics, 22,* 247-258.

Zenga, M. (1984). Proposta per un indice di concentrazione basato sui rapporti fra quantili di popolazione e quantili reddito. *Giornale Degli Economisti e Annali di Economia 48,* 301–326.

Spatial analysis and assessment of effectiveness of selected social services

Maciej Jewczak[1], Jadwiga Suchecka[2], Karol Korczak[3],
Marek Melaniuk[4]

Abstract

The aim of the article is to implement a new approach to assess the effectiveness of selected social services on the example of statistical data from the Social Welfare Centre in the city of Zgierz (Poland). The analysis was conducted in two ways – it concerned: 1) spatial differentiation analysis and 2) the assessment of the effectiveness of granting social services. In most analyses of spatial distribution for the assessment of the intensity of the social assistance, the indicators compare the structure of people receiving benefits in relation to the total population. The Authors propose the use of the Multiplicative Indicator of Poverty Intensity (*MIPI*), which consists of 4 components: the number of people living on a given street, the number of beneficiaries of social assistance and the number and the amount of social services granted. Assessment of the effectiveness of social assistance is treated in an extended way than the commonly-used approach expressing the ratio of the number of beneficiaries in relation to the actual number of people in need. The dynamic approach was also taken into account, which allowed determining changes in the effectiveness of social assistance and residence allowances. In the conclusions we have indicated the directions for further research that may be the basis for creating a more effective social policy.

Keywords: *Social welfare, benefits, social services, residence allowance, effectiveness.*
JEL Classification: I32, I38
DOI: 10.14659/SEMF.2018.01.19

1 Introduction

The period of transformation of Polish economy after 1989 caused a rapid pauperization of a part of society. This was related to the failure in adaptation to changes in both economy and social reality (Warzywoda-Kruszyńska and Grotowska-Leder, 1996; Cyrek, 2017; Krzysztofik et al., 2017; Rapeli et al., 2018). The phenomena of poverty, social exclusion and social support have become a significant problem within the framework of social policy

[1] Corresponding author: University of Lodz, Department of Spatial Econometrics, 90-214 Lodz, 37 Rewolucji Street 1905, e-mail: maciej.jewczak@uni.lodz.pl.
[2] University of Lodz, Department of Operational Research, 90-214 Lodz, 37 Rewolucji Street 1905, e-mail: suchecka@uni.lodz.pl.
[3] University of Lodz, Department of Computer Science in Economics, 90-214 Lodz, 37 Rewolucji Street 1905, e-mail: karol.korczak@uni.lodz.pl.
[4] University of Lodz, Department of Computer Science in Economics, 90-214 Lodz, 37 Rewolucji Street 1905, e-mail: marek.melaniuk@uni.lodz.pl.

(Grotowska-Leder, 2010). In Poland, the poverty can be classified according to its three categories:

- relative poverty – total household expenditure less than half of average households' expenditure (in 2016, it covered 13.9% of population),
- statutory poverty – covers expenditures lower than the amount entitling to apply for social assistance benefits, resulting from the Act on Social Assistance (in 2016, it covered 12.7% of population),
- extreme poverty – lack of satisfaction of the needs resulting from the existence minimum (in 2016, it covered 4.9% of population).

An important area of research on poverty is the spatial analysis of this phenomenon. This direction of research allows identifying the poverty enclaves (Sapiro, 2016). Moreover, from a practical point of view, a simultaneous and significant problem is the effectiveness of social policy. Analysis and evaluation of the performance of social policy could support decision making to counteract more efficient the poverty.

The purpose of this paper is to apply the *Multiplicative Indicator of Poverty Intensity* (*MIPI*). The proposed indicator includes: a) the number of people living on a given street, b) the number of beneficiaries and c) the value of the aid granted. The second aim of the research was the assessment of the effectiveness of granting social assistance and housing allowances. The analyses were mostly made on the example of the beneficiaries of the city of Zgierz (Lodz province) in years 2010-2016, however the estimates of *MIPI* concerned the final period of analyses.

The paper consists of two sections. First section contains the analysis of poverty intensity. In this section we present our own *MIPI* indicator. The assessment of the effectiveness of granted social assistance and housing allowances is presented in following section.

2 Analysis of poverty intensity

The analysis of the intensity of poverty was made on the basis of databases provided by the Social Welfare Center in the city of Zgierz (Poland). The released data is from 2016. Only the non-permanent benefits were considered for analysis: designated benefits, periodical benefits (usually granted for 3 months) and meals benefits. However, the benefits granted for a longer period of time, i.e. permanent benefits, care services and placement in social assistance homes have been omitted.

For defining the poverty enclaves, most commonly the relation of beneficiaries (B) benefiting from social assistance to the total number of citizens (C) – (B/C) indicator is used

(Warzywoda-Kruszyńska and Grotowska-Leder, 1996). However, this approach reveals its biggest drawback. Several associated with and at the same time important factors, such as: number of social services granted and the monetary value of services are excluded from the analysis. The issue mentioned emphasis the necessity for more complexed approach, which at the same time could give possibilities for adjustments.

In order to analyze the poverty intensity the *Multiplicative Indicator of Poverty Intensity* was constructed. There were 4 input data to calculate *MIPI* estimates:

- number of citizens (*C*) living in a certain location,
- number of beneficiaries receiving the benefits living in a certain location (*B*),
- number of social services for the beneficiaries living in a certain location (*S*),
- amount of money spent on social service for beneficiaries living in a certain location (*A*).

Then the following 3 indicators were calculated, as follows:

- *B*/*C* – percentage of beneficiaries (*B*) benefiting from social assistance in relation to the total number of citizens (*C*),
- *S*/*C* – number of benefits (*S*) admitted *per* citizen,
- *A*/*C* – amount (*A*) of the benefits provided *per* citizen.

The *Multiplicative Indicator of Poverty Intensity* (*MIPI*) is the expression of 3 indicators:

$$MIPI = \frac{B}{C} * \frac{S}{C} * \frac{A}{C}.$$

Let us consider the case of Szeroka Street, with C = 62 citizens, B = 34 beneficiaries, S = 498 social services and A = 55 123 PLN, the three components amounted to: *B*/*C* = 54.8%, *S*/*C* = 8, A/C = 889.1 and *Multiplicative Indicator of Poverty Intensity* reached: *MIPI* = 3 916.2.

Table 1 contains information on calculated indicators for selected 25 streets of Zgierz, which number of beneficiaries *per* citizen (*B*/*C*) exceeded in 2016 the ratio of 6%. All the streets were classified according to the decreasing value of the *MIPI* indicators which differs from the *B*/*C* order. If the intensity of poverty is higher (column with *MIPI* values), the lower street number was issued (*MIPI* order column). Only for 11 out of 25 streets, *B*/*C* order equaled the order of the *MIPI* indicator.

The *MIPI* values were further plotted on the map of Zgierz (Fig. 1). This allowed for analyzing the concentration of the spatial distribution of locations for the street layout. As a consequence, simultaneously it was possible to determine the poverty intensity map with the assessment of phenomenon's intensity.

Table 1. The Multiplicative Indicators of Poverty Intensity (MIPI) for the streets of Zgierz.

Streets	***B/C***	***B/C*** order	***S/C***	***A/C***	***MIPI***	***MIPI*** order
Szeroka	54.8%	1	8.0	889.1	3 916.2	1
Wspólna	32.3%	2	4.4	561.5	800.4	2
Słowackiego	24.9%	3	3.7	498.7	457.7	3
Narutowicza	22.1%	4	2.4	287.4	153.0	4
Plac Jana Pawła II	16.9%	8	2.4	290.5	117.4	5
Dąbrowskiego	18.8%	6	2.3	268.7	117.3	6
Plac Kilińskiego	19.9%	5	2.0	268.5	105.8	7
Popiełuszki	17.3%	7	2.0	230.9	79.8	8
Łęczycka	12.4%	11	1.7	203.2	41.5	9
Koszarowa	15.6%	10	1.6	156.2	38.2	10
Sieradzka	8.0%	22	1.9	231.8	34.9	11
Skargi	16.5%	9	1.8	106.4	31.2	12
Mielczarskiego	11.7%	12	1.3	161.1	25.0	13
Cezaka	11.4%	13	1.3	150.9	22.6	14
Pułaskiego	9.8%	15	1.3	149.6	19.5	15
Piłsudskiego	8.8%	18	1.3	170.3	18.9	16
Klonowa	9.3%	16	1.1	136.5	13.9	17
Piątkowska	8.8%	19	1.0	142.7	12.0	18
Śniechowskiego	8.0%	23	1.1	130.9	11.8	19
3 Maja	8.4%	20	1.0	130.8	11.4	20
Pawińskiego	8.4%	21	0.9	153.4	11.4	21
Aleksandrowska	9.8%	14	1.0	114.1	11.0	22
Konstantynowska	9.2%	17	0.9	105.1	8.9	23
Długa	7.8%	24	0.9	112.5	7.9	24
Rembowskiego	6.8%	25	0.7	94.8	4.8	25

The obtained higher poverty levels (exceeded values of *MIPI*) were located in the city center (the Old Town – with *MIPI* streets no: 1, 2, 5, 7, 11, 12). It should be also noted that the lower street number indicated higher intensity of poverty.

The poverty intensity should be of particular interest to social workers dealing with social assistance for beneficiaries or families that require such assistance (financial and non-financial support).



Fig. 1. Spatial distribution of *MIPI*. Size of symbol indicate the level of *MIPI*.

The major advantage of using the complexed *MIPI* indicator to analyze the intensity of poverty in a given city is a fact that it includes not only the number of beneficiaries (*B*), but also the number of received benefits (*S*) and the financial value of total assistance (*A*). The number of benefits is the main factor determining the frequency of granted assistance, while the value of material support can inform (at least partially) about the degree of satisfying the social needs. The combination of these three factors in one indicator (*MIPI*) showed adjusted (while compared with typical (*B*/*C*) analysis) intensity of the poverty phenomenon in selected location.

3 Assessment of the effectiveness of social services

In addition to spatial distribution analysis indicating poverty enclaves, one of the most important principles for the social benefits system functioning and its quality is the effectiveness issue (Suchecka and Jewczak, 2014). Many indicators can be proposed to assess

effectiveness, while the concept itself needs to be clarified (Golinowska, 2012; Golinowska and Topińska, 2002; Hryniewicka, 2011).

Effectiveness analysis is used not only for beneficiaries, but also for the units providing social services such as: socials welfare centers (Szatur-Jaworska, 2010). The term effectiveness should be understood as the degree to which the activities carried out contribute to the objectives' achievement. However, this definition might be perceived as too general statement. In practice of social welfare, this concept can be interpreted two ways:

1. either as the number or share of people covered by assistance in relation to the actual number of people in need for the assistance – this recognition of effectiveness is commonly adopted into practice;
2. or, it relates to people (or households) who have already received the aid (e.g.: in a certain year) and the share of those who receive aid in subsequent periods (years). The effectiveness defined within this framework was analyzed in the paper basing on the example of provided social assistance (in years 2012-2016) and granted housing allowances (for 2010-2013 period) by the Social Welfare Centre in Zgierz.

3.1 Assessment of the effectiveness of social assistance

Three types of social assistance allowances were taken into consideration: permanent, temporary and designated. The analyzed data related to the period from 2012 to 2016. Permanent benefits, placement in nursing houses and care services were omitted, while they represent assistance provided most often in long-term periods, and at the same time it was not possible to investigate their effectiveness in case considered. It was assumed that if the effectiveness of social benefits is to be high, beneficiaries using social assistance do not continued applying for the benefits in the next times period. The assumption allowed for constructing the temporary condition of effectiveness for social benefits, as follows:

$$Effectiveness\ Indicator = \frac{B_{t+s}}{B_t},$$

where: B_t – beneficiaries in base period (total number of people with granted social allowance), B_{t+s} – beneficiaries in comparable period (s – indicates the duration of service given, i.e.: for 1 year). The social benefit should be classified as effective when none of the beneficiaries from period t applied for the same services in period $t+s$, in this case the *Effectiveness Indicator* should equal zero ($Eff.Ind > 0$ indicates the ineffectiveness of social aid).

Table 2 contains numbers of people and the effectiveness indicators for individuals, who benefit from social assistance in 2012 and continued to receive assistance benefits in following years. In order to calculate the effectiveness ratios, complete data for the city of Zgierz was used.

Table 2. Numbers of beneficiaries and effectiveness indicators for social assistance in 2012 and following years.

Specification	2012	2013	2014	2015	2016
No. of beneficiaries	3 099	2 345	1 828	1 594	1 307
Effectiveness indicator (base period: 2012)	1	0.76	0.59	0.51	0.42

Fig. 2 indicates that the percentage of people who have benefited from social assistance since 2012 and shows a declining trend with significant time coefficient parameter (*p-value*<0.0000).

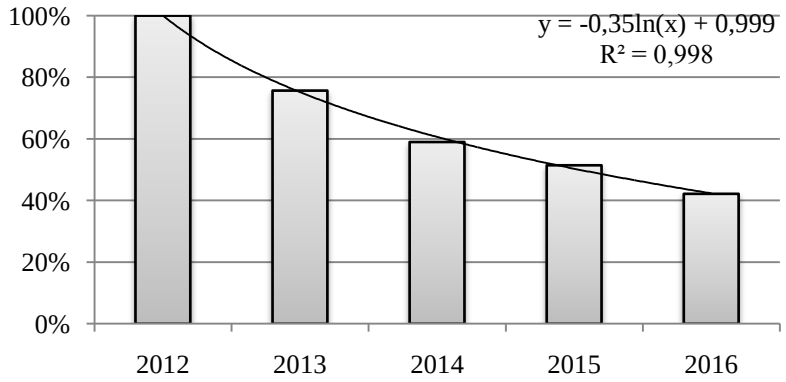


Fig. 2. The effectiveness indicator for social assistance between 2012 and 2016.

As it was already mentioned, the effectiveness level of social assistance was gradually decreasing, however the highest reduction took place in 2013 when compared with 2012 (the number of people dropped by 24%), while in the following years the rate of decrease was clearly slowing down and amounted to: 22% in 2014 compared to 2013; 13% in 2015 compared to 2014; 18% in 2016 compared to 2015.

Faster reduction in the number of beneficiaries between the first and the second period resulted from the fact that for approx. 10% of beneficiaries the assistance was successful and further support no longer required – the changes should be there conclued as positive. Similar

relationships were clearly visible when comparing the following periods (subsequent columns in Table 2). In order to verify the observed trends in the frequency use of social assistance in the current period (2012-2016), estimates were made for the control period, years 2002-2006 (considering only targeted benefits and for 2002). The obtained results were convergent for both periods and detailed estimates for 2002-2006 were presented in Table 3.

Table 3. The effectiveness indicator for designated benefits of social assistance between 2002 and 2006.

Specification	2002	2003	2004	2005	2006
No. of beneficiaries	1 865	1 315	1 068	953	845
Effectiveness indicator (base period: 2002)	1	0.71	0.57	0.51	0.45

A rapid decrease of 29% in the number of people benefiting from designated benefits occurred in 2003 (compared to 2002) and in subsequent years the rate of decline decreased significantly, which indicated that a significant number of people still needed assistance (the rates of decline amounted for: $d_{04/03}$=-18%, $d_{05/04}$=-11% and $d_{06/05}$=-11%).

3.2 Assessment of the effectiveness of housing allowances

In the paper, we analyzed the effectiveness of assistance provided in the area of housing allowances in Zgierz for the period 2010-2013 (Melaniuk, 2014). The calculation principles were identical to those in the case of social assistance. The calculated effectiveness indicators were presented in Table 4.

Table 4. The effectiveness indicator for beneficiaries receiving housing allowances in Zgierz in the years 2010-2013.

Specification	2010	2011	2012	2013
No. of beneficiaries	2 076	852	604	428
Effectiveness indicator (base period: 2010)	1	0.31	0.29	0.21

In the first column, we assumed the number of citizens receiving housing allowances in 2010 as base period. Of this number, in 2011 31.4% of families benefited from the allowances

and in the following two years – 29.1% and 20.6%, respectively. Naturally, in the following years, the housing allowances of these families, which go on to the next year have been omitted and only the newly started additions were taken into account. The indicators in the second and third columns are calculated analogically. For the interpretation purposes: if the housing allowance would be fully effective in the following years, the individuals would not have to re-apply for the allowance. The number of 21% in 2013 should be considered satisfactory in generally, since almost 80% of families after three year period, no longer required continuing applying for housing allowances. However, it should be also noted and highlighted that in our analysis there were cases that some individuals received a housing allowance for over a decade, without any interruption, which could indicate for ineffectiveness of social allowance policy.

Conclusions

The article presented new, advanced and more complexed approach towards assessing the effectiveness of social services and social policy as well. An important part of analyses was dedicated to identifying the enclaves of poverty in the city Zgierz, where we proposed the *Multiplicative Indicator of Poverty Intensity*. *MIPI* allowed for integrating not only the percentage of people benefiting from social services, but also the value of the aid received and the number of benefits granted directly to people living in certain locations (city streets). We treated the effectiveness of the social services as a dynamic phenomenon, taking into account the percentage of people profit from social assistance benefits and housing allowances in subsequent years in relation to the base year. The results of the research indicated that the effectiveness of social services is not satisfactory, while in five year period after being granted with the assistance, half of the individuals still received benefits (which by definition should be of permanent state). More positively one should consider the effectiveness of social services in form of housing allowances – in this very case for city of Zgierz, after 4 year period the number of people still receiving the allowance amounted to one fifth (20%) only.

In the paper, *Multiplicative Indicators of Poverty Intensity* were estimated for individual streets in one period. On the other hand, effectiveness indicators concerned the city area and were presented in dynamic approach. Further research should concentrate on establishing a comprehensive approach to the problem of poverty combining a dynamic analysis of social services effectiveness with a multi-criteria *MIPI* indicator.

References

Cyrek, M. (2017). Social efficiency of employment in three sectors — a comparison of Polish regions. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *12*(3), 417-432.

Golinowska, S. (2012). O biedzie i polityce jej zwalczania współcześnie. Podejście porównawcze w świetle Europejskiego Roku Zwalczania Ubóstwa i Wykluczenia Społecznego. In: Kubiak, H. E. (ed.), *Polska bieda w świetle Europejskiego Roku Walki z Ubóstwem i Wykluczeniem Społecznym.* Kraków: Oficyna Wydawnicza AFM.

Golinowska, S. & Topińska, I. (2012). *Pomoc społeczna – zmiany i warunki skutecznego działania*. Warszawa: Centrum Analiz Społeczno-Ekonomicznych.

Grotowska-Leder, J. (2010). Analiza zasobów wsparcia społecznego łodzian, *Praca Socjalna,* (8-9)*:* 181-196.

Hryniewicka, A. (2011). *Różne wymiary skuteczności w pomocy społecznej*. Warszawa: Instytut Rozwoju Służb Społecznych.

Melaniuk, M. (2014). Logistyka dystrybucji świadczeń społecznych. *Logistyka* (5)*:* 1968-1976.

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

Rapeli, M., Cuadra, C., Dahlberg, R., Eydal, G. B., Hvinden, B., Ómarsdóttir, I. L. & Salonen, T. (2018). Local social services in disaster management: Is there a Nordic model? *International Journal of Disaster Risk Reduction*, (27), 618-624.

Sapiro, P. (2016). Beyond the strictly orthodox/mainstream divide: Applying geodemographic analysis to a small nationwide sub-population. *Computers, Environment and Urban Systems*, (56), 36-47.

Suchecka, J. & Jewczak, M. (2014). Cost-Benefit Analysis in the study of the effectiveness of health programs. In: Papież M. & Śmiech S. (eds), *Proceedings of the 8th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena*, 194-203.

Szatur-Jaworska, B. (ed.). (2010). *Ewaluacja w służbach społecznych*. Warszawa: Mazowieckie Centrum Polityki Społecznej.

Warzywoda-Kruszyńska, W. & Grotowska-Leder, J. (1996). *Wielkomiejska bieda w okresie transformacji (zasiłkobiorcy pomocy społecznej)*. Łódź: Instytut Socjologii Uniwersytetu Łódzkiego.

New proposal of robust classifier for functional data

Daniel Kosiorowski[1], Dominik Mielczarek[2], Jerzy P. Rydlewski[3]

Abstract

A variety of economic research hypotheses may be translated into language of statistical discrimination analysis. The company's ability to adapt to changing environmental circumstances may be expressed in terms of a quality of a classifier used in a decision process by the company's management. The specific classifier is developed basing on company's experience expressed in the, so called, training sample. In practice, however, training samples contain outliers of various kinds, which influence the classifier quality. For this reason, robust classifiers, which are able to cope with various data imperfections, are especially desired. This paper focuses on robust classification issues for functional data. We present the state of art and indicate its consequences for the robust economic analysis. We propose an original classification rule appealing to the support vector machines methodology. We show its selected properties and apply it to an empirical issue related to monitoring of electricity market in Denmark.

Keywords: *classifier for functional objects, electricity market, robustness, support vector machines classifier*
JEL Classification: C14, C38, C44
DOI: 10.14659/SEMF.2018.01.20

1 Introduction

A variety of economic research hypotheses may be translated into a language of statistical discrimination analysis. Consider, for example, the company's ability to adapt to changing environmental circumstances, which may be expressed in terms of a quality of a classifier used in a decision process by the company's management. The specific classifier is developed basing on company's experience expressed in the, so called, training sample. In practice, however, training samples contain outliers of various kinds, which may adverse influence on the classifier quality. This fact motivates our studies. The paper focuses on robust classification issues for functional data related to the robust economic analysis of electricity market in Denmark in 2016. The recently developed statistical methodology named functional data analysis (FDA) enables for functional generalizations of well-known one and multivariate statistical techniques like analysis of variance, kernel regression or classification

[1] Corresponding author: Cracow University of Economics, Department of Statistics, ul. Rakowicka 27, 31-510 Kraków, e-mail: daniel.kosiorowski@uek.krakow.pl.
[2] AGH University of Science and Technology, Faculty of Applied Mathematics, al. A. Mickiewicza 30, 30-059 Krakow, Poland, e-mail: dmielcza@wms.mat.agh.edu.pl.
[3] AGH University of Science and Technology, Faculty of Applied Mathematics, al. A. Mickiewicza 30, 30-059 Krakow, Poland, e-mail: ry@agh.edu.pl.

techniques (see Horvath and Kokoszka, 2012; Górecki et al., 2014). The FDA enables us for effective analysis of economic data streams i.e., e.g., analysis of non-equally spaced observed time series, prediction of a whole future trajectory rather than single observations (see Kosiorowski, 2016). In the recent years several interesting for economic application procedures for functional data have been proposed (see Horvath and Kokoszka, 2012). The proposed techniques are not robust however.

In a context of functional data classification analysis, several issues are still unsolved. It should be stressed that a commonly acceptable definition of robustness for a classification procedure even in a multivariate case does not exist up to now (see Hubert et al., 2016). The robustness concept in this case should take into account a local nature of classification procedure (maybe robustness should be defined with respect to specified class rather than regarding the whole data set). One can propose a useful variant of qualitative robustness however: small changes of input should lead to small changes of output or a measure of quality of output. Therefore, we can adapt qualitative definition of robustness. It is possible to adapt Hampel's influence function as well. More recent and formal approaches may be found in Christmann et al. (2013). The main aim of the paper is to propose a new nonparametric statistical methodology, appealing to the Support Vector Machines method (SVM) (Schoelkopf and Smola, 2002), for classifying functional objects. We show its selected properties and apply our classifier to an empirical issue related to monitoring of electricity market in Denmark. Certain aspects of electricity prices modeling and forecasting have been described by Weron in his review paper (2014).

The rest of the paper is organized as follows. Section 2 sketches the basic concepts of classification in a functional setup. Section 3 introduces our procedure for classifying a functional data, and discusses the properties of the procedure through numerical simulations and tests the applicability of the proposed methodology on empirical examples. Section 4 describes the analysed electricity dataset. Section 5 conducts a short robustness analysis and a conclusion is provided.

2 A classifier for functional data

We are given a training sample, that is, *n* observations: U_1, U_2,…U_n and each observation can be described as $U_i=(X_i, Y_i)$, where $Y_i=-1$ or $Y_i=1$. The X_i are called patterns, cases, inputs or instances, and Y_i are called labels, outputs or targets. We would like to classify a new object *X* into one of the two labelled groups, basing on knowledge included in the training sample. A

classification rule defines a certain partition of the feature space into nonempty disjoint subsets, which are summed up to a whole feature space.

Classification methods for functional data include k-nearest neighbors (kNN) methods, reproducing kernel of a Hilbert space (RKHS) methods (see Schoelkopf and Smola, 2002), methods based on depth measures (see Cuevas and Fraiman, 2009; Kosiorowski et al., 2018), and neural networks methods. Functional outliers detection on the example of air quality monitoring has been recently described by Kosiorowski et al. (2018).

3 Our proposals

Let $X_1,X_{2,\ldots},X_m$ be any functional data from Hilbert space $L^2(\Omega)$ with the usual inner product defined by $< f,g >= \int_\Omega f(w)g(w)dw$, where Ω is a bounded subset of Euclidean space, and let numbers $Y_1,Y_{2,\ldots},Y_m$ be labels, i.e., $Y_i=1$ or $Y_i=-1$ for i=1,2,…,m. Patterns X_i are functions mapping set Ω into real numbers. We assume in the whole paper, that the set Ω is bounded, and then the space $L^2(\Omega)$ is separable, what is important in our considerations. Hence, there exists an orthonormal basis $\{Z_1,Z_2,\ldots\}$and every function X from the space $L^2(\Omega)$ can be described as the following series $X=\sum_{n=1}^{\infty}\langle X,Z_n\rangle\cdot Z_n$, where the series convergence is a convergence in the sense of norm of the space $L^2(\Omega)$.

In our further considerations we set $\Omega=[0,T]$. Patterns X_i belong to $L^2([0,T])$ space with the usual inner product defined by $< f,g >= \int_0^T f(t)g(t)dt$, and the orthonormal basis is either a standard Fourier basis or a spline basis.

In practice, we fix a natural number K and we determine a vector $c=(c_1,c_2,...,c_K)$ such that $\hat{X}=\sum_{n=1}^{K}c_n\cdot Z_n$, so that they minimize a real function ϕ given by the following formula $\phi(c)=\left(X-Zc\right)^T\cdot\left(X-Zc\right)$, where $X^T=(X(t_1),\ldots,X(t_M))$ and Z is a matrix of the form $\left[Z_j(t_i)\right]_{i=1,\ldots,M}^{j=1,\ldots,K}$ and $t_i\in[0,T]$ are knots. We propose a classifier for functional data of the form

$$f(X)=\int_\Omega X(\omega)W(\omega)d\omega+b$$

where b is any real number and weight function W is essentially bounded, i.e. $W\in L^2(\Omega)$ and chosen so that affine functional f be data-consistent i.e. $Y_i f(X_i)=1$, for any i=1,2,…,m.

In other words, we are given empirical data $(X_1, Y_1), (X_2, Y_2),\ldots,(X_m, Y_m)$. Basing on the data, we classify a new functional observation X into one of the groups looking only on *sgn* (f(X)). The classifier doesn't work, if f(X) = 0 (for practical purposes one may assume zero probability for such an event). Existence of the weight function W, as can be shown, is guaranteed with linear independence of the random functions $X_1,X_{2,\ldots},X_m$. Using functional analysis apparatus, i.a. Hahn-Banach Theorem, we have proved the following theorem.

Theorem. For any real number *b* there exists a function $W \in L^2(\Omega)$ such that

$$Y_i\left(\int_\Omega X_i(\omega)W(\omega)d\omega + b\right) = 1, \text{ for } i = 1,2,...,m$$

has a solution.

Note that, the weight function *W* satisfies

$$\int_\Omega X_i(\omega)W(\omega)d\omega + b = Y_i, \text{ for } i = 1,2,...,m.$$

It is now obvious, that in order to solve the classification problem it suffices to determine a weight function *W*, or equivalently to find a functional g such, that

$$g(X_i) = \int_\Omega X_i(\omega)W(\omega)d\omega = Y_i(1 - Y_i b) \text{ for } i = 1,2,...,m.$$

We show now, how to determine the hyperplane separating for functional data. Let fix an index i=1,2,…,m. We construct a bounded linear functional $g_i : L^2(\Omega) \to R$ such that $g_i(X_j) = \delta_{ij}$, where δ_{ij} is a Kronecker delta. Then, a functional g given by the formula $g = \sum_{j=1}^{m} Y_j(1 - Y_j b)g_j$ satisfies the equations (1).

In order to solve the set of equations (1) it suffices to indicate functionals g_i. Let us denote for any set of vectors $W_1,W_2,\ldots,W_m$ from Hilbert space with inner product $<\cdot,\cdot>$

$$M(W_1,...,W_m) = \det\left[< W_i, W_j > \right]_{i=1,...,m}^{j=1,...,m}$$

The number $M(W_1,...,W_m)$ is a Gram matrix determinant for a set of vectors $W_1,W_2,...,W_m$. Recall, that a set of vectors is linearly independent if and only if $M(W_1,...,W_m) > 0$. The functionals g_i we are looking, for any $Y \in L^2(\Omega)$ are given by the formula:

$$g_i(Y) = \frac{M(X_1,...,X_{i-1},Y,X_{i+1},...,X_m)}{M(X_1,...,X_m)}$$

We now give a formula for a separating hyperplane. After conducting some simple computations, the weight function for the functional g is given by the following formula:

$$W = \sum_{i=1}^{m} \frac{X_i - P_i(X_i)}{\left\| X_i - P_i(X_i) \right\|^2}$$

where $P_i : L^2(\Omega) \to V$ is an orthogonal projection on $V = span\{X_1, ..., X_{i-1}, X_{i+1}, ..., X_m\}$.

It is the most right place here to remark, that as $L^2(\Omega)$ space is not a RKHS (Reproducing Kernel Hilbert Space), then the proposed classifier is not of that kind as well. Moreover, it can be easily seen that the classifier is affine invariant, i.e. invariant with respect to the mapping $A: X \to L(X) + s$, where L is a linear mapping and s is a translation.

4 Robustness of a classification rule for functional data

The robustness of the classifying rule toward outliers depends on the functional outliers type. It should be different for the functional shape outliers, functional amplitude outliers and for functional outliers with respect to (w.r.t.) the covariance structure. That's why it is not easy to approximate breakdown point or influence function. It should be stressed, that there is no agreement as to the breakdown point or influence function concepts in the functional classification case (see Hubert et al., 2015 and 2016, and references therein). Many of the classical robust classification methods assumes multivariate normality or elliptical symmetry, which is a simplification of the more complex problem. Hubert et al. (2016) discusses some robust approaches that can deal with functional data. They make use of the concept of depth and present in the article a new technique - classification in distance space. They carry out a distance transformation and use a bagdistance to obtain a robust classification rule.

Cuevas and Romo (1993) studied qualitative robustness of bootstrap approximations when the estimators are generated by a statistical functional *T*. They showed that the uniform continuity of statistical functional *T* is a sufficient condition for a qualitative robustness of the bootstrap estimator. Denote now $L_n(F)=L_n(T;F)$ is the sampling distribution of the statistic $T_n(X_1,...,X_n)$ where the sample comes from *F*, and $L[L_n(F_n)]$ is the sampling distribution of the generalized statistic $L_n(F_n)$ in the space of relevant probability measures. Cuevas and Romo (1993) definition states that „given a sequence $\{T_n\}$ of statistics generated by a statistical functional *T*, the sequence of bootstrap approximations $\{L_n(F_n)\}$ is said to be qualitatively robust at *F* when the sequence of transformations $\{G \to L[L_n(G_n)]\}$ is asymptotically equicontinuous at *F*." Their concept is followed by i.e. Christmann et al. (2013) and by us. *Qualitative robustness* is thus defined as equicontinuity of the distribution of the considered statistic as the sample size is growing, and hence the concept of qualitative robustness is

related to continuity of the statistic in the relevant space, which is now considered as a function in the weak* topology (see Rudin, 1991).

Remark. It can be proved, following lines of Christmann et al. (2013) proof, that our classifier is qualitatively robust as well. In the proof we exploit the fact, that Gram matrix used in classifier construction, as a matrix, is a continuous mapping. Notice, however, that our classification rule is not a kernel classification rule, so their proof cannot be directly applied.

Tarabelloni (2017) defined Max-Swap Algorithm, that can be used to obtain more robust classification rule. He suggests to compute the covariances for two groups obtained by including the unit either in the first or second group. The distances between the two groups are computed then. The new observation is attributed to the group for which the considered distance *is* the least. Finally, he suggests to choose the swapping units in such a way that the distance between the estimated covariance operators at the next step is higher than the distance between the estimated covariance operators at the preceding step. Controlling numbers of units in each groups affects the robustness of his algorithm (for details see Tarabelloni, 2017).

5 Empirical analysis of the electricity dataset

We consider data from an electricity market in Denmark, where each day is represented as a function. We use an electricity consumption data retrieved from www.nordpoolgroup.com. Using classifying methods, we would like to classify a new functional object into one of the considered groups: working day or weekend. Fig. 1 presents functional observations of Danish electricity consumption in working days in 2016, while Fig. 2 presents functional observations of Danish electricity consumption in weekend in 2016.

Training sample consists of the functional data of electricity consumption representing first 20 days of each month in 2015. Subsequently, separating hyperplanes have been computed for each month and a classification has been conducted for functional observations of electricity consumption representing each day of 2016.

Let now g denote our classification rule. The distribution of (X, Y) is unknown, so we estimate the empirical risk of misclassification for our rule

$$\hat{L}(g) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{g(X_i)\neq Y_i\}},$$

where 1_S denotes the indicator function of the set S. We have compared our classification rule for the analyzed empirical data with: RKHS kernel methods with Gaussian kernel and

polynomial kernel (see Febrero-Bande and de la Fuente (2012) R package fda.usc). Empirical risk for our rule g is 26%, while empirical risk for Gaussian kernel RKHS method is 66%, for polynomial kernel RKHS method is 58%. Our method turned out to be computationally as intensive as the compared methods.

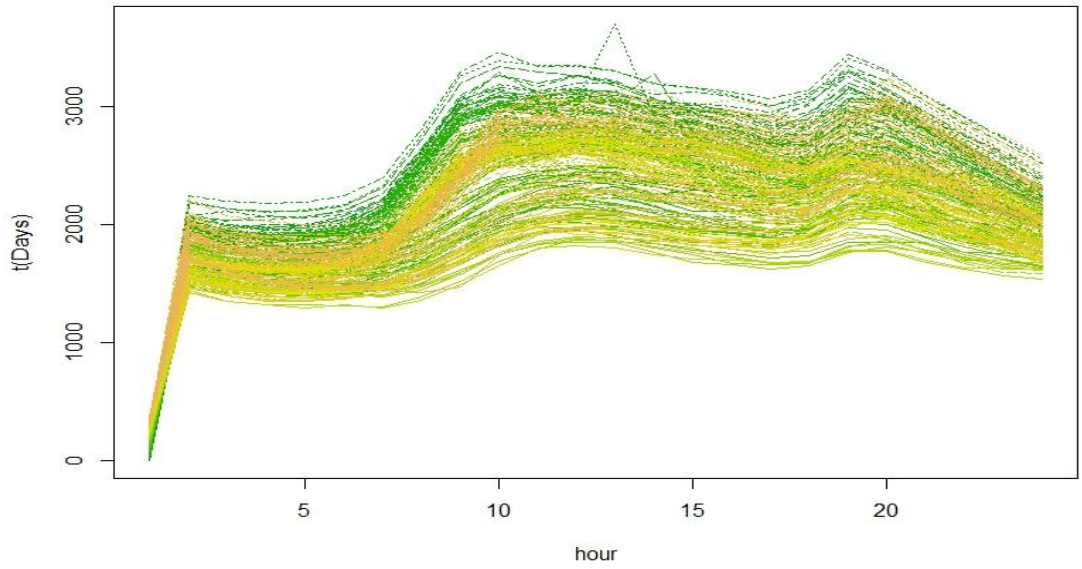


Fig.1. Functional observations electricity consumption in working days in 2016 for Denmark.

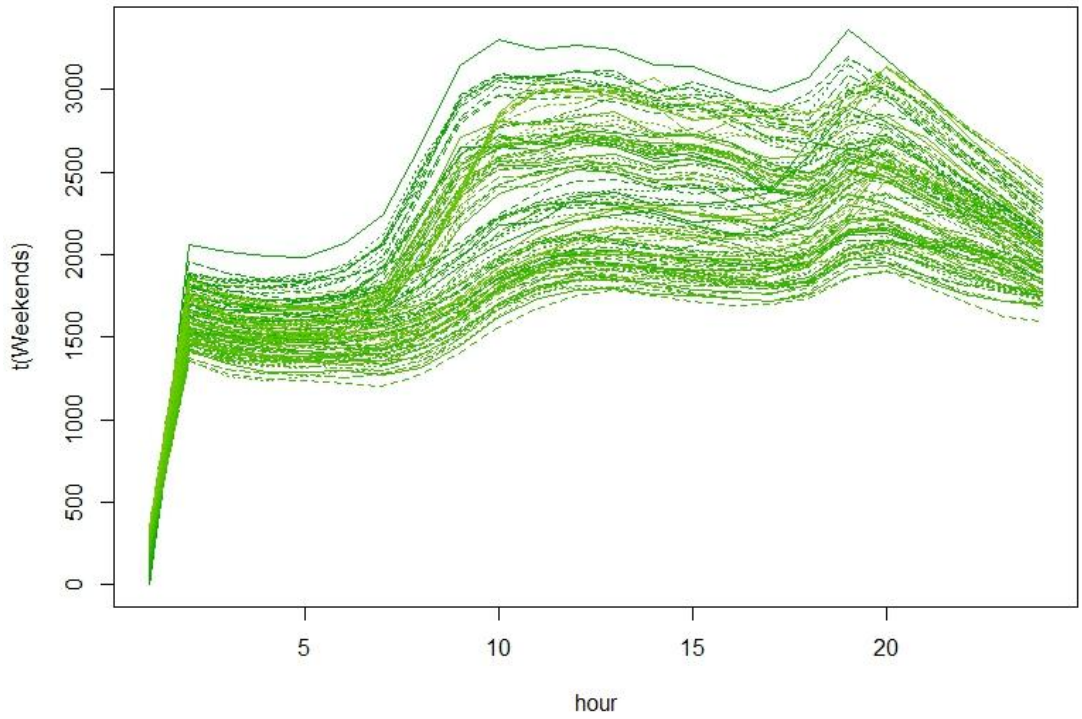


Fig.2. Functional observations electricity consumption in weekends in 2016 for Denmark.

Conclusions

This paper proposes new classification method for functional data. The presented method allows for monitoring phenomena appearing within the new economy, which are described by means of functions of a certain continuum. We show on a real data set, related to electricity consumption, some advantages and disadvantages of our proposal. In the future, we plan some further studies of the proposal involving gamma-regression or beta-regression.

Our method is significantly better w.r.t. empirical risk than the considered kernel RKHS methods. Nonetheless, note that our method is computationally as intensive as other methods.

Our proposals can be applied to different fields of an e-economy, i.e., (Web site management, protection of computer systems against hacking, spam filtering, etc. - as an e-economy provides a great deal of functional data) or to optimization of electricity production.

Acknowledgements

JPR and DM work was partially supported by the Faculty of Applied Mathematics AGH UST statutory tasks within subsidy of Ministry of Science and Higher Education, grant no. 11.11.420.004 and DK's research by the grant awarded to the Faculty of Management of CUE for preserving scientific resources for 2017 and 2018.

References

Christmann, A., Salibián-Barrera, M. & Van Aelst, S. (2013). Qualitative Robustness of Bootstrap Approximations for Kernel Based Methods. In: Becker, C., et al. (eds.), *Robustness and Complex Data Structures*, Springer, Berlin Heidelberg, 263-278.

Cuevas, A. & Romo, R. (1993). On robustness properties of bootstrap approximations. *Journal of Statistical Planning and Inference*, *37*(2), 181–191.

Cuevas, A. & Fraiman, R. (2009). On depth measures and dual statistics. A methodology for dealing with general data. *Journal of Multivariate Analysis*, *100*(4), 753-766.

Febrero-Bande, M. & de la Fuente, M. (2012). Statistical computing in functional data analysis: The R package fda.usc. *Journal of Statistical Software, 51*(4), 1-28.

Górecki, T., Krzyśko, M., Waszak, Ł. & Wołyński W. (2014). Methods of Reducing Dimension for Functional Data. *Statistics in Transition, 15*(2), 231-242.

Horvath, L. & Kokoszka, P. (2012). *Inference for functional data with applications*. Springer.

Hubert, M., Rousseeuw, P. & Segaert, P. (2015). Multivariate functional outlier detection. *Statistical Methods & Applications, 24*(2), 177–202.

Hubert, M., Rousseeuw, P. & Segaert, P. (2016). Multivariate and functional classification using depth and distance. *Advances in Data Analysis and Classification*. Doi: 10.1007/s11634-016-0269-3

Kosiorowski, D. & Zawadzki, Z. (2017). DepthProc: An R Package for Robust Exploration of Multidimensional Economic Phenomena. Retrieved from https://arxiv.org/abs/1408.4542v9

Kosiorowski, D. (2016). Dilemmas of robust analysis of economic data streams. *Journal of Mathematical Sciences, 1*(2), 59-72.

Kosiorowski, D., Rydlewski, J. P. & Zawadzki, Z. (2018). Functional outliers detection by the example of air quality monitoring. *Statistical Review* (in Polish, to appear).

Rudin, W., (1991). *Functional Analysis*. McGraw-Hill.

Schoelkopf, B. & Smola, A. J. (2002). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.* MIT Press.

Tarabelloni, N., (2017), *Robust Statistical Methods in Functional Data Analysis,* PhD thesis and roahd R package, Politecnico di Milano.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting, 30,* 1030–1081.

Outliers in functional time series – challenges for theory and applications of robust statistics

Daniel Kosiorowski[1], Dominik Mielczarek[2], Jerzy P. Rydlewski[3]

Abstract

This paper critically discusses the most promising, known from the literature, approaches to analysis of robustness in the functional time series setup. We also propose our own method of detecting functional outliers appealing to the generalized Young inverse function. The method is dedicated for detecting a certain kind of shape outliers. In empirical example we study day and night air pollution with PM10 in Katowice in 2016.

Keywords: *functional outliers detection; functional data analysis; robust economic analysis*
JEL Classification: C12, C13, C14
DOI: 10.14659/SEMF.2018.01.21

1 Introduction

Robust statistics studies various relationships between majorities and influential minorities of observations within data sets and within underlying statistical models. These studies involves behaviour of estimators or tests when samples come from data generating processes, which are in neighbourhoods (minority) of assumed models (majority). Having at our disposal a reasonable definition of majority as a by-product one may define an outlier as observation departing from the majority in a certain justified way (Maronna et al., 2006). Although main aim of robust statistics is to propose robust procedures, i.e., techniques having good properties at the assumed model as well as in case of a slight departure from the model, a detection of outliers may be even more important issue in a context of developing new theories explaining economic phenomena. Notice that departure from the main pattern may signal problems with an old theory or a completely new phenomenon. In practice, outliers detection procedures may be important in a context of safety systems development or e-economy monitoring.

Many economic phenomena may be treated as functional time series (FTS) – series of functions of a certain continuum representing, for example, time, temperature, interest rate, an aversion to risk, age of a credit applicant etc. Assuming a certain degree of regularity as to

[1] Corresponding author: Cracow University of Economics, Department of Statistics, ul. Rakowicka 27, 31-510 Kraków, e-mail: daniel.kosiorowski@uek.krakow.pl.
[2] AGH University of Science and Technology, Faculty of Applied Mathematics, al. A. Mickiewicza 30, 30-059 Krakow, Poland, e-mail: dmielcza@wms.mat.agh.edu.pl.
[3] AGH University of Science and Technology, Faculty of Applied Mathematics, al. A. Mickiewicza 30, 30-059 Krakow, Poland, e-mail: ry@agh.edu.pl.

data generating the phenomena, one can effectively model, conduct statistical inference and predict them within a framework offered by relatively new discipline of multidimensional statistics called functional data analysis (FDA). Unfortunately, observed empirical data sets very often manifest departure from the assumed regularity in a form of existence of outliers within data. These facts may rise doubts as to credibility of applications of functional generalizations of standard time series techniques like ARMA modelling (Horváth and Kokoszka, 2012; Górecki et al., 2017). It should be stressed however, that although there are known effective techniques of coping with outliers in one- and multidimensional cases, the situation is far from satisfactory in classical time series setting (Maronna et al., 2006; Galeano et al., 2006). Within the FDA we additionally face new challenges for development of robust statistical procedures. We have to cope with completely new classes of outliers. Unfortunately, straightforward generalizations of simple outliers detection tools, like the boxplot or the quantile-quantile plot, do not exist.

This paper critically discusses the most promising approaches, known from the literature, to outliers detection appearing in the FDA and focuses on selected challenges for the FTS framework. We also present our own proposal related to the generalized Young inverse function. The method is designed for detecting a certain kind of shape outliers. In an empirical example we study day and night air pollution with PM10 in Katowice in 2016. A reasonable outliers detecting algorithm in this context should indicate anomalies which may give us an insight into relationships between various factors and events influencing air quality and in consequence leading to a better local pro-ecological regulations. The rest of the paper is organized as follows. In Section 2 we list basic types of outliers in the FDS setup and briefly indicate the best detection procedures known from literature. We also discuss their drawbacks and challenges. In Section 3 we introduce our original procedure for detecting of certain kind of shape outliers. In Section 4 selected challenges in the FTS setup are discussed. Section 5 presents a comparison of the procedures in the empirical example. Section 6 consists of some conclusions and summary.

2 Outliers in functional data analysis setup

Within the FDA, functional data are considered as sample observations of random functions, i.e., random elements of certain function space. In general, such spaces are typically considered to be infinite dimensional, real and separable Banach or Hilbert spaces (Horváth and Kokoszka, 2012). In empirical studies we observe discrete noisy data, often appearing in unequally spaced time points. A first step in the analysis is to transform these data into

regular, smooth functions (see Horváth and Kokoszka, 2012, and references therein). Another important task that usually occurs during the pre-processing stage is the analysis of variability and splitting it into what is called *phase variability* and *amplitude variability*. The distinction among these two kinds of variability is a significant feature of FDA, as it does not have any counterpart in univariate or multivariate analysis (see Tarabelloni, 2017, and references therein). Referring to these two kinds of variability, one defines two types of outliers namely *shape outliers* and *magnitude outliers*, having in general different effect on the FDA procedures. The first taxonomy of functional outliers was proposed in Hubert et al. (2015). Although a formal and commonly accepted definition of functional outlier is still missing in the literature, a discrimination between amplitude and phase variability inspired the main and widely accepted distinction between the two kinds of outliers, i.e. magnitude and shape outliers. The first are related to amplitude, and are a direct analogue of the outlyingness concept in the multivariate context, while the second are related to phase variability, hence they are completely new and do not have a direct counterpart in classical statistics. The different nature of such outliers motivates researchers to propose various tools to detect and handle with them.

The most popular method designed for shape outliers detection is based on a concept of outliergram proposed in Arribas-Gil and Romo (2014) and intensively studied, for example, in Tarabelloni (2017). The method is rooted in a certain interesting property of the modified band depth of functional observations (López-Pintado and Romo, 2009). The best method of magnitude outliers detection is based on the functional boxplot proposed in Sun and Genton (2011), and further improved in Tarabelloni (2017). It is worth stressing that in practice a dataset may be affected by both magnitude and shape outliers, and a first step of an analysis is to separate them.

The boxplot, a well-known one-dimensional visualization technique, is used for detecting outliers when assuming normality. Under normality, within an interval [Q1-F*IQR;Q3+F*IQR], where Q1,Q3,IQR are correspondingly the first quartile, the third quartile and the interquartile range, and for F=1.5 we have about 0.99 probability mass. Points outside that interval are treated as outliers. Clearly, atypical observations can be either genuine but rare outcomes of the random process generating data, or can be corrupted data, due to a possible contamination of the sample. In a relation to this idea in a functional case, Sun and Genton (2011) proposed to assume a Gaussian process and choose F so that only fraction of 1% of the functions was flagged as outliers. Unfortunately, as in the functional case an analytic expression for F cannot be directly derived, a relevant resampling procedure

must be used. Since Gaussian functional data are far more complex than standard normal random variables, the procedure should be designed to take into account the first and the second moment of the dataset, i.e., the mean and covariance functions. Therefore, the adjustment process must be data-driven and has to be repeated for each dataset. Tarabelloni (2017) has recently proposed to use certain robust estimators and then to use bootstrap method for estimating appropriate quantiles and value for F. His method has been supported with a free roahd R package.

3 Our proposal using boxplot for Young inverse functions

Although the outliergram indicates well outliers belonging to a rich family of shape outliers, it lacks a sufficient precision of identification of type of departure of observation from a majority of observations, as it has been pointed out in Nagy et al. (2017). The sensitivity for types of outlyigness is especially important for developing new economic theories and applications, i.e., the outlying trajectory of a country development may be treated as incentive for developing a more general development theory that explains the root-cause of that trajectory. This drawback of the outliergram motivates us to propose an original method underlying a certain kind of shape "outlyingness in variation", frequently studied in economics (business cycle analysis). We aim at proposing a method, which is less computationally sophisticated than methods proposed in Nagy et al. (2017).

Let $\phi:[a,b]\to[c,d]$ be any real function (called a parent function). A function $\Phi:[c,d]\to[a,b]$ defined for any number $y\in[c,d]$ with a formula

$$\Phi(y)=\min\{x\in[a,b]:\phi(x)\geq y\}$$

we call generalized Young inverse function of function ϕ.

Let a functional data ϕ_j be defined on a finite number of knots, i.e.,

$$\phi_j:\{x_1,\ldots,x_{j_i}\}\to\{y_1,\ldots,y_{j_m}\}.$$

Therefore, an empirical generalized Young inverse function ϕ_j, namely, $\Phi_j:\left[\min\{y_1,\ldots,y_{j_m}\},\max\{y_1,\ldots,y_{j_m}\}\right]\to\{x_1,\ldots,x_{j_i}\}$ can be effectively defined for any real number $y\in\left[\min\{y_1,\ldots,y_{j_m}\},\max\{y_1,\ldots,y_{j_m}\}\right]$ with a formula:

$$\Phi(y)=\min\left\{x\in\{x_1,\ldots,x_{j_i}\}:\phi_j(x)\geq y\right\}.$$

Furthermore, we have a functional sample $Y^n=\{y_1,...,y_n\}$. We would like to show that a mapping of the following form

$$D(y \mid y_1,...,y_n)(z) = \{z \in [c,d] : \min_{i=1,...,n} y_i^{-1} \le y^{-1}(z) \le \max_{i=1,...,n} y_i^{-1}\},$$

where y^{-1} denotes generalized Young inverse function of function y, is a functional depth of a function y with respect to the sample Y^n.

An illustration of Young inverse functions is presented in Fig. 1-2.

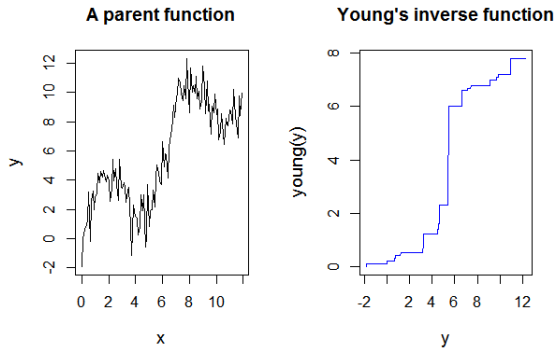


Fig. 1. An illustration of Young inverse function.

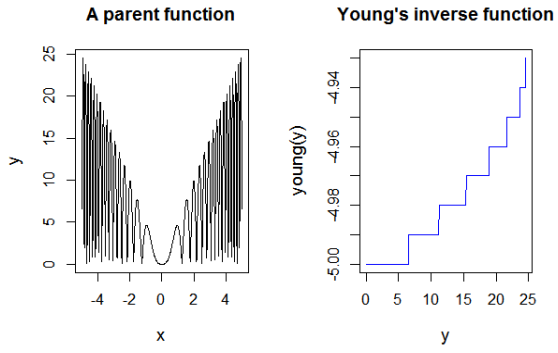


Fig. 2. An illustration of Young inverse function.

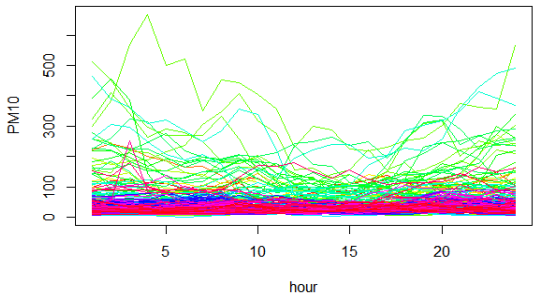


Fig. 3. Raw data on day and night air pollution with PM10 in Katowice in 2016.

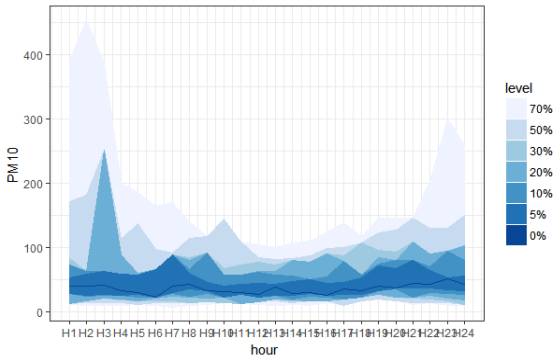


Fig. 4. Functional boxplot for air pollution data with PM10 in Katowice in 2016.

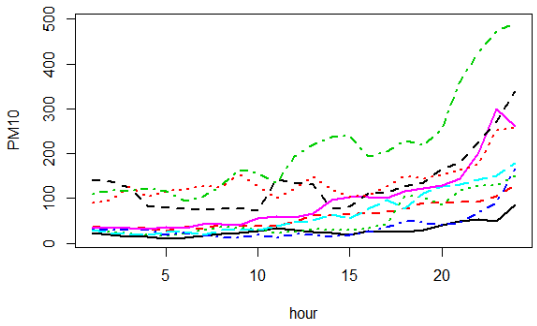


Fig. 5. Outlying day and night PM10 pollution curves indicated by Young inverse function method.

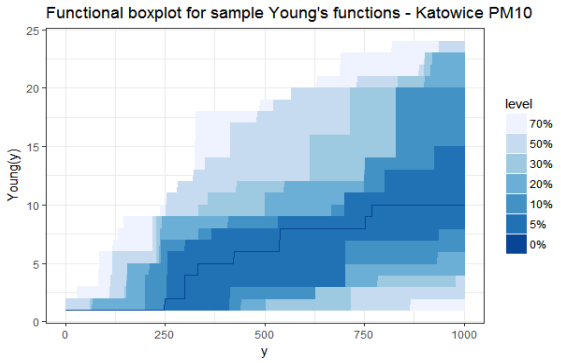


Fig. 6. Functional boxplot for Young inverse functions of air pollution curves in Katowice in 2016.

PROPOSAL: Let *MBD* denote a sample modified band depth and let *FM* denote a sample Frainman and Muniz depth (see López-Pintado and Romo, 2009). For a sample of functions $Y^n = \{y_1,...,y_n\}$ take the following steps (we fix thresholds of a_{MBD}, a_{Young}):

1. Verify the null hypothesis that Y^n has been generated by Gaussian law using statistic T_4 of Jarque-Berra type test for functional data proposed in Górecki et al. (2017). If the null hypothesis is rejected, then go to step 2, if it is not rejected, then go to step 3.

2. Calculate Tarabelloni (2017) version of the adjusted boxplot, and place outliers into a set O^T, indicating amplitude outliers.

3. Calculate the functional boxplot using the *MBD*, and place functions with $MBD < a_{MBD}$ into a set O^{MBD}, indicating amplitude outliers.

4. Calculate the outliergram, and place outliers into a set O^{AG}, indicating ”roughly” shape outliers.

5. Calculate the generalized Young inverse functions for the sample, i.e., $\left(Y^n\right)^{-1}$ and then calculate the functional boxplot for $\left(Y^n\right)^{-1}$. Place observations with *FM* depth $< a_{Young}$ into a set O^{Young}, indicating “outliers in variation”.

The final outliers set is $(Y^n \setminus (O^T \check{C} O^{MBD})) \text{ Ç } O^{AG} \text{ Ç } O^{Young}$.

Notice that one of the difficulties, when using a functional boxplot for outlier detection, is to indicate a number c such that the following formula is fulfilled: $P(MBD(y,Y^n) < c) = \alpha_{MBD}$,

where P denotes probability. In order to indicate a number c, we have to know P, which is a sampling distribution of functional depth and it is generally an unknown object.

4 Challenges in functional time series setup

Despite of many scientists efforts, robust statistics in area of time series analysis has much more gaps in comparison to uni- and multivariate statistics. The eighth chapter of the influential book by Maronna et al. (2006) may be treated as an introduction in this context. Outliers may cause bias in the model parameter estimates, and then, distort the size and power of statistical tests based on biased estimates. Secondly, outliers may increase the confidence intervals for the model parameters. Thirdly, as a consequence of the previous points, outliers strongly influence predictions. The best developed theory is available for ARIMA models. However, recent developments involve advances in coping with outliers when the series is generated by a general nonlinear models including as particular cases the bilinear model, the self-exciting threshold autoregressive (SETAR) model, the exponential autoregressive model, and the generalized autoregressive conditional heteroscedasticity (GARCH) models. Outliers in multivariate time series have been much less analysed than in the univariate case. For more recent researches related to our studies, let us only indicate that multivariate outliers were introduced in Tsay et al. (2000). Galeano et al. (2006) used projection pursuit methods to develop a procedure for detecting outliers, showing in particular, that testing for outliers in certain projection directions can be more powerful than testing the multivariate series directly. In view of these findings, an iterative procedure to detect and handle multiple outliers (based on a univariate search in these optimal directions) was proposed. The main advantage of this procedure is identification of outliers without prespecifying a vector ARMA model for the data. These ideas are in a close relation to the typical FDA framework, where due to infinite dimensional nature of the objects and richness of possible models, investigation of appropriate projections of data is a natural research tool. Note only, that even in case of outliers in an ARCH(1) model, we encounter two different scenarios, because an outlier can affect the level of a process as well as the volatility of the process. The case of a functional ARCH(1) model introduces the next level of complications.

In an analogy to the one dimensional setup (Maronna et al., 2006), one may introduce functional additive outliers (FAO). The FAO may have adverse influence on all the steps of the time series analysis. Functional innovative outliers (FIO) affect a single innovation, and correspond to an internal change of a single innovation of the time series and are often associated with isolated incidents. The effects of an FIO on a series are different depending on

whether the series is "stationary" or not (stationarity in the FTS setup is still an open issue). Maronna et al. (2006) introduced also functional level shift (FLS) which is a change in the level of the series, and the functional temporary change (FTC) which is an exponentially decreasing change in the level of the series. This shows a great number of possible functional outliers definitions.

The research challenge here is that the effect of a functional outlier not only depends on the model and the outlier size, as in the univariate case, on the interaction between the model and size as in multivariate case, but also on the outlier type. Additionally, as in the classical time series setup, the masking and swamping effects are still present, if a sequence of outlier patches is present in the functional time series (Galeano et al., 2006), i.e., a few outliers of the same magnitude appearing one after another (and forming a path).

5 Empirical study

The first application of functional outliers detection tools to air quality monitoring may be found in Febrero et al. (2008). In our empirical studies we, among others, have studied day and night air pollution with PM10 in Katowice in a period 01.09.2016 to 28.02.2017, using the data from the automatic measurement station located on Kossuth's street in Katowice and available from WIOŚ (http://powietrze.katowice.wios.gov.pl). The dataset consisted of 181 curves. Fig. 3 presents raw data on day and night air pollution with PM10 in Katowice in 2016. The same dataset has been exploited in the paper by Kosiorowski et al. (2018), where the authors have analyzed PM10 concentration in the air for five measurement stations in Silesian Region. The air pollution in the Silesian Region has been treated as a realization of a hierarchical functional random variable. Double functional median method has been used to compute a robust aggregate representing air pollution in the Silesian Region.

Fig. 4 presents the functional boxplot for air pollution data with PM10 in Katowice in 2016. Fig. 5 presents the outlying day and night PM10 pollution curves indicated by the generalized Young inverse function method and Fig. 6 shows the functional boxplot for the generalized Young inverse functions of air pollution curves in Katowice in 2016. Fig. 7 presents an adjusted functional boxplot, and air pollution data. Fig. 8 presents outliergram, and shape outliers indicated by it in the air pollution data. The adjusted functional boxplots and the outliergram were prepared using roahd R package (Tarabelloni, 2017), whereas functional boxplots for raw data and the generalized Young curves using DepthProc R package (Kosiorowski and Zawadzki, 2017).

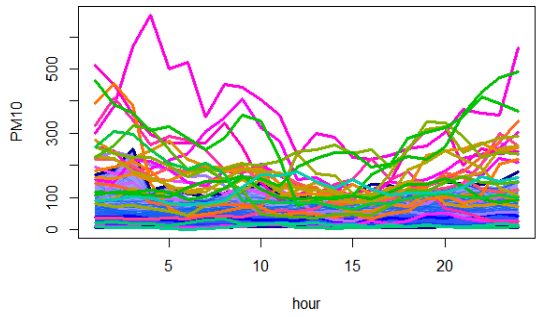


Fig. 7. Adjusted functional boxplot, air pollution data.

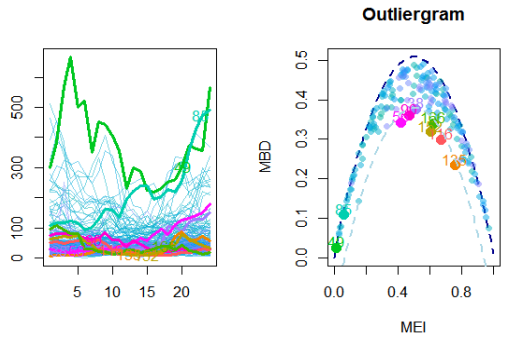


Fig. 8. Outliergram and outliers in air pollution data.

Conclusions

Reliable algorithms of functional outliers detection are desirable in a context of analysing various anomalies in economic phenomena described by functions. In the paper we have proposed a new functional outliers method and incorporated it into a decision procedure, which uses a known outlier detection apparatus i.e., the outliergram, the functional boxplot and the adjusted functional boxplot. Our approach is recommended for "outliers in variation" detection. Our future studies involve developing a theory for the proposed procedure as well as applications in optimization of a local pro-ecological policy.

Acknowledgements

DK thanks for the support related to CUE grant for the research resources preservation 2018. JPR and DM work was partially supported by the Faculty of Applied Mathematics AGH UST statutory tasks within subsidy of Ministry of Science and Higher Education, grant no. 11.11.420.004.

References

Arribas-Gil, A. & Romo J. (2014). Shape Outlier Detection and Visualization for Functional Data: the Outliergram. *Biostatistics*, *15*(4), 603-619.

Cuevas, A., Febrero, M. & Fraiman, R. (2006). On the Use of the Bootstrap for Estimating Functions with Functional Data. *Computational Statistics & Data Analysis*, *51*(2), 1063-1074.

Febrero, M., Galeano, P. & González-Manteiga, W. (2008). Outlier detection in functional data by depth measures, with application to identify abnormal NOx levels. *Environmetrics*, *19*: 331-345.

Galeano, P., Peña, D. & Tsay, R. S. (2006). Outlier detection in multivariate time series by projection pursuit. *Journal of the American Statistical Association, 101*(474), 654-669.

Górecki, T., Hörmann, S., Horváth, L. & Kokoszka, P. (2017). Testing Normality of Functional Time Series. *Journal of Time Series Analysis*. Doi:10.1111/jtsa.12281.

Horváth, L. & Kokoszka P. (2012). *Inference for Functional Data with Applications*. Springer Verlag, New York.

Hubert, M., Rousseeuw, P. & Segaert, P. (2015). Multivariate Functional Outlier Detection. *Statistical Methods and Applications, 24*(2), 177-202.

Kosiorowski, D. & Zawadzki, Z. (2017). DepthProc An R Package for Robust Exploration of Multidimensional Economic Phenomena, arXiv:1408.4542.

Kosiorowski, D., Mielczarek, D. & Rydlewski, J. P. (2018). Forecasting of a Hierarchical Functional Time Series on Example of Macromodel for the Day and Night Air Pollution in Silesia Region - A Critical Overview. *Central European Journal of Economic Modelling and Econometrics, 10*(1), 26-46.

López-Pintado, S. & Romo J. (2009). On the Concept of Depth for Functional Data. *Journal of the American Statistical Association, 104*(486), 718-734.

Maronna, R.A., Martin, R. D. & Yohai, V. J. (2006). *Robust Statistics - Theory and Methods.* John Wiley & Sons, Chichester.

Nagy, S., Gijbels I. & Hlubinka D. (2017). Depth-Based Recognition of Shape Outlying Functions. *Journal of Computational and Graphical Statistics, 26*(4), 883-893.

Sun, Y. & Genton M. (2011). Functional Boxplots. *Journal of Computational and Graphical Statistics, 20*(2), 316-334.

Tarabelloni, N., (2017). *Robust Statistical Methods in Functional Data Analysis*. Phd thesis and roahd R package, Politecnico di Milano.

Tsay, R.S., Peña, D. & Pankratz A.E. (2000). Outliers in Multivariate Time Series. *Biometrika, 87*(4), 789-804.

The logistic regression in predicting spike occurrences in electricity prices

Jadwiga Kostrzewska[1], Maciej Kostrzewski[2]

Abstract

Electricity supply and demand are subject to weather conditions (temperature, wind speed, precipitation) as well as daily, weekly or yearly seasonality due to e.g. an intensity of business activities. These features have a significant impact on the market and price behaviour. As a result of the lack of storage capacity sharp movements of electricity prices are often observed. An ability of modelling and forecasting jumps and spikes plays the crucial role in risk management. In the paper, the logistic regression is employed to predict spike occurrences. We investigate the impact of fundamental variables such as demand, weather and seasonal factors, on spikes occurrences. The point and interval theoretical probabilities are calculated. The classification accuracy is assessed by means of the *sensitivity*, *specificity*, *accuracy* and *AUC* measures. In our research we detect spikes using a quantile technique and a Bayesian *DEJD* model. We state that the logistic regression is a quite good tool to forecast moments of a spike occurrence. The logistic regression model is a well-known specification which seems to be reasonable tool of spike prediction.

Keywords: *electricity price, jumps, spike prediction, logistic regression, DEJD*
JEL Classification: C22, C53, Q02, Q41, Q47
DOI: 10.14659/SEMF.2018.01.22

1 Introduction

Energy markets are different from other commodity and financial markets due to the difficulty in storing large quantities of electricity. At the same time, power system stability requires a constant balance between production and consumption. Electricity supply and demand are subject to weather conditions (temperature, wind speed, precipitation, solar radiation), daily, weekly or yearly seasonality due to e.g. an intensity of business activities (working hours, peak hours, weekdays, holidays, near holidays). These features have a significant impact on the market and price behaviour. They might result in extreme spot price volatility and may in consequence bring the existence of sharp price movements, i.e. spikes and jumps. An ability of modelling and forecasting them plays the crucial role in electricity price forecasting and risk management.

In comparison to the extensive literature devoted to forecasting electricity prices, relatively less attention is paid to forecasting price spikes (see e.g. Weron, 2014). One of the

[1] Corresponding author: Cracow University of Economics, Department of Statistics, Rakowicka 27, 31-510 Kraków, Poland, e-mail: jadwiga.kostrzewska@uek.krakow.pl.
[2] Cracow University of Economics, Department of Econometrics and Operations Research, Rakowicka 27, 31-510 Kraków, Poland, e-mail: maciej.kostrzewski@uek.krakow.pl.

first work about forecasting spikes is Christensen's et al. (2012) paper, where the ACH model and logit model are employed. Eichler et al. (2012, 2014) compare the ACH and the logit model indicating the advantage of the latter.

The main objective of the article is the forecasting of electricity price spikes by means of the logistic regression. We investigate the impact of fundamental variables such as demand, weather and seasonal factors on spike occurrences, and determine point forecasts and 95% prediction intervals for a spike occurrence probability. The analysis is conducted for a standard level of a cut point (0.5) as well as a cut point calculated on the basis of the sensitivity-specificity plot. The quality of the classification is examined by means of the *accuracy*, *sensitivity*, *specificity* and the *AUC* measures.

2 Electricity spot prices and spikes

The term of a spike is in common use, but there is no unique and broadly accepted definition of it. In a certain simplification sudden and sharp movements of values in time series may be treated as spikes. While analysing a given time series we can easily discern the values which are spikes 'for sure' on account of their distinct outlying with respect to other observations. However, at the same time we do not know how to classify the remaining data points at hand – that is where to put the border line between observations which should and the ones which should not be classified as spikes. There are various methods of detecting spikes. Values surpassing a fixed threshold (see e.g. Christensen et al., 2012) or a threshold identified by some method or model are frequently classified as spikes. Actually, different methods lead to different moments of spike occurrences (see e.g. Janczura et al., 2013).

The research is based on the series of hourly electricity prices on Nord Pool market. Nord Pool is a leading power market and the largest one in Europe. About 380 members from 20 countries are active at Nord Pool. We focus on a day-ahead market (*Elspot*) on which 98% of electricity volume handled by the Nord Pool is traded. We model the hourly system price (EUR/MWh) which is an unconstrained market clearing reference price. Most standard financial contracts traded in the Nordic region use the system price as the reference price. At 12.00 a.m. all purchase and sell orders are aggregated into two curves for each delivery hour, and then the system prices are calculated for each delivery day-ahead hour.

In our research, the period of the analysis ranges from 2014/12/29 to 2017/07/02 and is divided into an in-sample (2014/12/29–2016/09/11) and an out-of-sample (2016/09/12–2017/07/02) period. The series of spot electricity prices is very volatile and 'spiky'. The highest prices are more than four times greater than the median price over analysed period.

Moreover, the prices are subject to intra-daily (24-hourly) and intra-weekly seasonality (see Fig. 1): higher electricity prices appear during on-peak hours, and lower prices during off-peak hours as well as during weekends.

Notably due to the lack of storage capacity sharp movements of prices are very frequent and violent on energy market. The power production, and hence the electricity price, is related to weather conditions. There are some seasonality patterns, especially in hydropower, but also in wind power production. The power consumption is intensified during business hours (on-peak hours) and also depends on the weather conditions (temperature, seasons). All of these factors may impact on prices and their abnormal sharp movements i.e. spike occurrences.

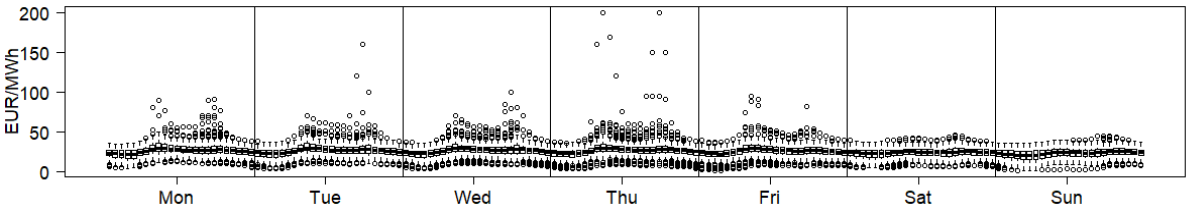


Fig. 1. The boxplots for electricity prices for each hour of a week: 2014/12/29–2017/07/02.

3 Applied methods and empirical results

In our research we adopt two methods of spike detection. The first technique identifies 2.5% of the highest values as spikes – likewise the variable price threshold method in (Janczura et al., 2013) or the quantile analysis in (Kostrzewska et al., 2016). We refer to it as the *QUA* technique. The second one is based on the Bayesian *DEJD* model considered by Kostrzewski (2015). We concentrate on the upwards spikes. Before applying the spike detection techniques we pre-process the series of the spot electricity prices as follows. Firstly, we get rid of the long-term seasonality component (LTSC, seasonal patterns in seasons, months etc.) by means of the Hodrick-Prescott filter[3], and thereafter, the short-term seasonal component (STSC, intra-weekly and intra-daily seasonal patterns) by means of the filter based on medians of the prices for each hour of a week[4]. Ultimately, we identify spikes in the remaining irregular component i.e. after applying the HP filter and the median STSC technique.

[3] Weron and Zator (2015) proposed using the HP filter for identifying the LTSC in electricity spot prices, which is less computationally complex and gives similar results to the wavelet technique.

[4] See (Janczura et al., 2013) for details about the filter based on the means. We apply more robust medians instead of the means.

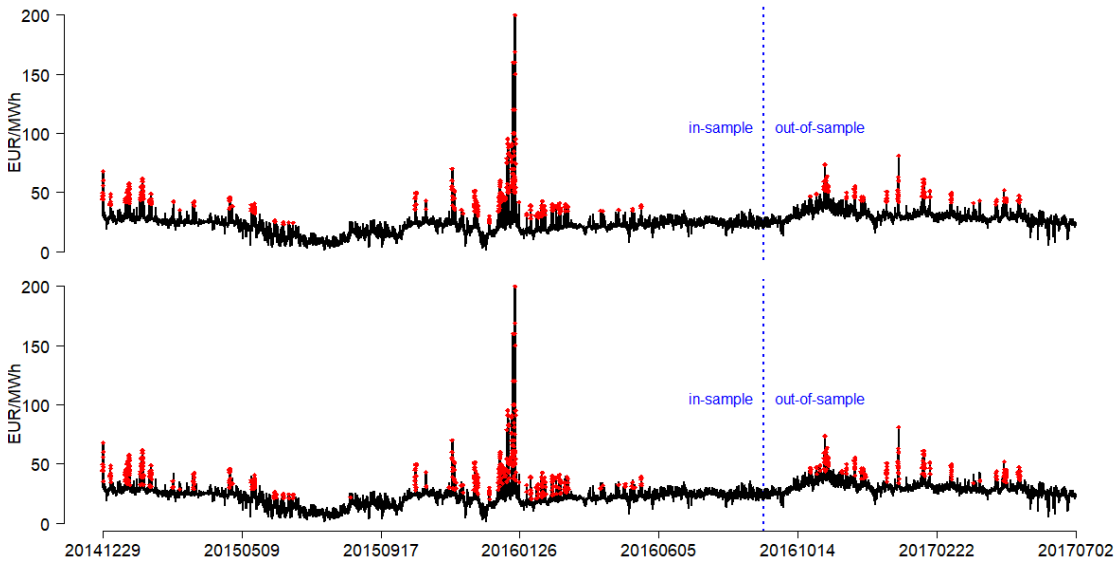


Fig. 2. The series of the hourly spot electricity prices with spikes (red) identified by the *QUA* (top) and the *DEJD* technique (bottom) over the period 2014/12/29–2017/07/02.

The electricity price spikes detected by means of the quantile technique *QUA* (2.5%) and the *DEJD* model (3.04%) are presented in Fig. 2. Not all spikes identified by the *QUA* are spikes according to the *DEJD*. Applying the HP and median filters enables to detect spikes in electricity prices regardless of the seasonal nature of prices. The blue dotted line indicates the border between in-sample and out-of-sample periods used in the paper.

Next, we employ the logistic regression in order to forecast moments of the spot electricity price spike occurrences. The dependent variable distinguishes between spikes (1's) identified by means of the one of the mentioned techniques and 'ordinary' values (0's). The in-sample period covers the period from 2014/12/29 to 2016/09/11 (14,952 hourly observations). We apply the extending window algorithm by 24 hours at a time with model re-estimation. Initially, we estimate the model and calculate the first day-ahead forecast covering 24 hours of a next day. Secondly, we extend the in-sample data set by 24 hours, estimate the model and calculate the next 24-hourly day-ahead forecast, and so on. The forecast is conducted in the out-of-sample period (2016/09/12–2017/07/02, 7,056 hourly observations), the forecast horizon covers 294 days and 294 logistic models are estimated.

In the logistic regression model, following exogenous variables are considered[5]: six dummy variables indicating days in a week (except Saturday – a reference day), wintertime (*winter*, dummy variable), on-peak hours (#8–#20) (dummy variable *peak*), electricity

[5] The variables *cons*, *minprice*, *lagprice*, *wind* and *hydro* are transformed so that the LTSC and STSC or only the LTSC is filtered out.

consumption forecasts (*cons*, MWh), wind power forecasts (*wind*, MWh), reservoir water levels (*hydro*, GWh), the minimum of the day-before hourly prices (*minp*, EUR/MWh), lagged by 48 hour prices (*lagp*, EUR/MWh), lagged by 48 hours failures of power plants (*fail*). The stepwise method is applied to select a final set of exogenous variables. Table 1 shows how frequently each variable occurs in the logistic regression models. The value '1' means that the variable is present in each of the 294 models, '0' – in none of them.

For each spike detection technique the variables such as: *Mon–Fri*, *peak*, *cons*, *wind* and *minp* explain the moments of spike occurrences for all logistic regression models. The results depict that only the dummy variable *Sun* does not explain the spike occurrences (with Saturdays as the reference). That means spikes on Saturdays behave almost the same as on Sundays, but differently than on the other days. The variable *winter* is present in 95.6% of 294 models (*QUA*) or in all models (*DEJD*), thus a distinction between winter and summer seasons is important in the majority of the models. The same conclusion can be made for the lagged by 48 hours failures of power plants (*fail*). The reservoir water level (*hydro*) is not so important and stays only in 12.9% (*QUA*) or 70.7% (*DEJD*) of 294 models. On the other hand, lagged by 48 hour prices (*lagp*) are important for all models in the case of the spike identification by the *QUA* and none of the models in the case of the *DEJD*.

Table 1. Frequency of an occurrence of each variable among 294 logistic regression models – spikes are identified by the quantile analysis (*QUA*) or the *DEJD* model.

	Mon	***Tue***	***Wed***	***Thu***	***Fri***	***Sun***	***peak***	***winter***	***cons***	***wind***	***hydro***	***minp***	***lagp***	***fail***
QUA	1	1	1	1	1	0	1	0.956	1	1	0.129	1	1	1
DEJD	1	1	1	1	1	0	1	1	1	1	0.707	1	0	0.980

We calculate the theoretical (point) probability and 95% prediction intervals for a probability of a spike occurrence for each hour (see Fig. 4-5). As mentioned before, the analysis is conducted for the standard level of a cut point (0.5) and the cut point calculated on the basis of the sensitivity-specificity plot (see Hilbe, 2016), named as the s-s cut point later in the paper. Fig. 3 presents the values of the s-s cut point for each of 294 estimated models. The values vary for individual models, and in the *QUA* approach are lesser than in the *DEJD*. In addition, we can compare the s-s cut point with the one calculated as the mean of the predicted values, which is often used in practice if the dependent variable has substantially more or less 1's than 0's as mentioned by Hilbe (2016). The average value of the forecasted probabilities of a spike occurrence is equal to 0.030 (*QUA*) or 0.037 (*DEJD*), which indicates

that the s-s cut point based on the sensitivity-specificity plot may be more appropriate than the standard cut point equal to 0.5.

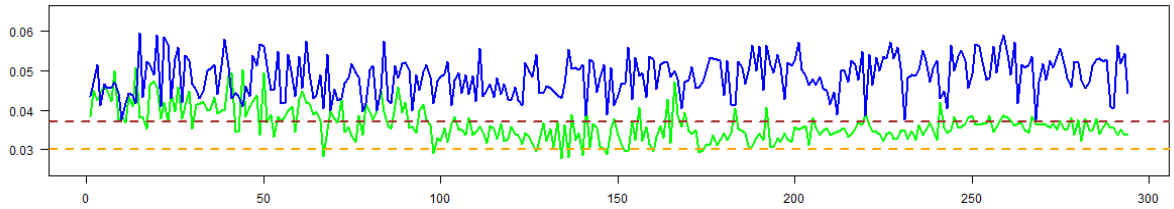


Fig. 3. The values of the s-s cut point (green – the *QUA*, blue – the *DEJD*) for each of 294 logistic regression models and the average predictive values of probability of spike occurrence (orange dashed line – the *QUA*, brown dashed line – the *DEJD*).

The quality of the classification is assessed by means of the *sensitivity*, *specificity*, *accuracy* and *AUC* measures over the in-sample and out-of-sample periods, separately, in the case of the cut point equal to 0.5 and the s-s cut point (see Table 2). The *sensitivity* is a percentage of correctly classified spikes among all spikes that occurred (i.e. identified by the method). The *specificity* is a percentage of correctly classified 'ordinary' prices among all 'ordinary' prices that occurred. The *accuracy* is a percentage of correctly classified spikes and 'ordinary' prices together. The *AUC* is an area under the receiver operator characteristic (*ROC*) curve.

Table 2. The in-sample and out-of-sample assessment of the quality of the classification: the average *sensitivity*, *specificity*, *accuracy* and *AUC* measures for 294 logistic regression models with the cut point equal to 0.5 and the s-s cut point – the *QUA* or *DEJD* approach.

		In-sample				**Out-of-sample**			
	Cut point:	***Sens***	***Spec***	***Accur***	***AUC***	***Sens***	***Spec***	***Accur***	***AUC***
QUA	0.5	55.5	**99.5**	**98.3**	77.5	57.4	**99.0**	**98.2**	78.2
	s-s	**95.5**	93.5	93.5	**94.5**	**96.3**	90.7	90.8	**93.5**
DEJD	0.5	48.7	**99.4**	**97.8**	74.0	56.2	**99.0**	**97.9**	77.6
	s-s	**90.8**	92.4	92.4	**91.6**	**94.6**	89.5	89.7	**92.1**

On the basis of the results reported in Table 2 we conclude that the assessment of the models by means of the *sensitivity* and *AUC* measures is higher when using the s-s cut point

than the cut point equal to 0.5, for both the *QUA* as well as the *DEJD* technique. This is due to the percentage of spikes identified by the method (2.5% – *QUA*, 3.04% – *DEJD*) is closer to the mean optimal cut point (0.036 – *QUA*, 0.049 – *DEJD*) than to the value 0.5. However, we note the lower values of the *specificity* and *accuracy* measures when the s-s cut point is employed, but still suitably high. This may be caused by the higher number of 'false alarms' (incorrectly predicted spikes) when using the lower cut point.

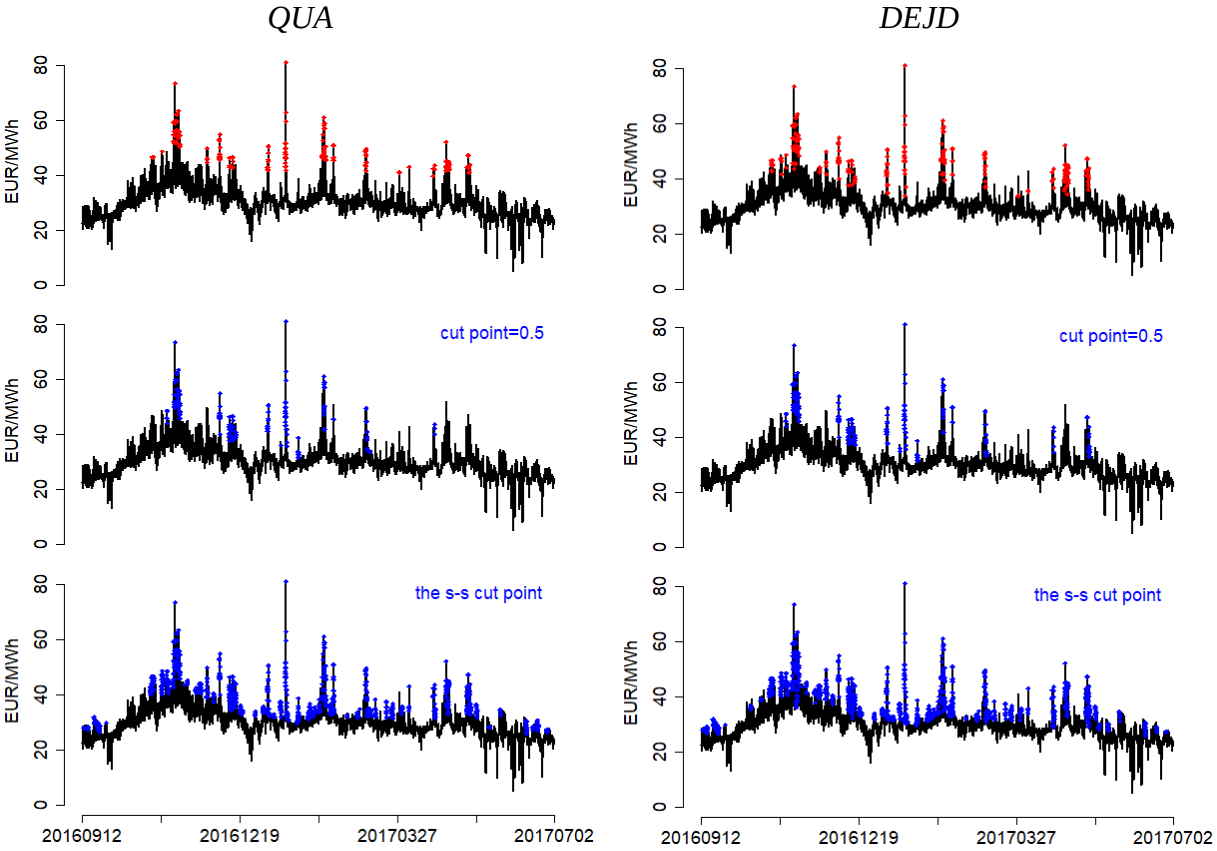


Fig. 4. The electricity price series over the out-of-sample period with moments of spike occurrences (red) detected by means of the *QUA* (top left) or the *DEJD* (top right) technique and the forecasted spikes (blue) with the 0.5 cut point (middle left and right) or the s-s cut point (bottom left and right).

Fig. 4 presents the moments of spike occurrences identified by means of the quantile technique and two variants of spike forecasts: using the cut point equal to 0.5 or the s-s cut points calculated for 294 models. The visual analysis indicates that the spike forecast with the s-s cut point is more reasonable than with the cut point equal to 0.5, although in the first case there are more 'false alarms' than in the later one. Such forecasts might be interesting for retailers who prefer to hedge on the financial market due to the 'false alarm' than set loss

down. The conclusions coincide with those drawn from the analysis of the quality of the classification made before.

In the research, we also determine the 95% prediction intervals for the probability of the spike occurrence for each hour (see Fig. 5). The prediction intervals enable to define forecast rules in a more or less restrictive way. In the more restrictive approach, the spike occurrence is forecasted if the entire prediction interval is above the cut point. In that case the number of 'false alarms' is reduced in comparison to the forecasts based on the point theoretical probabilities. On the other hand, in the less restrictive approach, the spike occurrence is forecasted when only the upper bound of the prediction interval is above the cut point. We adopt both approaches and calculate the *sensitivity*, *specificity*, *accuracy* and *AUC* measures with different cut points in order to compare the results with those in the case of the point forecasts of spike occurrences (see Table 3).

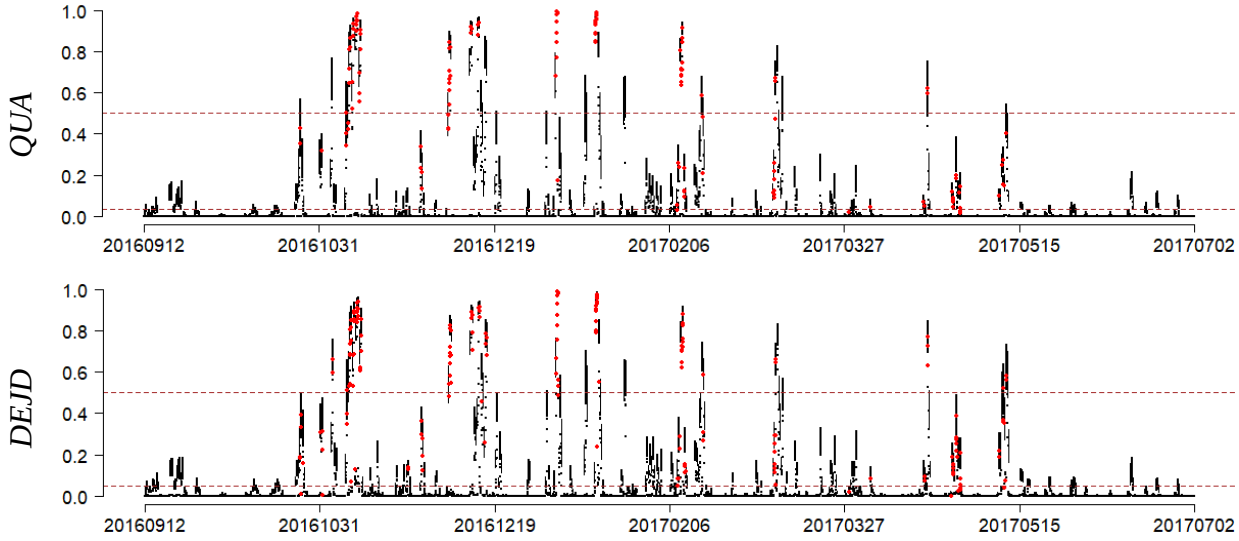


Fig. 5. The 95% prediction intervals of spike occurrences (black) constructed on the basis of the logistic regression. Horizontal dashed lines indicate the average s-s and 0.5 cut point. The theoretical probabilities for spikes detected by means of the *QUA* (top) or the *DEJD* technique (bottom) are depicted in red. If the red points are above the dashed line the spikes predicted by means of the logistic model coincide with those detected by means of the *QUA* or *DEJD*.

The results in Table 3 indicate that the highest values of the *sensitivity* and *AUC* measures are observed for the less restrictive rule of spike forecasts (upper bound), next for the forecasts based on the point theoretical probability and the lowest values – for the more restrictive approach (entire interval). On the other hand, the highest values of the *specificity*

and *accuracy* measures are noted for the more restrictive approach, then for the point forecasts and for the less restrictive approach – the lowest. That is, the less restrictive rule of spike forecasting brings a higher percentage of correctly predicted spikes in combination with a higher percentage of 'false alarms', and conversely. It is worth mentioning that for all approaches the AUC measure is above 90% with the s-s cut point and above 75% with the cut point equal to 0.5. These results indicate that the logistic regression model is a good tool to classify and forecast spike occurrences in the spot electricity prices.

Table 3. Out-of-sample assessment of the quality of the classification: the average *sensitivity*, *specificity*, *accuracy* and *AUC* measures for 294 logistic regression models with the cut point equal to 0.5 and the average s-s cut point – the results for the more restrictive approach (Upper), point forecast (Point) or less restrictive approach (Entire).

	QUA						***DEJD***					
Cut	**Upper**		**Point**		**Entire**		**Upper**		**Point**		**Entire**	
point:	**s-s**	**0.5**	**s-s**	**0.5**	**s-s**	**0.5**	**s-s**	**0.5**	**s-s**	**0.5**	**s-s**	**0.5**
Sens	98.5	64.7	96.3	57.4	91.2	52.2	96.2	58.9	94.6	56.2	90.3	48.1
Spec	88.9	98.6	90.7	99.0	93.4	99.4	87.3	98.6	89.5	99.0	92.4	99.4
Accur	89.1	97.9	90.8	98.2	93.3	98.4	87.5	97.5	89.7	97.9	92.3	98.1
AUC	93.7	81.7	93.5	78.2	92.3	75.8	91.8	78.7	92.1	77.6	91.3	73.8

Conclusions

In the research, the time series of electricity prices is an imbalanced dataset – there are more 'ordinary' values than the spikes. In consequence, adoption the standard cut point equal to 0.5 leads to worse results than the s-s cut point. The value of the s-s cut point is similar to the percentage of observations identified as spikes by means of the quantile (*QUA*) as well as the *DEJD* technique. The more restrictive approach based on 95% prediction intervals forecasts correctly less 'real' (i.e. identified by the method) spikes and less 'false alarms' in comparison with the forecast based on the point theoretical probability of spike occurrences. On the other hand, the less restrictive approach predicts correctly more spikes and more 'false alarms' than the point forecast method.

The exogenous variables have strong impact on the spikes prediction. We consider two variables corresponding to renewable energy sources – the wind power forecasts and the reservoir water levels. The wind power forecasts are important for all models. However, the reservoir water levels are employed only in 38 (in the case of the *QUA* technique) and 208

(the *DEJD*) out of 294 models. It might be explained by lower volatility of the latter variable which is constant during each week. Moreover, we would like to note that the failures of power plants variable is employed for almost all logistic regression models.

The logistic regression model is a well-known specification which seems to be reasonable tool of spike prediction. In our future research we are going to employ and compare other methods of spike detection and prediction.

Acknowledgements

The research is funded by the National Science Centre, Poland (grant no. 2016/23/B/HS4/03018).

References

Christensen, T. M., Hurn, A. S. & Lindsay, K. A. (2012). Forecasting spikes in electricity prices. *International Journal of Forecasting, 28*(2), 400-411.

Eichler, M., Grothe, O., Manner, H. & Türk, D. D. T. (2012). Modeling spike occurrences in electricity spot prices for forecasting. *METEOR*, RM/12/029.

Eichler, M., Grothe, O., Manner, H. & Türk, D. D. T. (2014). Models for short-term forecasting of spike occurrences in Australian electricity markets: a comparative study. *Journal of Energy Markets, 7*, 55-81.

Hilbe, J. M. (2016). *Practical guide to logistic regression*. CRC Press.

Janczura, J., Trück, S., Weron, R. & Wolff, R. C. (2013). Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling. *Energy Economics, 38*, 96-110.

Kostrzewska, J., Kostrzewski, M., Pawełek, B. & Gałuszka, K. (2016). The Classical and Bayesian Logistic Regression in the Research on the Financial Standing of Enterprises after Bankruptcy in Poland. In: Papież M., Śmiech S.(red.). *The 10th Professor Aleksander Zeliaś International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings*, 72-81.

Kostrzewski, M. (2015). Bayesian DEJD model and detection of asymmetry in jump sizes. *Central European Journal of Economic Modelling and Econometrics, 7*(1), 43-70.

Weron, R. & Zator, M. (2015). A note on using the Hodrick–Prescott filter in electricity markets. *Energy Economics, 48*, 1-6.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting, 30*(4), 1030-1081.

Standard of living in Poland at regional level - classification with Kohonen self-organizing maps

Marta Kuc[1], Aneta Sobiechowska-Ziegert[2]

Abstract

The standard of living is spatially diversified and its analyzes enable shaping regional policy. Therefore, it is crucial to assess the standard of living and to classify regions due to their standard of living, based on a wide set of determinants. The most common research methods are those based on composite indicators, however, they are not ideal. Among the current critiques moved to the use of composite indicators is the normative nature of the weights, the drawbacks of the linear method of aggregation and the fact that composite indicators flatten alternatives and differences among analyzed objects. That is why more objective alternative solutions are nowadays gaining popularity.

The aim of this article is the regional classification of Polish NUTS-4 regions due to their standard of living using SOM - self-organizing maps - the model of a neural network mapping multidimensional space into a two-dimensional map of neurons. In this study, the set of selected 45 variables describing the standard of living in Polish NUTS-4 regions in 2016 was used.

Keywords: *standard of living, neural networks, Kohonen self-organizing maps, regional analysis*
JEL Classification: C45, I31, R12
DOI: 10.14659/SEMF.2018.01.23

1 Introduction

For many years the standard of living has been the subject of interest not only to scientists but as well to politicians, journalists, and public opinion. As numerous studies show the standard of living is spatially diversified, not only at international level (Wawrzyniak, 2016) but as well at intra-country level (Kuc, 2017). Therefore monitoring the spatial diversity of the standard of living is important from the point of view of sustainable and coherent development, and may be used in shaping the regional policy at both country and the European Union level (Chrzanowska et at., 2017).

The main goal of this research is the implementation of Kohonen self-organizing maps to the regional classification of Polish NUTS-4 regions due to their standard of living in 2016. The analysis covers 314 counties (urban counties were not taken into consideration).

[1]Corresponding author: Gdansk University of Technology, Faculty of Management and Economics, Department of Economic Science, Traugutta 79, 80-233 Gdansk, marta.kuc@zie.pg.gda.pl.
[2]Gdansk University of Technology, Faculty of Management and Economics, Department of Economic Science, Traugutta 79, 80-233 Gdansk, as@zie.pg.gda.pl.

2 Standard of living measurement

The methodology of research on the standard of living has evolved from GDP, through measures 'adjusting' GDP (i.e. MEW, ISEW or Green GDP), measures 'replacing' GDP (i.e. HDI, HPI, The where to be born index or Gallup Index), measures 'supplementing' GDP (i.e. Decoupling Indicators) up to soft modelling, fuzzy analysis and artificial neural networks. Nowadays the most popular way of analyzing standard of living is the one in which author begins by constructing a conceptual scheme of some sort describing his understanding on the standard of living, including its constituents and determinants. There is no one, widely accepted, set of standard of living determinants. Eurofound is currently working on the preparation of a set of indicators that can be used in standard of living and social cohesion analysis, both at international and regional level. Based on the literature review one can prepare the following Mandala of the standard of living (Fig. 1). Where the standard of living is the core of Mandala and three circles represent resources shaping standard of living.

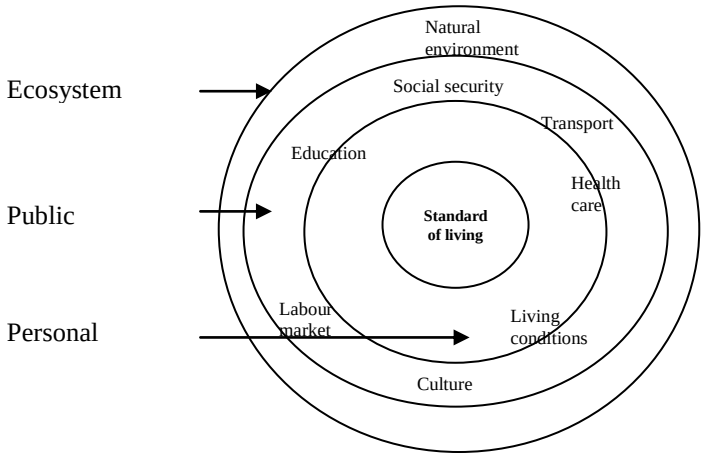


Fig. 1. The Mandala of Standard of Living.

Source: Authors' own study based on Goossens ed. (2007) and Michalos et al. (2011).

3 Pros and cons of composite indicators

Nowadays composite indicators are not only used in the standard of living and well-being analysis (Becker et al., 2017). They are also used in analysis of countries' competitiveness (Kruk and Waśniewska, 2017), socioeconomic development (Sobiechowska-Ziegert and Mikulska, 2013), quality of institutions (Balcerzak and Pietrzak, 2017), sustainable development (Pietrzak et al., 2017) and many others.

In the literature (Saltelli, 2007; Michalos et al., 2011) one can find following advantages of composite indicators: ease of interpretation; possibility to summarise multi-dimensional phenomenon; they are excellent communication tool to be used by media; make it easier to

evaluate and present an overall trends of phenomena over time and across geographic regions; help to reduce the visible size of a set of indicators without dropping the underlying information base; are flexible by including desired or excluding undesired variables.

However, composite indicators are not ideal. Among the current critiques of using composite indicators the following disadvantages can be listed (Saltelli, 2007; Michalos et al., 2011; Becker et al., 2017): lack of transparency in construction; subjective selection of indicators; ad hoc selection of weights and aggregation methods; risk of misinterpretation; oversimplification of complex phenomenon; risk of misusing by supporting desire policy; risk of giving misleading policy directions; occurrence of redundant variables and double-counting; normative nature of the weights; drawbacks of the linear method of aggregation; flattening alternatives and differences among analyzed objects; lack of clear meaning; variability of domains and indicators over time; sensitiveness to the choice of normalization and aggregation methods.

4 Kohonen self-organizing maps for the classification problem

Bearing in mind the above-mentioned set of disadvantages, in this research we want to use Kohonen self-organizing maps (SOM) to classify Polish NUTS-4 regions due to their standard of living. SOMs are exploratory data analysis technique projecting multi-dimensional data onto a two-dimensional space. That procedure allows clear visualization of the data and easy identification of groups with similar characteristics. In this sense, Kohonen maps can be thought of as a factor analysis combined with a cluster analysis (Deichmann et al., 2013). The advantage of self-organizing maps over composite indicators in the classification procedure is revealed by excluding such problems as normative nature of weight, sensitivity to used normalization and aggregation method, and also a subjective selection of indicators. Another advantage of Kohonen maps is the self-organizing property of the map which makes estimated components varies in a monotonic way across the map (Deichmann et al., 2007). The utility of SOMs in cluster analysis of multidimensional phenomenon has been presented, among others in Pisati et al. (2010).

The basic idea for constructing Kohonen maps is to reduce the complexity of matrix X consisting of input vectors x_i representing the coordinates of observation i in the d-dimensional input space, by projecting it onto a lower dimensional output space. This space usually takes the form of a two-dimensional grid arranged in a square or hexagonal lattice. Each grid cell called ‘node’ or ‘neuron’ is specialized in attracting observations that possess certain combinations of attributes. It means that each SOM node is characterized by a unique

1xd weight vector that belongs to the same coordinate space as the input vectors x_i. Input vectors are compared in the learning process of the SOM with the weight vectors and each observation *i* can be properly allocated to the best matching neuron. The detailed description of SOM algorithm is widely presented and can be found in Kohonen (2013), Pisati et. al (2010) and others.

5 Empirical analysis

To classify Polish NUTS-4 regions[3] due to their standard of living in 2016 Kohonen self-organizing maps were applied. In this research we used 45 variables describing the following standard of living dimension[4]:

- education: percentage of kids aged 3-5 participating in pre-school education (17), children aged 3-5 years per one place in a pre-school education center (26), students per one branch in primary schools (21), students per one branch in secondary schools (19), students per one branch in upper secondary schools (36), net (30) and gross (32) primary education ratio, net (31) and gross (33) junior high school education ratio;
- social security: percentage of population receiving social benefits (23), crime detection rate (27), number of crimes per 1000 inhabitants (28), divorces per 1000 inhabitants (35), poverty rate (38), percentage of children up to 17year old, for which parents received child benefit (39), old age dependency ratio (15);
- culture and recreation: number of people per 1 library (3), number of books borrowed by one reader (4), number of people per 1 place in cinema (5), number of accommodation places per 100000 inhabitants (12), number of hotels per 100000 inhabitants (11);
- healthcare: number of people per 1 pharmacy (1), number of doctors per 1000 inhabitants (10), number of nurses per 1000 inhabitants (20), deaths on cancer as a % of total deaths (43), deaths on cardiovascular diseases as a % of total deaths (44), infant mortality rate (45);
- labour market: unemployment rate (34), long-term unemployment rate (2), youth unemployment rate (16), number of job offers per 1 unemployed person (18), employment rate (24), average monthly wage (25), number of accidents at work per 1000 workers (42);

[3] All urban counties were excluded from the analysis due to avoid classification problems of other counties due to large differences in the level of neuron activation.
[4] Number in brackets is the same as input ID on Fig. 4.

- natural environment: percentage of protected area (22), industrial water sewage rate (37);
- transport and communication: number of cars per 1000 inhabitants (29), length of paved roads per 100km^2 (40), car accidents per 100000 inhabitants (42);
- living conditions: percentage of households connected to water supply (6), percentage of households connected to sewage system (7), percentage of population using gas installation (8), percentage of population connected to sewage treatment plants (9), average flat area per 1 person (13), number of flats per 1000 inhabitants (14).

The main variable selection criterion was data availability at NUTS-4 level. The sample size was 314 counties. To prepare U-matrix, batch-learning version of the SOM was used. After several trials, the most satisfying results[5] gave the SOM consisting of 49 nodes arranged in a 7x7 hexagonal lattice. From the U-matrix and the previous taxonomic analysis[6], it was possible to point six main groups of regions which show similarities in the scope of selected variables (see Fig. 2).

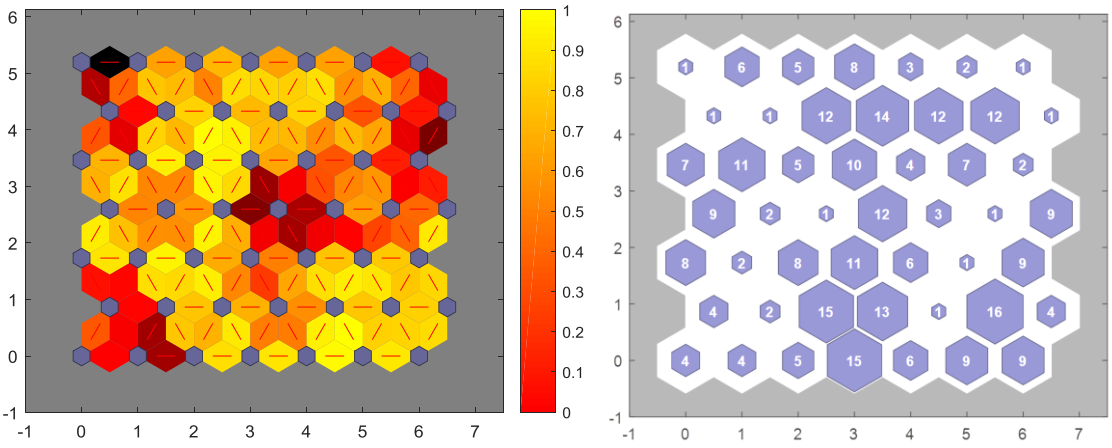


Fig. 2. The U-matrix of 7x7 SOM (left picture). Light shades indicate large similarity whereas dark shades show dissimilarities and cluster borders. The neuron locations in the SOM topology and the number of hits are shown in the right picture.

The light colour on the map represents smaller distances between neurons and hence the high intra-cluster homogeneity. The darker colour on the map, the larger distances between neurons and hence the smaller inter-cluster homogeneity. Based on the cells colours and

[5] There were also constructed other SOMs arranged in a rectangular grid as well as hexagonal lattice from 3x10 up to 10x10. When choosing the model, the aim was to ensure that the network spreads as well as possible on the examined objects and does not contain empty neurons.
[6] Tree Diagram for Ward's method and k-means clustering results indicated the existence of 6 clusters.

number of regions associated with each of the neurons, NUTS-4 regions were grouped into clusters which spatial arrangement is shown in Fig. 3.

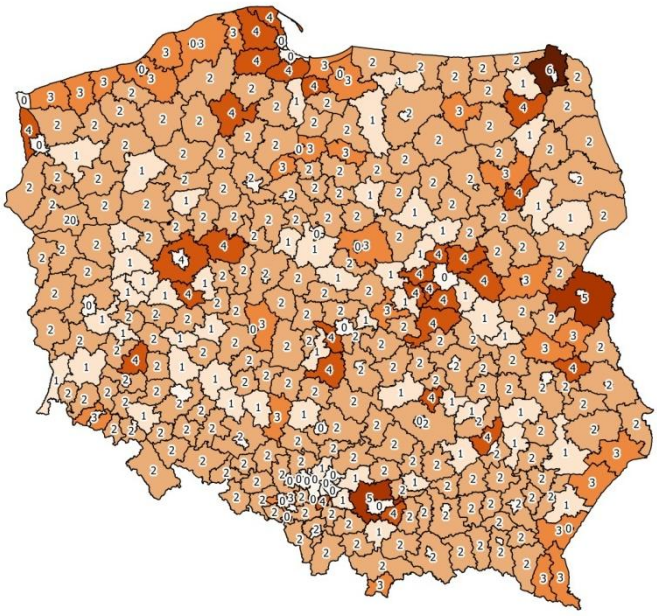


Fig. 3. The spatial arrangement of clusters according to NUTS-4 classification. 0 means urban counties excluded from the research.

According to this arrangement, biggest, suburban counties are mostly described with cluster 4, these are the regions with the highest average values for health care and almost the highest values for education, culture, labour market, transport and living conditions (Table 1). Which allows to define them as counties with the highest standard of living. Counties qualified as cluster 1 are regions further away from the cities, with no clearly developed service and production facilities (relatively low average values for the labour market and living conditions - Table 1). Majority of counties classified in cluster 2 are similar regardless of geographic location. Cluster 3 stretches along the coastline, in the south-eastern border areas and in some other counties which can be regions perceived as, among the others, developing tourism. In cluster 5 there were two counties classified - bialski and krakowski, which according to Table 1 can be perceived as the one with the lowest average values in almost all standard of living dimensions. In contrast to county suwalski which creates separate cluster no. 6 - with the highest values of variables presented in Table 1 (the exception is health care - in this area suwalski turned out to have the lowest value).

Table 1 presents the average values of variables with the strongest influence on neurons (Fig. 4) from each standard of living dimension (Fig. 1).

Table 1. Cluster statistics - group mean values for chosen variables.

Cluster no.	Input 17	Input 15	Input 4	Input 44	Input 2	Input 37	Input 29	Input 6
	S	D	S	D	D	S	S	S
1	75.77	31.78	18.80	47.12	39.26	97.56	586.18	89.16
2	72.67	30.57	19.20	47.04	39.36	96.73	574.20	84.16
3	67.45	29.42	18.18	45.78	38.43	97.75	570.99	84.94
4	83.85	28.89	18.65	43.34	36.86	99.83	564.01	87.80
5	74.33	29.85	18.09	51.50	41.65	96.22	608.83	84.00
6	47.13	28.70	21.18	41.53	30.70	100.00	604.11	87.50

The shades of colours and corresponding numbers of clusters on the map should not be associated directly with the standard of living. What they mean is the similarities and differences in the counties due to the 45 variables used for describing the standard of living. According to the results, one can notice that despite visible differentiation, counties do not form clusters solely on the basis of the geographical neighbourhood. The former division into Poland A and B related to the socio-economic development is also not observable here. Instead, the similarities can be sought rather in the proximity to large agglomerations and tourist attractiveness of the regions. SOM weight planes presented in Fig. 4 demonstrate the influence of each variable on a different cluster node. The darker the colour, the smaller the impact of a given variable. It can be noticed for instance that variables describing culture and recreation group, had influence only on the cluster in the left lower corner of the map which represents cluster number 3.

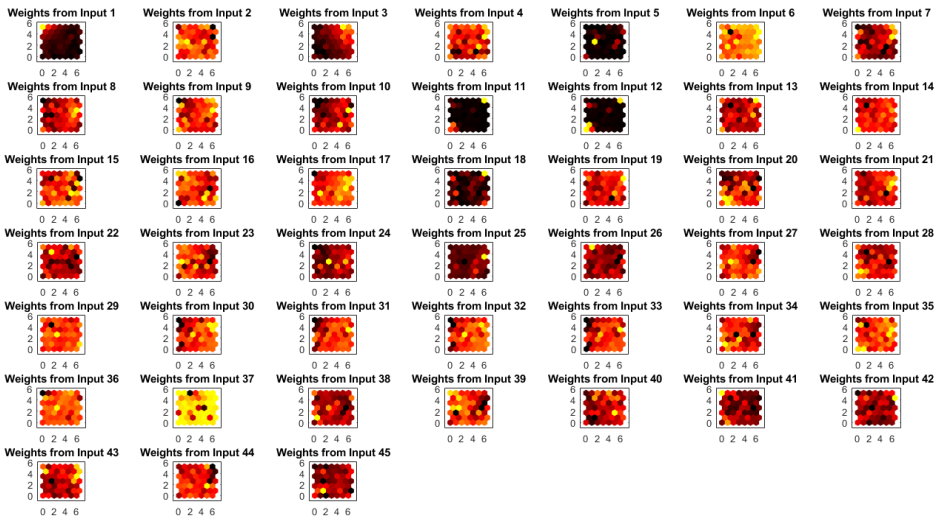


Fig. 4. Component planes for all variables.

To check the comparability of clusters obtained using SOM we also grouped analyzed counties based on composite indicator (CI) value. Three different methods were used to calculate CI (quotient transformation by max was used to normalize all variables):

- without a pattern: $CI_i = \frac{\sum_{i=1}^{m} z_{ij}}{m}$,
- with a pattern: $CI_i = 1 - \frac{d_i^+}{d_-}$,
- TOPSIS: $CI_i = \frac{d_i^-}{d_i^+ + d_i^-}$.

Where: z_{ij} - normalized values of j[th] variable in i[th] region ($i = 1,..,n; j = 1,\dots,m$); d_i^+- Euclidean distance between each region and a pattern (a pattern is an object with maximum values for stimulants and minimum for destimulants); $d_- = \bar{d} - 2s_d$; d_i^-- Euclidean distance between each region and an anti-pattern (an anti-pattern is an object with minimum values for stimulants and maximum for destimulants).

CI values were a basis to group counties into 4 groups (I - $CI_i \geq \overline{CI} + s_{CI}$; II - $\overline{CI} \leq CI_i < \overline{CI} + s_{CI}$; III - $\overline{CI} - s_{CI} \leq CI_i < \overline{CI}$ and IV - $CI_i < \overline{CI} - s_{CI}$. Also k-means algorithm was used as grouping tool[7]. The results of V-Cramer statistics are presented in Table 2.

Table 2. V-Cramer statistics.

Clustering method	CI without a pattern	CI with a pattern	CI TOPSIS	k-means 3 clusters	k-means 4 clusters	k-means 6 clusters
SOM	0.569	0.628	0.603	0.816	0.744	0.763

As can be seen in Table 2 SOM clustering results are much closer to those obtained using k-means algorithm than to those based on composite indicator value.

Conclusions

In this paper, the standard of living in Polish counties was considered based on 45 determinants. The first part of the paper discussed briefly standard of living measurement and pros and cons of composite indicators. At the empirical part, regions were clustered using self-organizing maps approach, which is a more objective clustering method. A number of clusters, however, has been obtained using U-matrix and the taxonomic analysis. The analysis showed that there is visible differentiation among the counties, however, they form clusters neither on the basis of geographical neighbourhood nor historical conditions. Two interesting

[7] $\overline{CI}$ - an average value of CI value, s_{CI} - standard deviation of CI value.

clusters of similar objects can be distinguished: the first one consist of counties adjacent to large agglomerations (cluster no. 4 - which contains counties with the highest standard of living); and cluster no. 3 that consists of counties attractive for tourism. Results presented in Table 1 shows that SOM has good discriminatory abilities not only in terms of structures similarity but as well values similarity, as they maximise between-group dissimilarities and minimalize within-group dissimilarity. They are also a useful tool to visualize high-dimensional data.

Further analysis should include dynamics of the described phenomenon as well as its determinants. Well-constructed SOM model will allow to predict and assess whether changes taking place in particular counties will improve or worsen their standard of living.

References

Balcerzak, A. P. & Pietrzak, M.B., (2017). Human Development and Quality of Institutions in Highly Developed Countries. In: Bilgin, M. H., Danis, H., Demir, E. & Can U. (Eds.). *Financial Environment and Business Development. Proceedings of the 16th Eurasia Business and Economics Society*. Springer International Publishing, 231-241.

Becker W., Saisana, M., Paruolo, P. & Vandecasteele, I. (2017), Weights and Importance in Composite Indicators: Closing the Gap. *Ecological Indicators, 80*, 12-22.

Chrzanowska, M., Dudek, H. & Landmesser J. (2017), Standard of Living in Selected Countries in Central Europe in the Period 2000-2014. In: Jedlicka, P., Maresova, P. & Soukal, I. (eds.) *Proceedings of the International Scientific Conference Hradec Economic Days 2017, 7*(1), 325-331.

Deichman, J., Eshghi, A., Haughton, S. & Woolford, S. (2007). Measuring the international digital divide: An application of Kohonen self-organizing maps. *International Journal of Knowledge and Learning*. *3*(6), 552-575.

Deichman, J., Haughton, D., Malgwi, C. & Soremekun, O. (2013). Kohonen Self-Organizing Maps as a Tool for Assessing Progress toward the UN Millennium Development Goals. *Journal of Human Development and Capabilities, 14* (3), 393-419.

Goossens, Y. (ed.), (2007). *Alternative progress indicators to Gross Domestic Product (GDP) as a means towards sustainable development*. European Parliament Study.

Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks, 37*, 52–65.

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

Kuc, M. (2017), Social convergence in Nordic countries at regional level. *Equilibrium Quarterly Journal of Economics and Economic Policy, 12* (1), 25-41.

Michalos, A. C., Smale, B., Labonté, R., Muharjarine, N., Scott, K., Moore, K., Swystun, L., Holden, B., Bernardin, H., Dunning, B., Graham, P., Guhn, M., Gadermann, A.M., Zumbo, B.D., Morgan, A., Brooker, A.-S. & Hyman, I. (2011). *The Canadian Index of Wellbeing. Technical Report 1.0. Waterloo.* On: Canadian Index of Wellbeing and University of Waterloo.

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

Pisati, M., Whelan, Ch., Lucchini, M. & Maitre, B. (2010). Mapping patterns of multiple deprivation using self-organising maps: An application to EU-SILC data for Ireland. *Social Science Research, 39* (3), 405-418.

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research, 81*(1), 65-77.

Sobiechowska-Ziegert, A. & Mikulska, A. (2013). Measure of the level of socio-economic development in provinces. *Quantitative methods in economics*, *XIV*(2), 200–209.

Wawrzyniak, D. (2016), Standard of Living in The European Union. *Comparative Economic Research - Central and Eastern Europe*, *19* (1).

Bayesian inference for deterministic cycle with time-varying amplitude

Łukasz Lenart[1]

Abstract

The main goal of the paper is to obtain posterior distribution for frequency in some generalization of deterministic cycle model proposed by Lenart and Mazur (2016), where the autoregressive model with time-varying almost periodic mean function was investigated with constant amplitude and frequencies. The assumption concerning constant amplitude in such model seems to be too strong to describe the changing nature of the business cycle. Hence, in this paper we assume that the mean function depends on unknown frequencies (related to the length of the cyclical fluctuations) in a similar way as for the almost periodic mean function proposed in Lenart and Mazur (2016), while the assumption concerning constant amplitude was relaxed. More specifically, we assume that the amplitude associated with a given frequency is time-varying and is a linear spline. We obtain the explicit marginal posterior distribution for vector of frequency parameters in the approximate model. We consider real data example concerning monthly production in industry in Poland. The main conclusion is that the posterior for frequency is still likely to be multimodal, but it seems that this multimodality is not as strong as in the deterministic cycle model with constant frequency proposed in Lenart and Mazur (2016).

Keywords: *deterministic cycle, time-varying amplitude, Bayesian inference*

JEL Classification: C22, E32

DOI: 10.14659/SEMF.2018.01.24

1 Introduction

The concept of stochastic cycle is very well known (see Harvey, 2004; Harvey and Jaeger, 1993; Harvey and Trimbur, 2003; Pelagatti, 2016; Trimbur, 2006; Koopman and Shephard, 2015; Azevedo et al., 2006; Harvey et al., 2007 and many others). This concept assumes the stationarity of cyclical fluctuations with zero mean function. The models with a deterministic cycle are not so popular as models with stochastic cycle. Following Harvey (2004) the concept of deterministic cycle is based on almost periodic function at time $t \in \mathbb{Z}$ with one frequency $\lambda \in (0, \pi)$ of the form

$$f(t) = a\sin(\lambda t) + b\cos(\lambda t).$$

It is widely known that above function is not flexible enough to describe the variable in time dynamics of business cycle. Therefore the more flexible concepts were considered. In Lenart and Pipień (2013) the nonparametric inference were considered under assumption that the conditional expectation of observed process contains almost periodic component with

[1] Cracow University of Economics, Department of Mathematics, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: lukasz.lenart@uek.krakow.pl.

more than one frequency. In Lenart et al. (2016) the parametric and nonparametric inference were considered under assumption that the mean function of cyclical process is almost periodic function with few frequencies. Finally in Lenart and Pipień (2015, 2017a, 2017b) authors consider the nonaprametric test based on subsampling approach to test the deterministic cycles. In Mazur (2016, 2017a, 2017b) the parametric model containing deterministic cycle was considered. In all above approaches the amplitude of considered deterministic cycle is assumed to be constant in time. This assumption seems to be strong taking into consideration the variable nature of the business cycle. Note that in recent few years the deviation cycle in Poland has flattened.

Therefore, we investigate in this paper the time varying amplitude by considering the following time-varying function:

$$g(\lambda, t) = a(t)\sin(\lambda t) + b(t)\cos(\lambda t) \tag{1}$$

of integer $t \in \{1,2,\dots,n\}$, where $a(t)$ and $b(t)$ are a linear splines with $r+1$ knots $\{(t_i, a_i), i = 0,1,\dots,r\}$ for $a(t)$ and $\{(t_i, b_i), i = 0,1,2,\dots,r\}$ for $b(t)$. We assume $t_0 = 1$ and $t_r = n$. Hence

$$a(t) = \sum_{i=1}^{r} \mathrm{I}_{\{t_{i-1} \le t < t_i\}} \left[a_{i-1} \frac{(t_i - t)}{t_i - t_{i-1}} + a_i \frac{(t - t_{i-1})}{t_i - t_{i-1}} \right], \qquad t \in [t_0, t_r), \quad a(t_r) = a_r,$$

$$b(t) = \sum_{i=1}^{r} \mathrm{I}_{\{t_{i-1} \le t < t_i\}} \left[b_{i-1} \frac{(t_i - t)}{t_i - t_{i-1}} + b_i \frac{(t - t_{i-1})}{t_i - t_{i-1}} \right], \qquad t \in [t_0, t_r), \quad b(t_r) = b_r.$$

Paper is organized as follows. In section 2 we introduce the model with time-varying amplitude of deterministic cycle. In section 3 we investigate the Bayesian inference for such model and we show the closed form for marginal posterior distribution for frequency vector. In the last section we consider real data example concerning monthly industrial production in Poland. Note that in the empirical analyses of the business cycle the choice of industrial production series of monthly frequency seems standard.

2 Model formulation

We consider the following autoregressive model of order p:

$$\Psi(L)(Y_t - g(\lambda, t) - \mu) = \varepsilon_t, \tag{2}$$

with time-varying mean function $g(\lambda, t) + \mu$, where $\Psi(L) = 1 - \sum_{k=1}^{p} \psi_k L^k$ is a standard polynomial in autoregressive part, $g(\lambda, t)$ is of the form (1) and ε_t is a white noise. Note that equivalently:

$$Y_t = \sum_{k=1}^{p} \psi_k Y_{t-k} + \Psi(L)[g(\lambda, t) + \mu] + \varepsilon_t.$$

In the next part of this section we show that the above model can be approximated by

$$Y_t = \sum_{k=1}^{p} \psi_k Y_{t-k} + \tilde{g}(\lambda, t) + \tilde{\mu} + \varepsilon_t,$$

where $\tilde{g}(\lambda, t)$ is of the general form (1).

To show this we consider a linear function $s(t)\colon \mathbb{Z} \to \mathbb{R}$ passing through the points (x, z_s) and (y, w_s). In such a case we have

$$s(t) = \frac{t(-(w-z))}{x-y} - \frac{yz - wx}{x-y},$$

$$\begin{aligned}\Psi(L)s(t)\sin(\lambda t) &= \left[1 - \sum_{k=1}^{p} \psi_k L^k\right] s(t)\sin(\lambda t) = s(t)\sin(\lambda t) - \sum_{k=1}^{p} \psi_k s(t-k)\sin(\lambda(t-k)) \\ &= w_s \frac{(x-t)[\sin(\lambda t) - \sum_{k=1}^{p} \psi_k \sin(\lambda(t-k))] - \sum_{k=1}^{p} k\sin(\lambda(t-k))\psi_k}{x-y} \\ &+ z_s \frac{(t-y)[\sin(\lambda t) - \sum_{k=1}^{p} \psi_k \sin(\lambda(t-k))] + \sum_{k=1}^{p} k\sin(\lambda(t-k))\psi_k}{x-y}.\end{aligned}$$

Using elementary trigonometric identities we obtain

$$\sum_{k=1}^{p} \psi_k \sin(\lambda(t-k)) = \sin(\lambda t) f_{ss} + \cos(\lambda t) f_{sc}$$
$$\sum_{k=1}^{p} k\psi_k \sin(\lambda(t-k)) = \sin(\lambda t) g_{ss} + \cos(\lambda t) g_{sc}$$
$$\sum_{k=1}^{p} \psi_k \cos(\lambda(t-k)) = \sin(\lambda t) f_{cs} + \cos(\lambda t) f_{cc}$$
$$\sum_{k=1}^{p} k\psi_k \cos(\lambda(t-k)) = \sin(\lambda t) g_{cs} + \cos(\lambda t) g_{cc},$$

where $f_{ss}, f_{sc}, f_{cc}, f_{cs}, g_{ss}, g_{sc}, g_{cc}, g_{cs}$ depends only on $\lambda, \psi_1, \psi_2, \dots, \psi_p$. In the same way we can decompose $\Psi(L)c(t)\cos(\lambda t)$, where $c(t)\colon \mathbb{Z} \to \mathbb{R}$ is a linear function, passing through the points (x, z_c) and (y, w_c). Using now elementary algebra we get

$$\begin{aligned}&\Psi(L)[s(t)\sin(\lambda t) + c(t)\cos(\lambda t)] = \tfrac{x-t}{x-y}\sin(\lambda t)[1 - f_{ss} - f_{cs}](w_s + w_c) \\ &+ \tfrac{t-y}{x-y}\sin(\lambda t)[1 - f_{ss} - f_{cs}](z_s + z_c) + \tfrac{z_s - w_s}{x-y}[\sin(\lambda t) g_{ss} + \cos(\lambda t) g_{sc}] \\ &+ \tfrac{x-t}{x-y}\cos(\lambda t)[1 - f_{cc} - f_{sc}](w_s + w_c) + \tfrac{t-y}{x-y}\cos(\lambda t)[1 - f_{cc} - f_{sc}](z_s + z_c) \\ &+ \tfrac{z_c - w_c}{x-y}[\sin(\lambda t) g_{cs} + \cos(\lambda t) g_{cc}].\end{aligned} \quad (3)$$

Note that

$$\tfrac{z_s - w_s}{x-y}[\sin(\lambda t) g_{ss} + \cos(\lambda t) g_{sc}] = \left(\tfrac{x-t}{x-y} + \tfrac{t-y}{x-y}\right)\tfrac{z_s - w_s}{x-y}[\sin(\lambda t) g_{ss} + \cos(\lambda t) g_{sc}]$$
$$\tfrac{z_c - w_c}{x-y}[\sin(\lambda t) g_{cs} + \cos(\lambda t) g_{cc}] = \left(\tfrac{x-t}{x-y} + \tfrac{t-y}{x-y}\right)\tfrac{z_c - w_c}{x-y}[\sin(\lambda t) g_{cs} + \cos(\lambda t) g_{cc}].$$

Hence, $\Psi(L)[s(t)\sin(\lambda t) + c(t)\cos(\lambda t)]$ can be equivalently written as

$$\Psi(L)[s(t)\sin(\lambda t) + c(t)\cos(\lambda t)] = \frac{x-t}{x-y}\sin(\lambda t)\widetilde{w}_s + \frac{t-y}{x-y}\sin(\lambda t)\tilde{z}_s + \frac{x-t}{x-y}\cos(\lambda t)\widetilde{w}_c + \frac{t-y}{x-y}\cos(\lambda t)\tilde{z}_s, \quad (4)$$

with bijective transformation $(w_s, w_c, z_s, z_c) \to (\widetilde{w}_s, \widetilde{w}_c, \tilde{z}_s, \tilde{z}_c)$, under constant $\lambda, \psi_1, \psi_2, \dots, \psi_p$. Based on (4) we can approximate the model (2) as

$$Y_t = \sum_{k=1}^{p} \psi_k Y_{t-k} + \tilde{g}(\lambda, t) + \tilde{\mu} + \varepsilon_t,$$

where $\tilde{g}(\lambda, t)$ is of the general form (1). We called the above model approximation since (4) holds on integer line, while we consider in the model a linear spline. The above model can be generalized in natural way to multi-frequency case

$$Y_t = \sum_{k=1}^{p} \psi_k Y_{t-k} + \sum_{k=1}^{m} \tilde{g}_k(\lambda_k, t) + \tilde{\mu} + \varepsilon_t, \quad (5)$$

where $\tilde{g}_k(\lambda_k, t)$, for $k = 1,2, \dots, m$ are of the general form (1). In such a case we use notation $\{(t_{i,k}, a_{i,k}), i = 0,1, \dots, r_k\}$ for $a_k(t)$ and $\{(t_{i,k}, b_{i,k}), i = 0,1,2, \dots, r_k\}$ for $b_k(t)$, where k corresponds to frequency λ_k, $k = 1,2, \dots, m$. To claryfy the model with time-varying amplitude we consider the following ilustrative example, where the dynamics of $g(\lambda, t)$ are ilustrated.

Example 1. We consider $n = 156$, $p = 0$, one frequency $\lambda = 0.15$, $r \in \{2,3,4,6\}$ and with equally spaced $t_0 = 1, t_1, t_2, \dots, t_r = n$. For fixed r we draw each $a_0, a_1, a_2, \dots, a_r$ from uniform distribution on the interval $[2,15]$ and $b_0, b_1, b_2, \dots, b_r$ from uniform distribution on the interval $[-5,0]$.

Table 1. Parameters used in example.

$\boldsymbol{r}$	$\{\boldsymbol{a_0}, \boldsymbol{a_1}, \boldsymbol{a_2}, \dots, \boldsymbol{a_r}\}$	$\{\boldsymbol{b_0}, \boldsymbol{b_1}, \boldsymbol{b_2}, \dots, \boldsymbol{b_r}\}$
$r = 2$	{8.9,2.2,5.8}	$\{-0.4, -2., -4.5\}$
$r = 3$	{14.7,8.3,3.2,14.}	$\{-0.8, -3.9, -2.4, -0.7\}$
$r = 4$	{4.7,10.,8.9,2.6,2.9}	$\{-4.5, 0., -1.4, -2.2, -4.1\}$
$r = 6$	{10.2,8.8,8.,4.7,3.4,8.5,6.8}	$\{-1., -1.4, -3.4, -2.3, -2.1, -3.6, -4.2\}$

The main finding from presented example (see Fig. 1) is that the cycle based on (1) with time-varying amplitude and with one frequency is much more flexible than deterministic cycle with one frequency and constant amplitude. Hence, the proposed deterministic cycle model with time-varying amplitude may be useful from practical point of view in statistical inference concerning amplitude and length of the cycle.

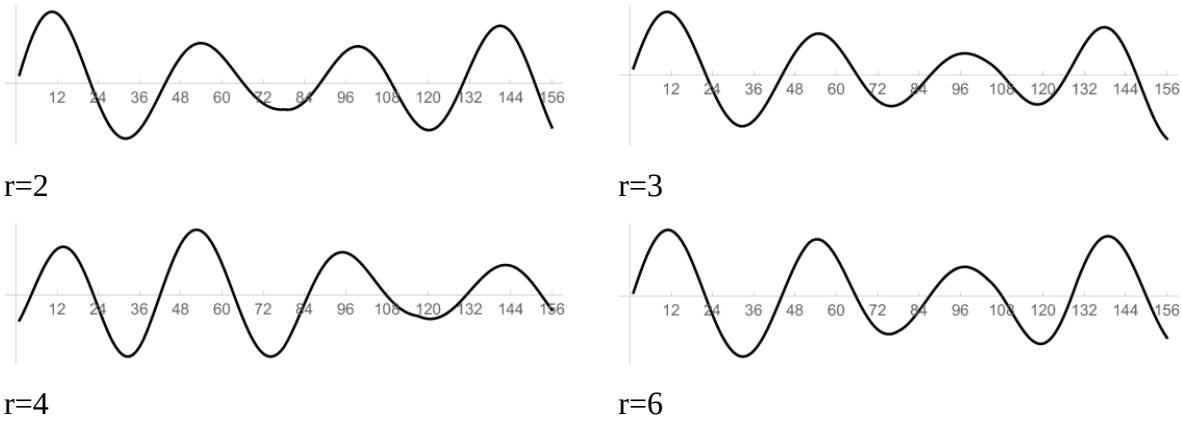


Fig. 1. Paths for $g(\lambda, t)$ for different r and $\{a_0, a_1, a_2, \dots, a_r\}$, $\{b_0, b_1, b_2, \dots, b_r\}$.

3 Bayesian inference for frequency

In this section we obtain closed form of marginal posterior distribution for vector of frequency $\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$. Note that model (5) can be equivalently written as:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon, \tag{6}$$

where $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^{\mathrm{T}}$, the matrix $\mathbf{X}$ depends on $\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)$ and the first coordinates of knots, β is a $\left(1 + p + 2(r_1 + r_2 + \cdots + r_{\mathrm{m}} + m)\right)$ x 1 vector of parameters:

$$\beta = [\mu \ \psi_1 \ \psi_2 \ \dots \ \psi_p \ a_{0,1} \ a_{1,1} \ \dots \ a_{r_1,1} \ b_{0,1} \ b_{1,1} \ \dots \ b_{r_1,1} \ \dots \ a_{0,m} \ a_{1,m} \ \dots \ a_{r_m,m} \ b_{0,m} \ a_{1,m} \ \dots \ b_{r_m,m}]^{\mathrm{T}}.$$

Moreover we assume that $\varepsilon_t \sim N(0, \tau^{-1})$, for $t = 1,2, \dots, n$, where $\varepsilon = [\varepsilon_1 \ \varepsilon_2 \ \dots \ \varepsilon_n]^{\mathrm{T}}$. Let us denote $\theta = (\beta, \tau, \lambda_1, \lambda_2, \dots, \lambda_m)$. The likelihood function is of the form:

$$p(\mathbf{y}|\theta) = \frac{1}{\sqrt{(2\pi)^n}} \tau^{\frac{n}{2}} \exp\left\{-\frac{\tau}{2}(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta)\right\}.$$

We assume the standard conjugate family of distributions:

$p(\theta) = p(\beta, \tau)p(\varphi) = p(\beta|\tau)p(\tau)p(\varphi)$,

where $\beta|\ \tau \sim N(\mathbf{b}, (\tau\mathbf{B})^{-1})$ and $\tau \sim G\left(\frac{n_0}{2}, \frac{s_0}{2}\right)$, with prior hyperparameters $\mathbf{b}, \mathbf{B}, n_0, s_0$. Under such notation:

$$p(\beta|\tau) = (2\pi)^{-k/2}(\det(\mathbf{B}))^{1/2}\tau^{k/2}\exp\left\{-\frac{\tau}{2}(\beta - \mathbf{b})'\mathbf{B}(\beta - \mathbf{b})\right\},$$

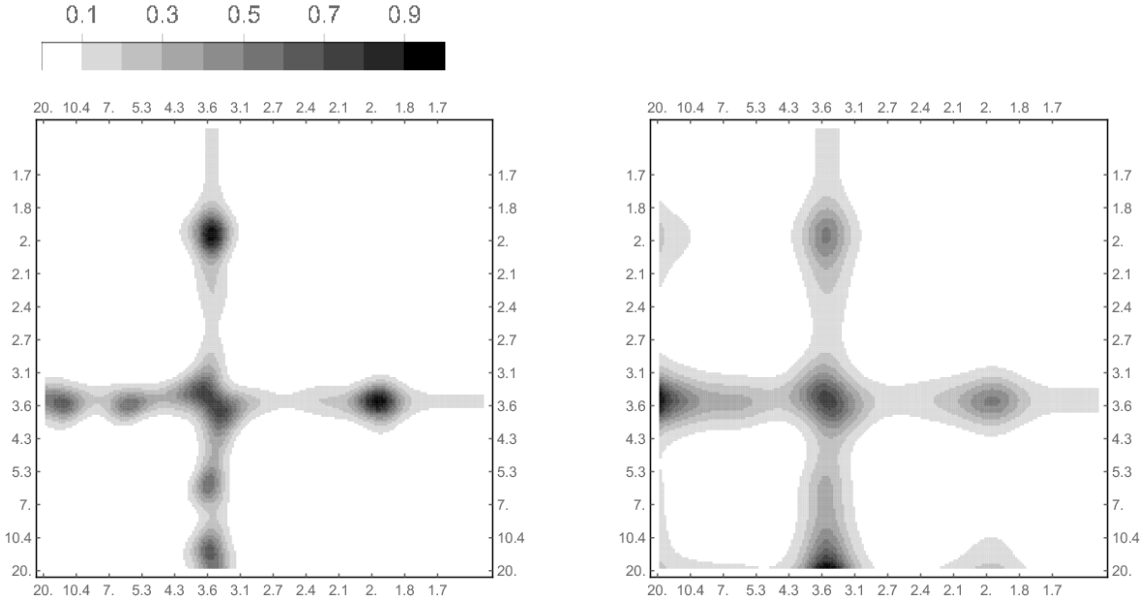
$$p(\tau) = \frac{(s_0/2)^{\frac{n_0}{2}}}{\Gamma\left(\frac{n_0}{2}\right)} \tau^{\frac{n_0}{2}-1} \exp\left(-\frac{s_0\tau}{2}\right).$$

For the frequency vector $(\lambda_1, \lambda_2, \ldots, \lambda_m)$ we assume uniform distribution on $[\lambda_L, \lambda_U]^m$. Using now the same arguments as in in Lenart and Mazur (2016) we get the marginal posterior distribution for Λ of the form

$$p(\Lambda|\mathbf{y}) \propto (\det(\mathbf{X}'\mathbf{X}+\mathbf{B}))^{-1/2}(\mathbf{y}'[\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X}+\mathbf{B})^{-1}\mathbf{X}']\mathbf{y}+s_0)^{-\frac{n+n_0}{2}}. \quad (7)$$

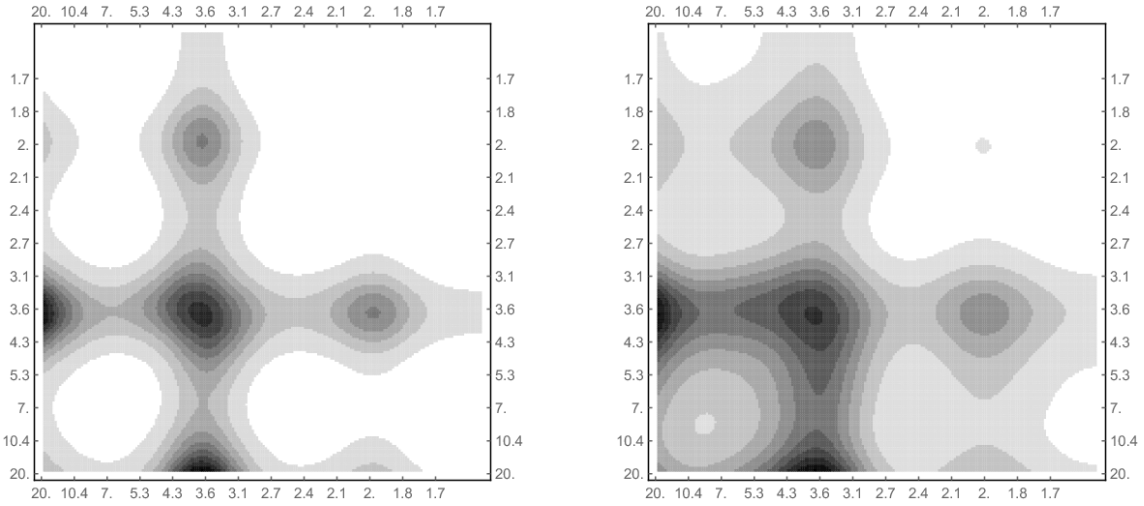
4 Real data example

We consider production in industry in Poland[2] (mining and quarrying; manufacturing; electricity, gas, steam and air conditioning supply – percentage change compared to same period in previous year, calendar adjusted data, not seasonally adjusted data) from Jan. 2001 to Oct. 2017. We restrict attention to the set of frequencies $[\frac{\pi}{120}, \frac{\pi}{9}]$. It means that we consider only the cycles which are not shorter than one and a half year and longer than 20 years. We fix $n = 180$ and we consider last 15 years for the data set in our empirical analysis. Fig. 2 presents the marginal posterior distributions (7) for bivariate case ($\Lambda = (\lambda_1, \lambda_2)$) for constant amplitude, $r = 1$, $r = 2$, $r = 3$ and equally spaced knots. Only the case $p = 10$ and $m = 2$ is presented.



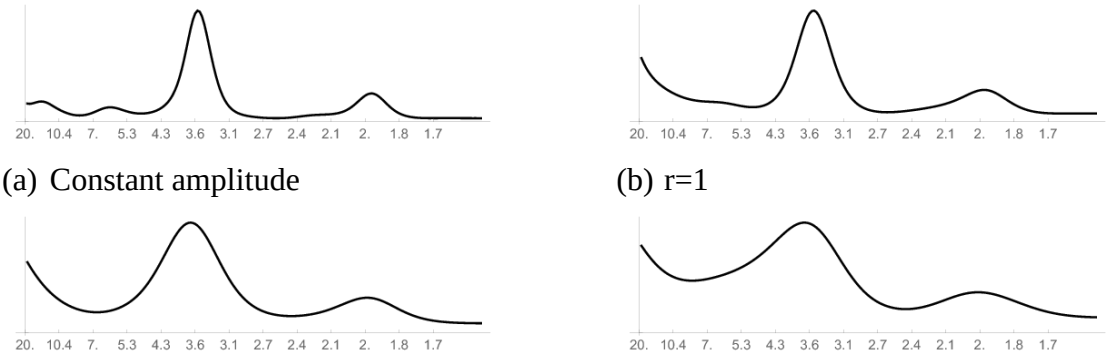
(a) Constant amplitude (b) r=1

[2]Source: Eurostat.



(c) r=2 (d) r=3

Fig. 2. Marginal posterior distribution (kernel) for $\Lambda = (\lambda_1, \lambda_2)$. Horizontal and vertical axis - length of the cycle in years, $p = 10$. Shades of grey corresponds to kernel value.



(a) Constant amplitude (b) r=1

(c) r=2 (d) r=3

Fig. 3. Marginal posterior distribution (kernel) for one frequency λ_1 in bivariate case $\Lambda = (\lambda_1, \lambda_2)$. Horizontal axis - length of the cycle in years, $p = 10$.

Marginal distribution (kernel) under constant amplitude (see Fig. 2 (a)) show two predominant length of the deterministic cycle. Cycle with length approx. 3.5 year and 2 years (see also Fig. 3). The mass is concentrated relativly close these two frequencies compared to time-varying amplitude (see Fig. 2 (b)-(d)). In the case of time-varying amplitude the mass is still concentrated near these frequencies, but with greater dispersion, especially for $r = 3$. Hence, the bigger r the weaker multimodality of posterior distribution. The main findings is that for each considered r the results clearly support the length of the business cycle approx. 3.5 years (see Fig. 3). The similar conclusions can be found in Lenart et al. (2016), where the same length of the cycle was detected using industrial production in Poland to Dec. 2014.

Conclusions

The closed form of marginal posterior distribution for frequencies in case of time-varying amplitude was shown. This gives the opportunity to expand the statistical inference proposed in Lenart and Mazur (2016). The real data example shows that in the case of last 15 years of industrial production in Poland (covering the period 2002-2017) the predominant length of the cycle is approx. 3.5 year. This conclusion is supported by both constant amplitude case and time-varying case. An open problem is the construction of the forecast assuming a time-varying amplitude of the deterministic cycle.

Acknowledgements

This research was financed from the funds granted to the Faculty of Finance and Law at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

Azevedo, J. V., Koopman, S. J. & Rua, A. (2006). Tracking the business cycle of the euro area: a multivariate model-based band-pass filter. *Journal of Business & Economic Statistics*, *24*(3), 278–290.

Harvey, A. C. (2004). Tests for cycles. In: Harvey, A.C., Koopman, S.J. & Shephard, N. (eds), *State space and unobserved component models*, *CUP*, pages 102–119.

Harvey, A. C. & Jaeger, A. (1993). Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics*, *8*, 231–247.

Harvey, A. C. & Trimbur, T. M. (2003). General model-based filters for extracting cycles and trends in economic time series. *Review of Economics and Statistics*, *85*(2), 244–255.

Harvey, A. C., Trimbur, T. M. & Van Dijk, H. K. (2007). Trends and cycles in economic time series: A bayesian approach. *Journal of Econometrics*, *140*, 618–649.

Koopman, S. J. & Shephard, N. (2015). *Unobserved Components and Time Series Econometrics*. Oxford university Press, Oxford.

Lenart, Ł. & Mazur, B. (2016), On Bayesian estimation of almost periodic in mean autoregressive models. *Przegląd Statystyczny*, *63*(3), 255–271.

Lenart, Ł., Mazur, B. & Pipień, M. (2016). Statistical analysis of business cycle fluctuations in Poland before and after the crisis. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *11*(4), 769–783.

Lenart, Ł. & Pipień, M. (2013). Almost Periodically Correlated Time Series in Business Fluctuations Analysis. *Acta Physica Polonica A*, *123*, 567–583.

Lenart, Ł. & Pipień, M. (2015). Własności empiryczne cyklu finansowego - analiza porównawcza Czech, Polski, Węgier, Wielkiej Brytanii i USA. *Folia Oeconomica Cracoviensia*, *56*, 81–112.

Lenart, Ł. & Pipień, M. (2017a). Non-Parametric Test for the Existence of the Common Deterministic Cycle: The Case of the Selected European Countries. *Central European Journal of Economic Modelling and Econometrics*, *9*(3), 201–241.

Lenart, Ł. & Pipień, M. (2017b). The Dynamics of the Credit Cycle in Selected Asian Countries. *Folia Oeconomica Cracoviensia*, *58*, 127-133.

Mazur, B. (2016). Growth cycle analysis: the case of Polish monthly macroeconomic indicators. *Folia Oeconomica Cracoviensia*, *57*, 37–54.

Mazur, B. (2017a). Probabilistic predictive analysis of business cycle fluctuations in Polish economy. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *12*(3),435–452.

Mazur, B. (2017b). Probabilistic Prediction Using Disaggregate Data: The Case of Value Added in Poland. *Folia Oeconomica Cracoviensia*, *58*, 85–103.

Pelagatti M. M. (2016). *Time Series Modelling with Unobserved Components*. Taylor & Francis Group, Boca Raton.

Trimbur T. M. (2006). Properties of higher order stochastic cycles. *Journal of Time Series Analysis*, *27*, 1–17.

Nonlinear stochastic cycle model

Łukasz Lenart[1], Justyna Wróblewska[2]

Abstract

The aim of this paper is to construct a stationary nonlinear stochastic cycle model by utilizing the idea of the linear innovations state space model. We construct our nonlinear stochastic cycle model using the idea of the deterministic cycle based on an almost periodic function. The main properties in time domain and frequency domain for proposed nonlinear stochastic cycle model were shown. The amplitude and phase of such cycle are stochastic and are directly defined by the amplitude and phase components. Extensions to multivariate case are possible. Our proposition gives an alternative for the popular stochastic cycle model, which can be reduced to linear ARMA model. For statistical inference, we do not need to use Kalman filter to compute the likelihood function, if we use state space models with single sources of error. However, the extension to the classical state space model with few sources of error is possible. The real data example has been presented in a Bayesian framework.

Keywords: *nonlinear stochastic cycle, linear innovations state space model, ARMA model*
JEL Classification: C22, E32
DOI: 10.14659/SEMF.2018.01.25

1 Introduction

The concept of stochastic cycle which can be reduced to an ARMA model is very well investigated in the literature (see Harvey, 1989; Harvey and Trimbur, 2003; Trimbur, 2006; Pelagatti, 2016; and many others). It is the most popular model in practice.

The alternative concept, known as the deterministic cycle concept is not popular. In relation to business cycle this is a consequence of conviction that the deterministic cycle is not flexible enough for description of the complex nature of business cycles. Harvey (2004) presents basic idea of deterministic cycle with one frequency. This idea was extended to multiple-frequency case in Lenart and Pipień (2013b), Lenart and Pipień (2015) and Lenart et al. (2016) where the existence of deterministic cycle was tested using the subsampling approach and the Fourier representation of an almost periodic function. Results presented in these papers show that in some cases the concept of deterministic cycle with multiple-frequency seems to be enough for the description of the dynamics of the business cycle.

[1] Corresponding author: Cracow University of Economics, Department of Mathematics, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: lukasz.lenart@uek.krakow.pl.
[2] Cracow University of Economics, Department of Econometrics and Operations Research, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: eowroble@cyf-kr.edu.pl.

In this paper we introduce the nonlinear stochastic cycle model by utilizing the idea of linear innovations state space model and idea of deterministic cycle model with one frequency. The main properties of such model were shown in time and frequency domain.

Taking into account the nonlinearity of the proposed model and the possibility of the multimodality in the likelihood function, we propose a Bayesian approach. There are several advantages of using Bayesian methods, but in this work we find that the ability to generate the posterior distributions of estimated parameters are one of the most important as the posterior distributions can inherit the multimodality. In contrast to the results obtained within the frequentist framework (usually point estimates and standard deviations) posterior distributions contain all the relevant information about quantities of interest. Of course the above mentioned multimodal possibility is also challenging in the Bayesian framework, as the MCMC methods employed to obtain posterior samples should be carefully settle in order to explore the whole posterior distributions' supports.

In section 2 we introduce the idea of the nonlinear stochastic cycle and we show the main properties of such model. In the section 3 the illustrative real data example is shown in a Bayesian framework.

2 The genesis of the model

2.1 The deterministic cycle model

Let us consider the deterministic cycle model of the form

$$Y_t = A\sin[\lambda(t+T)] + \varepsilon_t, \tag{1}$$

where $A \in \mathbb{R}$ is the amplitude, $\lambda \in (0, \pi)$ is the frequency, $T \in \mathbb{R}$ is the phase shift and ε_t is white noise. Note that the expectation function of Y_t is an almost periodic function $A\sin[\lambda(t+T)]$ with one frequency λ. This model is not flexible enough to describe most of cyclical patterns, in particular the dynamics of business cycle. The extension of this model was considered in both nonparametric and parametric case in Lenart and Pipień (2013a), Lenart and Mazur (2017), Lenart et al. (2016), Lenart and Pipień (2017), Mazur (2017). In this paper we generalize model (1) by considering a more flexible model with stochastic amplitude and stochastic phase shift by utilizing the idea of innovations state space models.

2.2 A model based on common white noise

By utilizing the idea of deterministic cycle and the idea of nonlinear innovations state space models (see Hyndman et al.,2008) we investigate the model of the following form:

$$\begin{cases} Y_t = & (A + A_{t-1})\sin[\lambda(t + T + T_{t-1})] + \mu + \varepsilon_t \\ A_t = & \psi_A A_{t-1} + \alpha_A \varepsilon_t \\ T_t = & \psi_T T_{t-1} + \alpha_T \varepsilon_t \end{cases} \quad \begin{matrix} \\ \text{deviations from amplitude } A \\ \text{deviations from phase shift T,} \end{matrix} \tag{2a}$$

where $A, T, \mu, \alpha_A, \alpha_T \in \mathbb{R}$ and ε_t is a white noise with variance σ^2. Equivalently, model (2a) can be represented as

$$\begin{cases} Y_t = & (A + A_{t-1})\sin[\lambda(t + T + T_{t-1})] + \mu + \varepsilon_t \\ A_t = & \psi_A A_{t-1} + \alpha_A [Y_t - (A + A_{t-1})\sin[\lambda(t + T + T_{t-1})] - \mu] \\ T_t = & \psi_T T_{t-1} + \alpha_T [Y_t - (A + A_{t-1})\sin[\lambda(t + T + T_{t-1})] - \mu], \end{cases} \tag{2b}$$

The component A_t can be interpret as stochastic deviations from the amplitude A, white the T_t as the stochastic deviation from phase shift. Then under the assumption that A_t is stationary and T_t is a random walk we formulate the following new theorem.

Theorem 2.1 *If $|\psi_A| < 1$, $\psi_T = 1$ and ε_t is the Gaussian white noise with variance σ^2 then $Y_t - \mu$ is zero-mean stationary time series with "cyclical" autocovariance function $\gamma_Y(\tau) = cov(Y_t, Y_{t+\tau})$ of the form*

$$\gamma_Y(\tau) = \begin{cases} \frac{1}{2} e^{-\frac{|\tau|\alpha_T^2 \lambda^2 \sigma^2}{2}} \left(\cos(\lambda\tau)(\gamma_A(\tau) + A^2) - \sin(\lambda|\tau|) \frac{A\lambda\alpha_A\alpha_T\sigma^2(1-\psi_A^{|\tau|})}{1-\psi_A} \right) & \text{for } \tau \neq 0 \\ \frac{1}{2}(\gamma_A(0) + A^2) + \sigma^2 & \text{for } \tau = 0. \end{cases} \tag{3}$$

The above theorem shows that, under the assumption that T_t is a random walk, $Y_t - \mu$ is zero-mean stationary process with cyclical autocovariance function. If we assume that T_t is stationary, then the mean function of $Y_t - \mu$ is almost periodic function with one non-zero frequency λ (see Theorem below).

Theorem 2.2 *If $|\psi_A| < 1$, $|\psi_T| < 1$ and ε_t is a Gaussian white noise with variance σ^2 then $E(Y_t)$ is a "cyclical" almost periodic function with one (non-zero) frequency λ:*

$$E(Y_t) = \mu + Ae^{-\frac{1}{2}\lambda^2\sigma_T^2}\sin\big(\lambda(T + t)\big) + \cos\big(\lambda(T + t)\big)e^{-\frac{1}{2}\lambda^2\sigma_T^2}\sigma_{AT}\lambda,$$

where:

$$\sigma_{AT} = \operatorname{cov}(A_t, T_t) = \alpha_A\alpha_T \frac{\sigma^2}{1-\psi_A\psi_T}, \tag{4}$$

$$\sigma_A^2 = \operatorname{var}(A_t) = \alpha_A^2 \frac{\sigma^2}{1-\psi_A^2}, \tag{5}$$

$$\sigma_T^2 = \operatorname{var}(T_t) = \alpha_T^2 \frac{\sigma^2}{1-\psi_T^2}. \tag{6}$$

Note that the advantage of the above model is the simplicity of the interpretation where amplitude and phase shift are modeled by separate components. In addition, using one source of errors we do not need to use the Kalman filter to compute the likelihood, which makes the statistical inference simpler.

2.3 A model based on unobserved components

In the same way the model based on unobserved components can be introduced. Assuming different sources of errors we propose,

$$\begin{cases} Y_t = & (A + A_{t-1})\sin[\lambda(t + T + T_{t-1})] + \mu + \varepsilon_t \\ A_t = & \psi_A A_{t-1} + \xi_t & \text{deviations from amplitude } A \\ T_t = & \psi_T T_{t-1} + \zeta_t & \text{deviations from phase shift T,} \end{cases} \quad (7)$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, $\xi_t \sim N(0, \sigma_\xi^2)$, $\zeta_t \sim N(0, \sigma_\zeta^2)$ are mutually independent.

Theorems 2.1 and 2.2 can be adopted to the above case.

3 The Bayesian model

In the empirical example the model considered in Theorem 2.1 (i.e. the one with $|\psi_A| < 1, \psi_T = 1$ and ε_t being the Gaussian white noise with variance σ^2) will be analysed. Additionally, due to identification issues, we assume that $T = 0$ and $\lambda T_0 < \Pi$. To complete the Bayesian model we specify the prior distribution of the parameter vector $\theta = (\sigma^2, \mu, A, A_0, \lambda, T_0, \psi_A, \alpha_A, \alpha_P)'$:

$$p(\theta) = p(\mu|\sigma^2)p(A|\sigma^2)p(A_0|\sigma^2)p(\sigma^2)p(T_0|\lambda)p(\lambda)p(\psi_A)p(\alpha_A)p(\alpha_T),$$

where the priors are as follows:

- $p(\mu|\sigma^2) = f_N(\mu|0, c_\mu\sigma^2)$ – the normal distribution with mean 0 and variance $c_\mu\sigma^2$,
- $p(A|\sigma^2) = f_N(A|0, c_A\sigma^2)$ – the normal distribution with mean 0 and variance $c_A\sigma^2$,
- $p(A_0|\sigma^2) = f_N(A_0|0, c_0\sigma^2)$ – the normal distribution with mean 0 and variance $c_{A_0}\sigma^2$,
- $p(\sigma^2) = f_{IG}(\sigma^2|n_\sigma, s_\sigma)$ – the inverse gamma distribution with mean $\frac{s_\sigma}{n_\sigma - 1}$ (for $n_\sigma > 1$) and variance $\frac{s_\sigma^2}{(n_\sigma-1)^2(n_\sigma-2)}$ (for $n_\sigma > 2$),
- $p(T_0|\lambda) = I_{(0,\Pi/\lambda)}(T_0)$ – the uniform distribution over the interval $(0, \Pi/\lambda)$,
- $p(\lambda) = I_{(0,2\Pi)}(\lambda)$ – the uniform distribution over the interval $(0,2\Pi)$,
- $p(\psi_A) = I_{(-1,1)}(\psi_A)$ – the uniform distribution over the interval $(-1,1)$,

- $p(\alpha_A) = f_G(\alpha_A|n_A, s_A)$ – the gamma distribution with mean $n_A s_A$ and variance $n_A s_A^2$,
- $p(\alpha_T) = f_G(\alpha_T|n_T, s_T)$ – the gamma distribution with mean $n_T s_T$ and variance $n_T s_T^2$.

The Bayesian model reads:

$$p(\theta, Y) \propto \lambda \alpha_A^{n_A-1} \alpha_T^{n_T-1} \sigma^{2\left(-n_\sigma-\frac{3}{2}-\tilde{T}\right)} \exp\left(-\frac{\alpha_A}{s_A}-\frac{\alpha_T}{s_T}\right) \times$$

$$\times \exp\left\{-\frac{1}{\sigma^2}\left[s_\sigma + \frac{1}{2}\left(\frac{\mu^2}{c_\mu}+\frac{A^2}{c_A}+\frac{A_0^2}{c_0}+\varepsilon'\varepsilon\right)\right]\right\},$$

where $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{\tilde{T}})'$ with $\varepsilon_t = Y_t - (A + A_{t-1})\sin[\lambda(t + T + T_{t-1})] - \mu =$

$$= Y_t - \left[A + \psi_A^{t-1} A_0 + \alpha_A\left(\sum_{i=0}^{t-2} \psi_A^i \varepsilon_{t-1-i}\right)\right] \sin\left(\lambda\left(t + T_0 + \alpha_T\left(\sum_{i=0}^{t-2} \varepsilon_{t-1-i}\right)\right)\right) - \mu,$$

$\tilde{T}$ denotes the number of observations.

The conditional posterior of σ^2 is the inverse gamma distribution, $p(\sigma^2|\cdot, y) = f_{IG}(\sigma^2|\bar{n}_\sigma, \bar{s}_\sigma)$, where $\bar{n}_\sigma = n_\sigma + \frac{\tilde{T}+3}{2}$, $\bar{s}_\sigma = s_\sigma + \frac{1}{2}\left(\frac{\mu^2}{c_\mu}+\frac{A^2}{c_A}+\frac{A_0^2}{c_{A_0}}+\varepsilon'\varepsilon\right)$, whereas the conditional posteriors of the remaining parameters do not belong to any of the known classes of distributions, and their kernels are as follows:

- $p(\mu|\cdot, y) \propto \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{\mu^2}{c_\mu}+\varepsilon'\varepsilon\right)\right\}$,
- $p(A|\cdot, y) \propto \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{A^2}{c_A}+\varepsilon'\varepsilon\right)\right\}$,
- $p(A_0|\cdot, y) \propto \exp\left\{-\frac{1}{2\sigma^2}\left(\frac{A_0^2}{c_0}+\varepsilon'\varepsilon\right)\right\}$,
- $p(T_0|\cdot, y) \propto \exp\left\{-\frac{\varepsilon'\varepsilon}{2\sigma^2}\right\}$,
- $p(\lambda|\cdot, y) \propto \lambda \exp\left\{-\frac{\varepsilon'\varepsilon}{2\sigma^2}\right\}$,
- $p(\psi_A|\cdot, y) \propto \exp\left\{-\frac{\varepsilon'\varepsilon}{2\sigma^2}\right\}$,
- $p(\alpha_A|\cdot, y) \propto \alpha_A^{n_A-1} \exp\left\{-\frac{\alpha_A}{s_A}-\frac{\varepsilon'\varepsilon}{2\sigma^2}\right\}$,
- $p(\alpha_T|\cdot, y) \propto \alpha_T^{n_T-1} \exp\left\{-\frac{\alpha_T}{s_T}-\frac{\varepsilon'\varepsilon}{2\sigma^2}\right\}$.

In order to obtain the sample from the joint posterior distribution we employ the Gibbs sampler for σ^2 and the Metropolis-Hastings procedure within the Gibbs sampler for the remaining parameters. In order to sample from $p(\lambda|\cdot, y)$ we suggest using a normalized Fourier transform as a proposal density (see e.g. Lenart and Mazur, 2016). According to our experience a gamma distribution, with a shape parameter 2 and a scale parameter equal to the previous state, works well as a proposal in the Metropolis-Hastings steps for α_A and α_T. For μ, A and A_0 random walk Metropolis-Hastings could be used, whereas for $p(\psi_A|\cdot, y)$ and

$p(T_0|\cdot, y)$ uniform distributions over $(-1, 1)$ and $\left(0, {}^{\pi}/_{\lambda}\right)$, respectively, seem to be a good choice.

Conclusions

In this paper we show the stationary nonlinear stochastic cycle model. We investigate the properties of such model. The closed form of the autocovariance function was shown. The autocovariance function is a pseudo periodic function. The proposed model contains the amplitude and phase shift components which allow direct interpretation of this components. The Bayesian inference requires MCMC simulation tools. For most parameters we use the Metropolis-Hastings procedure.

Acknowledgements

We would like to thank two anonymous referees for invaluable suggestions, which greatly improved the paper.

Łukasz Lenart was financed from the funds granted to the Faculty of Finance and Law at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

Justyna Wróblewska would like to acknowledge financial support from research funds granted to Faculty of Management at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

Appendix

Proof of Theorem 2.1 and 2.2. These proofs are based on below Lemmas and are available upon request.

Lemma 3.2 *If* $[X \;\; Y] \sim N\left(\begin{bmatrix}\mu \\ 0\end{bmatrix}, \Sigma\right)$ *where* $\Sigma = \begin{bmatrix}\sigma_{XX} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY}\end{bmatrix}$, *then*

- $E[X\sin Y] = \sigma_{XY} e^{-\frac{1}{2}\sigma_{YY}}$
- $E[X\cos Y] = \mu e^{-\frac{1}{2}\sigma_{YY}}$
- $E[X\exp(iY)] = (\mu + i\sigma_{XY}) e^{-\frac{1}{2}\sigma_{YY}}$
- $E[X^2\exp(iY)] = e^{-\frac{\sigma_{YY}}{2}}(\sigma_{XX} + (\mu + i\sigma_{XY})^2)$.

Proof of Lemma. Since i) and ii) is a natural consequence of iii) we show firstly iii). Note that $X|Y = y \sim N(\mu + y\sigma_{XY}/\sigma_{YY}, \sigma_{XX} - \sigma_{XY}^2/\sigma_{YY})$. Hence,

$$E[X\exp(iY)] = E(E(X\exp(iY)|Y)) = \int_{-\infty}^{\infty} \frac{\exp(iy)\left(e^{-\frac{y^2}{2\sigma_{YY}}(y\sigma_{XY}+\mu\sigma_{YY})}\right)}{\sqrt{2\pi}\sigma_{YY}^{3/2}}\, dy$$
$$= e^{-\frac{\sigma_{YY}}{2}}(\mu + i\sigma_{XY}).$$

The proof of iv) is analogical and therefore it is omitted.

Lemma 3.3 *If* $[X\ \ Y\ \ Z]\sim N\left(\begin{bmatrix}\mu_1\\ \mu_2\\ 0\end{bmatrix}, \Sigma\right)$ *with* $\Sigma = \begin{bmatrix}\sigma_{XX} & \sigma_{XY} & \sigma_{XZ}\\ \sigma_{XY} & \sigma_{YY} & \sigma_{YZ}\\ \sigma_{XZ} & \sigma_{YZ} & \sigma_{ZZ}\end{bmatrix}$, *then*

- $E[XY\sin Z] = e^{-\frac{\sigma_{ZZ}}{2}}(\mu_2\sigma_{XZ} + \mu_1\sigma_{YZ})$
- $E[XY\cos Z] = e^{-\frac{\sigma_{ZZ}}{2}}(\mu_1\mu_2 + \sigma_{XY} - \sigma_{XZ}\sigma_{YZ})$
- $E[XY\exp(iZ)] = e^{-\frac{\sigma_{ZZ}}{2}}\left(\sigma_{XY} + (\mu_1 + i\sigma_{XZ})(\mu_2 + i\sigma_{YZ})\right)$.

Proof of Lemma. Note that $E[XY\exp(iZ)] = E[\exp(iZ)E(XY|Z)]$ and $[X\ \ Y|Z]\sim N(\bar{\mu}, \bar{\Sigma})$, where $\bar{\Sigma} = \begin{bmatrix}\bar{\sigma}_{11} & \bar{\sigma}_{12}\\ \bar{\sigma}_{21} & \bar{\sigma}_{22}\end{bmatrix} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ and $\bar{\mu} = [p\ \ q]^T = [\mu_1\ \ \mu_2]^T + Z\Sigma_{12}\Sigma_{22}^{-1}$ and $\Sigma = \begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}$ is block matrix with Σ_{11} of size 2×2. Hence

$$E(XY|Z) = pq + \bar{\sigma}_{12} = \sigma_{XY} + \left(\mu_1 + \frac{Z\sigma_{XZ}}{\sigma_{ZZ}}\right)\left(\mu_2 + \frac{Z\sigma_{YZ}}{\sigma_{ZZ}}\right) - \frac{\sigma_{XZ}\sigma_{YZ}}{\sigma_{ZZ}}.$$

The next steps are elementary and therefore are omitted.

References

Harvey, A. C. & Trimbur, T. M. (2003). General model-based filters for extracting cycles and trends in economic time series. *Review of Economics and Statistics, 85(2)*, 244–255.

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

Harvey, A. C. (2004). Tests for cycles. *State Space and Unobserved Component Models: Theory and Applications*, 102-119.

Hyndman, R., Koehler, A. B., Ord, J. K. & Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach*. Springer Science & Business Media.

Lenart, Ł., Mazur, B. & Pipień, M. (2016). Statistical analysis of business cycle fluctuations in poland before and after the crisis. *Equilibrium, 11(4)*, 769–783.

Lenart, Ł. & Mazur, B. (2017). Business cycle analysis with short time series: a stochastic versus a non-stochastic approach. In: *The 11th Professor Aleksander Zelias International*

Conference on Modelling and Forecasting of Socio-Economic Phenomena, Conference Proceedings, Cracow: Foundation of the Cracow University of Economics (pp. 212-222).

Lenart, Ł. & Mazur, B. (2016). On Bayesian Inference for Almost Periodic in Mean Autoregressive Models. *Przegląd Statystyczny, 63*(3), 255-271.

Lenart, Ł. & Pipień, M. (2013a). Almost Periodically Correlated Time Series in Business Fluctuations Analysis. *Acta Phisica Polonica A, 123*, 567—583.

Lenart, Ł. & Pipień, M. (2013b). Seasonality revisited-statistical testing for almost periodically correlated stochastic processes. *Central European Journal of Economic Modelling and Econometrics*, *2*(5), 85-102.

Lenart, Ł. & Pipień, M. (2015). Empirical Properties of the Credit and Equity Cycle within Almost Periodically Correlated Stochastic Processes - the case of Poland, UK and USA. *Central European Journal of Economic Modelling and Econometrics*, 7, 168–186.

Lenart, Ł. & Pipień, M. (2017). Non-Parametric Test for the Existence of the Common Deterministic Cycle: The Case of the Selected European Countries. *Central European Journal of Economic Modelling and Econometrics*, *9*(3), 201–241.

Mazur, B. (2017). Probabilistic predictive analysis of business cycle fluctuations in Polish economy. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(3), 435-452.

Pelagatti, M. M. (2015). *Time series modelling with unobserved components*. CRC Press.

Trimbur, T. M. (2006). Properties of higher order stochastic cycles. *Journal of Time Series Analysis*, *27*(1), 1-17.

Automatic identification of competences expected by employers with the use of exploratory text analysis

Paweł Lula[1], Renata Oczkowska[2], Sylwia Wiśniewska[3]

Abstract

Exploratory text analysis allows to identify semantic components present in processing documents. For every component it is possible to describe its character and to evaluate its importance. Using the approach presented above for automatic analysis of job offers it is possible do discover components which are common for all texts and to estimate their importance in every offer. Unfortunately, semantic components obtained with the help of text mining algorithms, usually do not reflect competences within the meaning of specialists from the HR area.

In the paper authors are going to present a method which will be able to identify in a set of job offers semantic components corresponding to professional, social, personal or managerial competences. Also the method of competence description and evaluation will be proposed. The computational model used for the analysis is composed of the two parts. The first is formed by the Latent Dirichlet Allocation Model and identifies latent semantic components. The second element of the model maps semantic components calculated by the *LDA* approach into the set of components related to employers' expectations towards candidates for employment. The proposed method was used for analysis of the corpus containing job offers related to the field of human resources management.

Keywords: *competences, the Latent Dirichlet Allocation Model, job offer exploratory analysis*
JEL Classification: J24, C81
DOI: 10.14659/SEMF.2018.01.26

1 Introduction

Continuous analysis of employers' expectations towards candidates for employment and competences of candidates who are actively looking for a new job can be helpful for all institutions that have impact on modern labour market, particularly for universities which are responsible for development of professional education essential for future employees.

The analysis of employers' expectations can be performed with the use the exploratory analysis methods (Manning and Schütze, 1999), (Salton et. al., 1975). Latent Dirichlet Allocation (*LDA*) method (Blei, 2003), (Alghamdi and Alfalqi, 2015) seems to be useful for identification of semantic components (so called *topics*) in a given set of offers.

[1] Corresponding author: Cracow University of Economics, Department of Computational Systems, 27 Rakowicka Street, 31-510 Kraków, Poland, pawel.lula@uek.krakow.pl
[2] Cracow University of Economics, Department of Human Resources Management, 27 Rakowicka Street, 31-510 Kraków, Poland, renata.oczkowska@uek.krakow.pl
[3] Cracow University of Economics, Department of Human Resources Management, 27 Rakowicka Street, 31-510 Kraków, Poland, sylwia.wisniewska@uek.krakow.pl

Unfortunately, components provided by the *LDA* algorithm not always directly represent features which are taken into account by employers and by people responsible for employment policy. Therefore it is necessary to map *LDA* components into the space of competences which is defined by a set of features significant for labour market researchers and analysts.

The method proposed in the paper assumes that components calculated by the use of the *LDA* method are transformed by the classification model with class overlapping which generates labels and evaluates competences' importance by calculation assignment coefficients for every competence. In our approach a classification model has been built in a supervised mode, with the use of a training set retrieved from the *pracuj.pl* web site.

In the second section of the paper some definitions of competences are discussed. Also a competence taxonomy is presented. Next, in the third section, a proposed method of competences' identification is proposed. Some exemplary analysis are shown in the fourth part of the paper. Final conclusions are formulated in the last part of the paper. All calculations presented in the article were performed in R language.

2 The essence of competences

Studies on the subject literature indicates that currently prevail two approaches, according to which the concept of employee competences is interpreted. According to the first approach, competences are seen as people's characteristics, which account for a basis of desired behaviour at work, allowing them to achieve intended results. Competences are therefore understood as the ability to implement specific patterns of behaviour. According to the second approach, competences are understood as characteristics of a professional position. This interpretation defines competences as the ability to effectively perform professional duties in compliance with standards established by the organisation or to achieve desired results (Whiddett and Hollyforde, 2003).

The first approach is presented by Whiddett and Hollyforde (2007). They define competences as "the behaviours that individuals demonstrate when undertaking job relevant tasks effectively within a given organisational context". According to the second approach, the concept of competence is defined by Wright et. al. (2003). In their opinion, competence is "the ability to perform activities within an occupation or function to the standards expected in employment".

Analysis of the subject literature point at the diversity of typologies of competences. Among the most popular is the classification of Filipowicz (2014). He distinguishes the following types of competences:

1. social – determining the quality performed tasks associated with contacts with people (e.g. commercial contacts). The level of these competences determines the effectiveness of communication, cooperation and influencing other people. Sample social competences include, among others: oral communication, written communication, teamwork, building of relations, sharing of knowledge and experience;
2. personal – related to performance of tasks by the employee, and their level affects the quality of the performed tasks. They also determine the adequacy in actions and the speed of their performance. Personal competences include, among others: innovativeness, entrepreneurship, flexibility, organization of own work, time management, problem-solving, and stress management;
3. managerial – involve human resource management, both with soft areas of management, work organisation, as well as with strategic aspects of management. The effectiveness of management of the organisation's employees is determined by the level of these competences. Exemplary managerial competences include, among others: analytic thinking, strategic thinking, motivating, delegation of tasks, team-building;
4. professional (specialist, technical) – concern specialist tasks set for particular groups of positions. They are often connected with specific scopes of knowledge or skills. The level of these competences is reflected in the effectiveness of implementation of tasks typical of a given profession, position or performed function. Examples of professional competences include: process management, project management, professional knowledge, the ability to use modern information technologies.

The classification of competences presented above will be used as the basis of the computational model described in the next section.

3 Identification of competences in textual job offers

Automatic mapping job offer content into the space defined by competences relevant to human resources management specialists and labour market researchers and analysts is the main aim of the study (Oczkowska et. al., 2017). The proposed model has compound character and is composed of two sub-models:

- the Latent Dirichlet Allocation model used as a filter for erasing irrelevant elements from documents and for documents mapping into the space of semantic components (*topics*). However, topics identified by the *LDA* algorithm usually are not consistent with competences which are studied in the area of human resources management,
- the classification model with overlapping classes which is responsible for mapping documents from the space defined by *LDA* topics to the space of competences.

The structure of the model is presented in the Fig. 1.

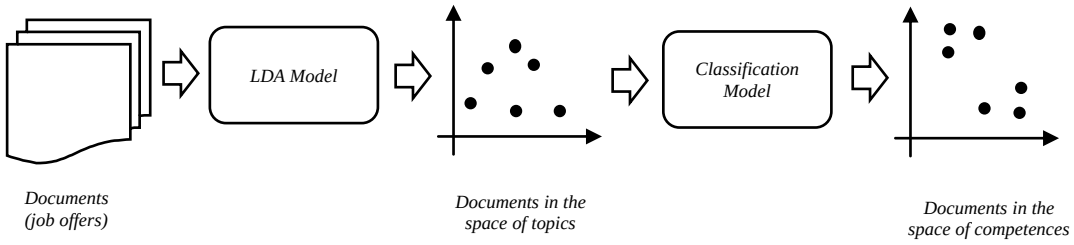


Fig. 1. The structure of the model mapping job offers into the space of competences.

3.1 The Latent Dirichlet Allocation Model

The process of analysis starts with the job offers processing with the *Latent Dirichlet Allocation* (*LDA*) model.

The *LDA* algorithm analyses the corpus which is composed of *LD* documents (it was assumed that one document from the corpus contains one job offer):

$$\boldsymbol{D} = [d_1 \quad \dots \quad d_{LD}]^T. \tag{1}$$

Words used in documents form the dictionary $\boldsymbol{V}$ containing *LV* terms:

$$\boldsymbol{V} = [v_1 \quad \dots \quad v_{LV}]^T \tag{2}$$

The *LDA* model identifies LT semantic components (topics) which reflect main issues appearing in the corpus. Topics constitute $\boldsymbol{T}$ structure:

$$\boldsymbol{T} = [t_1 \quad \dots \quad t_{LT}]^T \tag{3}$$

The results of the *LDA* models are presented as two matrices $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$. The $\boldsymbol{\Phi}$ matrix defines all topics identified by the *LDA* algorithm. Every topic is represented by the probability distribution over words taken from the dictionary $\mathbf{V}$. The $\boldsymbol{\Phi}$ matrix has a following form:

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi_{1,1} & \cdots & \phi_{1,LV} \\ \cdots & \cdots & \cdots \\ \phi_{LT,1} & \cdots & \phi_{LT,LV} \end{bmatrix} \tag{4}$$

where $\phi_{i,j}$ represents the probability of occurrence of the j-th word from the dictionary $\mathbf{V}$ in the i-th topic.

The $\mathbf{\Theta}$ matrix informs about the contribution of every topic to every document and has a following form:

$$\mathbf{\Theta} = \begin{bmatrix} \theta_{1,1} & \cdots & \theta_{1,LT} \\ \cdots & \cdots & \cdots \\ \theta_{LD,1} & \cdots & \theta_{LD,LT} \end{bmatrix} \tag{5}$$

where $\theta_{i,j}$ indicates the probability of occurrence of the j-th topic in the i-th document.

3.2 The Classification Model

The classification model should map a given document represented as a point in the space of *LDA* topics into the space of competences. Taking into account that one job offer includes requirements related to many competences, the classification model should calculate assignment coefficients representing the contribution of every competence to a given document.

Establishing the set of competences which should be analysed by the model constitutes the first step in the process of classification model building. The set of competences includes LC elements and can be defined as:

$$\boldsymbol{C} = \{C_1, C_2, \ldots, C_{LC}\}. \tag{6}$$

In contrast to the *LDA* model, the classification model was estimated in the supervised mode with the use of LC text files, where file f_i contains exemplary sentences derived from job offers related to the competence C_i. For every competence, on the basis of a file with description of requirements related to this competence, the probability distribution over words has been calculated. The result forms the $\mathbf{\Gamma}$ matrix:

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{1,1} & \cdots & \gamma_{1,LV} \\ \cdots & \cdots & \cdots \\ \gamma_{LC,1} & \cdots & \gamma_{LC,LV} \end{bmatrix} \tag{7}$$

where $\gamma_{i,j}$ is the probability of occurring of the j-th word in the description of the i-th competence.

Next the similarity of every *LDA* topic to every competence was calculated. In the current study the cosine similarity was used.

$$\delta^*_{i,j} = \frac{\sum_{v=1}^{LV} \phi_{i,v}\gamma_{j,v}}{\sqrt{\sum_{v=1}^{LV} \phi_{i,v}^2}\sqrt{\sum_{v=1}^{LV} \gamma_{j,v}^2}} \tag{8}$$

Similarity coefficients form a matrix with LT rows and LC columns. Dividing every element of this matrix by the sum of row elements, the $\mathbf{\Delta}$ matrix was obtained:

$$\mathbf{\Delta} = \begin{bmatrix} \delta_{1,1} & \cdots & \delta_{1,LC} \\ \cdots & \cdots & \cdots \\ \delta_{LT,1} & \cdots & \delta_{LT,LC} \end{bmatrix} \tag{9}$$

where:

$$\delta_{i,j} = \frac{\delta^*_{i,j}}{\sum_{k=1}^{LC} \delta^*_{k,j}} \tag{10}$$

The $\delta_{i,j}$ value represents the distribution of the j-th competence in the i-th *LDA* topic.

Using the document representation in the topic space, the contribution of the j-th competence in the i-th job offer can be expressed as:

$$s_{i,j} = \sum_{k=1}^{LT} \theta_{i,k} \delta_{k,j} \tag{11}$$

Finally, the description of documents in the space of competences is defined as the **K** matrix:

$$\mathbf{K} = \begin{bmatrix} \kappa_{1,1} & \cdots & \kappa_{1,LC} \\ \cdots & \cdots & \cdots \\ \kappa_{LD,1} & \cdots & \kappa_{LD,LC} \end{bmatrix} \tag{12}$$

where element:

$$\kappa_{i,j} = \frac{s_{i,j}}{\sum_{k=1}^{LC} s_{i,k}} \tag{13}$$

defines the importance of the j-th competence in the i-th offer.

4 The analysis of the competence expectations based on the exploratory analysis of job offers related the HR management area

During the study, a set of job offers related to the human resources management area was analysed. All offers were retrieved from the *pracuj.pl* web portal. For offers retrieving the *rvest* package was used (Wickham, 2016). The corpus was prepared with the help of the *tm* package (Feinerer et. al., 2008).

The following set of competences was taken into account:

$$\boldsymbol{C} = \{M_1, P_1, P_2, P_3, S_1, S_2, T_0, T_1, T_2, T_3, T_4, T_5\} \tag{14}$$

where:

- M_1 – managerial competences,
- P_1 – personal competences (innovativeness, problem solving ability, dealing with stress, ...),
- P_2 – possession of a university diploma,
- P_3 – possession of a driving licence,
- S_1 – communication competences (oral and written, also in foreign languages),
- S_2 – social competences,
- T_0 – general competences and experience in HR area,

- T_1 – legal and organizational aspects of HR management,
- T_2 – competences in recruitment and training,
- T_3 – IT competences,
- T_4 – competences in management, accounting and logistics,
- T_5 – competences in sales and customer relationship management.

Next, with the use of randomly chosen offers prepared in Polish, the matrix **Γ** was calculated. As an example, in the Fig. 2 the distribution of the most important words for the T_1 competence was presented.

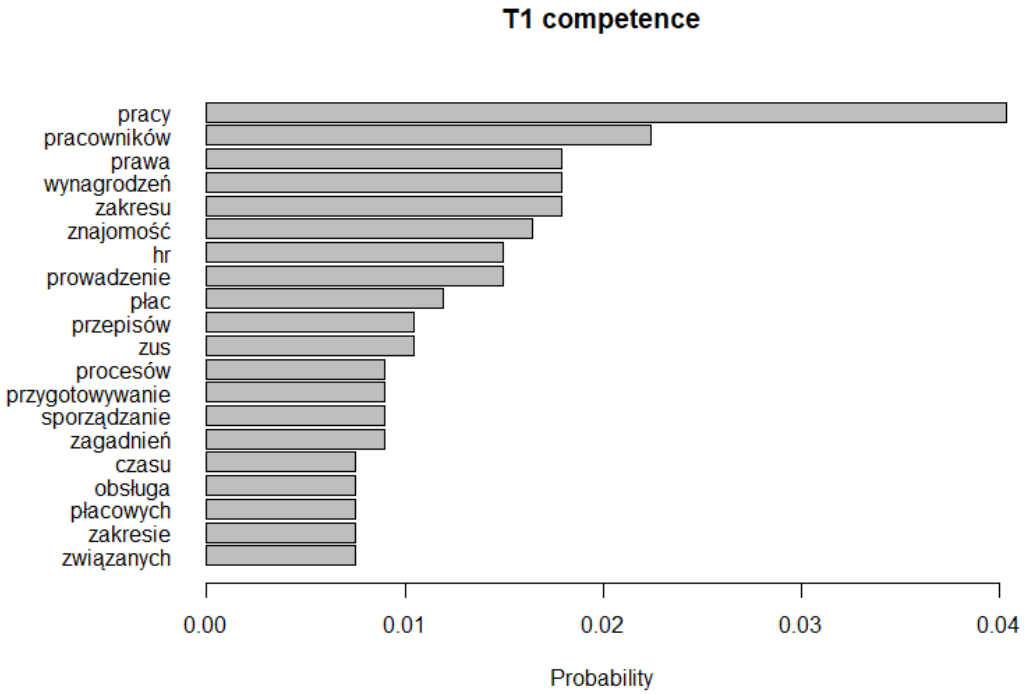


Fig. 2. The distribution of the most important words for the competence T_1.

Next the full set of job offers related to HR area was analysed with the use of the *LDA* algorithm. Calculations were performed with the help of the *topicmodels* package (Grün, Hornik, 2011). During the calculation process 40 topics were identified. Next the topic-competence similarity matrix was estimated with the use of the cosine formula. In the last step of calculation the **K** matrix was generated. To show the results the analysis of an exemplary job offer was performed. The test was prepared in Polish:

"Zadania: Zarządzanie bazą kandydatów. Tworzenie i publikacja ogłoszeń rekrutacyjnych. Selekcja aplikacji. Kontakt z kandydatami. Zarządzanie kalendarzem rekrutacji. Udział w rozmowach rekrutacyjnych. Generowanie raportów rekrutacyjnych. Wymagania: Znajomość języka angielskiego na poziomie pozwalającym bezproblemową komunikację. Zainteresowanie tematyką z zakresu rekrutacji. Bardzo dobre umiejętności interpersonalne. Bardzo dobra organizacja pracy. Mile widziane doświadczenie w pracy w dziale HR. Pracę w młodym, dynamicznym i otwartym na niekonwencjonalne rozwiązania zespole. Pracę w branży nowych technologii. Możliwość rozwoju w ramach wewnętrznych struktur organizacyjnych Integrację w i poza pracą"

The significance of competences in the above offer is presented in the Fig. 3.

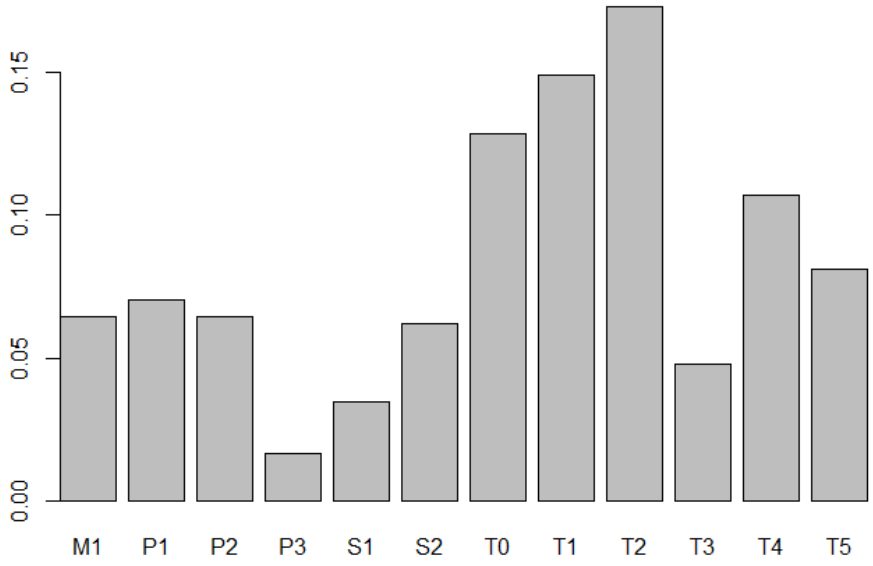


Fig. 3. The significance of competences in the exemplary job offer.

The results shows that competences T_0, T_1 and T_2 were recognized as the most important in an offer.

Conclusions

In the paper the method for job offer mapping from the space of semantic components (topics) identified by the *LDA* algorithm to the space of competences related to the HR area

was proposed. Experience gained from the study allows to point out main advantages of the proposed procedure. Using the method proposed in the paper job offer are presented in the space of competences instead of the space of *LDA* topics. It facilitates the process of automatic interpretation of employers' expectations towards candidates for employment. The proposed model is composed of two sub-models: *LDA* model for topics identification and classification model for mapping documents into the space of competences. These two elements are mutually independent. It means that changes made in one of them do not results in changes of the other. The *LDA* model works as a filter which reduces information noise in the input data by transforming original job offer from the space of words to the space on topics. It seems that the proposition presented in the paper can be helpful for analysis of processes existing on the labour market and can refine education policy of universities and other educational institutions.

Unfortunately, also some disadvantages were noted. Firstly, it should be underline that model creation process may be time-consuming. It is caused by the necessity of manual creation of text files into training set used for classification model building. Secondly, the adjustment of main model parameters (including established taxonomy of competences, the number of topics for the *LDA* model and the formula for similarity calculation between topics and competences) may be demanding and may extent the time necessary for model preparation. All calculations presented in the paper can be performed with the use of R software.

Acknowledgements

The project has been funded by the Cracow University of Economics within the Statutory Fund of the Faculty of Management.

References

Alghamdi, R. & Alfalqi, K. (2015). A Survey of Topic Modeling in Text Mining. *International Journal of Advanced Computer Science and Applications*. *6*(1), 147-153.

Blei, D., Ng, A. & Jordan, M., (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, (3), 993–1022.

Feinerer, I., Hornik, K. & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, *25*(5), 1-54.

Filipowicz, G. (2014). *Zarządzanie kompetencjami. Perspektywa firmowa i osobista*. Warszawa: Oficyna Wolters Kluwer Business.

Grün, B. & Hornik, K., (2011). Topicmodels: An R Package for Fitting Topic Models. *Journal of Statistical Software*, *40*(1), 1–30.

Manning, C. D. & Schütze H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA. MIT Press.

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.

Salton, G., Wong, A. & Yang, C., (1975). Vector-Space Model for Automatic Indexing. *Communications of the ACM*. *18*(11), 613-620.

Whiddett, S. & Hollyforde, S. (2003). *A Practical Guide to Competencies. How to Enhance Individual and Organisational Performance*. London: CIPD.

Whiddett, S. & Hollyforde, S. (2007). *Competencies*. London: CIPD.

Wickham, H., (2016), *Package 'rvest'*, https://cran.r-project.org/web/packages/rvest/rvest.pdf

Wright, M., Turner, D. & Horbury, C. (2003). *Competence Assessment for the Hazardous Industries. RR086*. Sudbury: HSE Books. http://www.hse.gov.uk/research/rrpdf/rr086.pdf

Identification of factors affecting the popularity of cross-border online shopping in the European Union countries

Radosław Mącik[1]

Abstract

The popularity of cross-border online purchases among consumers in EU countries is strongly diversified (between 3% and 81% of Internet users). The level of Internet penetration in a given country and the level of involvement in online shopping measured by the percentage of online shoppers do not adequately explain this diversification at EU countries level. Paper goal is to explain the level of the percentage of cross-border online shoppers at 34 European countries level, including 28 EU countries, 2 EEA members and 4 EU candidates, using the Classification And Regression Trees (CART) approach. Data used include the Eurostat quantitative data, supplemented with mystery shopping study data commissioned by the European Commission, indicators for local B2C e-commerce markets, and qualitative variables - e.g. sharing an official language, being insular country. The main identified factors affecting the level of cross-border online shopping are geo-blocking practices of online sellers, level of online-shopping from national sellers, having a common/similar language with another country, and EU membership. Unexpectedly, the existence of the local language version of AliExpress.com website coincides with the higher level of cross-border online shopping.

Keywords: *online cross-border shopping, B2C e-commerce, geo-discrimination, EU countries, CART analysis*
JEL Classification: D12, O33, C38
DOI: 10.14659/SEMF.2018.01.27

1 Introduction

During last 20 years, B2C and B2B e-commerce transformed the retail landscape substantially all over the world. Merchants adopted the internet technologies successfully and created online channel not only substituting the physical one but also making both channels complementary (Mącik, 2015). Easy access to various goods, easy comparison of prices and products, extended time to return goods, are main advantages of online shopping for consumers. Globalisation and economic integration made e-commerce activities not limited to domestic markets (Schu et al. 2016), allowing the retailers operate globally, also via global marketplaces (Jiang et al. 2011), influencing international logistics growth.

The cross-border online shopping still faces many obstacles, mainly legal, financial and logistics related ones. Even in economically integrated zones like the EU Member States, the consumers are frequently unable to shop online abroad, being discriminated in their consumer

[1] Corresponding author: Maria Curie-Sklodowska University/Faculty of Economics, Chair of Marketing, 5 Maria Curie-Sklodowska Square, 20-031 Lublin, Poland, radoslaw.macik@umcs.pl.

rights. Those practices conflict with the fundamental European Union freedoms, particularly with free movement of goods. Some difficulties can arise when the shopper's country of residence and the seller location are in countries without free trade agreements or customs unions, but as Chinese and United States sellers prove, those issues are easy to overcome, in contrast to prolonged time for delivery (Cho and Lee, 2017).

The aim of this paper is gaining a better understanding of country-level differences in online cross-border shopping for European countries including 28 EU Member States and a few EU candidates as well as two members of European Economic Area (EEA). Determinants of online cross-border shopping based on cross-national data are explored and explained using the Classification And Regression Trees (CART) approach.

2 Cross-border online shopping – a brief literature review

The term cross-border online shopping is the analogue to offline (traditional) cross-border shopping, differing only in the channel used (online instead offline). Such activity involves cross-border transactions, based on selling online physical or digital products to consumers with (legal) residence in another country than seller's (legal) location (Ballard and Lee, 2007). Similar terms such: overseas direct purchases, overseas online shopping, foreign online shopping, and cross-border e-sales are also used (Cho and Lee, 2017). The author proposes to distinguish between using "cross-border" term (for purchases in the neighbouring country or within the continent – e.g. Europe) from the word "overseas" fitting better (when real overseas shipping is involved – e.g. Asia to Europe). This is more than a semantic difference because shopping online is typically perceived easier when countries have similar geographic and cultural characteristics (Gomez-Herrera et al., 2014).

In effect, the level of cross-border online shopping varies substantially from country to country, even within Europe and EU-members. Online consumer markets within EU are far from the common market idea (European Commission, 2015). The differences in popularity of online cross-border purchases are the result of demand-side and supply-side factors. For demand side, cost related factors and connected with the time of delivery as well as mentioned legal difficulties (taxes and duties), possible language difficulties and perception of being discriminated by the foreign seller on the base of nationality or location, are favouring shopping from national sellers. From the supply side, high consumers involvement in cross-border online shopping may come from the size/structure of the national market, e.g. in a case of small (Luxembourg) or insular countries (Malta, Cyprus), where operating a separate national e-commerce infrastructure is not feasible. Also countries sharing a common language

are expected to have a higher involvement in online shopping abroad (Gomez-Herrera et al., 2014).

On the microeconomic level, the still common trade agreements between producers and resellers lead to granting territorially limited rights for distribution of products and brands (leading to grey online markets and lateral import). Even for digital goods, where instant delivery is possible, the intellectual property rights and different taxation (e.g. for movies, music), make international sales impossible or unwanted. Legal and logistic barriers in cross-border e-commerce arise for physical goods involving higher shipping cost, increasing the risk of damaging parcels in transit (Kawa, 2017), and requiring international payments handling, with the duties and taxes deduction (Gomez-Herrera et al., 2014). Also, sellers' self-limitation practices (treating foreign purchases as unwanted from technical, logistics, cost or communication reasons) lead to geo-discrimination practices experienced by consumers. Geo-discrimination takes the form of geo-blocking - the practices of automatically limiting access (or changing the terms of sale) by geo-localisation of the user's device connecting to seller's website. Less automated practices based on address filled in delivery form also apply. Mącik (2017) provides a detailed discussion about geo-discrimination of consumers.

3 Method

Paper uses the exploratory approach with secondary data analysis. The research question is:

RQ: *What factors are explaining the differences in popularity of cross-border online shopping on a country level in Europe, with the focus of European Union member states?*

Data used include quantitative secondary data retrieved from the Eurostat (data series: isoc_ec_ibuy), including the percentages of internet users buying online from national sellers and from abroad. It was supplemented with geo-blocking prevalence data retrieved from mystery shopping study (European Commission, 2016), and other selected indicators including qualitative variables, dichotomously coded (Yes/No) – e.g. the insular country location, small population country (<1 million of inhabitants), EU membership, common or similar official language to another country, and also having Amazon.com branch with national website and/or having AliExpress.com website in local language.

The dependent variable to explain was the level of popularity of cross-border online shopping on country level measured as the percentage of Internet users declaring making at least one cross-border transaction to buy goods or services over the internet. Collected data refer to 28 EU Member States and six other countries: two EEA members (Iceland and Norway), and four EU Candidates (Turkey, Serbia, Montenegro, FYR of Macedonia).

The Classification And Regression Trees (CART) approach (Breiman et al., 1984) is a primary analytical method used (as the number of objects is low in relation to the independent variables, not linear dependencies are observed, and no natural clusters are identified). CART analysis is nonparametric and robust to outliers. Technically the algorithm grows a large tree and then prunes it to a size having lowest cross-validation error (Loh, 2014).

4 Results

Buying online activities level (regardless domestic or abroad) varies substantially from country to country within Europe and EU-members (Fig. 1).

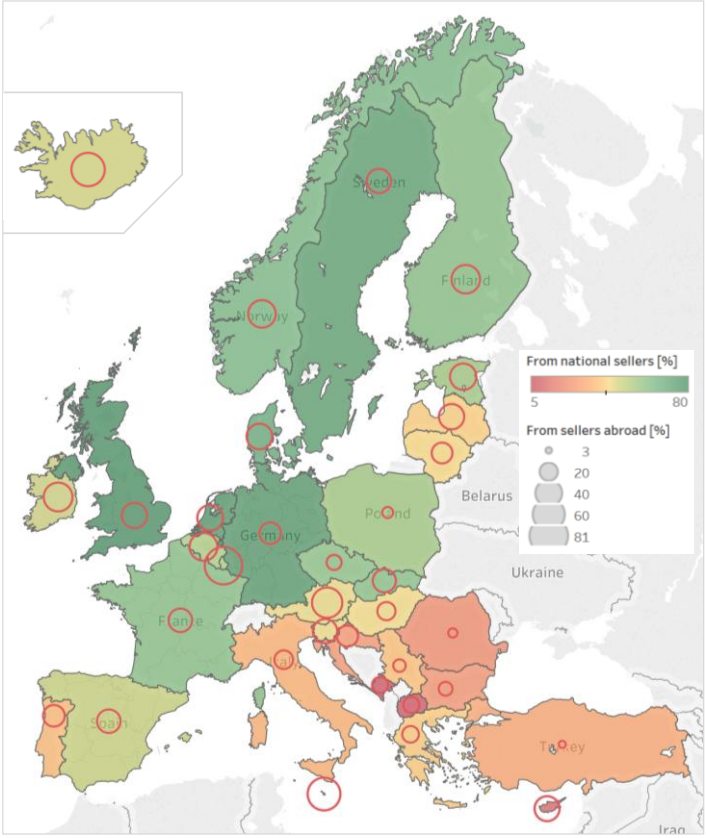


Fig. 1. Online purchases (from national sellers and from abroad) in 2017 – map.

The popularity of cross-border online shopping in 2017 in European countries according to Eurostat estimation ranged between 3% and 81% (for Turkey and Luxembourg respectively). Mean for this variable was 31.5% with standard deviation equal 17.9%, and median equal 32%. In four countries (Serbia, Poland, Romania, Turkey) cross-border online shopping is practised by less than 10% of Internet users. At the same time between 5% and 80% of internet users (for Montenegro and United Kingdom respectively) were shopping

online from national sellers. For UE members mean and median for such activities were 45.5% with standard deviation 20.3%, being the lowest for Cyprus (11%) and Malta (18%).

Both variables are assumed as positively correlated on a country level, but according to Eurostat, this correlation weakens with time, for 2017 being not significant (Pearson r=0,3, p=0,085). Also, the not linear relationship between online purchases from national sellers and abroad seems to be weak (Fig. 2), suggesting seeking other explanatory variables.

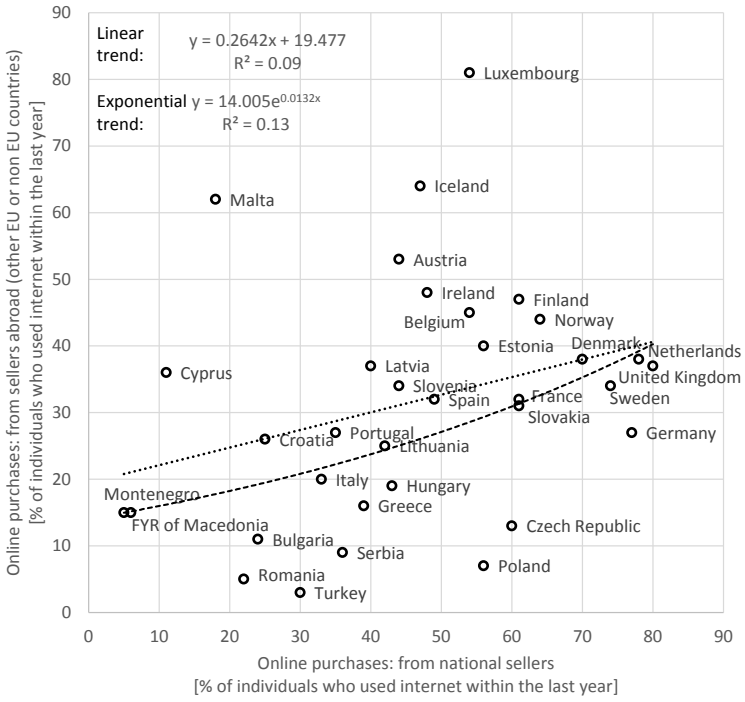


Fig. 2. Online purchases (from national sellers and from abroad) in 2017.

As Gomez-Herrera et al. (2014) and Mącik (2017) research suggest, among possible variables explaining the level of cross-border online shopping can be the prevalence of geo-blocking (data available for EU Members only), and qualitative variables described earlier. The prevalence of geo-blocking the purchases of consumers from abroad is very high – in average in 68.5% of cases making cross-border online purchase encountered difficulties (with standard deviation = 6.2%, median 69%, min. 57% (for the UK), and max. 78% (for Latvia)).

As the number of observations is low, and most of the independent variables are qualitative, the Classification And Regression Trees (CART) analysis with V-fold cross-validation has been performed including for EU members the geo-blocking prevalence, and shopping online from national sellers for all 34 countries as quantitative explanatory variables, and set of qualitative variables. As analysis goal was to explain mainly the lowest and the highest values for cross-border online shopping, the response variable has been coded

into three classes: “low” (<15%), “average” (16-41%) and “high” (>=42%). Optimal trees using Breiman et al. (1984) procedure involving costs of classification were found (Fig. 3, Fig. 6).

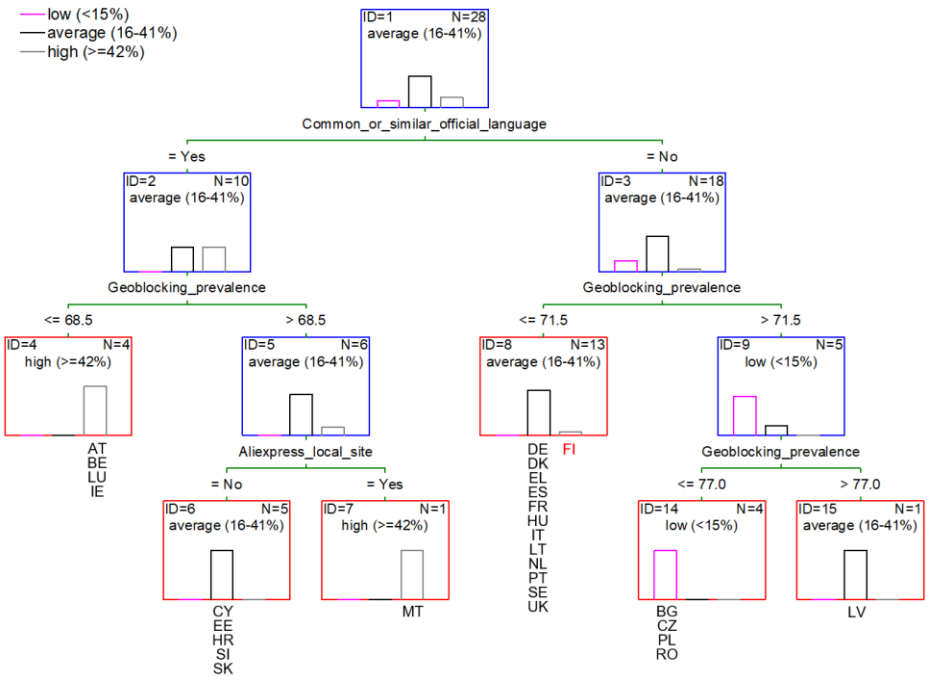


Fig. 3. Classification tree for EU Member countries – dependent variable: Online purchases from sellers abroad (other EU or non-EU countries) as % of individuals who used internet within the last year (country name shortcuts according to ISO 3166 alpha-2 format).

The classification for EU Member countries is efficient, only one case (Finland) is misclassified (to “average” class instead “high” one - Fig. 4). That gives classification accuracy of 96.4%. The “high” value of cross-border online shopping popularity is connected mainly with having common or very similar official language with another country and lower than 68.5% prevalence of geo-blocking (cases of Austria, Belgium, Luxemburg and Ireland). Alternatively sharing a common language, higher geo-blocking prevalence than 68.5% and the existence of Aliexpress.com site in local language also gives the same classification result (case of Malta). For “low” class of cross-border online shopping popularity most likely leads having own national official language and geo-blocking prevalence between 71.5% and 77% (Bulgaria, Czech Republic, Poland and Romania are classified to this group).

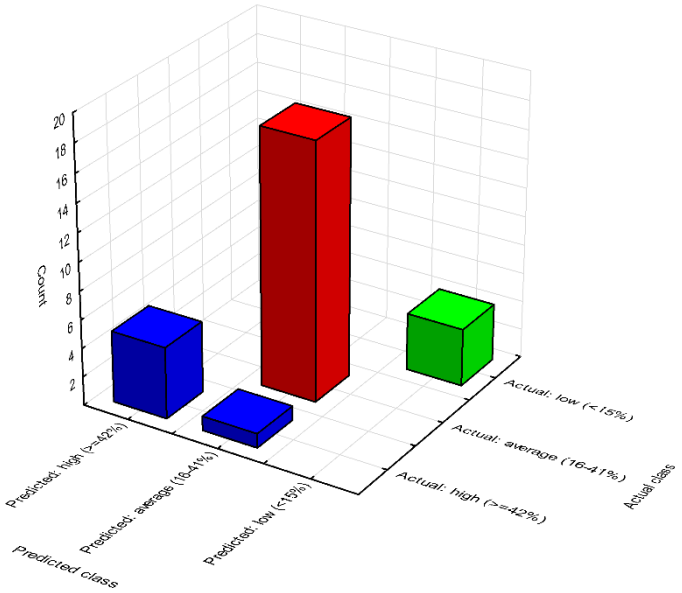


Fig. 4. Classification effectiveness for the tree from Fig. 3.

In classification, main role played numerical variable the prevalence of geo-blocking, and the existence of localised websites of global online marketplaces (particularly AliExpress.com) (Fig. 5). Smaller importance had being very small country (<1 million of inhabitants), insular location and common or similar official language with another country.

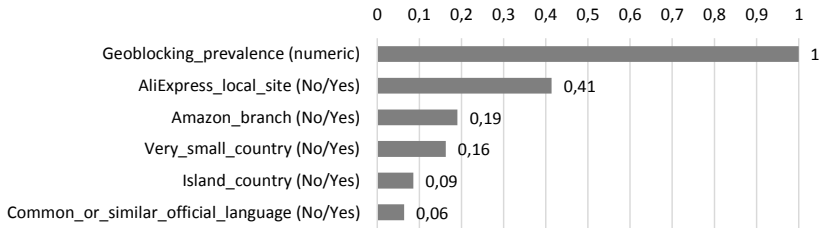


Fig. 5. The importance of predictors in classification from Fig. 3.

In the second case (34 countries) the percentage of internet users shopping online from national sellers substituted the prevalence of geo-blocking because of lack of data for six non-EU countries. This classification is less efficient (with an accuracy of 82.4%) – only the "average" group is classified well. More often countries were misclassified to "average" class instead "high" one, than to "average" class instead of "low" (Fig. 7).

For 34 countries case, the EU countries are firstly separated from non-EU ones. In a non-EU group, the level of cross-border online shopping depends on the online shopping from national sellers. Values >41,5% result in "high" activity in cross-border e-commerce (Iceland, Norway), values <41,5% create "low" group (Montenegro, FYRoM, Serbia and Turkey). For

EU countries lower values of the of cross-border online shopping from higher ones are separated by common/similar official language. Lack of common language connected with domestic online shopping <28.5% leads to "low" class for cross-border online purchases (Bulgaria, Romania). For common/similar language criterion met, the result is "average" or "high" (more likely if AliExpress.com site in the local language exists (Ireland, Malta)).

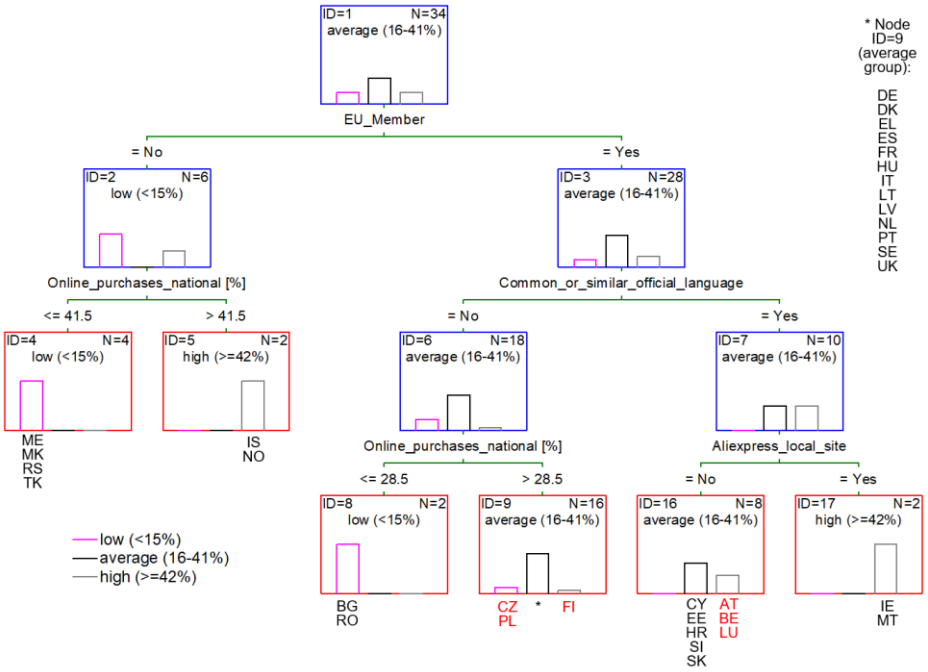


Fig. 6. Classification tree for 34 European countries – dependent variable: as in Fig. 3 (country name shortcuts according to ISO 3166 alpha-2 format).

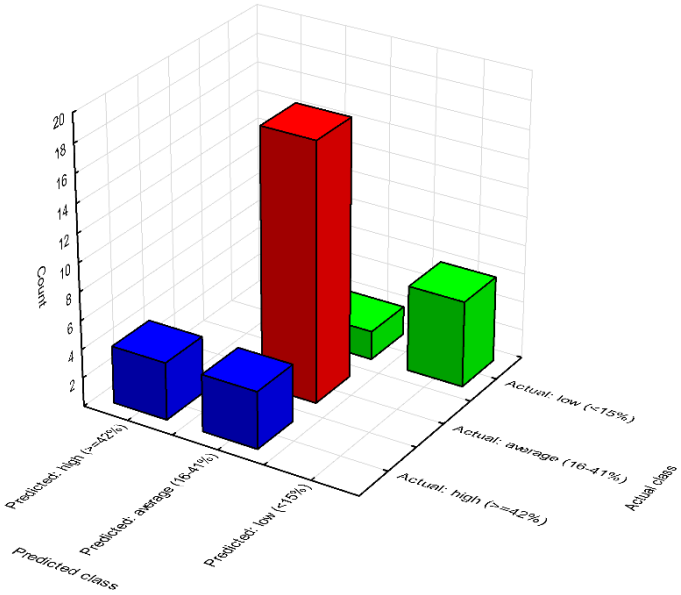


Fig. 7. Classification effectiveness for the tree from Fig. 6.

Unfortunately, three smaller countries with a common language (Austria, Belgium and Luxembourg) are classified in "average" group instead "high" one. Also, Poland and the Czech Republic are classified as "average" instead "low", because of having quite a high level of domestic online purchases activity. Finland is misclassified in both analyses.

The primary role in classification for 34 countries case played numerical variable: the percentage of Internet users shopping online from national sellers. Categorical variables: being insular country and the existence of localised websites of global e-commerce marketplaces (in that case more likely Amazon.com), as well as sharing common language, and being EU Member country were less important for classification (Fig. 8).

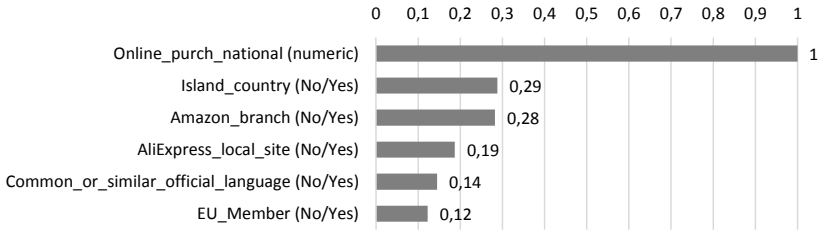


Fig. 8. The importance of predictors in classification from Fig. 6.

Conclusions

Obtained results for EU countries confirm the supposition that higher geo-blocking prevalence discourages consumers from engagement to cross-border online shopping (the case of larger Central European countries). Lower geo-blocking level with sharing common/similar official language with another country favour higher level of cross-border online shopping on a country level (particularly for smaller countries). In case of 34 European countries, the accuracy of classification is worse, although it is confirmed that the level of online shopping from national sellers leads to higher percentage of Internet users buying online abroad, particularly for six non-EU countries. Having common/similar official language, as well EU membership allows predicting the level of cross-border online shopping on the country level as "average" or "high". An unexpected finding is that the access to Aliexpress.com website in official country's language may lead to the substantially higher use of cross-border online shopping (for Malta and Ireland) – this requires further examination.

Acknowledgements

Presented research has been made without any external support to the author.

References

Ballard, C. L. & Lee, J. (2007). Internet Purchases, Cross-Border Shopping, and Sales Taxes. *National Tax Journal*, *60*(4), 711–725.

Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Cho, H. & Lee, J. (2017). Searching for Logistics and Regulatory Determinants Affecting Overseas Direct Purchase: An Empirical Cross-National Study. *The Asian Journal of Shipping and Logistics*, *33*(1), 11–18.

European Commission. (2015). *Consumer attitudes towards cross-border trade and consumer protection* (Flash Eurobarometer No. 397). European Commission. Retrieved from http://ec.europa.eu/COMMFrontOffice/publicopinion/index.cfm/ResultDoc/download/DocumentKy/69149

European Commission. (2016). *Mystery shopping survey on territorial restrictions and geo-blocking in the European Digital Single Market. Final Report*. European Commission. Retrieved from http://ec.europa.eu/consumers/consumer_evidence/market_studies/docs/geoblocking_final_report_2016_en.pdf

Gomez-Herrera, E., Martens, B. & Turlea, G. (2014). The drivers and impediments for cross-border e-commerce in the EU. *Information Economics and Policy*, *28*, 83–96.

Jiang, B., Jerath, K. & Srinivasan, K. (2011). Firm Strategies in the "Mid Tail" of Platform-Based Retailing. *Marketing Science*, *30*(5), 757–775.

Kawa, A. (2017). Supply Chains of Cross-Border e-Commerce. In: Król, D., Nguyen, N. T. & Shirai, K. (eds.), *Advanced Topics in Intelligent Information and Database Systems* (Vol. 710, pp. 173–183). Cham: Springer International Publishing.

Loh, W.-Y. (2014). Fifty years of classification and regression trees. *International Statistical Review*, *82*(3), 329–348.

Mącik, R. (2015). Substytucyjność i komplementarność fizycznego i wirtualnego kanału zakupu. *Annales Universitatis Mariae Curie-Skłodowska. Sectio H. Oeconomia*, *49*, 99–111.

Mącik, R. (2017). Geo-discrimination in Online Shopping. The Consumer's Perspective. *Handel Wewnętrzny*, *63*(6), 214–223.

Schu, M., Morschett, D. & Swoboda, B. (2016). Internationalization Speed of Online Retailers: A Resource-Based Perspective on the Influence Factors. *Management International Review*, *56*(5), 733–757.

The health transition and an ageing society – example of selected European countries

Justyna Majewska[1], Grażyna Trzpiot[2]

Abstract

Population ageing is a natural and inevitable process in each country with a highly developed economy. Countries from the western part of Europe have been struggling with this challenge from a long time. Countries from eastern and central Europe have been preparing for this challenge in the last few decades. In most of the European countries over 20% of people are over age 60. According to national statistical offices projections by three decades this ratio is going to double. Such an age structure of population has serious economic and social consequences – for pension system and health protection above all. Changes in mortality and disease profiles in conjunction with progress in health and socio-economic development allow for analysis of an epidemiological transition, and in broader notion – health transition. The aim of this article is to analyze the health status and mortality causes among people aged 65 years and over in selected European countries that are in different stages of health transition. In the article selected health measure and method of decomposing differences in expected life expectancy are used.

Keywords: *health transition, ageing, life expectancy, DALYs*
JEL Classification: J10, J11, C51
DOI: 10.14659/SEMF.2018.01.28

1 Introduction

We are observing an increase in the share of people older people in relation to the general population in Europe, so the discussion on demographic, epidemiological and health transition is extremely timely and necessary. The demographic transition is said in the situation of the transition of society from a high to a low reproduction rate and from a high to a low mortality rate. The demographic transition is accompanied by an epidemiological transition, which led to a change in the distribution of causes of death (more often we deal with degenerative diseases than infectious diseases) and an increase in mortality in later life. However, the health transition is related to the improvement of medical care (its effectiveness, change of orientation - from treatment to prevention), lifestyle change and positive economic

[1] University of Economics in Katowice / Department of Demography and Economic Statistics, ul 1 Maja 50, 40-287 Katowice, grazyna.trzpiot@ue.katowice.pl
[2] Corresponding author: University of Economics in Katowice / Department of Demography and Economic Statistics, ul 1 Maja 50, 40-287 Katowice, justyna.majewska@ue.katowice.pl

and social changes resulting in longer life span, but more importantly - life expectancy in health.

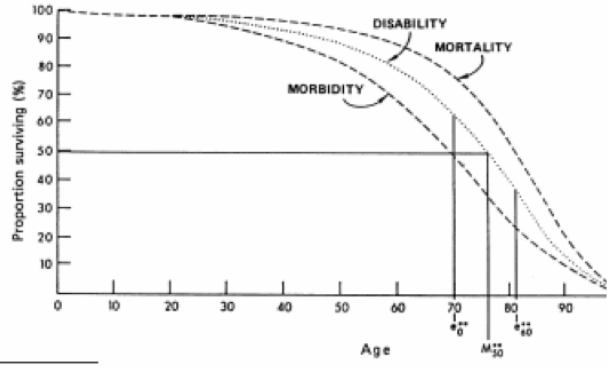
An aging society is a challenge for European economies. The aging population will have financial consequences for the public finance system in the form of an increase in healthcare costs and care for older people. What is needed is an innovative approach and modern models of health care, including a departure from a hospital-based system for integrated healthcare, support for the evaluation of health technologies and a more effective use of online solutions. This is why analyzes and empirical studies are important, indicating trends in mortality, its causes and diseases most often affecting the society.

Analysis of the health transition process allows to answers the questions related primarily to the health of older people. Particularly interesting is the answer the question: does life expectancy go hand in hand with a longer and longer life in health and shorten the years lived with functional limitations caused by poor health, or does it entail extending the life span with diseases and disability? Due to the high degree of complexity and difficulty of the problems at work, selected aspects of the health transition process are presented. The theoretical part describes the health transition model. The empirical part focuses on the diagnosis of the current state and changes in the mortality and health status of older people over the past 26 years (1990-2016) in three European countries. The main aim of the study is to analyze the health status and causes of mortality in people aged 70 and above in relation to one of the theory of health transition - the general theory of aging.

2 Model of health transition

The general model of health transition was proposed by WHO in 1984, which presents the relationship between mortality and broadly understood morbidity and disability is called the general model of health transition (Fig. 1). The upper line “mortality” represents the survival curve that determines the fraction of people in the surviving population in a given age. The “disability” curve represents the percentage of people of a given age with no disability or unproductive disease. The field below the “morbidity” curve represents the time of life in good health, i.e. free of chronic diseases. The distance between the lines of morbidity and mortality is a “health gap” and is the difference between the actual state of health in the population and the state determined as perfect or full, in which everyone would live to death without serious diseases and disabilities. The area above the survival curve is the loss resulting from premature mortality and mortality before reaching the age considered optimal or maximum for a given population. In the literature there are several theories describing the

possible relations between health and disability curves and the survival curve occurring during the period of rapid increase in life expectancy and the aging of the population.



e_{0}^{**} and e_{60}^{**} are the number of years of autonomous life expected at birth and at age 60, respectively.
M_{50}^{**} is the age to which 50% of females could expect to survive without loss of autonomy.

Fig. 1 The general model of health transition (Jagger and Reyes-Beaman, 2003).

Three basic theories of health transition are known in the literature: compression of morbidity (1980), expansion of morbidity (1979), dynamic equilibrium theory (1982), and a new approach called the general theory of aging of the population (2005). The last one is the subject of interest from the perspective of this article.

The authors of this theory (Michel and Robine, 2004) emphasize that in order to understand the ongoing relationship between the course of mortality trends and changes in disability and disease burden, it is necessary to take into account the stage of demographic and epidemiological transition where a country is located, as well as socio-economic, geographical, cultural and medical conditions, which determine the stage of advancement of the aging process in connection with the health condition. The basis for the theory of aging of the population is the assumption of cyclic successive stages, the components of which are processes included in the model of health transition.

The first stage concerns the increase in the life expectancy due to the decrease in mortality in the oldest age groups. The increase in the survival of people at an advanced age burdened with various diseases and ailments results in an increase in the number of years lived with disability and the share of people in the population who are burdened with chronic diseases (expansion of morbidity). The second stage corresponds to the theory of dynamic equilibrium. It is associated with further medical progress, which allows to stop or slow down the development of many diseases and delay the occurrence of severe symptoms of the disease. The third stage (compression of morbidity) occurs as a result of entering the aging process of the next generation. It is assumed that people from this cohort, thanks to pro-healthy

behaviors and increasingly better living conditions, will enjoy better health compared to the previous generation. The authors suppose that the process may be continued and after the completion of one cycle, the first stage will be re-performed, albeit with a different quality.

According to this theory, in the 1970s and 1980s the Western countries entered the second phase, which was largely due to progress in cardiology. This progress allowed for the treatment of many diseases of the circulatory system, which not only increased the chances of survival, but also reduced the extent of the occurrence of more severe disabilities. An important element of this stage is adaptation to the existing demographic and epidemiological changes in health policy activities, which was directed to the growing population of elderly people and diseases typical of this age.

3 Assessment of health status of population

Health measures are required in an assessment of the health status of the population. Summary measures of population health are measures that combine information on mortality and non-fatal health outcomes to represent the health of a particular population as a single number.

Over the last decades, several indicators have been developed to adjust mortality to reflect the impact of morbidity or disability. These measures fall into two basic categories (Murray et al., 2009): (1) health expectancies - measure years of life gained or years of improved quality of life (e.g. disability-adjusted life expectancy, healthy adjusted life expectancy, quality adjusted life expectancy), (2) health gaps - measure lost years of full health in comparison with some “ideal” health status or accepted standard (e.g. potential years of life lost, healthy years of life lost, quality adjusted life years, disability adjusted life years).

In this paper disability adjusted life years (DALYs) estimates are used to compare disease burden in populations. The DALYs are computed for a specific disease or a group of diseases such that, the total time lost due to premature mortality in a population is added to the total time lost due to a disability. Formally, DALYs expresse the equivalent of health years that have been lost due to years of illness or disability of a certain degree of severity and duration (YLD) and the number of years of death lost due to premature death (YLL). The YLL is estimated as the number of people who died at a particular age, multiplied by the total standard number of years to survive by persons of a certain age in which the death occurred. The estimation of YLD for a specific cause in a specific time period requires multiplying the number of cases of a given event in a given period by the average duration of the disease and by the weight factor. One year of DALY is understood as one year of lost healthy life. The

disadvantage of DALYs is the often missing data needed to calculate it, e.g. the incidence of diseases by age and sex, the number of cases of disabling illness, the average age of disability in the population, duration of disability, and the number of disabilities (Gromulska et al., 2008). For this reason, the DALY value is generally estimated. Details on DALYs estimation, methodological and ethical issues are described in a literature e.g. Gold et al. (2002).

4 Leading death causes among the elderly in selected European countries

Over the last quarter century, despite the evidence of the increase in the number of people aged 65+ and the increase in the number of deaths in this group, the intensity of mortality has significantly decreased. These changes are reflected in the lengthening of the life expectancy. Moreover, moving the largest percentage of deaths towards increasingly older age groups also indicates an improvement in the health of the population.

The analysis is carried out on the example of three countries. There are two main reasons for selecting only three European countries (Denmark, Poland, France) to the analysis. Firstly, according to some analysis (e.g. Majewska, 2017; Lazar et. al., 2016) they represent groups of countries with different trends of life expectancy and different pace of life expectancy rise. Secondly, the age structure of cardiovascular mortality – that plays the major role in the second stage of health transition – is much older in Denmark and France than in Poland (see Vallin and Mesle, 2005). Besides, according to existing research from late 90s French population completed the second stage of health transition and they enter the third stage, whereas Denmark actually has started to catch up on France.

Below, the mortality of the elderly (65+) are examined not only in the basic groups of causes of death (caused by diseases such as cardiovascular diseases and cancer), but also mortality caused by specific chronic diseases. Therefore, ten broad groups of causes of death are analyzed: (1) cardiovascular diseases, (2) respiratory diseases, (3) cirrhosis and other chronic liver diseases, (4) diabetes, blood diseases, and genitourinary system, (5) digestive diseases, (6) mental and substance use disorders, (7) musculoskeletal disorders, (8) neoplasms, (9) neurological disorders, (10) other non-communicable diseases.

According to the Global Burden of Disease (GDB, 2015) we conclude that in terms of the number of YLLs due to premature death in Denmark ischemic heart disease, trachea, bronchus, and lung cancers, and cerebrovascular disease were the highest ranking causes, in France – ischemic heart disease, trachea, bronchus, and lung cancers, and cerebrovascular disease, and in Poland – ischemic heart disease, cerebrovascular disease, and trachea,

bronchus, and lung cancers. Besides, the leading risk factor is tobacco smoking in Denmark, in France and Poland – dietary risks.

The main causes of DALYs and percent change between 1990 and 2016 are presented in Fig. 2 - 5. Bars going up show the percent by which DALYs have increased since 1990. Bars going down show the percent by which DALYs have decreased. In Poland mental and substance use disorders in male group, and neurological disorders are on the rise. Whereas in France – musculoskeletal disorders and neurological disorders are on the rise. It is noteworthy that the group of neurological disorders includes Alzheimer's and Parkinson's diseases. The interesting thing is observed in case of Denmark – the rise of DALYs in five groups of disease, which was not observed in this country at the end of the previous century (compare Vallin and Mesle, 2005). This is related with a slowdown in increase of life expectancy in the last years.

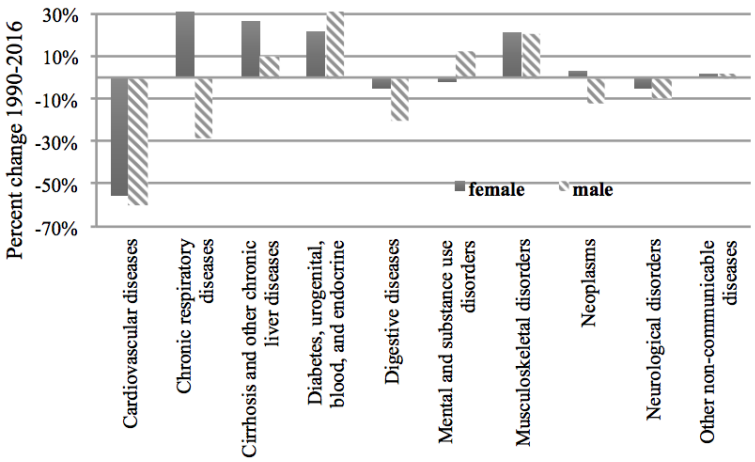


Fig. 2 Leading causes of DALYs and percent change 1990 to 2016 for Denmark.

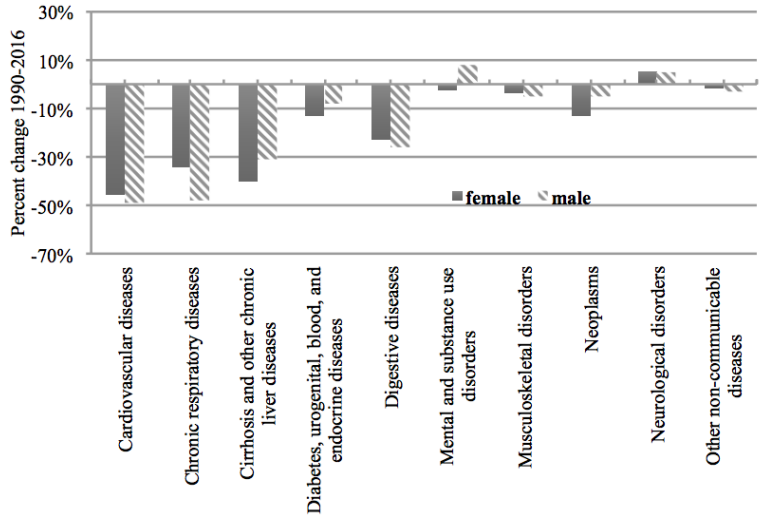


Fig. 3 Leading causes of DALYs and percent change 1990 to 2016 for Poland.

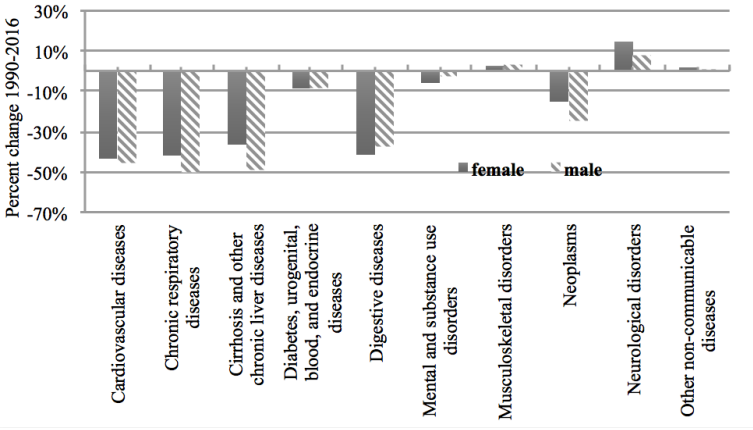


Fig. 4 Leading causes of DALYs and percent change 1990 to 2016 for France.

The above results are evidence for effective fight against diseases, mainly cardiovascular diseases (despite significant participation of this group in all analyzed causes of deaths).

5 Contribution of trends in age-specific mortality for the main groups of causes in two stages of changes in life expectancy

A necessary condition for entering the third phase of transition is an uninterrupted increase in life expectancy. However, has there been a slowdown in the aging process and a shift in the emergence of diseases over time?

Fig. 5-7 present the results of the decomposition of changes in life expectancy (in two periods 1990-2002 and 2003-2016, for seven age groups) taking into account the basic causes of deaths. Only the results for female are presented. Positive numerical values of the decomposition of changes in life expectancy signify an increase in life expectancy resulting from the improvement of health and reduction of mortality caused by various causes, while negative values mean a decrease in life expectancy associated with worsening mortality in a given group of deaths (Wróblewska, 2015). The decomposition was made by age using the Arriaga method (1984). The graph show the results of decomposition only for women, as the probability of entering the third phase of transition is much higher for this group.

The obtained results illustrate significant differences in the share of death cause groups in the increase in life expectancy for women between the periods. Although it can be seen that in the period 2003-2016 favorable changes occur, i.e. lifetime increases associated with a drop in mortality in the analyzed groups of causes of deaths, but the shift of deaths – for example for the group of cardiovascular diseases for Poland – from the age group 65-69 to 95+ is not visible. Therefore, it is not possible to talk about the entry of women into the third phase of health transition in case of Poland. But the trends are moving in the right direction. In case of

Denmark unfavorable changes that occur in recent decades in the mortality of older people are clearly observed. Neoplasms mortality is becoming an increasing problem, and this is very far from a new stage in the health transition devoted to the fight against ageing. France is very good example for shifting diseases toward older groups.

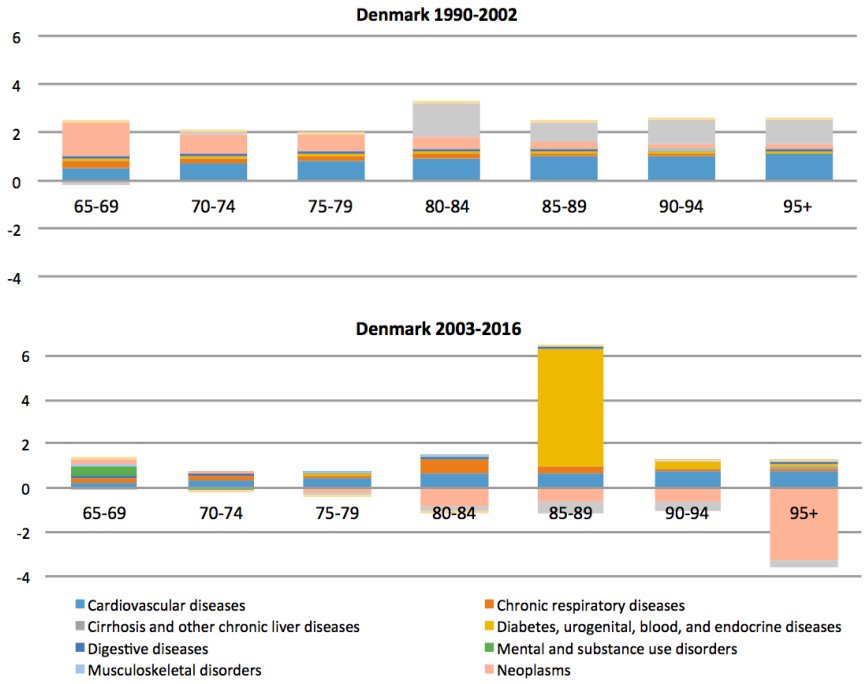


Fig. 5 Contribution of trends in age-specific mortality for 10 main groups of causes in two stages of changes in female life expectancy in Denmark.

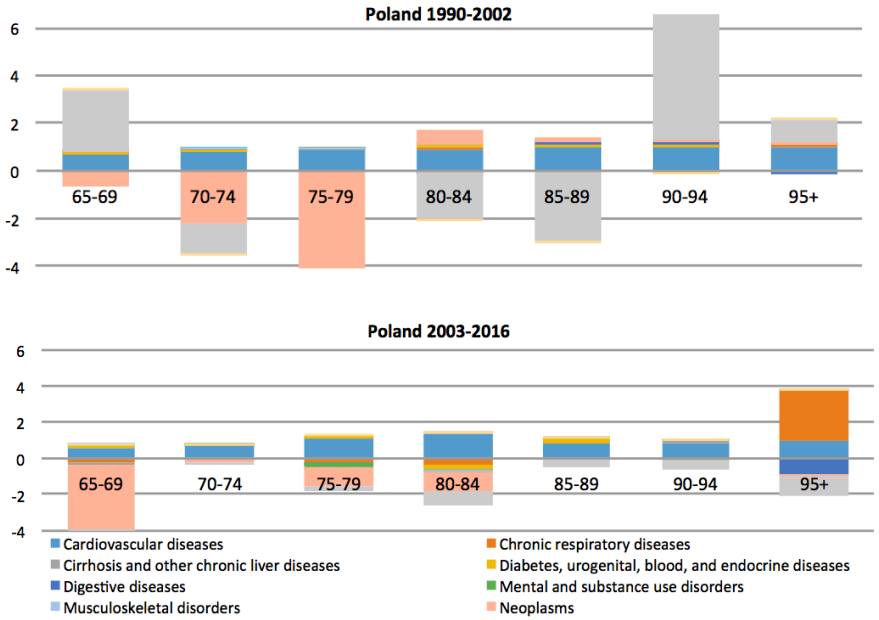


Fig. 6 Contribution of trends in age-specific mortality for 10 main groups of causes in two stages of changes in female life expectancy in Poland.

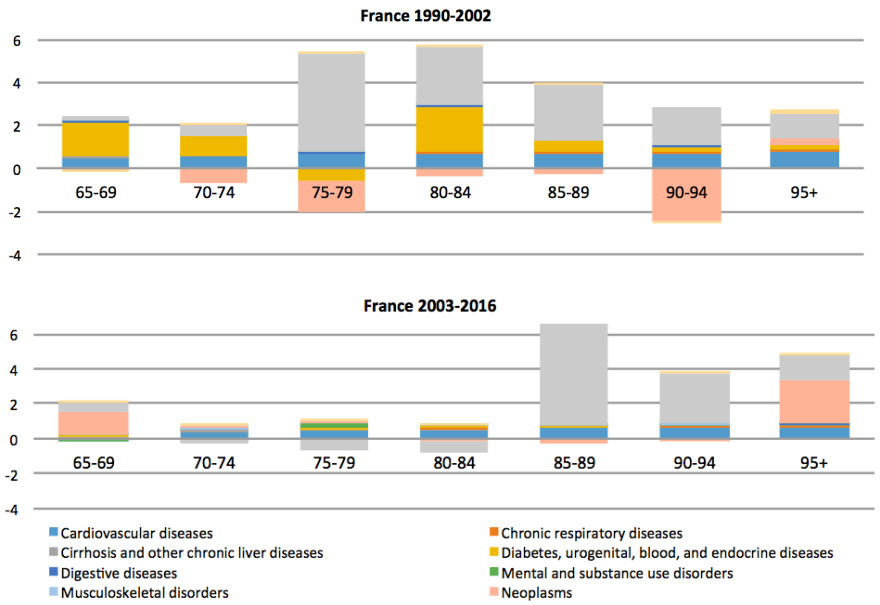


Fig. 7 Contribution of trends in age-specific mortality for 10 main groups of causes in two stages of changes in female life expectancy in France.

Conclusions

In the article the diagnosis of the changes in the mortality and health status of the older people over the past 26 years (1990-2016) were made. An analysis of mortality in three countries showed that issue of health transition and an assessment if a population entered the third stage is more complex than the theory describes. There is clear evidence that the life expectancy increases due to the effective fight against diseases, especially cardiovascular diseases. However, nowadays inference about the transition from the second to the third stage should also include the complex analysis of diseases strictly related with old people like neurological diseases. According to the analysis of Vallin and Meslé (2004), Denmark was a great example of country with perspectives on stepping to the third stage of health transition, however this analysis shows that increasing neoplasms mortality seems to distance from entering the third phase. For future analysis of health transition it is necessary to assess the impact of neurological diseases (like Alzheimer's, Parkinson's) and all diseases that depends on health behaviors and living conditions that are projected to greatly increase.

References

Arriaga, E. E. (1984). Measuring and explaining the change in life expectancies. *Demography*, *21*(1), 83-96.

GBD Data. Retrieved December 15, 2017, from http://www.healthdata.org/gbd/data

Gromulska, L., Wysocki, M. J. & Goryński, P. (2008). Lata przeżyte w zdrowiu (Healthy Life Years, HLY) – zalecany przez Unię Europejską syntetyczny wskaźnik sytuacji zdrowotnej ludności. *Przegląd Epidemiologiczny, 62*(4), 811-820.

Gold, M. R., Stevenson, D. & Fryback, D. G. (2002). HALYs and QALYs and DALYs, Oh My: Similarities and Differences in Summary Measures of Population Health. *Annual Review of Public Health, 23*(1), 115-134.

Jagger, C. & Reyes-Beaman, S. (2003, January). Monitoring Health by Healthy Active Life Expectancy – A User's Guide. Retrieved January 15, 2018, from https://www.researchgate.net/publication/263943580_Monitoring_Health_by_Healthy_Active_Life_Expectancy-A_User's_Guide.

Lazar, D., Buiga, A. & Deaconu, A. (2016). Common stochastic trends in European mortality levels: testing and consequences for modeling longevity risk in insurance. *Romanian Journal for Economic Forecasting, 2*, 152-168.

Majewska, J. (2017). An EU cross-country comparison study of life expectancy projection models. In: *Statistical working papers Eurostat, Selected papers from the 2016 Conference of European Statistics Stakeholders - Special issue* (pp. 83-94). Luxembourg: Publications Office of the European Union.

Michel, J. & Robine, J. (2004). A "New" General Theory of Population Ageing. *Geneva Papers on Risk and Insurance - Issues and Practice, 29*(4), 667-678.

Murray, C., Lopez, A. & Mathers, C. D. (2009). 2.10 Assessing health needs: The global burden of disease approach. *Oxford Textbook of Public Health.*

Vallin, J. & Meslé, F. (2004). Convergences and divergences in mortality. *Demographic Research, Special 2*, 11-44.

Vallin, J. & Meslé, F. (2005). Convergences and divergences: an analytical framework of national and sub-national trends in life expectancy. *Genus, 61*(1), 83-124.

Wróblewska, W. (2006). Analiza umieralności w Polsce w latach 1970−2003. Dekompozycja zmian w oczekiwanym trwaniu życia noworodka. *Studia Demograficzne, 1*(149), 28-48.

Wróblewska, W. (2015). Zmiany umieralności według przyczyn zgonów w przekroju wojewódzkim. *Wiadomości Statystyczne,* (11), 30-53.

Cyclical fluctuations of global food prices: a predictive analysis

Błażej Mazur[1]

Abstract

Forecasting of fluctuations in worldwide food prices is of considerable practical importance – however, it is also difficult, especially for longer horizons. Standard econometric models often fail to take into account non-seasonal cyclicality present in some of the series under consideration. In particular in many cases the models either assume mean-reversion (so the forecast paths stabilize at some value) or a random-walk type behaviour (so the forecast paths are determined by the last observation available). In other words, it is not easy to obtain forecasts that reveal future turning points (so are capable of forecasting of out-of-sample deviations from mean).In order to deal with the issue we make use of Bayesian deterministic cycle models based on Flexible Fourier type representation in order to analyse dynamic behaviour of FAO food price indexes. We focus on prediction of individual year-on-year growth rates in horizons ranging from one to twenty four months ahead. As the purpose of the paper is to analyse performance of density forecasts, we investigate log-predictive score as well as continuous ranked probability score. We find a clear pattern of fluctuations in dairy prices and a bit less-evident cyclical-like fluctuations in meat prices, so the deterministic-cycle models are relevant for the series.

Keywords: *Bayesian inference, food prices, density prediction, periodicity*
JEL Classification:E37; E31; C53
DOI: 10.14659/SEMF.2018.01.29

1 Introduction

The issue of food price forecasting is quite challenging.However, literature provides no clear conclusion indicating optimal techniques for probabilistic prediction of food prices in horizons of 1-2 years. The issue seems to be of considerable practical importance for a number of reasons – for example the recent increase in Polish inflation is to some extent due to transmission of high food prices (including the dairy products). Moreover, the development of some counter-cyclical policy is of vital importance for well-being and security of less-developed countries (see e.g. Myers, 2006; Kalkuhl et al., 2016). In the paper we tackle the related econometric issues asking whether the use of models that explicitly account for periodicity might improve the quality of density forecasts of food prices in longer horizons.

[1]Cracow University of Economics, Department of Econometrics and Operations Research, Rakowicka 27, 31-510 Kraków, Poland, e-mail: blazej.mazur@uek.krakow.pl.

2 The data: FAO food price indexes

The data considered here are obtained based on FAO price indexes (available at www.fao.org, as accessed on 07.06.2017). We make use of monthly deflated indexes (the total and five sub-indexes), converted into year-on-year growth rates, see Fig. 1. We analyse 317 monthly observations (1991 M1 – 2017 M5). In the paper we conduct a pseudo-real time analysis, hence we do not make an effort to control for data revisions – the recursive experiments are conducted using the same data vintage.

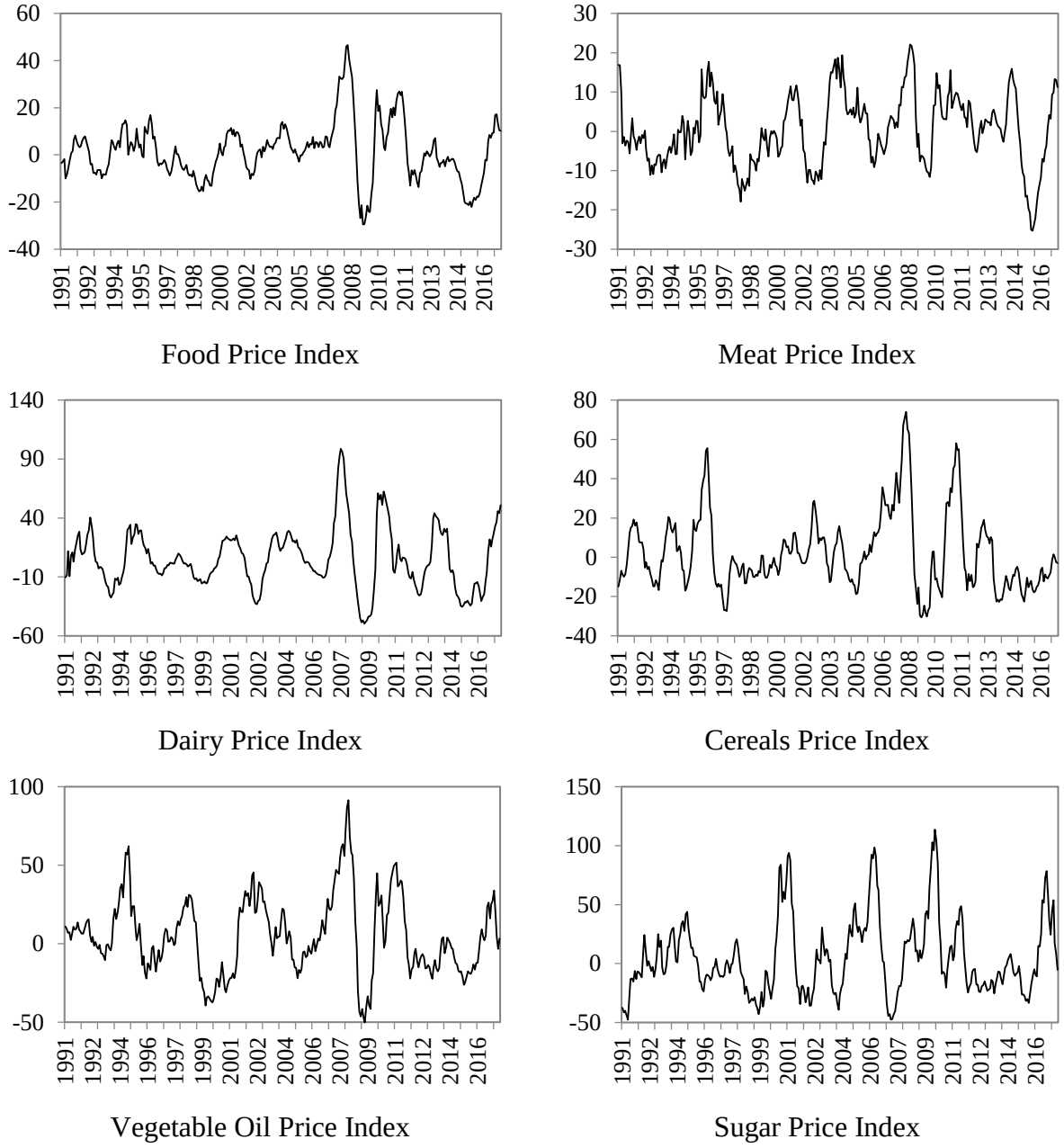


Fig. 1. The data: year-on-year growth rates of deflated FAO price indexes (as of 07.06.2017).

The series clearly display some sort of mean-reversion (which might take a regular, perhaps periodic pattern). There is also evident co-dependence (for example a noticeable influence of the global financial crisis) – however, in the paper we focus on univariate methods. Certain advanced problems of co-dependence between food prices and other prices are considered by e.g. Śmiech et al. (2017).

3 Models and methods for density predictions of series displaying periodicity

One of concepts used in order to avoid potential inadequacy of stationary time-series models without the need for accepting certain disadvantages of $I(1)$ nonstationarity is that of cyclostationarity. One of modelling approaches motivated by the idea is that of deterministic cycle models. Although the assumptions underlying the models used in the paper seem to be somewhat restrictive (as the pattern of cyclical fluctuations is assumed to be regular), the models easily allow for considering more than just one frequency of fluctuations.

Lenart et al. (2016) use Bayesian inference methods for in-sample analysis conducted with a time-series model defined as Gaussian AR(p) deviations from a periodic (hence time-varying) unconditional mean $\mu_t$defined as follows:

$$\mu_t = \delta + \sum_{f=1}^{F} \alpha_f \sin(\varphi_f t) + \beta_f \cos(\varphi_f t), \quad (1)$$

where F controls number of frequencies (or order of the Flexible Fourier approximation), α's and β's control phase shifts and amplitudes while φ's control frequency (and hence period length) of individual Fourier components. Inference on φ's and in particular obtaining density forecasts that account for estimation uncertainty is non-trivial and justifies the use of the Bayesian approach (details are given by Lenart and Mazur, 2017 and references cited therein). Crucially, one has to set F: $F = 0$ implies time-invariant mean in (1) and hence the model becomes a stationary AR(p) process with iid Gaussian errors. Large values of F result in a risk of overfitting (and might lead to very complicated posteriors). Here we consider models with $F = 0,1,2,3,4$ and $p = 12$ and a model with $F = 0$ and $p = 24$, which gives a total of 6 models (two being stationary Gaussian autoregressive processes, see Table 1).

We consider prediction horizons from $h = 1$ to $h = 24$ months ahead and analyse a sequence of 120 expanding subsamples (with sample size increasing from 317-120 to 317-1). Each model is re-estimated with each subsample, and a predictive distribution over the next 24 months is obtained using these estimates. Such an exercise, with six series considered here, represents quite a considerable numerical burden. The models are estimated using

MCMC techniques (with M-H steps within a Gibbs sampler), priors for φ's and the autoregressive parameters are uniform (with $0 < \phi_L < \varphi < \phi_U < \pi$).

Table 1. Ex-post predictive accuracy of density forecasts: LPS and CRPS scores.

	M1	**M2**	**M3**	**M4**	**M5**	**M6**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**
F	0	0	1	2	3	4	0	0	1	2	3	4
p	12	24	12	12	12	12	12	24	12	12	12	12
H				**LPS**						**CRPS**		
						Food Price Index						
24	-155.8	-165.5	-155.8	*-154.3*	-155.3	-156.8	7.4	9.0	7.2	*7.1*	7.3	7.8
20	-168.1	-176.4	*-166.4*	*-166.4*	-167.0	-169.5	8.3	10.3	*8.1*	*8.1*	*8.1*	8.7
16	-180.9	-186.3	*-180.7*	-182.3	-184.2	-187.1	9.5	10.6	*9.4*	9.7	9.9	10.2
12	*-196.7*	-200.5	-202.1	-204.4	-208.6	-211.7	*10.6*	11.4	11.3	11.6	12.0	12.2
8	-209.4	*-208.5*	-211.4	-218.3	-218.8	-227.2	10.6	*10.5*	11.0	11.2	11.6	12.2
4	-199.3	*-195.0*	-197.4	-199.4	-201.8	-206.7	7.2	*6.9*	7.0	7.1	7.3	7.7
1	-163.6	*-144.2*	-146.7	-146.9	-146.7	-147.3	2.5	*2.4*	*2.4*	*2.4*	*2.4*	2.5
						Dairy Price Index						
24	-187.1	-186.2	-180.1	-180.2	*-179.7*	-181.4	16.9	16.3	14.1	14.0	14.0	*13.9*
20	-196.6	-193.6	-191.7	*-191.4*	-192.2	-192.4	16.7	16.1	14.7	*14.5*	14.6	*14.5*
16	-202.1	-199.5	-199.2	*-198.9*	-199.2	-199.2	15.3	15.4	13.9	13.8	13.8	*13.7*
12	-207.9	-208.8	-204.2	-204.0	*-203.9*	-204.3	14.6	16.0	13.0	12.9	*12.8*	*12.8*
8	-217.6	-218.8	*-211.0*	-212.8	-212.7	-211.0	14.0	15.0	12.6	12.5	*12.4*	12.5
4	-218.3	-218.7	-211.3	-210.5	*-210.4*	-210.5	10.7	10.6	10.1	*10.0*	*10.0*	*10.0*
1	-164.5	-164.9	-188.7	*-164.1*	-164.4	-164.4	4.0	*3.9*	4.0	4.0	4.0	4.0
						Meat Price Index						
24	-150.5	-149.0	*-144.3*	-145.2	-146.4	-148.2	5.8	5.8	*5.3*	5.5	5.7	5.9
20	-154.2	-155.2	*-150.9*	-151.6	-153.3	-154.2	5.6	5.8	*5.4*	5.5	5.7	5.8
16	-159.7	-161.6	*-157.3*	-158.4	-159.4	-159.7	5.5	5.9	*5.4*	5.6	5.7	5.7
12	-171.1	-169.7	*-167.7*	-169.1	-170.2	-170.6	5.9	6.1	*5.7*	5.9	5.9	6.0
8	-182.8	*-176.6*	-180.9	-184.5	-184.7	-185.0	6.1	6.1	*6.0*	6.2	6.2	6.2
4	-177.2	*-169.8*	-176.7	-179.8	-179.7	-179.6	4.7	*4.5*	4.7	4.8	4.7	4.7
1	-131.4	*-128.1*	-131.0	-131.5	-131.7	-131.8	1.9	*1.8*	1.9	1.9	2.0	2.0

In order to evaluate the resulting density forecasts we make use of scoring rules: log-predictive score and continuous ranked probability score (see Gneiting and Raftery, 2007). In order to gain some insights into differences across models we informally examine sequences of point forecasts (predictive expectations). We also consider a future tendency index (FTI) discussed by Mazur (2017a, pp. 443) that provides information as to prevailing trends in the forecast period (taking into account the cross-horizon stochastic dependence).

4 Discussion of the empirical results

Despite the fact that the point of interest here is that of density prediction, we begin with analysis of paths of point forecasts. This is because such analysis sheds some light at differences across models in terms of out-of-sample behaviour.

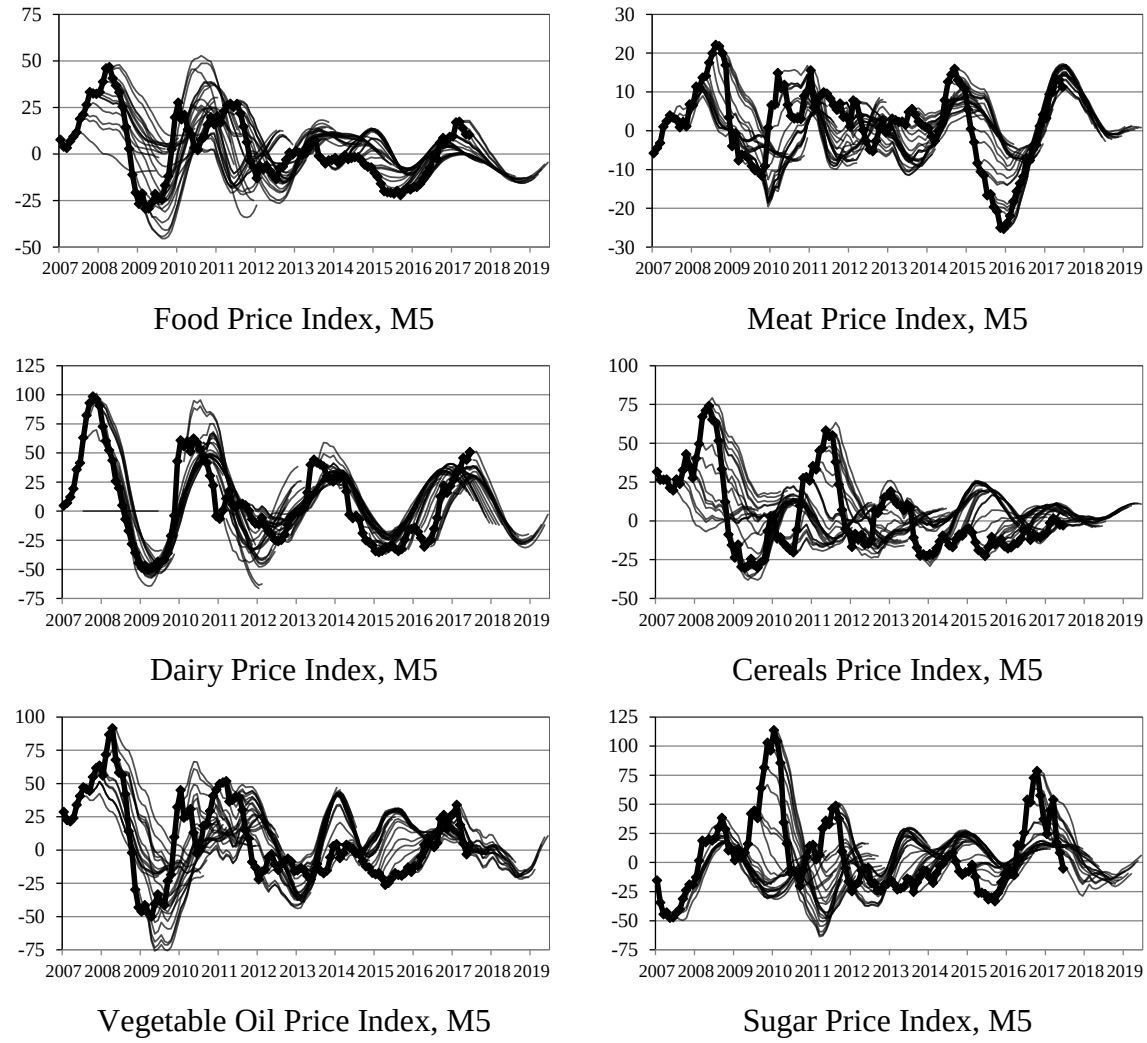
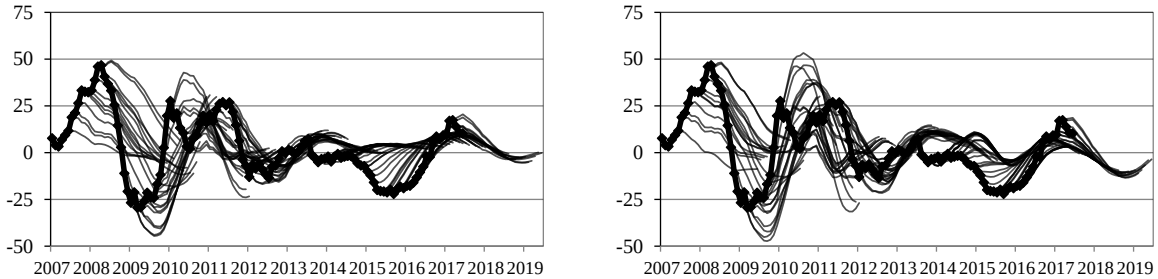


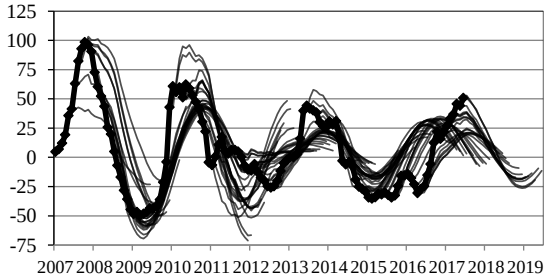
Fig. 2. Recursive sequences of point forecasts from a model with $F = 3$.

Fig. 2 presents recursively obtained sequences of forecast paths using a model with $p = 12$ and $F = 3$, hence allowing for quite a complicated pattern of quasi-periodic fluctuations. One might note that the pattern seems particularly adequate for the Dairy Price Index. However, for the remaining series the evidence is rather mixed: the forecasts display out-of-sample fluctuations even at longer horizons, but the fluctuations are not necessarily adequate.

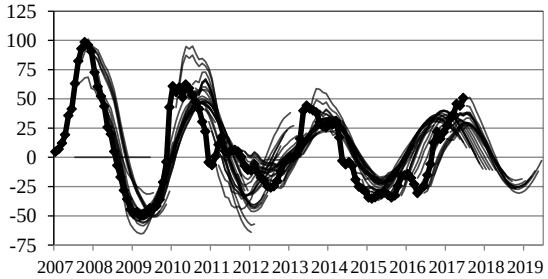


Food Price Index, M1 (F = 0)

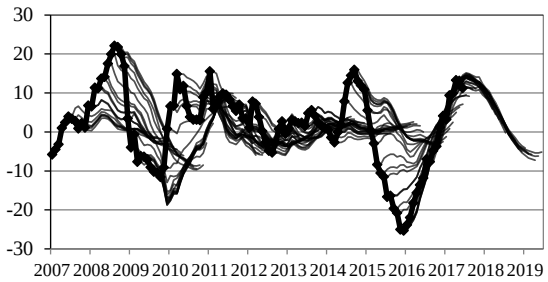
Food Price Index, M4 (F = 2)



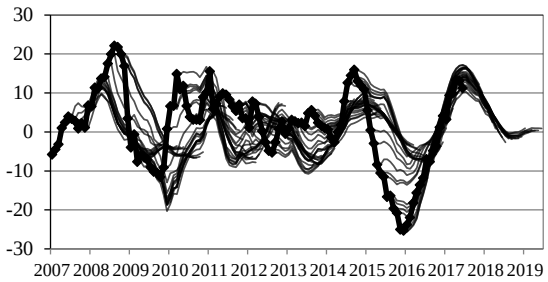
Dairy Price Index, M1 (F = 0)



Dairy Price Index, M4 (F = 2)



Meat Price Index, M1 (F = 0)



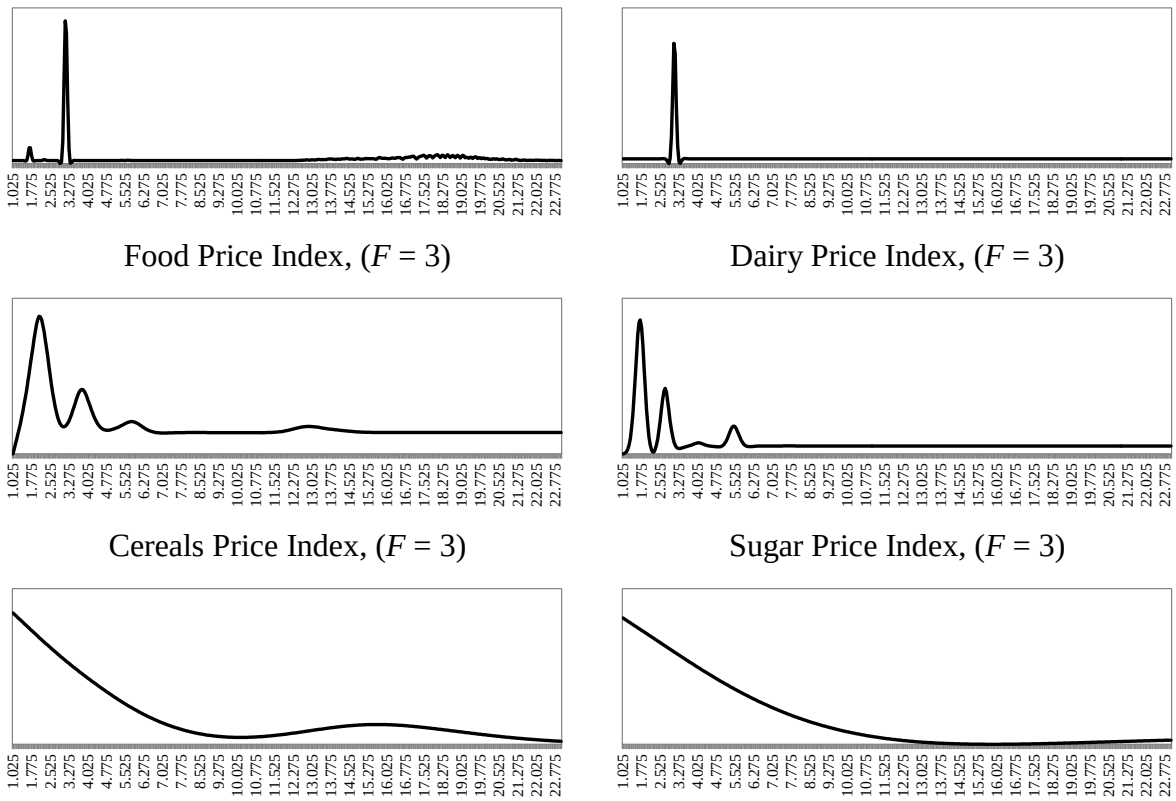
Meat Price Index, M4 (F = 2)

Fig. 3. Recursive sequences of point forecasts from models with F = 2 and F = 0.

The informal analysis is consistent with information conveyed by formal criteria: Table 1 contains ex-post characteristics of predictive accuracy for three selected series: the overall index as well as Dairy and Meat sub-indices. We report LPS and CRPS scores across models and selected horizons (please note that for each horizon we use the last 121-h realized values, ending with 2017 M5). The LPS values (the higher the better) are cumulated across realizations (computed with natural logs) hence differences between the scores for h = 1 can be interpreted in terms of predictive BayesFactors. The CRPS scores are averaged across realizations and the orientation is chosen to mimic that of mean absolute error: the smaller the better. For longer horizons both criteria tend to indicate models with F> 0, although the effect is most evident for the Dairy series; for short horizons (Food, Meat), the AR(24) seems to be the most successful one. The differences between models are not large. This is probably due to two reasons. Firstly, for variables other than Dairy the periodic pattern is not that regular.

Secondly, the lag length p = 12 (or 24) is considerable and hence the AR(p) models are capable of delivering a periodic-like behaviour out-of-sample (formally being stationary).

Fig. 3 presents a comparison of forecast paths of models with periodicity (F> 0) and a pure autoregressive processes (F = 0). The results seem to reinforce the above conclusion – even for the dairy series the model with F = 0 is capable of generating periodic-like forecast paths (however, the periodic effectsare somewhat more evident with F> 0).



Food Price Index, (F = 3) Dairy Price Index, (F = 3)

Cereals Price Index, (F = 3) Sugar Price Index, (F = 3)

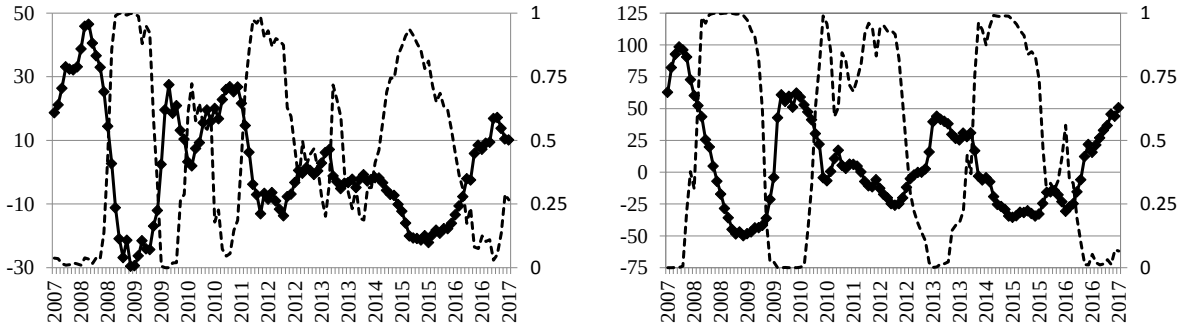
Vegetable Oil Price Index, (F = 3) Meat Price Index, (F = 3)

Fig. 4. Inference on period length (in years) based on posterior for φ_f from M5, full sample.

Fig. 4 contains (un-normalized) full-sample based posteriors for period length induced from marginal posteriors for $\varphi_f$parameters (the period length in months is given by $2\pi/\varphi_f$). The values of prior hyperparameters ϕ_L and ϕ_U are chosen to exclude cycles shorter than one year and longer than 24 years. One might note a single spike corresponding to approx. 3 years for the Dairy series. It is also visible in Food series, though there is one additional peak indicating shorter cycles and some probability mass distributed along values corresponding to cycles of 15-20 yrs. Please note that for Cereals and Sugar series the spikes are much less concentrated (leaving more estimation uncertainty) and the dominant period is shorter; there are two more evident modes and some less-elevated local maxima. However, for Meat and

Oils series, the results are quite different – as no distinct spikes are visible, the results might suggest rather different pattern of dynamics (presumably a more stochastic one).

We also consider a probabilistic index of future tendencies (FTI_M), as described in Mazur (2017a, p. 443).It goes beyond horizon-specific marginal density forecasts, making use of the joint predictive distribution (over all the horizons) instead.Each value of the FTI_{12} index in Fig. 5 represents the probability that (based on information up to agiven time T) the predicted quantity averaged over the second half of the prediction period(here: $12<h<24$) would be higher than the average over the first half of the prediction period ($0<h<13$), hence the index is supposed to lead actual changes of trends by about a year or so.



Food Price Index, M4 (F = 2) Dairy Price Index, M5 (F = 3)

Fig. 5. Realized data (solid line) vs. values of predictive tendency index FTI_{12}(dotted).

For the dairy series (with more evident periodicity) the index takes more extreme values. The index is close to zero in 2013 (just as the dairy prices seem to hit an upper turning point). Moreover, it gets close to one in 2014 and 2015, with increasing prices in 2016 and 2017. Interestingly, the index again hits the minimum level at the sample end, indicating that the upturn in dairy prices would revert in late 2017 or 2018. For the food prices the index hits the maximum just before a turning point in price dynamics in 2015. The values closer to the sample end indicate that it would be more likely to observe decreasing tendency in food prices.

Conclusions

We discuss practical importance of Bayesian deterministic-cycle models with many frequencies. In-sample applications have been considered by e.g. Lenartet al. (2016) or Mazur (2016). In this paper we focus on assessment of the predictive adequacy, with emphasis on density forecasting performance. The issue has certain empirical relevance as the models

display interesting out-of-sample properties, since the forecast paths differ from those of random-walk type models as well as from stationary ones.

In particular we examine density forecasts of y-on-y growth rates of deflated FAO food price indexes within a recursive expanding-window experiment. We find quite a considerable heterogeneity across the series under consideration. For the Dairy data there is evidence that deterministic cycle models have superior forecast accuracy. However, the periodic signal in the series is quite strong and might be captured to some extent by other models as well. The effect seems to carry over to the aggregate index to some extent. On the other hand, Meat prices and Vegetable Oil Prices either display strong time-inhomogeneity or the dynamic pattern is much more stochastic – hence, the deterministic cycle models seem to be of no use here. For the Cereals and Sugar series the evidence is mixed (as it is for the aggregate index) - there seems to be some periodicity, but it is weaker (compared to that of the Dairy series).

Therefore one might conclude that for the purpose of modelling of all the indexes one should consider a more heterogeneous model classes. Possibly stochastic cycle models or models with time-varying periodic behaviour could be adequate for the Cereals and Sugar series, while a fully stochastic models should be used for Meat and Vegetable Oils series. Moreover, it would be very interesting to see whether the aggregate index could be more accurately predicted based on heterogeneous disaggregate forecasts of the five sub-indexes. A similar case (where the aggregate forecast benefits from accounting for heterogeneity of the sub-aggregates) is discussed by Mazur (2017b) for Polish macroeconomic data at quarterly frequency.

Acknowledgements

This research was supported by the research grant 2017/25/B/HS4/02529 financed by the National Science Centre, Poland.

References

Gneiting, T. & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477), 359-378.

Harvey, A.C., Trimbur, T.M.& Van Dijk, H.K. (2007). Trends and cycles in economic time series: A Bayesian approach. *Journal of Econometrics*, *140*, 61-649.

Huerta, G. & West, M. (1999a). Bayesian inference on periodicities and component spectral structure in time series. *Journal of Time Series Analysis*, *20*(4), 401-416.

Huerta, G. & West, M. (1999b). Priors and component structures in autoregressive time series models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61*(4), 881-899.

Kalkuhl, M., Von Braun, J. & Torero, M. (Eds.). (2016). *Food price volatility and its implications for food security and policy*. Springer Open.

Lenart, Ł., Mazur, B. & Pipień, M. (2016). Statistical Analysis of Business Cycle Fluctuations in Poland Before and After the Crisis. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 11*(4), 769-783.

Lenart, Ł. & Mazur, B. (2017). Business cycle analysis with short time series: a stochastic versus a non-stochastic approach. In: Papież, M. & Śmiech, S. (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,212-222.

Mazur, B. (2016). Growth Cycle Analysis: the Case of Polish Monthly Macroeconomic Indicators. *Folia Oeconomica Cracoviensia, 57*, 37-53.

Mazur, B. (2017a). Probabilistic predictive analysis of business cycle fluctuations in Polish economy. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(3), 435-452.

Mazur, B. (2017b). Probabilistic Prediction Using Disaggregate Data: the Case of Gross Value Added in Poland. *Folia Oeconomica Cracoviensia*, *58*, 85-103.

McCoy, E. J. & Stephens, D. A. (2004). Bayesian time series analysis of periodic behaviour and spectral structure. *International Journal of Forecasting*, 20(4), 713-730.

Myers, R. J. (2006). On the costs of food price fluctuations in low-income countries. *Food Policy, 31*(4), 288-301.

Śmiech, S., Papież, M. & Fijorek, K. (2017). Volatility spillovers between food, energy, US dollar, and equity markets. Evidence from Diebold -Yilmaz's approach. In: Papież, M. & Śmiech, S. (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 , 407-416.

A family of non-standard bivariate distributions with applications to unconditional modelling in empirical finance

Błażej Mazur[1], Mateusz Pipień[2]

Abstract

We develop a class of parametric bivariate distributions that are capable of accounting for non-standard empirical properties that are evident in some financial time series. We aim at creating a parametric framework that allows for serious divergences from the multivariate Gaussian case both in terms of properties of marginal distributions and in terms of the dependence pattern. We are particularly interested in obtaining a multivariate construct that allows for considerable degree of heterogeneity in marginal properties of its components (like tail thickness and asymmetry). Moreover, we consider non-standard dependence patterns that go beyond a linear correlation-type relationship while maintaining simplicity, obtained by introducing rotations. We make use of marginal distributions that belong to generalized asymmetric *t* class analysed by Harvey and Lange (2017), allowing not only for skewness but also for asymmetric tail thickness. We illustrate flexibility of the resulting bivariate distribution and investigate its empirical performance examining unconditional properties of bivariate daily financial series representing dynamics of stock price indices and the related FUTURES contracts.

Keywords: *Bayesian inference, generalized asymmetric t distribution, skewness, orthogonal matrices, rotations*
JEL Classification: C52, C53, C58
DOI: 10.14659/SEMF.2018.01.30

1 Introduction

The empirical distributions of economic variables might display serious divergences from the multivariate normal case; see Genton (2004). For example pervasive heavy tails and volatility amplify studies that aim at searching for an appropriate families of multivariate probability distributions that would lead to successful modelling of empirical features in case of related financial returns. In particular since 2000's many authors tried to go beyond conditional normality assumed commonly in case of Multivariate GARCH (M-GARCH) models. For example conditionally elliptical distribution in DCC model was presented by Pelagatti and Rondena (2004). Some other non-Gaussian conditional distributions we analysed by Bauwens and Laurent (2005), Sahu et al. (2001), Pipień (2012) and others. However, commonly applied econometric strategy using the Maximum Likelihood estimation procedure might

[1]Corresponding author: Department of Econometrics and Operations Research, Cracow University of Economics, Rakowicka 27, Cracow, Poland; e-mail: eomazur@cyf-kr.edu.pl.
[2]Department of Econometrics and Operations Research, Cracow University of Economics, Rakowicka 27, Cracow, Poland; e-mail: eepipien@cyf-kr.edu.pl.

result with a considerable small-sample bias; see Iglesias and Phillips (2006).There is no doubt that proper modelling of empirical features observed for the case of related financial time series require construction of a flexible class of distributions; moreover, the development of alternative methods of statistical inference is necessary. In this paper we address these two issues. We propose a novel class of parametric bivariate distributions to model empirical properties that are evident in some financial time series. We depart from the multivariate Gaussian case both in terms of properties of the marginal distributions and in terms of the co-dependence pattern. In order to achieve flexibility we make use of marginal distributions that belong to generalized asymmetric Student-*t* class analysed by Harvey and Lange (2017), allowing not only for skewness but also for asymmetric tail thickness. We also develop methods of formal Bayesian inference and present posterior analysis within constructed class of sampling models. We also consider the issue of Bayesian (density) prediction.In the paper we illustrate flexibility of the resulting bivariate distribution and investigate its empirical performance examining unconditional properties of bivariate daily financial series representing dynamics of stock price indices and the related FUTURES contracts.

2 A family of non-standard bivariate distributions

For *m*-variate random variable $\mathbf{z} = (z_1, \dots, z_m)$ with independent coordinates, i.e. with $p(\mathbf{z}) = \prod_{i=1}^{m} p_i(z_i)$, one may consider a distribution resulting from a linear transformation:

$$\boldsymbol{\varepsilon} = A \cdot \mathbf{z} + \mathbf{b}.$$

For a nonsingulartransformation matrix $A_{[mxm]}$the distribution of $\boldsymbol{\varepsilon}$ is described by a well-defined density of the following form:

$$p(\boldsymbol{\varepsilon}) = \frac{1}{|\det(A)|} \prod_{i=1}^{m} p_i(A^{-1\,(i)}(\boldsymbol{\varepsilon} - \mathbf{b})),$$

where $A^{-1\,(i)}$ denotes *i*-th row of matrix A^{-1}.In what followswe make use of this result in a bivariate setting, going far beyond the standard scheme where *A* is defined as a root of symmetric and positive definite matrix generating the covariance structure.

To define the univariate distributions $p_i$we apply the generalised Student *t*-distribution proposed recently by Harvey and Lange (2017). It generalizes previous results by Zhu and Zinde-Walsh (2009), Zhu and Galbraith (2010)as well asFernández and Steel(1998) and Theodossiou (1998) among others, see the references and discussion by Harvey and Lange (2017).

However, the form used here is re-scaled to ensure unit variance. The resulting probability density function (with mode at 0 and nonzero mean in general) is:

$$f_{GAST}(z;h,\alpha,\upsilon_L,\upsilon_R,\eta_L,\eta_R)=\frac{K_{LR}\sqrt{f}}{\sqrt{h}}\begin{cases}\left[1+\frac{1}{\eta_L}\left(\frac{-z\sqrt{f}}{2\alpha\sqrt{h}}\right)^{\upsilon_L}\right]^{-\frac{1+\eta_L}{\upsilon_L}}, & z\le 0\\ \left[1+\frac{1}{\eta_R}\left(\frac{z\sqrt{f}}{2(1-\alpha)\sqrt{h}}\right)^{\upsilon_R}\right]^{-\frac{1+\eta_R}{\upsilon_R}}, & z>0\end{cases}$$

where GAST stands for 'generalized asymmetric skew *t*' and *h* denotes variance. The distribution is a two-piece version of a generalized *t* distribution; parameter $0<\alpha<1$ introduces skewness, with $\alpha = 0.5$ denoting the absence of skewing (symmetry requires also$\eta_L = \eta_R$ and$\upsilon_L = \eta_R$);υ's control shape around the mode (being more flat or spiked, in a GED-like manner, with $\upsilon = 2$ leading to t-type shape), while η's affect tail thickness (we require that η_L, $\eta_R > 2$ to ensure that variance is finite). However, the influence of η's and υ's on tail behaviour is not separated clearly. Setting $\upsilon_L = \upsilon_R = 2$ and $\eta_L = \eta_R$ leads to skew-*t* case, with skew-normal and normal distributions being the limiting ones with η_L and$\eta_R\to\infty$. Hence, the asymmetric and flexible distribution encompasses a number of well-known distributions, including the GED (η_L and $\eta_R\to\infty$). We assume η_L, $\eta_R > 2$ and υ_L, $\upsilon_R > 1$.K_{LR} and f denote rather complicated functions of shape parameters: K_{LR} and $\alpha^*$is given by Harvey and Lange (2017), note that P{Z < 0} = α^*,and$f = d - c^2$ with:

$$c=-\alpha^*2\alpha\frac{\eta_L^{\frac{1}{\upsilon_L}}\Gamma\left(\frac{2}{\upsilon_L}\right)\Gamma\left(\frac{\eta_L-1}{\upsilon_L}\right)}{\Gamma\left(\frac{\eta_L}{\upsilon_L}\right)\Gamma\left(\frac{1}{\upsilon_L}\right)}+(1-\alpha^*)2(1-\alpha)\frac{\eta_R^{\frac{1}{\upsilon_R}}\Gamma\left(\frac{2}{\upsilon_R}\right)\Gamma\left(\frac{\eta_R-1}{\upsilon_R}\right)}{\Gamma\left(\frac{\eta_R}{\upsilon_R}\right)\Gamma\left(\frac{1}{\upsilon_R}\right)},$$

$$d=-\alpha^*4\alpha^2\frac{\eta_L^{\frac{2}{\upsilon_L}}\Gamma\left(\frac{3}{\upsilon_L}\right)\Gamma\left(\frac{\eta_L-2}{\upsilon_L}\right)}{\Gamma\left(\frac{\eta_L}{\upsilon_L}\right)\Gamma\left(\frac{1}{\upsilon_L}\right)}+(1-\alpha^*)4(1-\alpha)^2\frac{\eta_R^{\frac{2}{\upsilon_R}}\Gamma\left(\frac{3}{\upsilon_R}\right)\Gamma\left(\frac{\eta_R-2}{\upsilon_R}\right)}{\Gamma\left(\frac{\eta_R}{\upsilon_R}\right)\Gamma\left(\frac{1}{\upsilon_R}\right)}.$$

Consider a product measure-typebivariate generalised Student *t* distribution of the form:

$$p(\mathbf{z})=p(z_1,z_2)=p_{Z_1}(z_1)p_{Z_2}(z_2),$$

where$p_{Z_i}(z_i)$ is the generalised Student *t* distribution of Harvey and Lange (2017) with individual shape parameters, transformed into the above GAST form (parametrized in terms of variance).The only restriction is that each coordinate in $\boldsymbol{z}$ has the same variance h_1, i.e. $V(\mathbf{z})=h_1\mathbf{I}$. The existence of variance could in principle be relaxed (in a scale-driven model) in order to allow for e.g. Cauchy-type tails. Now assume that the variable **z**is subject toa linear transformation, but the transformationmatrix is orthogonal; i.e.: $\mathbf{v}=\mathbf{R}_{(\varphi)}\mathbf{z}$**,** where:

$$\mathbf{R}_{(\varphi)}=\begin{bmatrix}\cos(\varphi) & \sin(\varphi)\\ -\sin(\varphi) & \cos(\varphi)\end{bmatrix}.$$

The matrix $\mathbf{R}_{(\varphi)}$imposes clockwise rotationby angle φ;$\mathbf{R}_{(\varphi)}^{-1}=\mathbf{R}_{(\varphi)}'$ with$\det(\mathbf{R}_{(\varphi)})=1$; the transformation might affect the density type but leaves the covariance structure

intact.Thus v_i's are uncorrelated with the same variances, but their marginal distribution might change (relative to those of z_i's). The density of the distribution of $\mathbf{v}$ is given by the formula:

$$p(\mathbf{v}) = \det(\mathbf{R}_{(\varphi)}) p_{Z_1}\left(\mathbf{R}_{(\varphi)}^{(1)'}\mathbf{v}\right) p_{Z_2}\left(\mathbf{R}_{(\varphi)}^{(2)'}\mathbf{v}\right) = p_{Z_1}\left(\mathbf{R}_{(\varphi)}^{(1)'}\mathbf{v}\right) p_{Z_2}\left(\mathbf{R}_{(\varphi)}^{(2)'}\mathbf{v}\right).$$

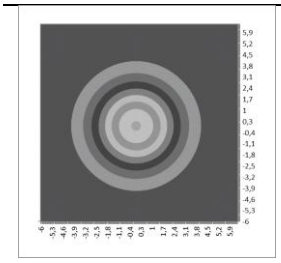
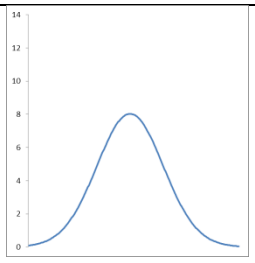
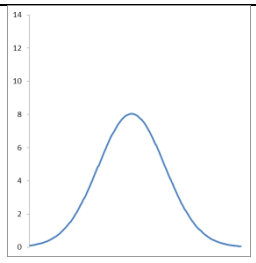
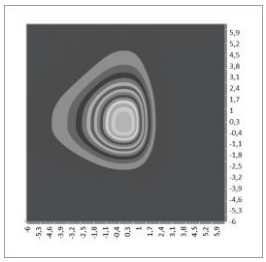
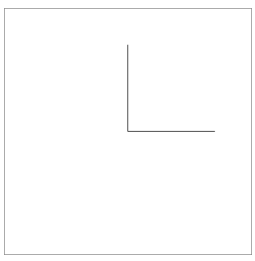
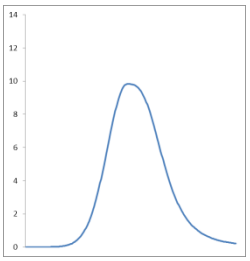
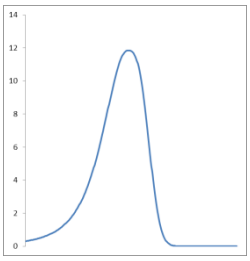
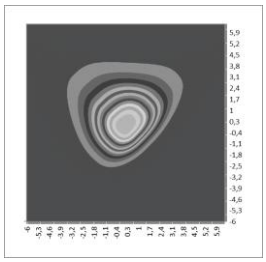
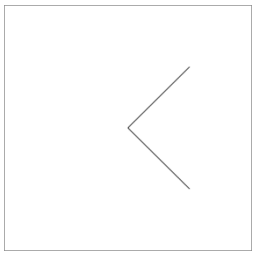
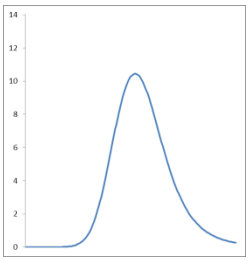
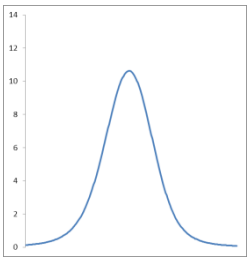
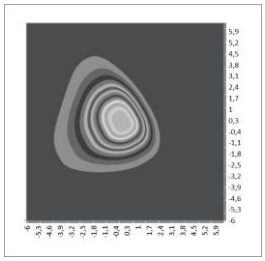
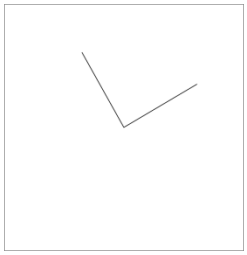
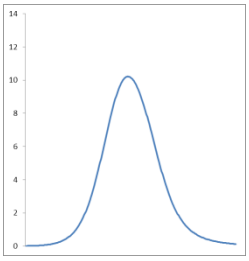
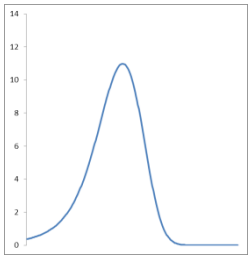
Isodensities	**Change of canonical basis**	**Distribution of vertical margin**	**Distribution of horizontal margin**
	invariance for any orthogonal transformation		
			
			
			

Fig.1. Plots of isodensities, transformation of the canonical basis and marginal distributions of coordinates in Gaussian case (first row) as well as in case of $\mathbf{v} = \mathbf{R}_{(\varphi)}\mathbf{z}$, for $\varphi = 0$ (second row), $\varphi = \frac{\pi}{4}$ (third row), $\varphi = -\frac{\pi}{6}$ (fourth row).

Fig. 1 shows how the shape of the isodensities of $\mathbf{v}$ varies with respect to different values of the shape and the asymmetry parameters. In each case we analyse distributions with variances for margins equal to 4. In the first row we plotted the reference case as the bivariate

Gaussian distribution, being a limiting case here. Possible directional asymmetry and different tail behaviour is presented by the isodensities in the second row. The effect of rotation by a different angle is shown in the third and fourth row. The most interesting property of the analysed distributions is that the dependence pattern assumes zero correlations.

Isodensities	**Change of canonical basis**	**Distribution of vertical margin**	**Distribution of horizontal margin**

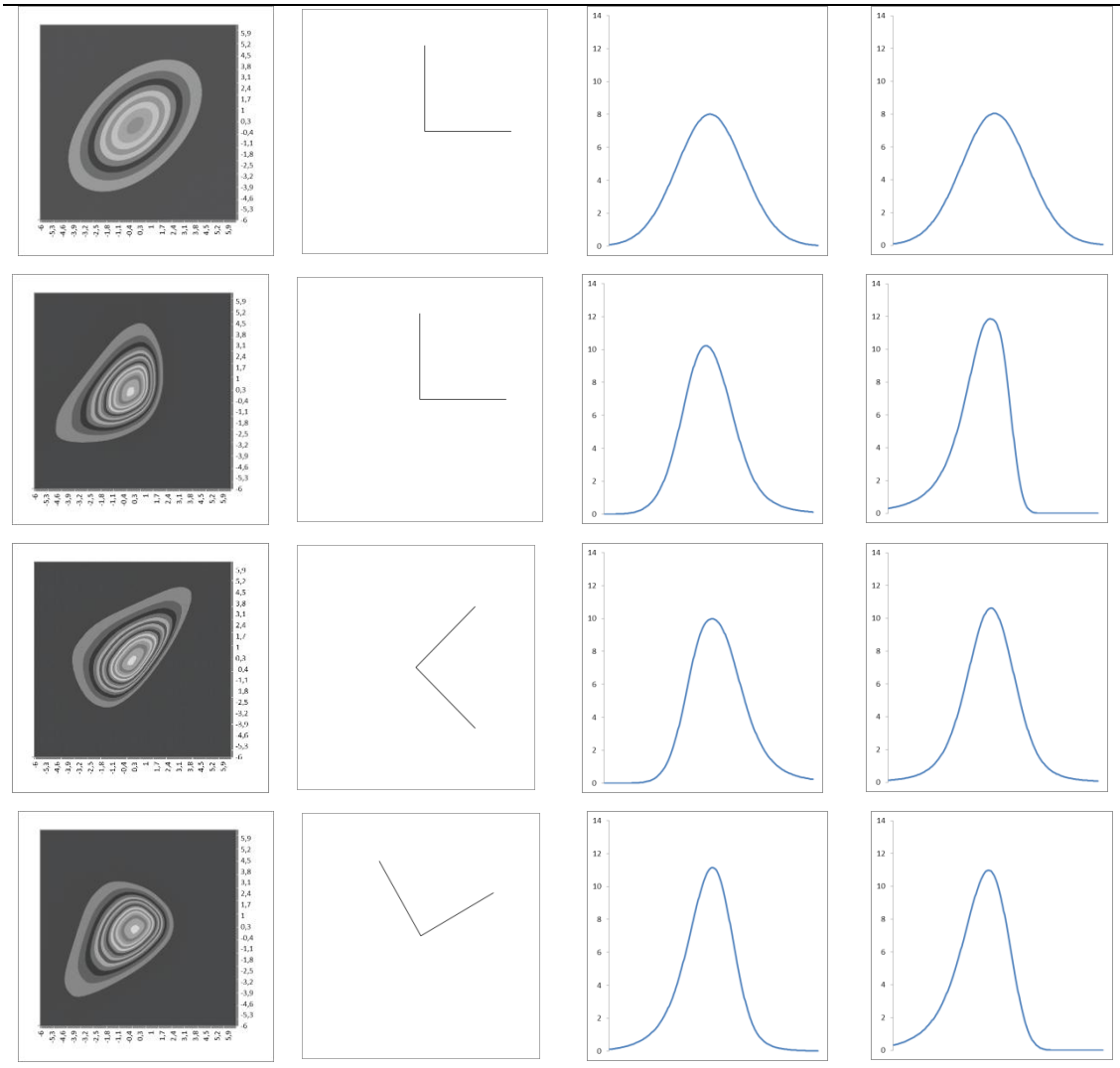


Fig.2. Plots of isodensities, transformation of the canonical basis and marginal distributions of coordinates in Gaussian case (first row) as well as in case of $\mathbf{y} = \mathbf{Cv}$for $\varphi = 0$ (second row), $\varphi = \frac{\pi}{4}$ (third row), $\varphi = -\frac{\pi}{6}$ (fourth row). In each case we keep the same variances for margins equal to 4 and correlation $\rho = 0.5$.

The (rotated) vector **v** is subject to a further linear transformation that imposes location (mode)**m**and covariance structure (i.e. correlation and differences in variances) upon **v**:

$$\mathbf{y} = \mathbf{C}\mathbf{v} + \mathbf{m}.$$

The matrix **C** can be parametrized using different concepts of matrix roots, though here we assume that it has the following form:

$$\mathbf{C} = \begin{bmatrix} \sqrt{1-\rho^2} & \rho \\ 0 & \frac{\sqrt{h_2}}{\sqrt{h_1}} \end{bmatrix}.$$

Then h_i are variances of coordinates and $\rho \in (-1,1)$ represents the correlation coefficient between y_1 and y_2. The density of the distribution of *y* is given by the following formula:

$$p(\mathbf{y}) = \det(\mathbf{C}^{-1}) p_{Z_1}\left(\mathbf{R}_{(\varphi)}^{(1)'}\mathbf{C}^{-1}(\mathbf{y}-\mathbf{m})\right) p_{Z_2}\left(\mathbf{R}_{(\varphi)}^{(2)'}\mathbf{C}^{-1}(\mathbf{y}-\mathbf{m})\right).$$

Fig. 2 depicts isodensities of some exemplary cases of distribution of $\mathbf{y} = \mathbf{C}\mathbf{v}$. We analyse correlated versions of distributions presented on Fig. 1. In each case we assumed correlation coefficient ρ = 0.5. Bivariate distributions presented on Fig. 2 show remarkable degree of flexibility in modelling structure of observables, though it does not exceed the case of a linear transformation of the product measure initially defined for a vector $p(\mathbf{z})$. Its flexibility results from the fact that all the shape parameters could be made dimension-specific (in the space of z's). Crucially, the original formulation of Harvey and Lange allows for a complicated asymmetry pattern which is here generalized to a higher dimension. The distribution is unimodal by construction (which is not necessarily true about some other flexible constructs like mixtures), however its mean is a complicated function of all the model parameters. The rotation angle φis identified if $p(\mathbf{z})$ defines a distribution class that is not closed under rotations, which holds almost everywhere in the parameter space considered here. However, e.g. for a (limiting) special case of bivariate Gaussian distribution, φ would be locally unidentified.

3 Empirical illustration

We analysed daily logarithmic returns of the S&P500 SPOT and FUTURES together with volumes traded, covering the period from 28.08.2001 till 12.12.2017; 4099 observations. We considered four bivariate datasets, namely the daily returns of the SPOT index with daily returns of the FUTURES volume traded (dataset A), the daily returns of the SPOT index with daily returns of the SPOT volume traded (dataset B), the daily returns of the SPOT index with daily returns of the FUTURES index (dataset C) and the daily returns of the SPOT volume

traded with the daily returns of the FUTURES volume traded (dataset D). The empirical; distribution of analysed bivariate series are presented on Fig. 3.

We applied the class of bivariate distributions, presented in the previous part to model the unconditional distribution of analysed series. In order to perform this task we constructed Bayesian models for each analysed series.The estimation was carried over using the Metropolis-Hastings Random Walk sampler; we assume prior independence across all the model parameters. The priors are informative though tailored to convey relatively weak information, for example for φ,ρ and α we assume uniform priors. A posteriori we find limited skewness (with α = 0.5 being rather likely) but strong shape asymmetry (e.g. clear evidence against $\upsilon_L = \upsilon_R$ or $\eta_L = \eta_R$ in some cases) which justifies the empirical relevance of shape-asymmetric distributions of Zhu and Zinde-Walsh (2009), Zhu and Galbraith (2010)as well as Harvey and Lange (2017). We find support for υ_L < 1.5 (e.g. using SPOT index returns), 2 <η_L< 3 e.g. using volume growth rates of FUTURES; full estimation results are available from the authors upon request.

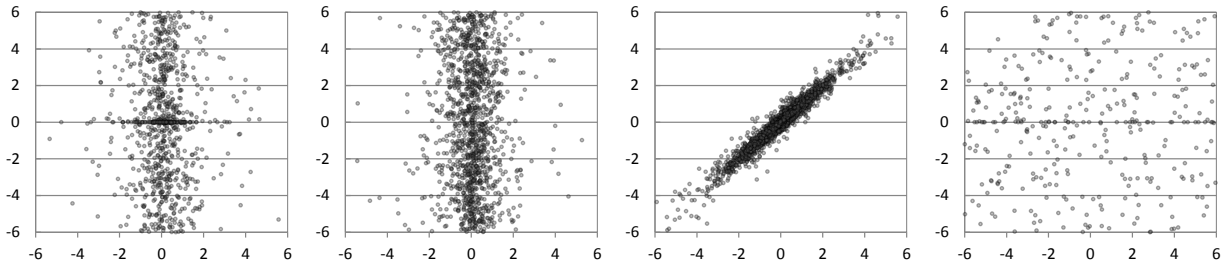


Fig.3. Analysed bivariate time series (the axes are adjusted to match Fig. 4).

The estimated unconditional distributions are presented on Fig. 4. For each dataset we plotted isodensities of the distribution corresponding to$p(\mathbf{y})$with posterior means of parameters used as plug-in estimates. We report the empirical importance of the rotation effect (relying on the posterior mean of parameter φ). The shapes of resulting marginal univariate distributions are also presented.In case of datasets A and B we see a little data support in favour of dependence. Also the rotation effect seems negligible. The posterior mean of parameter φ is equal to 0.068 in case of dataset A and to 0.050 in case of dataset B. Also both datasets support small negative correlation, indicating no substantial linear dependence between the variables. The posterior mean of the correlation ρ= - 0.089 in case of dataset A and ρ= – 0.057 for dataset B.

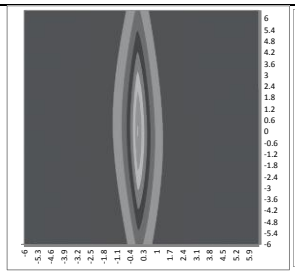
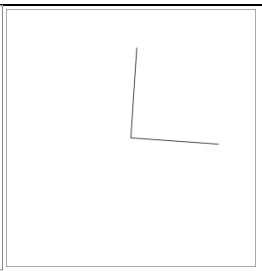
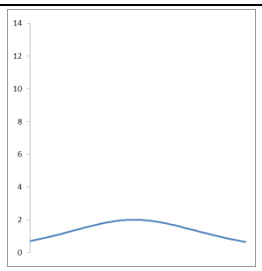
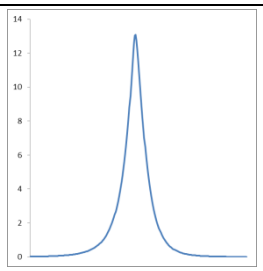
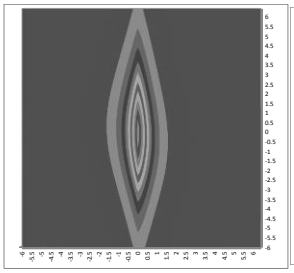
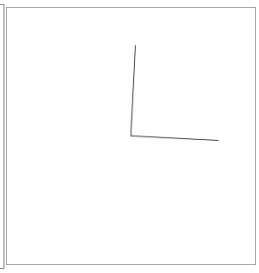
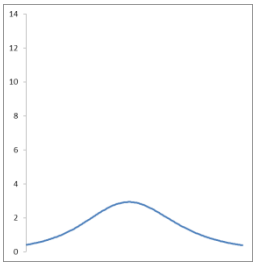
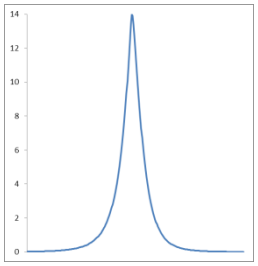
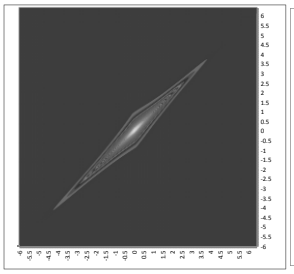
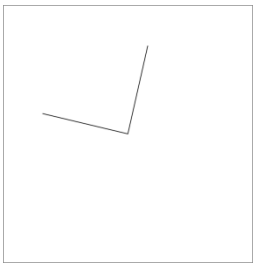
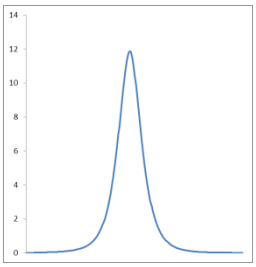
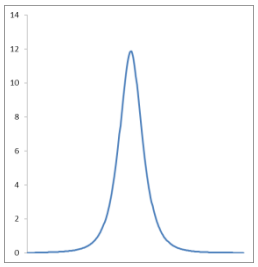
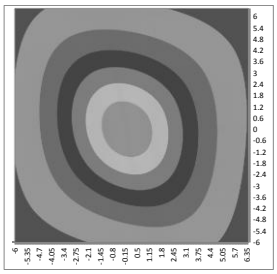
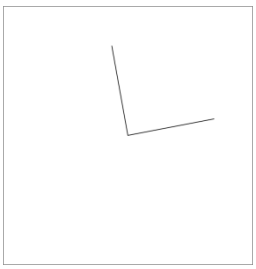
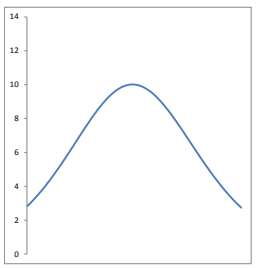
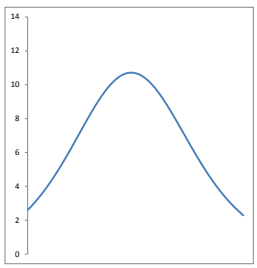
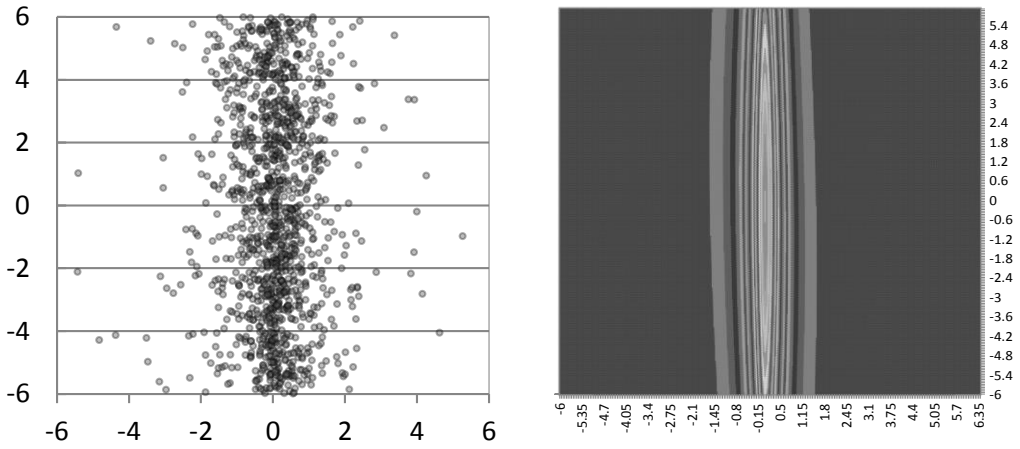
DATA	Estimated isodensities	Estimated orthogonal transormation	Estimated distribution of vertical margin	Estimated distribution of horizontal margin
A				
B				
C				
D				

Fig. 4. Estimation results: plots of isodensities and margins for the parameter values corresponding to posterior means.

The strong linear dependence as well as empirical importance of the rotation effect was obtained for the case of dataset C. Estimated posterior mean of parameter φ = -1.345 indicates strong counter clockwise rotation of coordinates, by more than 75 degrees. The posterior mean of correlation parameter ρ = 0.969 is rather high in this case. A moderate effect of dependence was obtained in case of dataset D. We report some evidence in favour of the rotation effect, as the posterior mean of φ = -0.185. It results with counter clockwise rotation of coordinates by about 10 degrees. The dataset D can be also described by small positive

correlation, since the posterior mean of ρ=0.236. The posterior-predictive distribution (that takes into account the estimation uncertainty) is depicted in Fig. 5 (for the dataset B).



observations posterior-predictive distribution

Fig. 5. Dataset B: the data versus the posterior-predictive density.

Conclusions

We develop a class of flexible multivariate distributions that differs from the usual ones in two particular aspects. Firstly, we allow for high degree of stochastic heterogeneity across variables (addressing asymmetry and tail thickness issues), allowing for 5 shape parameters per dimension. Secondly, we introduce the dependence not only via covariances but also via rotations (which is possible due to generality of the distribution). This approach differs from the alternative ones e.g. using the copula functions: here the form of marginal distribution is not directly controlled. However, the dependence structure is imposed by a well-defined transformation that goes beyond considering covariances only while it remains tractable also in higher dimensions. Hence we provide a practical generalization of a product measure which allows for high degree of heterogeneity, more complicated dependence while avoiding potential problems that arise within high-dimensional modelling using copula functions. The number of shape parameters increases linearly with dimension, but one could of course consider less-heavily parametrized special cases obtained by linear constraints. Therefore the model provides ageneral framework allowing for the search for empirically relevant (restricted) special cases. Importantly, the construct considered here could be used to define conditional distribution in a dynamic model, which will be subject to further research.

Acknowledgements

This research was supported by the research grant 2017/25/B/HS4/02529 financed by the National Science Centre, Poland.

References

Bauwens L. & Laurent S. (2005). A New Class of Multivariate Skew Densities, with Application to GARCH Models.*Journal of Business and Economic Statistics23*(3), 346-354.

Fernández, C. & Steel, M. F. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, *93*(441), 359-371.

Genton, M. G. (2004).*Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Edited vol., Chapman & Hall/CRC, Boca Raton.

Harvey, A. & Lange, R. J. (2017). Volatility Modeling with a Generalized t Distribution. *Journal of Time Series Analysis*, *38*(2), 175-190.

Iglesias, E. M., & Phillips, G.D.A. (2012). Estimation, Testing, and Finite Sample Properties of Quasi-Maximum Likelihood Estimators in GARCH-M Models.*Econometric Reviews31*(5) 532-557

Pelagatti, M. & Rondena S. (2006). Dynamic Conditional Correlation with Elliptical Distributions.*Working Papers*No 20060508, UniversitàdegliStudi di Milano-Bicocca, Dipartimento di Statistica.

Pipień M. (2012) Orthogonal Transformation of Coordinates in Copula M-GARCH Models - Bayesian Analysis for WIG20 SPOT and FUTURES Returns, *Folia OeconomicaCracoviensia 53*, 21-40.

Sahu, S. K., Dey, D. K. & Branco, M. D. (2003).A new class of multivariate skew distributions with applications to bayesian regression models.*Canadian Journal of Statistics31*(2), 129-150.

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution.*Management Science*, *44*(12-part-1), 1650-1661.

Zhu, D. & Galbraith, J. W. (2010). A generalized asymmetric Student-t distribution with application to financial econometrics.*Journal of Econometrics*, *157*(2), 297-305.

Zhu, D. & Zinde-Walsh, V. (2009). Properties and estimation of asymmetric exponential power distribution.*Journal of Econometrics*, *148*(1), 86-99.

Firms' duration analysis in Łódzkie Voivodship using a regression tree approach

Artur Mikulec[1], Małgorzata Misztal[2]

Abstract

Enterprises' duration analysis is a set of statistical techniques dealing with duration time, i.e. the time to the occurrence of an event of interest, for example the liquidation of an enterprise, from its inception. The objective of such analyses is to look for factors (covariates) that significantly affect the duration of the enterprise. The use of traditional statistical methods, e.g. regression analysis, is not possible in this situation due to the occurrence of the censored observations. The most popular statistical methods in duration analysis comprise: the Kaplan-Meier survival curves and the Cox proportional-hazards regression model. As an alternative, survival trees (a regression tree approach) can be used. The model building process is based on recursive partitioning of the set of enterprises into homogenous subsets and the Kaplan-Meier estimate of the survival function or the Cox regression model in each terminal node is reported. The article presents the results of enterprises duration time analysis in Łódzkie Voivodship on the basis of data from the National Official Business Register REGON in the years 2010-2015 with the use of survival trees (the GUIDE algorithm – *Generalized Unbiased Interaction Detectionand Estimation* – proposed by Loh, 2002). The results are compared with the analysis based on the Kaplan-Meier survival curves and the Cox proportional-hazards regression model.

Keywords: *enterprises, duration analysis, recursive partitioning method, survival trees*
JEL Classification:C10, C14, C41
DOI: 10.14659/SEMF.2018.01.31

1 Introduction

Survival analysis is a branch of statistics that concerns various techniques and statistical methods for examining processes, the common feature of which is their duration – from the initial moment to the final moment. Initially, these procedures were used in medicine and related sciences (Machin et al., 2006), currently they are used in many scientific disciplines. In the social sciences and economics, survival analysis is called the analysis of duration.

The key issue in this type of analysis for enterprises is the search for variables significantly affecting their duration. The occurrence of censored observations precludes the use of traditional statistical methods, e.g. regression analysis. The methods most commonly used in the analysis of duration are: the estimation of the Kaplan-Meier survival function and

[1]Corresponding author: University of Łódź, Department of Statistical Methods, 41 Rewolucji 1905 r. Street, 90-214 Łódź, Poland, e-mail: amikulec@uni.lodz.pl.
[2]University of Łódź, Department of Statistical Methods, 41 Rewolucji 1905 r. Street, 90-214 Łódź, Poland, e-mail: mmisztal@uni.lodz.pl.

the Cox proportional-hazardsregression model. As an alternative to the Kaplan-Meier method and the Cox proportional-hazards regression model, a recursive partitioning method, whose graphical representation are a survival trees, could be proposed. The literature of the subject does not provide examples of the application of the recursive partitioning method in the analysis of the duration of enterprises, hence the authors of the present work have attempted to combine these two groups of analysis methods.

The analysis was based on a dataset of enterprises established in the Łódzkie Voivodship in the years 2010-2015, including the liquidated ones. The aim of the conducted research was to identify the characteristics of enterprises that have a significant impact on the duration of enterprises and are associated with the risk of their liquidation.

2 Data characteristics

The individual data about 119.3 thous. of enterprises[3] established in the Łódzkie Voivodship in 2010-2015 were used in the work. Out of this number, 28.3 thous. were liquidated (23.7%) by the end of 2015, and 91.0 thous. (76.3%) continued their economic activity. These firms have been treated as censored data – see Table 1.

Data on enterprises established and liquidated in the Łódzkie Voivodship were obtained from the Statistical Office in Lodz and contained: REGON[4], location, ownership form, legal form, Polish Classification of Activity (PKD 2007), respectively the date of the firm's establishment and/or the date of liquidation of the business and enterprise size – measured with the number of employed persons (up to 9, 10-49, 50 and more).

In case of the ownership form, due to a large number of its variants, companies were aggregated into three sub-groups: private sector entities (pure ownership), public sector entities (pure ownership), private sector enterprises (mixed ownership). Both among liquidated and censored enterprises, private sector enterprises were dominant (pure ownership) – 99.9% and 98.9%, respectively.Taking into account the legal form of

[3] The enterprises were selected according to the methodology of survey *SP-3 Report on economic activity of enterprises* (annual survey of enterprises with up to 9 persons employed),conducted by the Central Statistical Office and the Statistical Office in Lodz. The authors omitted the criterion of the size of the enterprise (number of employed persons), because the analysis was to concern all enterprises conducting manufacturing, trade or service activity on the free market basis – for profit and for self-employment (regardless of their size).

[4] The REGON register is the only available source of data on established and liquidated enterprises – regardless of the legal form, as by definition it contains data on all entities of the national economy. Bearing in mind some imperfections of the REGON register, it should be stated that determining the exact size of a defined cohort of enterprises was not possible, hence the obtained data should be treated as a large research sample (Markowicz, 2012).

enterprises, natural persons conducting economic activity constituted the largest percentage (both among the liquidated and censored enterprises), 97.6% and 89.1%, respectively. From the point of view of enterprise size, microenterprises prevailed – in the group of liquidated and censored entities, the ones with up to 9 persons employed accounted for 99.6% and 99.0%, respectively.Among the analyzed enterprises established in the Łódzkie Voivodship, which were liquidated in 2010-2015, the largest part conducted their activity in Łódź (35.0%). The entities liquidated in the other poviats of the Łódzkie Voivodship constituted from 0.75% (skierniewicki) to 6.76% (zgierski) of all liquidated enterprises. On the other hand, the structure of enterprises liquidated in 2010-2015 by type of business activity was as follows. The highest percentage constituted firms from section G - Trade;repair and motor vehicle (39.0%) and from section F - Construction (13.3%). Entities liquidated in the Łódzkie Voivodship from other sections of the PKD 2007 constituted from 0.07% (B - Mining and quarrying) to 8.4% (C - Manufacturing) of the total liquidated enterprises.

Table 1. The number of enterprises in the Łódzkie Voivodship according to the year of creation.

Year of creation	Enterprises established		Enterprises liquidated	Enterprises censored
	N	%		
Total	**119297**	**100.0**	**28321**	**90976**
2010	22734	19.1	9947	12787
2011	19270	16.1	7219	12051
2012	19049	16.0	5490	13559
2013	19940	16.7	3729	16211
2014	19396	16.3	1507	17889
2015	18908	15.8	429	18479

3 Methods

The analysed dependent variable is the enterprise duration time in days, i.e. the time up to the liquidation of the enterprise. The data are right-censored: enterprises enter the study at different times within 6 years (2190 days), some of them experience the event (liquidation) before the end of the study and some are observed until December 31, 2015 (the end of the study) without the occurrence of the event (so their survival time is censored).The most popular statistical methods in duration analysis comprise: the Kaplan-Meier survival curves (Kaplan and Meier, 1958) and Cox proportional-hazards regression model (Cox, 1972).

The Kaplan-Meier (KM) estimation is a non-parametric procedure used for estimating the survival function of a homogeneous right-censored data. Graphical representation of the KM survival probability against time is the KM survival curve that can be used to estimate measures such as median survival time. To compare KM survival curves for different groups of objects the log-rank test can be applied. The KM estimation does not provide the possibility to model survival times (or hazard functions) as a function of a set of covariates. It can be done with the use of Cox proportional-hazards (PH) model:

$$h(t,\mathbf{x}) = h_0(t)\exp\left(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p\right), \tag{1}$$

where $\mathbf{x} = \left(x_1, x_2, \ldots, x_p\right)$is the vector of *p* explanatory variables (covariates) and$h_0\left(t\right)$is the baseline hazard function. The Cox model is reduced to the baseline hazard when all the explanatory variables are equal to zero or when there are no explanatory variables in the model.

As an alternative, to model the relationship between a survival time and a set of covariates, a regression tree approach can be used. The model building process is based on recursive partitioning of the set of objects into homogenous subsets and the Kaplan-Meier estimate of the survival function or the Cox regression model is reported in each terminal node. The graphical representation of the model takes the form of survival trees. A detailed review of the survival trees algorithms is presented by Bou-Hamad et al. (2011). The advantage of the survival trees over the Cox model lies in the great flexibility, no assumptions on distributions of the survival times and the possibility to automatically detect interactions between covariates without the need tospecify them beforehand. As it is also emphasized by Bou-Hamad et al. (2011): "a single tree can naturally group subjectsaccording to their survival behaviour based on their covariates".

In this paper we propose to use the GUIDE (*Generalized Unbiased Interaction Detectionand Estimation*) algorithm (Loh, 2002) to build the survival tree. GUIDE can fit a piecewise-constant, piecewise-simple linear or piecewise multiple linear proportional-hazards model to the right-censored response data. The general formula of the model fitted by GUIDE is:

$$h(t,\mathbf{x}) = h_0(t)\sum_i I(\mathbf{x} \in S_i)\exp\left(\boldsymbol{\beta}_i{}^T\mathbf{x}\right), \tag{2}$$

where $I(\bullet)$ is the indicator function, $S_i$is a set corresponding node I and $\boldsymbol{\beta}_i$is its associated coefficient vector (for more details see: Loh et al., 2015 and Loh, 2017). In our study the piecewise-constant model was used:

$$h(t, \mathbf{x}) = h_0(t) \sum_i I(\mathbf{x} \in S_i) \exp(\beta_{i0}). \tag{3}$$

All the calculations were done with the use of the following packages: STATISTICA ver. 12.5, SPSS ver.24 and GUIDE ver. 27.4[5].

4 Results

Fig. 1 shows the survival tree obtained using the GUIDE algorithm. Only three variables were used for splitting the nodes. There are 7 terminal nodes (leaves). Sample size (N) and mean relative risk RR (relative to sample average ignoring covariates) are presented for each leaf. Relative risk, i.e. the risk of the liquidation of the enterprise, is calculated using the formula:

$$RR = \exp(\boldsymbol{\beta}^T \mathbf{x} - \beta_*), \tag{4}$$

where **x**is the covariate vector of the observation, $\boldsymbol{\beta}$ is the estimated regression coefficient vector in the node and $\beta_*$is the coefficient in the constant model: $h_0(t)\exp(\beta_*)$, fitted to all the training cases in the root node (see Loh, 2017). In our research $\beta_* = -0.0001228$.

Estimated median survival time is also presented below each node. It is the survival time t such that: $\exp\{-h_0(t)\exp(\boldsymbol{\beta}^T \mathbf{x})\} = 0.5$, using linear interpolation of $h_0(t)$. The median survival times greater than the largest observed time are marked with plus (+) sign.

The survival tree model, presented in Fig. 1, shows that the risk of liquidation is the lowest (RR=0.0631 relative to the sample average for the whole data set) for legal persons (node 5). There are 5990 firms in that node and most of them (98.35%) are limited liability companies and joint stock companies. The group of enterprises with the highest liquidation risk (RR=1.4315 relative to average) entered the node no. 14. There are 12946 firms in that node. These are natural persons conducting economic activity, classified in sections: A, D, E, F, G, I, K, N, R, S (according to the Polish Classification of Activity,PKD2007) and locatedin one of the following poviats: kutnowski, opoczyński, radomszczański, sieradzki, wieluński and wieruszowski. The estimated median survival time in that node is 1739.74 days.

[5] Freely distributed by the Author at http://www.stat.wisc.edu/~loh/guide.html (access: 27.01.2018).

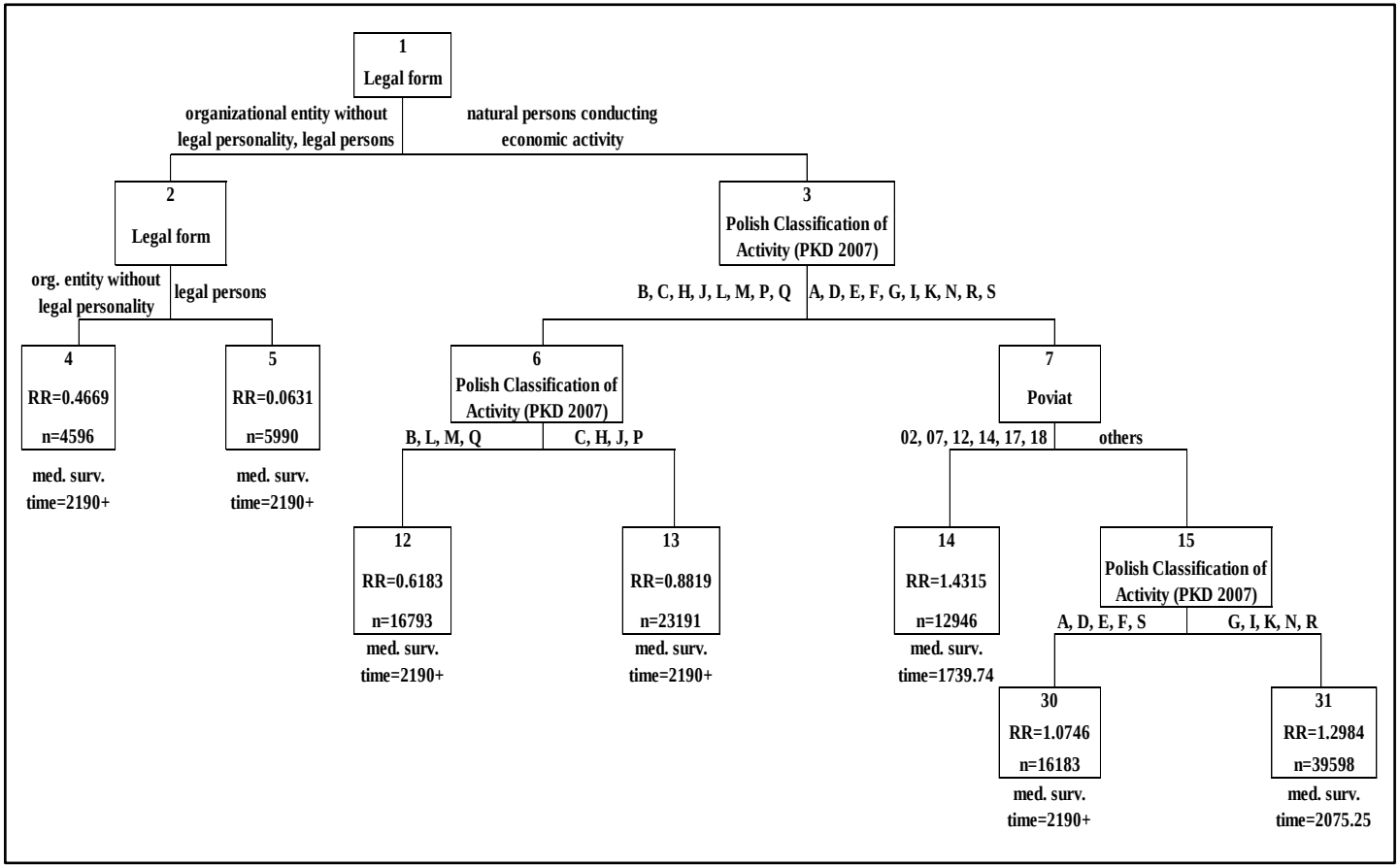


Fig. 1. Piecewise-constant relative risk survival tree (GUIDE algorithm).

Legend: **Polish Classification of Activity (PKD 2007)**:**A -** Agriculture, forestry and fishing; **B -** Mining and quarrying; **C -**Manufacturing; **D -** Electricity, gas, steam and air conditioning supply; **E -** Water supply; sewerage, waste management and remediation activities; **F -**Construction; **G -**Trade; repair of motor vehicles; **H -** Transportation and storage; **I -** Accommodation and food service activities; **J -** Information and communication; **K -** Financial and insurance activities; **L -**Real estate activities; **M -** Professional, scientific and technical activities; **N -** Administrative and support service activities; **P -** Education; **Q -** Human health and social work activities; **R -** Arts, entertainment and recreation; **S -** Other service activities; **Poviats**: **02** - kutnowski; **07** - opoczyński; **12** - radomszczański; **14** - sieradzki; **17** - wieluński; **18-**wieruszowski.

The legal form, Polish Classification of Activity and localization of business activity turned out to be the factors which significantly influenced the risk of enterprise liquidation also in the case of using the Cox proportional-hazards model. Due to the limited length of the paper, the presentation of detailed results has been omitted; only statistically significant results are briefly outlined. The risk of enterprise liquidation significantly increase for: natural persons conducting economic activity (Hazard Rate: HR=2.185; p<0.0001), firms from section K according to PKD 2007 (HR=1.316; p=0.0199) and location of the company in poviats: kutnowski(HR=1.147; p=0.0021), opoczyński (HR=1.123; p=0.0208), piotrkowski(HR=1.112; p=0.0277), radomszczański (HR=1.166; p=0.0002), sieradzki (HR=1.118; p=0.0080) and wieluński(HR=1.112; p=0.0222). Factors that significantly reduce the risk of enterprise liquidation include among others: conducting economic activity by legal persons (HR=0.014; p<0.0001), conducting business activity in the following PKD sections: B (HR=0.590; p=0.0358), C (HR=0.766; p=0.0223), H (HR=0.768; p=0.0255), J (HR=0.693; p=0.0024), L (HR=0.615; p=0.0004), M (HR=0.585; p<0.0001), Q (HR=0.454; p<0.0001 and location of the enterprise in the poviatłódzkiwschodni (HR=0.864; p=0.0018). These results are similar to the conclusions drawn from the analysis of the GUIDE survival tree.

It is also interesting to compare the Kaplan-Meier survival curves for 7 groups of enterprises separated in the leaves of the GUIDE tree (see Fig. 2). The survival probabilities for all the enterprises are also presented (the dotted line). A sharper decrease of the KM survival curve for all firms can be observed in the first 2 years (730 days). This is the result of the end of the period of preferential social security contributions for new enterprises. The probability of survival of 2190 days is 0.5845. The KM curve for all the data separates the curves for the nodes 4, 5, 12, 13 from the curves for the nodes 30, 31, 14.When comparing the survival curves for the GUIDE tree terminal nodes, it can be noticed that the KM curve for the node no. 5 is consistently higher than the KM curves for all other nodes. Only legal persons are observed in the node no. 5 and the probability of surviving 2190 days is 0.9586. The lowest-lying KM curve refers to the node no. 14. The probability of survival of 2190 days is 0.4793.The log-rank test (Mantel-Cox) was applied for overall comparison of KM curves. P-value<0.001 indicates that the null hypothesis (all survival curves are the same) should be rejected both for the joint analysis and for all pairwise comparisons. Since the Kaplan-Meier survival functions are obtained with the use of the traditional product limit formula, the median survival times are not equal to those shown below each node of the GUIDE tree in Fig. 1. The median survival time for the firms located in the node no. 14 is equal to 1808 days and for the firms in the node no. 31 it is greater than the largest observed time.

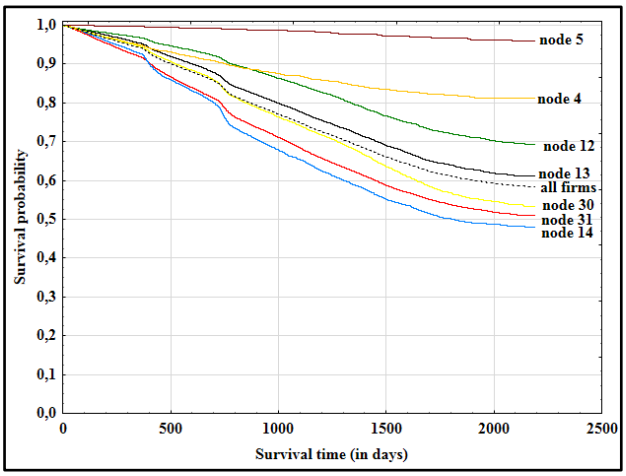


Fig. 2. The Kaplan-Meier survival curves for all data and 7 groups of enterprises separated in the leaves of the GUIDE tree.

Conclusions

Three variables were used to divide enterprisesinto tree nodes: legal form, Polish Classification of Activity (PKD 2007) and the location of the firm's headquarters (poviat).The influence of these variables on the survival of enterprises is confirmed by the literature of the subject.For example, Harhoff et al. (1998) provide evidence for the German market. They state that the joint-stock companies are less willing to leave the market than the natural person conducting economic activity (Harhoff et al., 1998). Also in Portugal, the companies with limited liability were less likely to be liquidated (Mata and Portugal, 2002). For the Polish market, some confirmation could also be found that legalform has an impact on the time of conductinga business activity. Enterprises with a complex legal form are characterized by higher-level survival curves (Szymański, 2011; Śmiech, 2011; Ptak-Chmielewska, 2016).The risk of liquidation of enterprises depends on the type of business. The results obtained by the authors of the work are consistent with other results for the Polish market (Markowicz, 2012). The greatest risk of liquidation (relatively short duration) was characteristic for enterprises from the section G - Trade; repair of motor vehicles and F - Construction. Similar results for enterprises from section G - Trade; repair of motor vehicles can be found in the papers(Śmiech, 2011; Ptak-Chmielewska, 2016).However, taking into account the location of the enterprise (poviat), as the factor affecting its survival, it is necessary to recall the work of Szymański (2011). The author showed in his research that firms located in “typical poviats” and “urban poviats” in Poland are characterized by a lower survival probability and firms located in cities have a higher probability of survival. This conclusion, was confirmed firstly,

in the analysis of the survival of enterprises for the ŁódzkieVoivodship, and secondly, it provides an explanation for the classification of firms to the node no. 14 on the GUIDE tree. They had the lowest Kaplan-Meier survival curve (the probability of survival of 2190 days was equal to 0.4793 for them). These are the poviats located farthest from the city of Lodz, on the outskirts of the ŁódzkieVoivodship, while in 2010-2015 Lodz as the "central city" focused over 39% of enterprises in the ŁódzkieVoivodship(Mikulec, 2017).The share of enterprises liquidated in these poviats was similar (on average around 32.0%) and these were natural persons conducting economic activity mainly from the sections G and F (75.8% of enterprises in total). The possible factors that could affect the relatively short duration of these entities are: a low demand for their products and services due to high unemployment rate (except for the wieruszowski poviat); a large number of super- and hypermarkets (except for the opoczyński and wieruszowski poviats); a significant number of agricultural holdings (except for the kutnowski and wieruszowski poviats) and a low level of the use of EU funds for the creation and development of microenterprises in the wieruszowski poviat. More detailed analysis of factors influencing the duration of enterprises in poviats will be the subject of further research.

As a result of the GUIDE tree algorithm for the division of firms, the size of the enterprise as measured by the number of employed persons and the form of ownershipwere not selected. In the work ofPtak-Chmielewska (2016), the influence of the size of enterprises on their duration was demonstrated, but the analysis was conducted on a small research sample.

The recursive partitioning method can be used as a complement to the classical methods – the estimation of the Kaplan-Meier survival function and the Cox proportional-hazards regression model. Its advantage is the ability to distinguish groups of similar enterprises in terms of survival time and transparent visualization of the results obtained in the case of individual models. Improvement in the stability of created survival tree models can be obtained with the use of ensemble methods. This will become the subject of further research.

References

Bou-Hamad, I., Larocque, D. & Ben-Ameur, H. (2011). A review of survival trees. *Statistics Surveys, 5,* 44-71.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B, 34,* 187-220.

Harhoff, D., Stahl, K. & Woywode, M. (1998). Legal form,growthandexitofwestGermanfirms - empirical results for manufacturing, construction, trade and service industries.*Journal of Industrial Economics, 46(4),* 453-488.

Kaplan, E. L. & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of American Statistical Association, 53,* 457-481.

Loh, W.-Y. (2002). Regression trees with unbiased variable selection and interaction detection.*StatisticaSinica, 12,* 361-386.

Loh, W.-Y. (2017). GUIDE User Manual (updated for version 26.0). Department of Statistics, University of Wisconsin–Madison, http://www.stat.wisc.edu/~loh/guide.html, (access: 27.01.2018).

Loh, W.-Y., He, X. & Man, M. (2015). A regression tree approach to identifying subgroups with differential treatment effects. *Statistics in Medicine, 34,* 1818-1833.

Machin, D., Cheung, Y. B. & Parmar, M. K. B. (2006). *Survival analysis: a practical approach. Second edition*. John Wiley& Sons, Ltd.

Markowicz, I. (2012). Statystyczna analiza żywotności firm. *Rozprawy i Studia t. (CMIX) 835*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Mata, J. & Portugal, P. (2002). The survival of new domestic and foreign-owned firms. *Strategic Management Journal, 23(4),* 323-343.

Mikulec, A. (2017). Analiza wskaźnikowa trwania przedsiębiorstw w województwie łódzkim. *Wiadomości Statystyczne, 11(678),* 29-55.

Ptak-Chmielewska, A. (2016). *Determinanty przeżywalności mikro i małych przedsiębiorstw w Polsce*. Oficyna Wydawnicza SGH, Warszawa.

Szymański, D. (2011). *Badanie żywotności nowo powstałych mikroprzedsiębiorstw w Polsce w latach 2002-2007*.Rozprawa doktorska, http://depotuw.ceon.pl/handle/item/177, (access: 26.01.2018).

Śmiech, S. (2011). Analiza przeżycia podmiotów gospodarczych w województwie małopolskim w latach 2002-2008.*Zeszyty Naukowe Uniwersytetu Ekonomicznegow Krakowie, 876,* 121-132.

The accuracy of trade classification rules for the Warsaw Stock Exchange

Paweł Miłobędzki[1], Sabina Nowak[2]

Abstract

We evaluate the accuracy of five trade classification rules for the Warsaw Stock Exchange: tick, reverse tick, quote, Lee and Ready and Ellis, Michaely and O'Hara. In doing so we use the transaction data on stocks from the large cap WIG20 index from the period May-September 2017. We find that the quote rule correctly classifies 100% of transactions initiated by buyers and sellers. Almost the same excellent job does the Lee and Ready rule. The Ellis, Michaely and O'Hara rule is less successful albeit its success rate exceeds 95% of the transactions assigned to both sides. The tick and the reverse tick rules exhibit a very low accuracy. The tick rule correctly classifies only 25.35% of transactions initiated by buyers and 25.95% of transactions initiated by sellers. The reverse tick rule performs even worse classifying as much as 16.66% and 16.67% of such transactions accordingly. The reason for their low accuracy is that the stock prices remain unchanged at the WSE at about 70% of all transactions. We also show that in case both classification rules are modified to account for either the preceding or the following transactions price changes their accuracy significantly increases.

Keywords: *accuracy of trade classification rules, market microstructure, Warsaw Stock Exchange*
JEL Classification: G10, G14, G15
DOI: 10.14659/SEMF.2018.01.32

1 Introduction

Information on parties to a trade who initiate transactions in financial markets plays an essential role in analysing the intraday price formation within the trade indicator models setup (Hagströmer et al., 2016). It is also helpful for determining the information content of trades, the order imbalance, the inventory accumulation of liquidity providers, the price impact of large transactions as well as the effective spread (Ellis et al., 2000). In case such information is not available which nowadays prevails because most public databases do not contain initiator flags, trade initiators and a trade direction are to be inferred using trade classification rules. These commonly employed in the empirical work include that of the tick (T), the reverse tick (RT), the quote (Q), the Lee and Ready (1991) (LR) as well as the Ellis et al. (2000) (EMO) (see Table 1 for their description). Their accuracy, albeit relatively high, varies across mature markets depending upon the trading price, the trade size and the time from previous trade, being rather low for short sales and trades inside the quotes (see Ellis et al., 2000; Odders-

[1] Corresponding author: University of Gdańsk, Department of Econometrics, ul. Armii Krajowej 101, 81-824 Sopot, Poland, e-mail: pawel.milobedzki@ug.edu. pl.
[2] University of Gdańsk, Department of Econometrics, ul. Armii Krajowej 101, 81-824 Sopot, Poland, e-mail: sabina.nowak@ug.edu.pl.

White, 2000; Finucane, 2000; Savickas and Wilson, 2003; Chakrabarty et al., 2007; Asquith et al., 2010; Rosenthal, 2012). To the best of our knowledge nothing has been known on the issue for emerging and developing markets so far except from the Taiwan Stock Exchange and the exchanges in Istambul and São Paulo (see Lu and Wei, 2009; Aktas and Kryzanowski, 2014; Perlin et al., 2014).[3]

Table 1. Trade classification rules.

Rule	Definition
TR	Current trade is a buy (sell) if its price is above (below) the closest different price of a preceding trade.
RT	Current trade is a buy (sell) if it is followed by a trade with a lower (higher) price.
QR	Current trade is a buy (sell) if its price is above (below) the mid-point of the bid and ask spread (mid-point price). Trades executed at the mid-point price are not classified.
LR	Current trade is a buy (sell) if its price is above (below) the mid-point price (as the QR). In case the trade price is equal to the mid-point price, the TR applies.
EMO	Current trade is a buy (sell) if it is executed at the bid (ask). In case it is executed at in between the quotes, the TR applies.

The aim of the paper is to overcome this deficiency for the WSE. To this end we use the transaction data on stocks from the large cap WIG20 index from the period 2 May-29 September 2017. The data come from Thomson Reuters.[4] We examine the stocks in the WIG20 because they account for approximately 80% in worth of all session trades in the analysed period. The period itself includes two quarterly revisions being held on 16 June and 15 September. Neither of them resulted in stock entries/exits into/from the index. The transaction data are stamped to the nearest millisecond. Each transaction record includes information on the trade price and the volume accompanied by the best bid and ask, the size of bid and ask

[3] Olbryś and Mursztyn (2015) estimated the fraction of trades classified as buys and sells at the WSE by the T, the Q, the LR and the EMO rules but they did not check for the accuracy of classification.

[4] The data are collected from the Thomson Reuters Eikon 4 database under the partnership agreement between the University of Gdańsk and the Thomson Reuters company. The system enables to download the data on trade no older than 3 months.

and the trade flag. The latter enables us to recognize the transaction type (regular, auction, cross, etc). Knowing that we identify a particular transaction as being initiated by a buyer (seller) if the resulting trade price is equal to the best ask (bid) or greater (lower) than that. Our definition of the trade initiator differs from those of Lee and Radhakrishna (1996) and Odders-White (2000) who recognize an investor as the initiator if he (she) places either a market order or places his (her) order chronologically last. We take into account only the normal trade, normal price and regular order transactions and exclude those of the open price, auction trade, trading at last, short sales as well as the cross ones. In the result our data set consists of 4658470 trades.

We find that the Q rule correctly classifies 100% of transactions initiated by the buyers and the sellers. Almost the same excellent job does the LR rule. The EMO rule is less successful albeit its success rate exceeds 95% of the transactions assigned to both sides. The T rule correctly classifies 25.35% of transactions initiated by the buyers and 25.95% of transactions initiated by the sellers. The RT rule performs even worse classifying 16.66% and 16.67% of such transactions accordingly. Their low accuracy is due to the fact that in the analysed period the stock prices at the WSE remain unchanged at about 70% of all transactions. In case we modify the T and the RT rules to account for either the preceding or the following transactions price changes their accuracy increases to reach 67.59% at most.

The remainder of the paper proceeds as follows. In Section 2 we sketch the present state of the art in examining the accuracy of common classification rules. In Section 3 we explain in more detail the nature of data we use in the empirical work and report on the accuracy of rules in question and their modification. The last Section briefly concludes.

2 Short Review of the Literature on the Classification Rules Accuracy

We gather the results of previous research on the accuracy of classification rules in Table 2.[5] Most of the findings invoked therein refer to the U.S. markets. Of a particular interest are those obtained for assets on data sets exhibiting a more detailed information on quotes and trades and referring to the multiplicity of rules.

Ellis et al. (2000) used the data on 313 newly traded NASDAQ stocks between 27 September 1996 and 29 September 1997 which contained the bid quote, the ask quote, the price, the trade volume, a trader identity code and a buy/sell indicator. By using the latter two they could determine whether a trade was a buy or a sell and documented 77.7%, 76.4%, 81.1%

[5] We report only on research performed on stocks and options written on them. The full list of papers regarding other assets is available to concerned readers on a request.

and 81.9% accuracy for the T, the Q, the LR and the EMO rules, respectively. They also find that all rules perform rather poorly in case of trades executed inside the quotes, large trades, trades during high volume periods and Electronic Communications Network trades.

Finucane (2000) based his research on the TORQ database which contained information on quotes, orders and trade direction for the sample of 144 stocks randomly selected from a size-stratified population of NYSE stocks for the three-month period November 1990 through January 1991. He compared the actual trade direction with that predicted by the T and the LR classification rules and found out that both methods were surprisingly about the same correctly identifying trade direction between 83% to 84% of the time. He suggested that the similar performance of both rules was due to the significant sample proportion of cross and stopped orders and trades between the quotes that might lower the accuracy of LR algorithm.

Lee and Radhakrishna (2000) on the same database noticed that approximately 40% of reported trades could not be unambiguously classified as either buys or sells mainly due to the presence of cross and stopped orders. Nevertheless 84.0% and 93.0% out of those classified agreed with the TORQ classification while identified with the use of the Q and the LR rules. The similar conclusions were reached by Odders-White (2000) who estimated the performance of the T, the Q and the LR rules at 78.6%, 74.9% and 85.0%, respectively. She also revealed that the rules systematically misclassified trades inside the quotes, small trades as well as trades in large or frequently traded stocks. Chakrabarty et. al (2007) concluded on trade classification rules accuracy using two samples consisted of 750 NASDAQ stocks traded on the INET ECN, the April and between April and June 2005 ones. The data contained information on the order, the trade history as well as on the buy/sell indicator. The major finding was that applying all rules to both samples resulted in almost equal success rate estimates around 74-77%. Rosenthal (2012) restricted attention to stocks included in the Russel 1000 large caps and 2000 small cap indices. His dataset covered transactions across 2836 stocks from three primary U.S. markets (AMEX, NASDAQ and NYSE) on two first trading days of December 2004. He showed that the introduction of delay models for estimating quotes and the modelling approach to trade classification led to a 1-2% improvement of the T, the LR and the EMO rule success rates in comparison to current methods.

Chakrabarty et al. (2015) using the true trade classification for equity transactions derived from the NASDAQ's Total View-ITCH order book and quotes and trades data from the Daily TAQ database provided by the NYSE showed that the success rate for the T and the LR rules were around 90-92%, but the latter rule was slightly more accurate across all strata (time, volume and trade bars for large, medium and small caps).

Table 2. Trade classification rules accuracy – summary of the literature review.

Study	Market, Instruments, Period	Rule			
		T	Q	LR	EMO
Ellis et al. (2000)	NASDAQ, 313 stocks, 96M9-97M9	77.7	76.4	81.1	81.9
Finucane (2000)	NYSE, 144 stocks, 90M11-91M1	83.0		84.4	
Lee and Radhakrishna (2000)	NYSE, 144 stocks, 90M11-91M1		84.0	93.3	
Odders-White (2000)	NYSE, 144 stocks, 90M11-91M1	78.6	74.9	85.0	
Savickas and Wilson (2003)	CBOE, 826 options, NASDAQ/OTC stocks, 95M7-95M12	59.4	82.8	80.1	76.5
Chakrabarty et. al (2007)	NASDAQ, 750 stocks, 05M4	75.6		75.8	76.8
	NASDAQ, 750 stocks, 05M5-M6	75.4		74.4	75.8
Asquith et al. (2010)	NASDAQ, 100 stocks, 05M3, M6, M12	41.3^{s}	37.5^{s}	39.1^{s}	
		60.3^{b}	61.0^{b}	62.9^{b}	
	NYSE, 100 stocks, 05M3, M6, M12	4.7^{s}	14.4^{s}	14.6^{s}	
		98.9^{b}	83.6^{b}	88.0^{b}	
Chakrabarty et al. (2012)	NASDAQ, 200 stocks, 05M6-M12			68.2^{st}	
				69.2^{lt}	
Rosenthal (2012)	NYSE, NASDAQ, AMEX, 2836 stocks, 04M12	66.2		71.7	72.8
Chakrabarty et al. (2015)	NASDAQ, 300 stocks, 11M5, M6, M7	90.8^{i}		92.6^{i}	
		89.7^{ii}		91.6^{ii}	
		89.2^{iii}		91.2^{iii}	
	NYSE, 300 stocks, 11M5, M6, M7	91.7^{i}		93.4^{i}	
		89.0^{ii}		90.9^{ii}	
		87.8^{iii}		90.1^{iii}	
Aitken and Frino (1996)	ASX, ECN, all stocks, 92M7-94M6	74.4			
Theissen (2001)	FSE, 15 stocks, 96M9-96M10	72.2	75.4	72.8	
Lu and Wei (2009)	TWSE, 684 stocks, 06M1-M6	74.2	92.8	96.5	95.0
Aktas and Kryzanowski (2014)	BIST, 30 stocks, 08M6-M12	90.4	95.0	96.4	86.9
Perlin et al. (2014)	BOVESPA, 15 stocks, 09M1-10M1	72.0			
Pöppe et al. (2016)	DB, 30 stocks, 12M10-M11	82.0		86.6	90.4

[b]Buyer initiated trades, [s]Seller initiated trades, [st]Short trades, [lt]Long trades, [i]One-hour time bar; [ii]100-trade bar; [iii]10000-volume bar. In case of Asquith et al. (2010) the maximum levels of accuracy are given.

Finally Savickas and Wilson (2003) compared the ability of the trade classification rules in question to classify options trades at the CBOE. Their dataset reported the trade direction on options underlain by 826 assets, mainly stocks from the NYSE/AMEX and the NASDAQ/ OTC. They estimated the rule success rate ranging from 59.4% (T rule) to 82.8% (Q rule). More interestingly, all rules happened to perform very poorly for the index options. The main source of their misclassification were outside-quote and reversed-quote trades.

The general conclusion stemming from the short review of accuracy rules performance on the U.S. markets is threefold. First, the estimates of accuracy apart from those reported in Asquith et al. (2000) for short sales range from mediocre to marvellous. Second, the LR rule and (or) the EMO rule perform better than their T and (or) Q counterparts on the NASDAQ, the NYSE and the AMEX, but on average the difference in their accuracy is rather slight (see Ellis et al., 2000; Finucane, 2000; Lee and Radhakrishna, 2000; Odders-White, 2000; Chakrabarty et al., 2007, 2015; Rosenthal, 2012). Third, trades in between the quotes and cross orders are deemed to have been the main sources of misclassification. The same applies to other international markets: the Deutsche Börse (Pöppe et al., 2016) and the Taiwan Stock Exchange (Lu and Wei, 2009). However at the CBOE, in Frankfurt and Istanbul the Q and the LR rules perform the best (see Theissen, 2001; Aktas and Kryzanowski, 2014; Savickas and Wilson, 2003).

3 Data and Empirical Results

Our primary dataset covers transactions on stocks from the WIG20 in the period 24 April-5 October 2017. It consists from trades labelled with 10 different trade flags indicating possible transaction types being executed at the WSE albeit it is the 'Normal Trade, Normal Price, Regular Order' trades that dominate the dataset.[6] Thus we clean it up dropping all residual flag trades as well as those of April and October to retain only those exhibiting the three-month regular trade in 20 largest cap stocks from May through September. Next we identify the trades as buys, sells and those executed inside the quotes (closer to the best ask/bid and not identified – traded at the mid-point price) comparing their trade prices with the best asks

[6] They constitute 96.63% of the dataset. The relevant frequency statistics are available to concerned readers on a request.

and the best bids that follow. We stack the results of cleaning up and identification of trades (trade initiators) in Table 3. They indicate that of all trades the buys at the best ask and the sells at the best bid prevail adding up to 99.35% of all cleaned trades. A tiny rest mainly consists of all inside the quotes trades including those not identified and those closer to the best ask and closer to the best bid. The fractions of buys above the best ask and sells below the best bid are residual.

Table 3. Trades included in the cleaned dataset by their type.

Trade type	No. of trades		Fraction	
Buy – Above Best Ask	1859		0.04	
Buy – At Best Ask	2323008	2324867	49.87	49.91
Inside – Closer to Best Ask	11756		0.25	
Inside – Not Ident	822	26366	0.02	0.56
Inside – Closer to Best Bid	13788		0.30	
Sell – At Best Bid	2305169		49.48	
Sell – Below Best Bid	2068	2307237	0.04	49.53
All trades		4658470		100.00

Table 4. Classification success rates for the T, the RT, the Q and the LR rules.

Trade type	T	RT	Q	LR	EMO
Buy – Above Best Ask	22.22	35.56	100.00	100.00	22.22
Buy – At Best Ask	25.35	16.64	100.00	100.00	95.82
Inside – Closer to Best Ask	20.58	15.24	100.00	100.00	20.58
Inside – Not Ident	87.10	66.79	100.00	87.10	87.10
Inside – Closer to Best Bid	21.78	12.98	100.00	100.00	21.78
Sell – At Best Bid	25.96	16.66	100.00	100.00	95.79
Sell – Below Best Bid	32.74	27.13	100.00	100.00	32.74
Buy	25.35	16.66	100.00	100.00	95.76
Inside	23.28	15.66	100.00	99.60	23.28
Sell	25.97	16.67	100.00	100.00	95.73
All trades	25.64	16.66	100.00	100.00	95.34

The results in Table 4 lay groundwork for the accuracy comparison of the rules in question. The accuracy statistics derived for all trades show that the Q rule performs the best cor-

rectly identifying all trades. The second is the LR rule which does almost the same excellent job. The next is the EMO which misclassifies only 4.66% of all trades. The remaining rules perform very poorly. The T rule misclassifies 74.36% of all trades while the RT rule misclassifies as much as 83.34% of them.

The accuracy statistics derived solely for buys, sells and inside the quotes trades indicate that the LR rule misclassifies only 0.4% of the latter. The EMO is slightly worse misclassifying 4.24% of buys, 4.27% of sells and 76.72% of inside the quotes trades. The accuracy of other rules for buys, sells and inside the quotes trades is about the same as their accuracy for all trades.

Table 5. Classification success rates for the modified T and RT rules.

Trade type	T1	T2	T3	T4	T5	RT1	RT2	RT3	RT4	RT5
Buy – Above BA[a]	31.79	36.47	39.27	40.77	41.64	51.86	59.33	64.17	67.24	69.34
Buy – At BA	40.62	50.55	57.45	62.45	66.18	26.26	32.40	36.63	39.70	42.00
Inside – Closer to BA	33.07	40.90	46.87	50.88	53.99	22.58	26.62	29.13	30.74	31.77
Inside – Not Ident	77.13	69.46	63.14	57.91	54.50	48.66	37.47	28.95	23.97	20.68
Inside – Closer to BB[b]	35.05	44.22	50.46	55.03	58.38	19.18	22.86	25.18	26.59	27.47
Sell – At BB	41.57	51.78	58.82	63.85	67.59	26.45	32.65	36.90	39.97	42.26
Sell – Below BB	49.85	59.04	64.94	69.29	72.10	39.51	46.23	50.39	53.29	55.27
Buy	40.61	50.54	57.43	62.43	66.16	26.28	32.42	36.65	39.72	42.03
Inside	35.48	43.53	49.26	53.27	56.30	21.62	24.99	27.06	28.36	29.18
Sell	41.58	51.79	58.82	63.86	67.59	26.46	32.66	36.92	39.98	42.27
All trades	41.06	51.12	58.07	63.09	66.81	26.35	32.50	36.73	39.78	42.07

[a]BA – best ask, [b]BB – best bid

The poor performance of the T and the RT rules (and the EMO for inside the quotes trades) is due to that in the period in question prices at the WSE remain unchanged at about 70% of all trades. But in case sequences of two and more consecutive trades are considered the fraction of them for which prices remain unchanged dramatically decreases. That is why we propose a modification of the T and the RT rules to account for either the preceding or the following transaction price changes. Their performance for different sequence lengths up to five is given in Table 5. In what follows the accuracy of both classification rules increases with the increasing length of sequence considered to reach at most 66.81% for the T5 rule. For all trades the estimates of accuracy statistics for the T rule are superior over those for the

RT rule. The difference in classification success rates ranges from 14.71% (R1 vs. RT1) to 24.74% (R5 vs. RT5). Almost the same applies to the buys, the sells and the trades executed inside the quotes. Nevertheless the accuracy rates for the modified RT rules are rather at unacceptable levels (below 50%). The accuracy of our classification results is comparable to those of Lu and Wei (2009), Aktas and Kryzanowski (2014), and Perlin et al. (2014) for TWSE, BIST and BOVESPA respectively, except for the T rule.

Conclusions

We estimate the accuracy rates of five common classification rules (T, RT, Q, LR, EMO) using the tick data on 20 large cap stocks listed at the WSE in the period May-September 2017. We find that the Q rule performs the best correctly identifying 100% of trades. The second is the LR rule which misclassifies only 0.4% of the inside the quotes trades. The next is the EMO misclassifying as much as 4.66% of all trades. The T and the RT rules perform very poorly misclassifying 74.36% and 83.34% of them, respectively. The reason is that in the analysed period prices at the WSE remain unchanged at about 70% of all trades. In case we modify the T and the RT rules to account for either the preceding or the following transactions price changes their accuracy increases to reach at most 66.81%. For all trades the estimates of accuracy statistics for the T rule are superior over those for the RT rule.

References

Aitken, M. & Frino, A. (1996). The accuracy of the tick test: Evidence form the Australian Stock Exchange. *Journal of Banking & Finance*, *20*, 1715-729.

Aktas, O. U. & Kryzanowski, L. (2014). Trade classification accuracy for the BIST. *Journal of International Financial Markets, Institutions & Money*, *33*, 259-282.

Asquith, P., Oman, R. & Safaya, C. (2010). Short sales and trade classification algorithms. *Journal of Financial Markets*, *13*(1), 157-173.

Chakrabarty, B., Li, B., Nguyen, V. & Van Ness, R. A. (2007). Trade classification algorithms for electronic communications network trades. *Journal of Banking and Finance*, *31*(12), 3806-3821.

Chakrabarty, B., Moulton, P. C. & Shkilko, A. (2012). Short sale, long sale, and the Lee-Ready trade classification algorithm revisited. *Journal of Financial Markets*, *15*(4), 467-491.

Chakrabarty, B., Pascual, R., & Shkilko, A. (2015). Evaluating trade classification algorithms: Bulk volume classification versus the tick rule and the Lee-Ready algorithm. *Journal of Financial Markets*, *25*, 52-79.

Ellis, K., Michaely, R. & O'Hara, M. (2000). The Accuracy of Trade Classification Rules: Evidence from Nasdaq. *Journal of Financial and Quantitative Analysis*, *35*(4), 529-551.

Finucane, T. J. (2000). A Direct Test of Methods for Inferring Trade Direction from Intra-Day Data. *Journal of Financial and Quantitative Analysis*, *35*(4), 553-576.

Hagströmer, B., Henricksson, R. & Nordén, L. L. (2016). Components of the Bid-Ask Spread and Variance: A Unified Approach. *Journal of Futures Markets*, *36*(6), 545-563.

Lee, C. M. C. & Radhakrishna, B. (2000). Inferring investor behavior: Evidence from TORQ data. *Journal of Financial Markets*, *3*, 83-111.

Lee, C. M. C. & Ready, M. J. (1991). Inferring Trade Direction from Intraday Data. *Journal of Finance*, *46*(2), 733-746.

Lu, Y. C. & Wei, Y. C. (2009). Classification of trade direction for an equity market with price limit and order match: evidence from the Taiwan stock market. *Investment Management and Financial Innovations*, *6*(3), 135-147.

Odders-White, E. R. (2000). On the occurrence and consequences of inaccurate trade classification. *Journal of Financial Markets, 3*, 259-286.

Olbryś, J. & Mursztyn, M. (2015). Comparison of selected trade classification algorithms on the Warsaw Stock Exchange. *Advances in Computer Science Research, 12,* 37-52.

Perlin, M., Brooks, C. & Dufour, A. (2014). On the performance of the tick test. *Quarterly Review of Economics and Finance, 54*, 42-50.

Pöppe, T., Moos, S. & Schiereck, D. (2016). The sensitivity of VPIN to the choice of trade classification algorithm. *Journal of Banking & Finance, 73*, 165-181.

Rosenthal, D. W. R. (2012). Modeling Trade Direction. *Journal of Financial Econometrics*, *10*(2), 390-415.

Savickas, R. & Wilson, A. J. (2003). On Inferring the Direction of Option Trades. *Journal of Financial and Quantitative Analysis*, *38*(4), 881-902.

Theissen, E. (2001). A test of the accuracy of the Lee/Ready trade classification algorithm. *Journal of International Financial Markets, Institutions and Money*, *11*(2), 147-165.

Dynamic measure of development

Iwona Müller-Frączek[1]

Abstract

The article presents a possible using of the author's method – normalisation with respect to the pattern – in the construction of synthetic measure. When a stimulant (destimulant) is normalized, for each object the share of its distances from the maximum (minimum) in the total distance from the maximum (minimum) of all objects is determined. Such transformation meets the requirements of normalisation - deprives variables their units and unifies their ranges. Normalisation with respect to the pattern has properties suggested in the literature - preserves skewness, kurtosis and the Pearson correlation coefficients. Moreover, although the current data are the sole data used to convert variables, normalized diagnostic variables are comparable across time. This feature gives an advantage of pattern normalisation over other methods in dynamic analysis of complex phenomena.

The article uses normalisation with respect to the pattern in construction of Hellwig's measure of development, in which Euclidean distances from an abstract ideal point are calculated. Since normalized diagnostic variables become destimulants with the minimum value equals 0, the ideal point used to construct a synthetic measure is constant over time. So, the values of modified measures are comparable both across objects and time. One can compare the positions of objects in the rankings as well as the values of the measures themselves (calculate the increments of values, descriptive characteristics, etc.).

Keywords: *synthetic measure, aggregate variable, composite indicator, normalisation, standardisation*
JEL Classification: C19, C38
DOI: 10.14659/SEMF.2018.01.33

1 Introduction

The article concerns the methods of comparing objects due to the level of a complex phenomenon, and more specifically, the linear ordering of these objects. Objects are identified with points in a multidimensional space. To order these points, their one-dimensional projections are constructed. In this way, a synthetic measure of a complex phenomenon is defined. The synthetic measure is also called the aggregate variable or the composite indicator.

The issues raised in the article are quite popular. For a selected complex phenomenon, many different examples of synthetic variables can be found in the literature. For example, in Booysen (2002), the history of creating various composite indicators of socio-economic

[1]Corresponding author: Nicolaus Copernicus University, Faculty of Economic Sciences and Management, Gagarina 13A, 87-100 Toruń, Poland, e-mail: muller@econ.umk.pl.

development of countries has been described. Synthetic variables are created for both science and politics. They are then also widely exploited in journalism.

Due to its simplicity, the synthetic approach to describing qualitative phenomena has many supporters. On the other hand, this type of simplification of complex phenomena is too great for many professionals. The advantages and disadvantages of composite indicators can be found among others in Saltelli (2007).

It is certain that the construction of a synthetic measure should be carried out with due diligence. Rules that should be followed and subsequent stages of construction are presented, for example, in Saisan and Saltelli (2011) or Zeliaś (2002a).

The last stage of constructing a synthetic variable is the aggregation of diagnostic variables. The most common methods are the simplest ones: arithmetic mean and geometric mean. In Poland, Hellwig's method (more advanced) is widespread. The method, called *measure of economic development*, has been described in Hellwig (1968 a, b) and quoted in Fanchette (1971). The measure of development bases on multidimensional Euclidean distances from a pattern - a point in the space whose coordinates are determined by the most favourable observations of diagnostic variables.

The article proposes a modification of the Hellwig's method. It mainly concerns the use of another type of variables normalisation. There are many methods of normalisation (see for example: Milligan and Cooper, 1988; Jajuga and Walesiak, 2000; Pawełek, 2008; Zeliaś, 2002a, 2002b). The article uses *normalisation with respect to the pattern*. This is a new method proposed in Müller-Frączek (2017b). This normalisation has advantages that allow us to construct a dynamic synthetic measure whose values are comparable both across objects and time.

The use of the pattern normalisation in another type of synthetic variable can be found in Müller-Frączek (2017a).

The layout of the article is as follows: Section2 reminds the original Hellwig's construction, Section 3 quotes the concept and main properties of the normalisation with respect to the pattern, Section 4 presents the construction of dynamic measures of objects development, the article ends with conclusions.

2 Hellwig's measure of development

Assume that our goal is to order $n \in \mathbb{N}$objects according to the level of the complex phenomenon. We know a collection of $r \in \mathbb{N}$ diagnostic variables, which characterize this phenomenon. Then the objects are identified with points in a r-dimensional space with

coordinates equal to the values of diagnostic variables. Letx_{ip}, $i = 1, \dots, n$,$p = 1, \dots, r$bethe corresponding data matrix.

Assume that diagnostic variables meet both substantive and statistical requirements (for more details see for example Zeliaś, 1982). Among themwe distinguish stimulants and destimulants. Stimulants positively influence the analyzed phenomenon, whereas the influence of destimulants is negative. The set of stimulants is marked by S, while the set of destimulants by D.

In the first step of the original Hellwig's method, diagnostic variables should be standardized to render them comparable. After standardisation they take the form:

$$x_{ip}^{'} = \frac{x_{ip} - \overline{x_p}}{S(x_p)}, \tag{1}$$

where$\overline{x_p}$is the average value of the variable x_p:

$$\overline{x_p} = \frac{1}{n}\sum_{j=1}^{n} x_{jp} \ , \tag{2}$$

whereas $S(x_p)$is the standard deviation of the variable x_p:

$$S(x_p) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(x_{jp} - \overline{x_p})^2} \ . \tag{3}$$

In the next step of the method, *the patternof economic development*$x^+ = (x_1^+, x_2^+, \dots, x_r^+) \in \mathbb{R}^r$is determined. The pattern is an idealabstract point in the multidimensional space, whose coordinates take the most favourable values of the considered diagnostic variables:

$$x_p^+ = \begin{cases} \max\limits_j x_{jp}^{'} & \text{if} \quad x_p \in S \\ \min\limits_j x_{jp}^{'} & \text{if} \quad x_p \in D \end{cases} . \tag{4}$$

The basis for the construction of Hellwig's development measure are the Euclidean distances between objects and the pattern:

$$d_i^+ = \sqrt{(x_{i1}^{'} - x_{i1}^+)^2 + (x_{i2}^{'} - x_{i2}^+)^2 + \cdots + (x_{ir}^{'} - x_{ir}^+)^2} \ . \tag{5}$$

These distances form a synthetic measure, which is a numerical characteristic of the analyzed qualitative phenomenon.The greateris value d_i^+, the worse is the situation of i-th object.

Since d_i^+ are not normalizedand the direction of the relationship between them and the phenomenon is opposite than expected,*the measure of economic development* of i-th object isgiven by:

$$m_i = 1 - \frac{d_i^+}{\overline{d^+} + S(d^+)}, \tag{6}$$

where $\overline{d^+}$ is the average distance between objects and pattern:

$$\overline{d^+} = \frac{1}{n}\sum_{j=1}^{n} d_j^+ , \tag{7}$$

while $S(d^+)$ is the standard deviation of these distances:

$$S(d^+) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(d_j - \overline{d^+}\right)^2} . \tag{8}$$

The denominator of the formula (6) guarantees limitation of development measure. Only in extreme, very rare cases, the values of the measure go beyond the interval $[0,1]$.

In the original, the Hellwig's method was described in a static situation - for a given unit of time. However, in practice it is also used in dynamic researches. In this case, to ensure the comparability of results across time, both observations for all objects and all time units are taken into account when determining the mean (2) and deviation (3) as well as the pattern (4) (see for example Müller-Frączek and Muszyńska, 2016). However, such stochastic approach raises doubts, especially in the case of regional research in which we work with all objects - that is, a population, not a sample (see also Zeliaś, 2002a, 2002b).

An additional disadvantage of the stochastic approach is the necessity of recalculating all results with the appearance of observations for the next unit of time.

3 Normalisation with respect to the pattern

An application of normalisation with respect to the pattern (Müller-Frączek, 2017b) in the construction of synthetic measures can be a solution of the problems indicated at the end of the previous section. A characteristic feature of this method is the comparability of normalized values of variables across time, although a deterministic and not stochastic approach is used for normalisation.

For simplicity, consider one diagnostic variable, which is observed for $n \in \mathbb{N}$ objects and $T \in \mathbb{N}$ time units. For each $t = 1, \ldots, T$ let $x^t = (x_1^t, x_2^t, \ldots, x_n^t) \in \mathbb{R}^n$ be the corresponding data vector.

For each unit of time $t = 1, \ldots, T$, we choose the most favourable value of the variable x^t, we call them the *pattern values* (or *patterns* for short):

$$x^{t+} = \begin{cases} \max\limits_i x_i^t & \text{if} \quad x \in S \\ \min\limits_i x_i^t & \text{if} \quad x \in D \end{cases}. \tag{9}$$

Note that pattern values change over time.

For $t = 1, \ldots, T$ the pattern values are used for normalisation of the variable x^t according to the formula:

$$u_i^{+t} = \frac{|x_i^t - x^{+t}|}{\sum_{j=1}^n |x_j^t - x^{+t}|} = \begin{cases} \dfrac{x^{+t} - x_i^t}{\sum_{j=1}^n \left(x^{+t} - x_j^t\right)} & \text{if} \quad x \in S \\ \dfrac{x_i^t - x^{+t}}{\sum_{j=1}^n \left(x_j^t - x^{+t}\right)} & \text{if} \quad x \in D \end{cases}. \tag{10}$$

The transformation (10), called *normalisation with respect to the pattern* or *pattern normalisation* for short, satisfies the requirements for normalisation - it deprives variables their units and unifies their ranges. Furthermore, pattern normalisation has properties suggested in the literature for such type of transformation (compare Walesiak, 2014; Jajuga and Walesiak, 2000): it preserves skewness and kurtosis of distributions, as well as does not change Pearson's linear correlation coefficients between variables (proofs and more others properties can be found at Müller-Frączek, 2017b).The Table 1. presents some descriptive characteristics of variables after normalisation with respect to the pattern, to simplify the notation, the indexes are omitted.

Table 1. Descriptive characteristics of the distribution of variables after pattern normalisation.

Name of characteristic	Value of characteristic
Arithmeticmean	$\overline{u^+} = \dfrac{1}{n}$
Standard deviation	$S(u^+) = \begin{cases} \dfrac{S(x)}{n(x^+ - \bar{x})} & \text{if} \quad x \in S \\ \dfrac{S(x)}{n(\bar{x} - x^+)} & \text{if} \quad x \in D \end{cases}$
Skewness	$A(u^+) = \begin{cases} -A(x) & \text{dla} \quad x \in S \\ A(x) & \text{dla} \quad x \in D \end{cases}$
Kurtosis	$K(u^+) = K(x)$
Pearson correlationcoefficient	$\mathrm{r}(u_1^+, u_2^+) = \begin{cases} \mathrm{r}(x_1, x_2) & \text{if } x_1, x_2 \in S \text{or} x_1, x_2 \in D \\ -\mathrm{r}(x_1, x_2) & \text{otherwise} \end{cases}$

Transformation (10), like standardisation (1), belongs to the group of normalisations, for which the value of the normalized variable for the i-th object is influenced by all the values of

the variable. Scaling (or unitarisation), which is very popular method of normalisation in empirical researches, does not have this feature. In this case, only the maximum and minimum influence the values after normalisation.

Transformation (10) is not just a technical operation. After pattern normalisationvariables have a clear interpretation. For i-th object, the value of normalized variable determines the share of its distance from the pattern in the total distance from the pattern of all objects. In the context of constructing synthetic variables, this means that after pattern normalisation diagnostic variables become destimulants, irrespective of their initial nature.

In dynamic studies, the most important advantage of pattern normalisation is the comparability of the values of normalized variables both across objects and time, although only values from the current unit of time are used for transformation. If the value of a normalized variable for the i-th object has increases, then the situation of this object has worsened.

Normalisation with respect to the pattern can be used, among others, for the construction of synthetic measures. An example is the additive synthetic measure presented in Müller-Frączek, 2017a. The present article proposes an application of pattern normalisation in the construction of *dynamic measure of development* based on the Hellwig's concept.

4 Dynamic synthetic measure

As in Section 2, consider a complex phenomenon, which is characterized by a set of $r \in \mathbb{N}$ diagnostic variables. These variables are observed for $n \in \mathbb{N}$ objects in the space and $T \in \mathbb{N}$ units of time. For each$t = 1, \dots, T$letx_{ip}^{t}bethe data matrixof dimension $n \times r$.

The steps of constructing dynamic measure of development are analogous to the original Hellwig's method. At the beginning, all diagnostic variables are brought to comparability, using pattern normalisation instead of standardisation.This transformation causes that all variables become destimulants, and their minimum values are equal to 0. Therefore, the pattern obteined in the next step takes the form $(0,0, \dots ,0) \in \mathbb{R}^{r}$. This pattern does not change over time, this is an important advantage from the point of view of constructing dynamic synthetic measures.

In the next step the multidimensional Euclidean distances between the objects and the pattern d_i^{0t} are determined. They take the form:

$$d_i^{0t} = \sqrt{(u_{i1}^{+t})^2 + (u_{i2}^{+t})^2 + \dots + (u_{ir}^{+t})^2}\,. \tag{11}$$

Distances (11) are the quantitative descriptions of the objects due to the analyzed qualitative phenomenon.

Because of the comparability of normalized variables, also distances from the pattern are comparable both across objects and time. If the value of d_i^{0t} is greater than $d_j^{0\tau}$, then the situation of the i-th object at the moment t is worse than the situation of the j-th object at the moment τ.Note that, one can compare not only the rankings, but also the values of the synthetic measure d (calculate its increments, descriptive characteristics, etc.).

Next, for $t = 1, \dots, T$, $i = 1, \dots, n$we define *dynamic measures of development*of the i-th object at the moment $t$by the formula:

$$\mu_i^t = 1 - \frac{d_i^{0t}}{\sqrt{r}}. \tag{12}$$

The higher is the value of the measure, the better is the situation of the object.

Similarly to Pluta (1976), the denominator in the formula (12) depends only on the number of variables. Such a form preserves the comparability of measure μ both across objects and time. Additionally, unlike the original method, the values of measures never go beyond the range [0,1].

Conclusions

The article presents the construction of dynamic measure of development, based on Hellwig's concept. The Hellwig's measure of development uses multidimensional Euclidean distances from a pattern - a point in the space whose coordinates are determined by the most favourable observations of diagnostic variables. In the original, the method applies to static situation.The article proposes a dynamic version of Hellwig's measure.The proposed modification consists primarily in the application of normalisation with respect to the pattern instead of standardisation. The values of the measures are comparable not only across space, but also across time, although only the current observations are used in their determination. In this way, the stochastic approach to normalisation and pattern determination is avoided. Such approach is controversial in regional research, in which we work with the whole population of objects.

In subsequent studies, an attempt will be made to include spatial relationships in the construction of measures, as in Pietrzak (2014).

References

Booysen, F. (2002). An overview and evaluation of composite indices of development. *Social Indicators Research, 59*(2), 115-151.

Fanchette, S. (1971). Synchronic and diachronic approaches in the UNESCO project on human resources indicators. Wroclaw Taxonomy and Bivariate diachronic analysis. In:*Workshop on Indicators of National Development*, International Social Science Council, European Consortium for Political Research, Lausanne.

Hellwig, Z. (1968a). Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr. *Przegląd Statystyczny*,*15*(4), 307-326.

Hellwig, Z. (1968b). *Procedure of Evaluating High-level Manpower Data*. UNESCO, Methods and Analysis Unit, Department of Social Sciences.

Jajuga, K.& Walesiak, M. (2000). Standardisation of Data Set under Different Measurement Scales. *Studies in Classification, Data Analysis, and Knowledge Organization Classification and Information Processing at the Turn of the Millennium,* 105-112.

Milligan, G. W.& Cooper, M. C. (1988). A study of standardisation of variables in cluster analysis. *Journal of Classification*,*5*(2), 181-204.

Müller-Frączek, I. (2017a). Propozycja miary syntetycznej. *Przegląd Statystyczny*,*64*(4), 413-428.

Müller-Frączek, I. (2017b). Normalisation with respect to pattern. *Repozytorium UMK*.http://repozytorium.umk.pl/handle/item/4769.

Müller-Frączek, I. & Muszyńska, J. (2016). Implementation of Europe 2020 Strategy: the taxonomic analysis, In: Simberova, I., Milichovsky, F. & Zizlavsky, O. (eds.), *Smart and efficient economy: preparation for the future innovative economy*, University of Technology Faculty of Business and Management, Brno, 346-355.

Pawełek, B. (2008). *Metody normalizacji zmiennych w badaniach porównawczych złożonych zjawisk ekonomicznych*. Wydawnictwo Uniwersytetu Ekonomicznego, Kraków.

Pietrzak, M. (2014). Taksonomiczny miernik rozwoju (TMR) z uwzględnieniem zależności przestrzennych. *Przegląd Statystyczny*, *61*(2), 181-201.

Pluta, W. (1976). Taksonomiczna procedura prowadzenia syntetycznych badań porównawczych za pomocą zmodyfikowanej miary rozwoju gospodarczego. *Przegląd Statystyczny*, *23*(4), 511-518.

Saisana, M. & Saltelli, A. (2011). Rankings and Ratings: Instructions for Use. *Hague Journal on the Rule of Law*,*3*(02), 247-268.

Saltelli, A. (2007). Composite Indicators between Analysis and Advocacy. *Social Indicators Research,81*(1), 65-77.

Walesiak, M. (2014). Przegląd formuł normalizacji wartości zmiennych oraz ich własności w statystycznej analizie wielowymiarowej. *Przegląd Statystyczny*, *61*(4), 363-372.

Zeliaś, A. (1982). Kilka uwag o kryteriach doboru zmiennych w modelach ekonometrycznych. *Folia Oeconomica Cracoviensia*, *24*, 21-36.

Zeliaś, A. (2002a). Uwagi na temat wyboru metody normowania zmiennych diagnostycznych, In: Kufel, T., Piłatowska, M. (eds.), *Analiza szeregów czasowych na początku XXI wieku*, Wydawnictwo Uniwersytetu Mikołaja Kopernika, Toruń.

Zeliaś. A. (2002b). Some notes on the selection of normalisation of diagnostic variables. *Statistics in Transition*,*5*(5), 787-802.

Assessing accuracy of trade side classification rules. Methods, data, and problems

Joanna Olbryś[1], Michał Mursztyn[2]

Abstract

Trade side classification algorithms enable us to assign the side that initiates a transaction and distinguish between the so-called buyer- and seller-initiated trades. According to the literature, such classification is essential to assess both market liquidity and dimensions of market liquidity based on high frequency intraday data. The main problem is that trade and quote data is not publicly available for many stock markets and researchers have to utilize indirect methods to infer trade side. The aim of this paper is to investigate major problems in assessing accuracy of trade side classification algorithms. We evaluate and compare four most frequently utilized procedures using intraday data for 105 companies from the Warsaw Stock Exchange (WSE). Moreover, an analysis of the robustness of the results is provided over the whole sample period from January 2, 2005 to December 30, 2016, and three consecutive sub-periods of equal size, covering the pre-crisis, crisis, and post-crisis periods. The empirical experiment shows that the Lee-Ready (1991) algorithm and tick rule perform better than other methods on the WSE, regardless of the choice of the sample.

Keywords: *market microstructure, high frequency data, trades, quotes, trade side classification procedures*
JEL Classification:C10, C55, C63, C81, G1
DOI: 10.14659/SEMF.2018.01.34

1 Introduction

High frequency data is important in studying a variety of issues related to trading processes and market microstructure. To measure both market liquidity and dimensions of market liquidity based on intraday data, it is essential to recognize the side that initiates a transaction, and to distinguish between the so-called buyer- and seller-initiated trades. In the absence of information regarding whether a trade is buyer or seller initiated, many researchers employ various classification procedures to infer trade side on financial markets in the world.

On the contrary to research conducted on international stock markets, there are only a few studies that make use of trade side classification algorithms for the WSE data. For instance, Nowak (2017) analyses the problem of asset pricing on the basis of high frequency data for the WSE. She uses the quote rule to classify the side that initiates a transaction. Olbryś (2017) conducts the study of interaction between market depth and market tightness on the WSE and

[1] Corresponding author: Bialystok University of Technology, Faculty of Computer Science, Wiejska 45 A, 15-351 Białystok, j.olbrys@pb.edu.pl.
[2]Bialystok University of Technology, Faculty of Computer Science, Wiejska 45 A, 15-351 Białystok, m.mursztyn@student.pb.edu.pl.

she uses the Lee and Ready (1991)algorithm to infer trade sides. Olbryś and Mursztyn (2017) estimate and analyse selected liquidity proxies derived from intraday data and supported by the Lee-Ready algorithm inferring the initiator of a trade.

As the subject of accuracy of a trade side classification is crucial for market microstructure research, the goal of this paper is to investigate major problems in assessing accuracy of selected algorithms inferring the initiator of a trade. We evaluate and compare four most frequently utilized procedures based on intraday data for 105 WSE-listed companies in the period from January 2, 2005 to December 30, 2016. Moreover, an analysis of the robustness of the obtained results is provided over the whole sample period and three consecutive sub-samples of equal size, namely the pre-crisis, crisis, and post-crisis periods. The empirical experiment shows that the Lee-Ready algorithm and the tick rule perform better than other methods on the WSE, regardless of the choice of the sample.

The remainder of the study is organized as follows. Section 2 specifies algorithms used in the research and presents a brief literature review in the context of assessing accuracy of trade side classification procedures. Section 3 describes the data. In Section 4, we present and discuss the empirical experiment for high frequency data from the WSE. The last section encompasses the conducted research with a brief summary.

2 Trade side classification procedures

The goal of trade side classification is to determine the initiator of a transaction and to classify trades as being either buyer or seller motivated. However, a formal definition of a trade initiator is rarely stated in the literature. Moreover, researchers use various definitions of a trade initiator based presumably on data availability. For example, the so-called 'immediacy' definition describes initiators as traders who demand immediate execution (Lee and Radhakrishna, 2000). Thus, it is usually possible to classify each trade as either buyer- or seller-initiated. Odders-White (2000) considers the last arriving order to be trade initiator (the so-called ‘chronological’ definition). These two definitions are equivalent in many cases. In both of them, the initiator is the person who caused the transaction to occur. Theissen (2001) proposes somewhat different definition based on a position taken by a specialist, which is appropriate for a hybrid equity market.

Table 1. Selected trade side classification procedures.

Rule	Conditions	
QR	Trade is classified as buyer-initiated if	Trade is classified as seller-initiated if
	$P_t > P_t^{mid}$	$P_t < P_t^{mid}$
	If $P_t = P_t^{mid}$ then a trade is not classified.	
TR	Trade is classified as buyer-initiated if	Trade is classified as seller-initiated if
	$P_t > P_{t-1}$	$P_t < P_{t-1}$
	When $P_t = P_{t-1}$, a trade is signed using the previous transaction price.	
	If the sign of the last non-zero price change is positive (negative) then the trade is classified as buyer-initiated (seller-initiated).	
LR	I stage	
	Trade is classified as buyer-initiated if	Trade is classified as seller-initiated if
	$P_t > P_t^{mid}$	$P_t < P_t^{mid}$
	If $P_t = P_t^{mid}$ then:	
	II stage	
	Trade is classified as buyer-initiated if	Trade is classified as seller-initiated if
	$P_t^{mid} > P_{t-1}$	$P_t^{mid} < P_{t-1}$
	When $P_t^{mid} = P_{t-1}$, the decision is made using the sign of the last non-zero price change. If $P_t > P_{t-k}$ then it is buyer-initiated. If $P_t < P_{t-k}$ then it is seller-initiated.	
EMO	I stage	
	Trade is classified as buyer-initiated if	Trade is classified as seller-initiated if
	$P_t = P_t(a) = P_t^L$	$P_t = P_t(b) = P_t^H$
	II stage	
	Trades with prices different from best ask and best bid prices are categorized by the first stage of the tick rule and P_t is compared to P_{t-1}.	
	If $P_t > P_{t-1}$ then trade is classified as buyer-initiated.	
	If $P_t < P_{t-1}$ then trade is classified as seller-initiated.	

Notes: P_t - the transaction price at the time *t*, approximated by the closing price.

$P_t(a) = P_t^L$ - the best ask price, approximated by the lowest price at the time *t*.

$P_t(b) = P_t^H$ - the best bid price, approximated by the highest price at the time *t*.

$P_t^{mid} = \frac{P_t(a)+P_t(b)}{2} = \frac{P_t^L+P_t^H}{2}$ - the so-called quoted midpoint at the time *t*.

Table 2. Assessing accuracy of selected trade side classification procedures.

Market	Procedures	Data	Results	Source
Australian Stock Exchange (ASX)	TR	SEATS[3] database	An accuracy of ≈74% (TR)	Aitken and Frino (1996)
NASDAQ	QR, TR, LR, EMO	TAQ[4] database	An accuracy of 76.4% (QR), 77.66% (TR), 81.05% (LR),81.9% (EMO)	Ellis et al. (2000)
NYSE	QR, TR, LR	TORQ[5]database	An accuracy of ≈75% (QR), ≈79% (TR), ≈85% (LR)	Odders-White (2000)
NYSE	TR, LR	TORQ database	An accuracy between 83% and 84% (both TR and LR)	Finucane (2000)
Frankfurt Stock Exchange (FSE)	TR, LR	Handels-überwachungsstelle	An accuracy of 72.8% (LR) and 72.2% (TR)	Theissen (2001)
Taiwan Stock Exchange (TWSE)	QR, TR, LR, EMO, and others	TWSE database	An accuracy of 92.75% (QR), 74.18% (TR), 96.46% (LR), 94.97% (EMO)	Lu and Wei (2009)
NYSE, NASDAQ	QR, TR, LR	TAQ, CRSP[6] databases	QR, TR, and LR classify a majority of short sales as buyer-initiated	Asquith et al. (2010)
NASDAQ	TR, LR, BVC	ITCH[7], TAQ databases	An accuracy of 90.8% (TR), 92.6% (LR), 79.7% (BVC)	Chakrabarty et al. (2015)

There are quite many trade side classification methods described in the literature, but the most frequently used are: the quote rule (QR), the tick rule (TR), the Lee-Ready (LR) algorithm, and the Ellis-Michaely-O'Hara (EMO) algorithm (see Table 1.)[8]. Easley et al.

[3]SEATS – the Stock Exchange Automated Trading System , the Australian Stock Exchange.
[4]TAQ –the Trades and Quotes database.
[5]TORQ–the Trades, Orders, Reports, and Quotes database.
[6] CRSP –the Center for Research in Security Prices.
[7]ITCH – the ITCH database provided by NASDAQ.
[8]The implementation of four trade side classification algorithms is presented in detail in the paper (Olbryś and Mursztyn, 2015).

(2013) propose the BVC (Bulk Volume Classification) procedure, but this approach apportions trades into buy volume and sell volume, and does not assign trades to be either buyer- or seller-initiated.

The accuracy of trade classification algorithms has been examined in many studies, however the results have been diverse for various stock markets. Table 2. summarizes the empirical results of assessing the accuracy of selected trade side classification procedures on some international stock markets. We focus on the QR, TR, LR, and EMO methods.

As presented in Table 2., there are some discrepancies in trade side classification results between markets. The possible explanation of them is that stock market structure, institutional differences between markets, and trading mechanisms may affect the accuracy of classification procedures. Specifically, the U.S. stock markets (NYSE, AMEX and NASDAQ) and the FSE are hybrid and primarily quote-driven markets in which market-makers/specialists play a prominent role, while the ASX, TWSE and WSE are order-driven markets with fully automated trading systems.

3 Data description

In this study, high frequency data 'rounded to the nearest second' from the WSE (available at www.bossa.pl) for the 105 WSE-listed companies was utilized. The dataset contains the opening, high, low and closing prices, as well as volume for a security over one unit of time. The whole sample covers the period from January 2, 2005 to December 30, 2016 (3005 trading days). To verify the robustness of the obtained results, the comparison of trade side classification rules is provided both for the whole sample and over three consecutive subsamples, each of equal size (436 trading days) (see Olbryś and Mursztyn, 2015):

1. The pre-crisis period from September 6, 2005 to May 31, 2007.
2. The crisis period from June 1, 2007 to February 27, 2009.
3. The post-crisis period from March 2, 2009 to November 19, 2010.

The crisis period on the WSE was formally defined based on the paper (Olbryś and Majewska, 2015), in which the statistical method for the formal identification of market states was employed.

When forming the database, we included only the securities that were listed on the WSE for the whole sample period since December 31, 2004, and were not suspended. The 139 WSE companies met these basic conditions, and they were initially selected. Next, to mitigate the non-trading problem on the WSE, we excluded the stocks that exhibited extraordinarily many non-traded days during the whole sample period, precisely, above 300 zeros in daily

volume, which constituted about 10% of all 3005 trading days. The database consists of 105 companies regarding the presented way of their selection. The dataset is large. Within the trading days during the whole sample period, the total number of records in the database is equal to 35 307 993.

4 The empirical experiment on the Warsaw Stock Exchange

We evaluate and compare four most frequently utilized trade side classification rules for the data from the WSE. An analysis of the robustness of the results is provided over the whole sample period and three adjacent sub-periods, each of equal size. The average percentage values of classified and unclassified trades in the case of all trade classification procedures, for the whole group of 105 WSE-listed stocks are presented in Table 3. The empirical findings indicate that usefulness of various trade side classification methods on the WSE is not qualitatively the same, whereas the results turn out to be robust to the choice of the period.

Specifically, the tick rule (TR) and the Lee-Ready (LR) algorithm are more appropriate compared to the quote rule (QR) and EMO procedure. In the case of the TR and LR methods, the percentage of unclassified transactions is relatively low and similar, which is consistent with the literature. For example, Theissen (2001) points out that the Lee-Ready method classifies transactions quite correctly, but the simpler tick test performs almost equally well. The amount of buyer- and seller-initiated trades is almost equal, with a little predominance of buyer-initiated in all investigated periods. This evidence is in accordance with the literature, as it is demonstrated in some papers that short sales are sometimes misclassified as buyer-initiated by some trade side classification algorithms (Asquith, 2010).

On the contrary, applicability and accuracy of the QR and the EMO procedures is rather low, with high percentage of unclassified trades for all companies, regardless of a firm's size and the choice of the period. In our opinion, this phenomenon on the WSE is caused by the problem of relatively many trades for which high and low prices are equal over one unit of time. Moreover, the EMO method was proposed for the NASDAQ, which is a hybrid market, while the WSE is a pure order-driven market. Hence, the WSE differs from the NYSE and the NASDAQ, and therefore the empirical results based on the U.S. stock markets are not comparable to Polish stock market.

Table 3. The average percentage values of classified and unclassified trades for the whole group of the 105 WSE-listed companies (the LR, TR, QR, and EMO procedures). The best results for the LR algorithm are marked in bold.

Rule	Period	Mean value of buyer-initiated trades (%)	Mean value of seller-initiated trades (%)	Mean value of unclassified trades (%)
LR	Whole sample	48.06%	45.58%	**6.36%**
	Pre-crisis	49.08%	45.41%	**5.51%**
	Crisis	46.47%	47.03%	**6.50%**
	Post-crisis	47.77%	44.85%	**7.38%**
TR	Whole sample	47.95%	45.60%	6.45%
	Pre-crisis	49.07%	45.35%	5.58%
	Crisis	46.28%	47.11%	6.61%
	Post-crisis	47.64%	44.86%	7.50%
QR	Whole sample	5.88%	6.35%	87.77%
	Pre-crisis	5.95%	6.44%	87.61%
	Crisis	5.85%	6.73%	87.42%
	Post-crisis	5.69%	5.89%	88.42%
EMO	Whole sample	6.35%	5.88%	87.77%
	Pre-crisis	6.44%	5.95%	87.61%
	Crisis	6.73%	5.85%	87.42%
	Post-crisis	5.89%	5.69%	88.42%

Table 4. The percentage of identically and differently classified trades for the pairs (TR, LR) and (QR, EMO) during the whole sample period from January 2, 2005 to December 30, 2016, for the whole group of 105 WSE-listed companies.

	(TR, LR)	**(QR, EMO)**
Both unclassified	1.79%	88.84%
The same classification	97.02%	0.04%
Mixed classification (includes opposite classification)	1.19%	11.12%

TR and LR procedures perform better than QR and EMO methods. For this reason, subsequent Table 4. includes more details concerning identically and differently classified

trades for the pairs (TR, LR) and (QR, EMO). One can observe that QR and EMO procedures work substantially worse than TR and LR on the WSE.

According to the literature, the main way of assessing accuracy of trade side classification procedures is to compare classification results with true trade directions. The problem with transaction data availability is vast for the WSE, and hence the following question has been formulated:

- Is it possible to directly test the accuracy of trade side classification algorithms for the intraday data from the Warsaw Stock Exchange?

Regardless of the choice of a formal definition of a trade initiator, assigning a trade direction requires presence of the trades and quotes that precede the trade. However, the transaction dataset for the WSE contains only trade prices, the best-bid and best-ask prices, volume, date, and order-entry time[9]. Unlike that, transaction data sets that are utilized in other studies contain more useful information. For example, Aitken and Frino (1996) test the accuracy of the tick rule for the data from the ASX, which is (like the WSE) a pure order-driven market (see Table 2.). They use the data that include all trades and quotes. Lu and Wei (2009) employ database provided by the TWSE, which is also an order-driven market (see Table 2.). The transactions data are preserved in the order, trade, and disclosure files. The trade file contains the date, trade time, order type (buy or sell) of transaction, trade volume, trade serial number, trade price, trade categories, and the identify of a trader. The order file includes the date, stock code, order type (buy or sell), and so on.

Comparing the available data with those used in other studies we conclude that the answer to the above question is rather negative. Unfortunately, we are not able to recognize true trade directions on the WSE. Therefore, we have to rely on the literature and utilize indirect methods for inferring the initiator of a trade in the research concerning various aspects of microstructure of the Polish stock market.

Conclusion

Analysing the accuracy of the trade side classification is obviously important as this accuracy determines the validity of empirical research based on classification methods. In practice, to measure both market liquidity and dimensions of market liquidity based on intraday data, it is essential to recognize the side that initiates a transaction, and to distinguish between the so-called buyer- and seller-initiated trades. However, the problem with transaction data

[9]Transaction data coming from Bloomberg under the license agreement between Bloomberg and Bialystok University of Technology (the grant No. 2016/21/B/HS4/02004).

availability is crucial. Our empirical results based on intraday data rounded to the nearest second indicate that the tick rule and the Lee-Ready algorithm outperforms the quote rule and EMO procedure on the WSE. The main advantage of the tick rule is that it requires only trades (quotes are not necessary). The advantage of the Lee-Ready algorithm is that it combines tick and quote rules, and therefore incorporates more information since it utilizes past quotes. Inasmuch as the TR and LR procedures perform well and the results are qualitatively and quantitatively much the same for the data from the WSE, we can essentially recommend both methods for further research on Polish equity market. Moreover, the empirical experiments confirm that the trade side classification results turn out to be robust to the choice of the period. Specifically, they are not significantly different during the crisis period compared to other investigated periods. The robustness of the TR and LR algorithms to a sample choice is certainly an important merit.

Acknowledgements

This research has been supported by the grant No. 2016/21/B/HS4/02004 from the National Science Centre in Poland.

References

Aitken, M. & Frino, A. (1996). The accuracy of the tick test: Evidence from the Australian Stock Exchange. *Journal of Banking and Finance, 20*, 1715-729.

Asquith, P., Oman, R. & Safaya, C. (2010). Short sales and trade classification algorithms. *Journal of Financial Markets, 13(1)*, 157-173.

Chakrabarty, B., Pascual, R. & Shkilko, A. (2015). Evaluating trade classification algorithms: Bulk volume classification versus the tick rule and the Lee-Ready algorithm. *Journal of Financial Markets, 25*, 52-79.

Easley, D., López-de Prado, M. & O'Hara, M. (2013). Bulk classification of tradingactivity. *Johnson School Research Paper 8-2012*.

Ellis, K., Michaely, R. & O'Hara, M. (2000). The accuracy of trade classification rules: Evidence from Nasdaq. *Journal of Financial and Quantitative Analysis, 35(4)*, 529-551.

Finucane, T. (2000). A direct test of methods for inferring trade direction from intraday data. *Journal of Financial and Quantitative Analysis, 35(4)*, 553-576.

Lee, C.M.C. & Radhakrishna, B. (2000). Inferring investor behavior: Evidence from TORQ data. *Journal of Financial Markets, 3*, 83-111.

Lee, C.M.C., & Ready, M. (1991). Inferring trade direction from intraday data. *Journal of Finance, 46(2)*, 733-746.

Lu, Y.C. & Wei, Y.C. (2009). Classification of trade direction for an equity market with price limit and order match: Evidence from the Taiwan stock market. *Investment Management and Financial Innovations, 6(3),*135-147.

Odders-White, E. (2000). On the occurrence and consequences of inaccurate trade classification. *Journal of Financial Markets, 3*, 259-286.

Nowak, S. (2017). Order imbalance indicators in asset pricing: Evidence from the Warsaw Stock Exchange. In: Jajuga, K, Orlowski, L and Staehr, K (eds), *Contemporary Trends and Challenges in Finance. Springer Proceedings in Business and Economics*, Springer, Cham, 91-102.

Olbryś, J. (2017). Interaction between market depth and market tightness on the Warsaw Stock Exchange: A preliminary study. In: Jajuga, K, Orlowski, L andStaehr, K (eds),*Contemporary Trends and Challenges in Finance. Springer Proceedings in Business and Economics*, Springer, Cham, 103-111.

Olbryś, J. & Majewska, E. (2015). Bear market periods during the 2007-2009financial crisis: Direct evidence from the Visegrad countries. *Acta Oeconomica, 65(4)*, 547-565.

Olbryś, J. & Mursztyn, M. (2015). Comparison of selected trade classification algorithms on the Warsaw Stock Exchange. *Advances in Computer Science Research, 12*, 37-52.

Olbryś, J. & Mursztyn, M. (2017). Measurement of stock market liquidity supported by an algorithm inferring the initiator of a trade. *Operations Research and Decisions, 27(4),* 111-127.

Theissen, E. (2001). A test of the accuracy of the Lee/Ready trade classification algorithm. *Journal of International Financial Markets, Institutions and Money, 11(2)*, 147-165.

A hybrid MSV-MGARCH generalisation of the *t*-MGARCH model

Jacek Osiewalski[1], Anna Pajor[2]

Abstract

Hybrid MSV-MGARCH models, in particular the MSF-SBEKK specification, proved useful in multivariate modelling of returns on financial and commodity markets. Here we propose a natural hybrid generalisation of conditionally Student *t* MGARCH models. Our new hybrid EMSF-MGARCH specification is obtained by multiplying the MGARCH conditional covariance matrix H_t by a scalar random variable g_t, which comes from a latent process with auto-regression parameter φ that, for $\varphi = 0$, leads to an inverted gamma distribution for g_t and thus to the *t*-MGARCH case. If $\varphi \neq 0$, the latent variables g_t are dependent, so (in comparison to the *t*-MGARCH specification) in the new model of the observed time series we get an additional source of dependence and one more parameter. We apply the scalar BEKK specification as the basic MGARCH structure. Using the Bayesian approach, equipped with MCMC simulation techniques, we show how to estimate the new hybrid EMSF-SBEKK model. We present an empirical example that serves to illustrate the hybrid extension of the *t*-SBEKK model and its usefulness, as well as to compare it to the MSF-SBEKK case.

Keywords: *Bayesian econometrics, Gibbs sampling, time-varying volatility, multivariate GARCH processes, multivariate SV processes*

JEL Classification: C11, C32, C51

DOI: 10.14659/SEMF.2018.01.35

1 Introduction

In volatility modelling of financial time series, hybrid MSV-MGARCH models were introduced by Osiewalski and Pajor (2007, 2009) and Osiewalski and Osiewalski (2016) in order to use relatively simple model structures that exploit advantages of both model classes: flexibility of the MSV class (where volatility is modelled by latent stochastic processes) and relative simplicity of the MGARCH class. In their first attempt, Osiewalski and Pajor (2007) used only one latent process and the DCC covariance structure proposed by Engle (2002). However, Osiewalski (2009) and Osiewalski and Pajor (2009) suggested an even simpler model, also based on one latent process, but with the scalar BEKK covariance structure. The parsimonious hybrid MSF-SBEKK specification has been recognized in the literature (see

[1] Corresponding author: Cracow University of Economics (Faculty of Management, Dept. of Econometrics and Operations Research); ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: eeosiewa@cyf-kr.edu.pl

[2] Cracow University of Economics (Faculty of Finance and Law, Dept. of Mathematics); ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: eopajor@cyf-kr.edu.pl

Teräsvirta, 2012; Amado and Teräsvirta, 2013; Carriero et al., 2016) and proved useful in multivariate modelling of returns on financial and commodity markets (see Pajor, 2010, 2014; Pajor and Osiewalski, 2012; Osiewalski and Osiewalski, 2012, 2013; Pajor and Wróblewska, 2017). Any MSF-MGARCH model amounts to using a conditionally normal MGARCH process and multiplying its conditional covariance matrix H_t by such positive random variable g_t that $\ln(g_t)$ follows a Gaussian AR(1) process with auto-regression parameter φ. If $\varphi = 0$, then such MSF-MGARCH case reduces to the MGARCH process with the conditional distribution being a continuous mixture of multivariate normal distributions with covariance matrices $g_t H_t$ and g_t log-normally distributed.

In this paper we propose a natural hybrid extension of very popular MGARCH models with the Student *t* conditional distribution. Our new models are obtained by multiplying H_t by random variable g_t coming from such latent process (with auto-regression parameter φ) that, for $\varphi = 0$, g_t has an inverted gamma distribution and leads to the *t*-MGARCH specification, where the conditional distribution can be represented as a continuous mixture of multivariate normal distributions with covariance matrices $g_t H_t$ and an inverted gamma distribution of g_t. If $\varphi \neq 0$, the latent variables g_t are dependent, so (in comparison to the *t*-MGARCH model) in the new model of the observed time series we get an additional source of dependence and one more parameter.

In the next section the general form of a MSV-MGARCH model as well as its special MSF-MGARCH structure are presented, and the EMSF-MGARCH class is formally defined. In Section 3 it is shown how to estimate the new hybrid EMSF-SBEKK model using the Bayesian approach and MCMC simulation techniques. In Section 4 an empirical example is presented; it serves to illustrate the hybrid extension of the *t*-SBEKK model and its validity, as well as to compare it briefly to the previous MSF-SBEKK specification.

2 Hybrid *n*-variate volatility specifications

Assume there are *n* assets. We denote by $r_t = (r_{t1} \ \ldots \ r_{tn})$ *n*-variate observations on their logarithmic return (or growth) rates, and we model them using the basic VAR(1) framework:

$$r_t = \delta_0 + r_{t-1}\Delta + \varepsilon_t; \quad t=1, ..., T; \tag{1}$$

where *T* is the length of the observed time series. The hybrid MSV-MGARCH model class for the disturbance term ε_t is defined by the following equality:

$$\varepsilon_t = \zeta_t \, H_t^{1/2} G_t^{1/2}, \tag{2}$$

where: $\{\zeta_t\}$ is a strict *n*-variate white noise with unit covariance matrix, $\{\zeta_t\} \sim iiD^{(n)}(0, I_n)$;

H_t and G_t are square matrices of order n, symmetric and positive definite for each t; H_t is a non-constant function of the past of ε_t and corresponds to the conditional covariance matrix of some MGARCH specification; G_t is a non-constant function of a (scalar or vector) stochastic latent process $\{g_t\}$, which is non-trivial (i.e., constituted of variables g_t dependent over time); see Osiewalski and Osiewalski (2016). Under (1) and (2), the conditional distribution of r_t (given the past of r_t and the current latent variable g_t) is determined by the distribution of ζ_t; it has mean vector $\mu_t = \delta_0 + r_{t-1}\Delta$ and covariance matrix $\Sigma_t = G_t^{1/2\prime} H_t G_t^{1/2}$, which depends on both g_t and the past of r_t.

Building upon an idea presented by Osiewalski and Osiewalski (2016), we define a new subclass of the MSV-MGARCH model class. This subclass corresponds to the Gaussian white noise $\{\zeta_t\}$ and positive-valued scalar latent processes $\{g_t\}$ such that $G_t = g_t I_n$ and

$$\ln g_t = \varphi \ln g_{t-1} + \ln \gamma_t, \tag{3}$$

where $\zeta_t \perp \gamma_s$ for all $t, s \in \{1, ..., T\}$, $0 < |\varphi| < 1$ and $\{\gamma_t\}$ is a sequence of independent positive random variables with the same distribution belonging to a specific parametric family. The simplest MSV structure, called MSF and used by Osiewalski and Pajor (2009) to construct hybrid models with Gaussian $\{\zeta_t\}$, is based on the assumption that $\{\ln \gamma_t\}$ is a Gaussian white noise with unknown variance σ^2. In MSF-MGARCH hybrid models, (3) represents a two-parameter family of stationary and causal Gaussian AR(1) processes. Now we extend this basic case by considering other latent processes (3), corresponding to different parametric distribution classes of γ_t. We use the term "Extended MSF-MGARCH (EMSF-MGARCH) model" for all MSV-MGARCH models based on Gaussian $\{\zeta_t\}$ and distributions other than log-normal for γ_t in (3). In EMSF-MGARCH cases $\{\ln \gamma_t\}$ needs not be a white noise process.

In the MSF-MGARCH and EMSF-MGARCH cases, the conditional distribution of r_t (given its past and g_t) is Normal with mean μ_t and covariance matrix $\Sigma_t = g_t H_t$. In the MSF-MGARCH model, due to properties of Gaussian autoregressive processes, the marginal distribution of g_t is log-normal, so the distribution of r_t given its past (only) is the scale mixture of $N(\mu_t, g_t H_t)$ distributions with log-normal g_t. The mixing distribution depends on φ, and obviously remains log-normal for $\varphi=0$; this value leads to the MGARCH model with a specific ellipsoidal conditional distribution (the log-normal scale mixture of normal distributions). In the EMSF-MGARCH model class, the distribution of r_t given its past (only) is also the scale mixture of $N(\mu_t, g_t H_t)$ distributions, but we are not able to derive the marginal distribution of g_t, which obviously depends on φ. However, for $\varphi=0$ (the value excluded in the

definition of our new hybrid models) $g_t=\gamma_t$, so the distribution of g_t is known by assumption. Since $\varphi=0$ corresponds to the MGARCH model with some ellipsoidal conditional distribution, i.e. the scale mixture of N(μ_t, g_tH_t) distributions with g_t distributed as γ_t, we may view any EMSF-MGARCH specification as a natural hybrid extension of the MGARCH model that is obtained for $\varphi=0$. In this paper we consider such EMSF-MGARCH extension of the popular MGARCH model with conditional Student *t* distribution. We focus on a particular, simple form of the MGARCH covariance matrix H_t, namely the scalar BEKK (SBEKK) form. It leads to the EMSF-SBEKK model based on an inverted gamma distribution of g_t; this is our hybrid extension of the *t*-SBEKK model.

3 Bayesian EMSF-SBEKK model and MCMC simulation of its posterior distribution

Assume that ε_t in (1) is conditionally Normal (given parameters and latent variables, jointly grouped in θ) with mean vector 0 and covariance matrix g_tH_t. The SBEKK form of H_t is as follows:

$$H_t = (1-\beta-\gamma)A + \beta(\varepsilon_{t-1}'\varepsilon_{t-1}) + \gamma H_{t-1}. \tag{4}$$

The univariate latent process $\{g_t\}$ fulfils (3) with γ_t^{-1} gamma distributed with mean 1 and variance $2/\nu$, i.e. $\{\gamma_t\} \sim iiIG(\nu/2, \nu/2)$.

In order to efficiently estimate our EMSF-SBEKK model, which is based on as many latent variables as the number of observations, we use the Bayesian approach equipped with MCMC simulation techniques. The Bayesian statistical model amounts to specifying the joint distribution of all observations, latent variables and "classical" parameters. The assumptions presented so far determine the conditional distribution of observations and latent variables given the parameters. Thus, it remains to formulate the marginal distribution of parameters (the prior or *a priori* distribution). We assume independence among almost all parameters and use the same prior distributions as Osiewalski and Pajor (2009) for the same parameters. The $n(n+1)$ elements of $\delta=(\delta_0\ (\text{vec}\ \Delta)')'$ are are assumed *a priori* independent of other parameters, with the $N(0, I_{n(n+1)})$ prior. Matrix A is a free symmetric positive definite matrix of order n, with an inverted Wishart prior distribution (for A^{-1}: Wishart prior distribution with mean I_n), and β and γ are free scalar parameters, jointly uniformly distributed over the unit simplex. As regards initial conditions for H_t, we take $H_0 = h_0 I_n$ and treat $h_0 > 0$ as an additional parameter (*a priori* Exponentially distributed with mean 1). φ has the uniform distribution over (-1, 1), for ν we assume the Exponential distribution with mean $1/\lambda_\nu$ truncated to $(2, +\infty)$.

We can write the full Bayesian model as

$$\begin{aligned} &p(r_1,...,r_T,g_1,...,g_T,\delta,A,\beta,\gamma,\varphi,\nu,h_0) = \\ &= p(\delta)p(A)p(h_0)p(\beta,\gamma)p(\varphi)p(\nu)\prod_{t=1}^{T} f_N^n(r_t \mid \mu_t, g_t H_t)\times \\ &\times\prod_{t=1}^{T}\frac{\left(\frac{\nu}{2}g_{t-1}^{\varphi}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)}\left(\frac{1}{g_t}\right)^{\frac{\nu}{2}+1} e^{-\frac{\frac{\nu}{2}g_{t-1}^{\varphi}}{g_t}}. \end{aligned} \tag{5}$$

The posterior density function, proportional to (5), is highly dimensional and non-standard. Thus Bayesian analysis is performed on the basis of a MCMC sample from the posterior distribution, which is obtained using Gibbs algorithm, i.e. the sequential sampling from the conditional distributions obtained from (5):

$$p(\delta \mid r_1,...,r_T,g_1,...,g_T,A,\beta,\gamma,\varphi,\nu,h_0) \propto p(\delta)\prod_{t=1}^{T} f_N^n(r_t \mid \mu_t, g_t H_t),$$

$$p(A \mid r_1,...,r_T,g_1,...,g_T,\delta,\beta,\gamma,\varphi,\nu,h_0) \propto p(A)\prod_{t=1}^{T} f_N^n(r_t \mid \mu_t, g_t H_t),$$

$$p(\beta,\gamma,h_0 \mid r_1,...,r_T,g_1,...,g_T,\delta,A,\varphi,\nu) \propto p(\beta,\gamma)p(h_0)\prod_{t=1}^{T} f_N^n(r_t \mid \mu_t, g_t H_t),$$

$$p(\varphi \mid r_1,...,r_T,g_1,...,g_T,\delta,A,\beta,\gamma,\nu,h_0) \propto e^{\varphi\frac{\nu}{2}\sum_{t=1}^{T}\ln g_{t-1}} \times e^{-\frac{\nu}{2}\sum_{t=1}^{T}\frac{g_{t-1}^{\varphi}}{g_t}} I_{(-1,1)}(\varphi),$$

$$p(\nu \mid r_1,...,r_T,g_1,...,g_T,\delta,A,\beta,\gamma,\varphi,h_0) \propto \left(\frac{\nu}{2}\right)^{\frac{T\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)^{-T} e^{-\kappa\nu},$$

$$\text{where } \kappa = -\frac{1}{2}\sum_{t=1}^{T}\ln\frac{g_{t-1}^{\varphi}}{g_t} + \frac{1}{2}\sum_{t=1}^{T}\frac{g_{t-1}^{\varphi}}{g_t} + \lambda_{\nu}$$

$$p(g_t \mid r_1,...,r_T,g_1,...,g_{t-1},g_{t+1},...,g_T,\delta,A,\beta,\gamma,\varphi,\nu,h_0) \propto$$

$$\propto f_{IG}(g_t \mid n/2+(\nu/2)(1-\varphi),(1/2)(r_t-\mu_t)H_t^{-1}(r_t-\mu_t)'+(\nu/2)g_{t-1}^{\varphi})\, e^{-\frac{\nu}{2}\frac{g_t^{\varphi}}{g_{t+1}}};\quad t=1,...,T-1;$$

$$\begin{aligned} &p(g_T \mid r_1,...,r_T,g_1,...,g_{T-1},\delta,A,\beta,\gamma,\varphi,\nu,h_0) \\ &\propto f_{IG}(g_T \mid n/2+(\nu/2)(1-\varphi),(1/2)(r_T-\mu_T)H_T^{-1}(r_T-\mu_T)'+(\nu/2)g_{T-1}^{\varphi}). \end{aligned}$$

Drawing from each conditional distribution above is done through Metropolis-Hastings steps.

4 An empirical example

In order to illustrate empirical validity of the EMSF–SBEKK model – in comparison to the pure GARCH, *t*-SBEKK specification – we use the same bivariate data sets as Osiewalski and Pajor (2009). The first data set consists of the official daily exchange rates of the National

Bank of Poland (NBP fixing rates) for the US dollar and German mark in the period 1.02.1996 – 31.12.2001. The length of the modelled time series of their daily growth rates (logarithmic return rates) is 1482. The second data set consists of the daily quotations of the main index of the Warsaw Stock Exchange (WIG) and the S&P500 index of NYSE. We model 1727 logarithmic returns from the period 8.01.1999–1.02.2006. In the case of exchange rates, both series are highly non-Normal and they are quite strongly positively correlated. The other data set shows smaller deviations from Normality and much weaker correlation.

Table 1. Posterior means (and standard deviations) of the parameters of the MSF–SBEKK and EMSF–SBEKK models for the exchange rates (T=1482).

parameter	**MSF–SBEKK**		**EMSF-SBEKK (λ_v = 1/10)**		**EMSF-SBEKK (λ_v = 1/30)**	
δ_{01}	0.044	(0.009)	0.040	(0.010)	0.040	(0.010)
δ_{02}	-0.005	(0.010)	-0.006	(0.010)	-0.006	(0.010)
δ_{11}	-0.020	(0.025)	-0.015	(0.025)	-0.014	(0.025)
δ_{12}	-0.012	(0.026)	-0.010	(0.026)	-0.009	(0.026)
δ_{21}	-0.012	(0.021)	-0.015	(0.021)	-0.015	(0.021)
δ_{22}	-0.040	(0.025)	-0.038	(0.025)	-0.039	(0.025)
a_{11}	0.153	(0.029)	0.079	(0.010)	0.079	(0.010)
a_{12}	-0.053	(0.018)	-0.024	(0.007)	-0.024	(0.007)
a_{22}	0.174	(0.034)	0.089	(0.012)	0.089	(0.012)
φ	0.411	(0.086)	**0.302**	**(0.096)**	**0.304**	**(0.096)**
σ^2 or v	0.540	(0.070)	**5.508**	**(0.573)**	**5.527**	**(0.575)**
β	0.084	(0.013)	0.068	(0.012)	0.068	(0.012)
γ	0.878	(0.015)	0.869	(0.016)	0.869	(0.016)
$\beta + \gamma$	0.962	(0.008)	0.937	(0.026)	0.937	(0.026)
h_0	0.053	(0.051)	0.038	(0.037)	0.039	(0.036)

In Tables 1 and 2 the posterior means and standard deviations of the MSF–SBEKK and EMSF–SBEKK parameters are presented for the exchange rates and stock indices data, respectively; the results for the MSF–SBEKK case are taken from Osiewalski and Pajor (2009). It is important to note that the posterior distribution of φ, the latent process auto-regression parameter, is further from zero in the MSF–SBEKK model for both data sets (especially for stock indices). It seems that the MSF–SBEKK model really needs the non-

trivial Gaussian AR(1) latent process in order to describe the data, so that the case $\varphi = 0$, i.e. the SBEKK model with log-normal scale mixture as the conditional distribution, is clearly excluded. The question whether our EMSF–SBEKK model can be reduced to the *t*-SBEKK case cannot be answered so easily. For the exchange rates data, the posterior probability that $\varphi < 0$ is approximately 0.001 only and $\varphi = 0$ is included in the highest posterior density (HPD) interval of probability content as high as 0.998. Thus the *t*-SBEKK model is inadequate. But for the stock data, the posterior probability that $\varphi < 0$ is 0.061 for λ_v=1/10 and 0.045 for λ_v=1/30, and $\varphi = 0$ is included in the HPD interval of probability content 0.863 or 0.949, depending on the prior hyper-parameter λ_v. The *t*-SBEKK model cannot be rejected for the stock data and its empirical relevance is sensitive to the prior specification.

Table 2. Posterior means (and standard deviations) of the parameters of the MSF–SBEKK models for the stock indices (T=1727).

parameter	**MSF–SBEKK**		**EMSF-SBEKK (λ_v = 1/10)**		**EMSF-SBEKK (λ_v = 1/30)**	
δ_{01}	0.072	(0.026)	0.067	(0.026)	0.066	(0.026)
δ_{02}	0.028	(0.022)	0.026	(0.023)	0.026	(0.023)
δ_{11}	0.015	(0.024)	0.011	(0.024)	0.011	(0.024)
δ_{12}	0.012	(0.020)	0.009	(0.020)	0.010	(0.020)
δ_{21}	0.302	(0.027)	0.297	(0.026)	0.297	(0.026)
δ_{22}	-0.022	(0.026)	-0.023	(0.026)	-0.023	(0.025)
a_{11}	1.127	(0.267)	0.668	(0.141)	0.658	(0.141)
a_{12}	0.159	(0.104)	0.088	(0.070)	0.088	(0.071)
a_{22}	0.729	(0.176)	0.469	(0.097)	0.461	(0.098)
φ	0.872	(0.156)	**0.319**	**(0.220)**	**0.309**	**(0.209)**
σ^2 or v	0.036	(0.041)	**14.607**	**(5.054)**	**14.110**	**(3.798)**
β	0.021	(0.006)	0.032	(0.006)	0.032	(0.006)
γ	0.970	(0.007)	0.953	(0.008)	0.953	(0.008)
$\beta + \gamma$	0.991	(0.003)	0.985	(0.005)	0.985	(0.005)
h_0	2.881	(1.026)	2.277	(0.812)	2.252	(0.807)

The dependence of the marginal posterior distribution of φ on λ_v (the hyper-parameter of the prior of v), observed for the stock data, is intriguing. The parameters φ, v are independent

a priori and only weakly dependent *a posteriori* in the case of exchange rates; their bivariate posterior distribution is almost the same for both values of λ_v (see Fig. 1). However, the joint posterior distribution of (φ, v) for the stock data, shown in Fig. 2, looks very different. It is bimodal, reveals a non-linear dependence between parameters and looks a bit different for λ_v=1/10 and for λ_v=1/30, with non-negligible second mode in the latter case. Note that the global mode corresponds to much lower values of the degrees of freedom and auto-regression parameter than the second mode. The larger φ (the stronger dependence in the latent process), the higher v (the thinner tail of the conditional distribution of the latent variable). Thus it is intuitive that λ_v=1/30, which gives high prior chances to thin tails (higher than λ_v=1/10 gives), may lead to the so high second mode. These subtle issues would require comparing different Bayesian models formally, which is not easy at all.

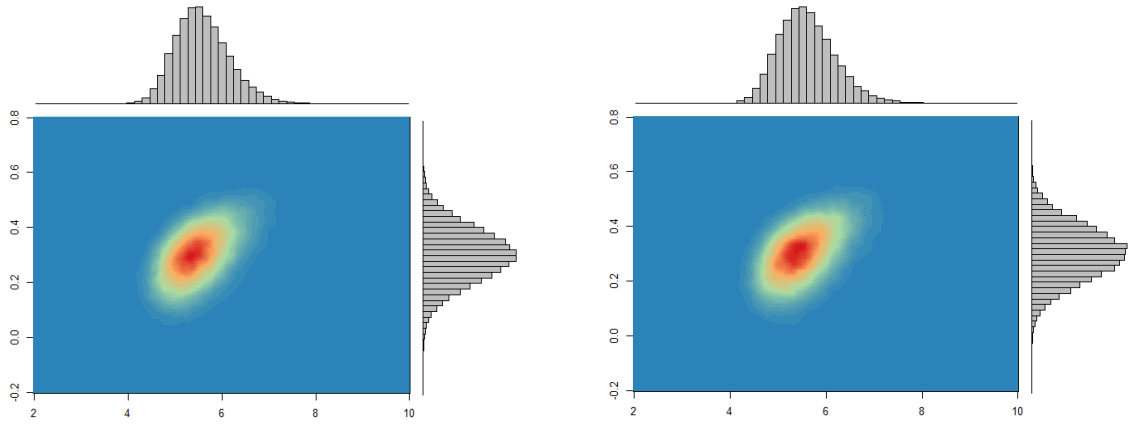


Fig. 1. The posterior distribution of (φ, v) in the EMSF-SBEKK model (exchange rates, T=1482; left panel: λ_v = 1/10; right panel: λ_v = 1/30).

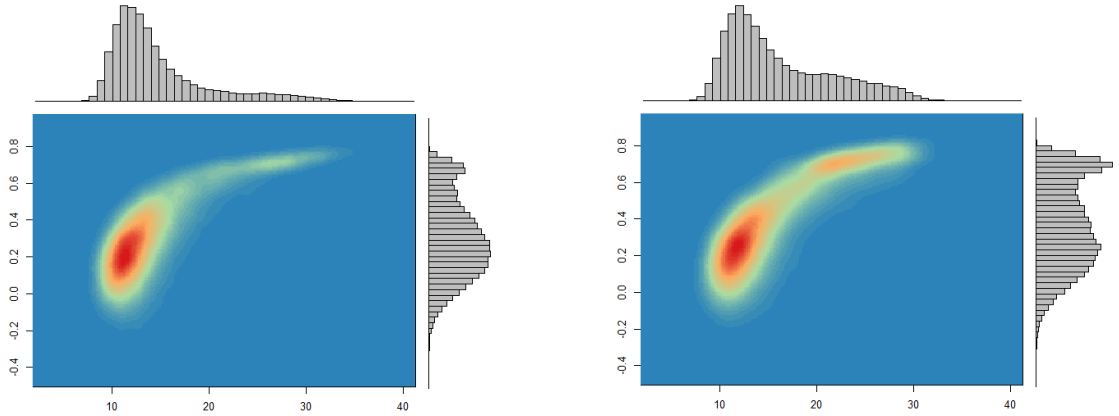


Fig. 2. The posterior distribution of (φ, v) in the EMSF-SBEKK model (stock indices, T=1727; left panel: λ_v = 1/10; right panel: λ_v = 1/30).

Note that the formal Bayesian model comparison (through Bayes factors and posterior odds) is computationally very difficult in our hybrid framework. The crucial issue is that of precisely calculating the numerical value (for the data at hand) of the marginal density of observations $p(r_1,..., r_T)$, which is the integral of the density (5) with respect to its all other arguments (i.e., latent variables and parameters). In order to approximate $p(r_1,..., r_T)$ within MCMC sampling from the posterior distribution, Osiewalski and Osiewalski (2013, 2016) used the harmonic mean estimator with a specific correction. Such approach does not have so good properties as the corrected arithmetic mean estimator (CAME) proposed by Pajor (2017). However, the use of CAME in dynamic models with latent processes is not numerically feasible yet, due to very high dimensions of the Monte Carlo simulation spaces. Thus, in this study we do not calculate the posterior model probabilities for the proposed EMSF-SBEKK model, its *t*-SBEKK limit case and the original MSF-SBEKK specification. The formal Bayesian comparison of alternative models is left for future research.

The aim of this paper was to show how to construct (and estimate within the Bayesian approach) a hybrid EMSF-MGARCH generalisation of the *t*-MGARCH model, focusing on the simple SBEKK structure. The Bayesian analysis of our EMSF-SBEKK model relies on Gibbs sampling with Metropolis-Hastings steps. Our illustrative empirical example suggests that the EMSF-MGARCH specification (that serves to generalise the *t*-MGARCH model) can relatively easily accommodate heavy tails – through latent process based on inverted gamma disturbances – in comparison to the MSF-MGARCH model, based on log-normal distribution and requiring larger values of the latent process auto-regression parameter φ.

Acknowledgement

We acknowledge support from research funds granted to the Faculty of Management (the first author) and the faculty of Finance and Law (the second author) at Cracow University of Economics, within the framework of the subsidy for the maintenance of research potential.

References

Amado, C. & Teräsvirta, T. (2013). Modelling volatility by variance decomposition. *Journal of Econometrics*, *175*, 142-153.

Carriero, A., Clark, T. & Marcellino, M. (2016). Common drifting volatility in large Bayesian vars. *Journal of Business and Economic Statistics*, *34*, 375-390.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized

autoregressive conditional heteroskedasticity models. *J Bus Econ Stat.*, *20*, 339-350.

Osiewalski, J. (2009). New hybrid models of multivariate volatility (a Bayesian perspective). *Przegląd Statystyczny (Statistical Review)*, *56*, 15-22.

Osiewalski, J. & Osiewalski, K. (2016). Hybrid MSV-MGARCH models – general remarks and the GMSF-SBEKK specification. *Central European Journal of Economic Modelling and Econometrics*, *8*, 241-271.

Osiewalski, J. & Pajor, A. (2007). Flexibility and parsimony in multivariate financial modelling: a hybrid bivariate DCC–SV model. In: Milo, W. & Wdowiński, P. (eds.), *Financial Markets. Principles of Modeling, Forecasting and Decision-Making* (FindEcon Monograph Series No.3), Łódź University Press, Łódź, 11-26.

Osiewalski, J. & Pajor, A. (2009). Bayesian analysis for hybrid MSF–SBEKK models of multivariate volatility. *Central European Journal of Economic Modelling and Econometrics*, *1*, 179-202.

Osiewalski, K. & Osiewalski, J. (2012). Missing observations in daily returns – Bayesian inference within the MSF-SBEKK model. *Central European Journal of Economic Modelling and Econometrics*, *4*, 169-197.

Osiewalski, K. & Osiewalski, J. (2013). A long-run relationship between daily prices on two markets: the Bayesian VAR(2)–MSF-SBEKK model. *Central European Journal of Economic Modelling and Econometrics*, *5*, 65-83.

Pajor, A. (2010). *Wielowymiarowe procesy wariancji stochastycznej w ekonometrii finansowej. Ujęcie bayesowskie (Multivariate Stochastic Variance Processes in Financial Econometrics. Bayesian approach, in Polish)*, Cracow University of Economics, Kraków.

Pajor, A. (2014). Konstrukcja optymalnego portfela w bayesowskim modelu MSF-SBEKK. Portfele funduszy inwestycyjnych PKO TFI. *Bank i Kredyt* (*Bank & Credit*), *43*, 53-78.

Pajor, A. (2017). Estimating the marginal likelihood using the arithmetic mean identity. *Bayesian Analysis*, *12*, 261-287 (https://projecteuclid.org/euclid.ba/1459772735).

Pajor, A. & Osiewalski, J. (2012). Bayesian Value-at-Risk and Expected Shortfall for a large portfolio (multi- and univariate approaches). *Acta Physica Polonica A*, *121*, no. 2-B, B-101–B-109.

Pajor, A. & Wróblewska, J. (2017). VEC-MSF models in Bayesian analysis of short- and long-run relationships. *Studies in Nonlinear Dynamics and Econometrics*, *21*(3), 1-22.

Teräsvirta, T. (2012). Nonlinear models for autoregressive conditional heteroskedasticity. In: Bauwens, L., Hafner, C. & Laurent, S. (eds.), *Handbook of Volatility Models and their Applications*, Wiley, New York, 49–69.

Is development of the renewable energy sector crucial for the electricity consumption-growth nexus in the EU countries?

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

Abstract

The aim of the study is to assess the impact of development of the renewable energy sector in the EU on the relationship between renewable and non-renewable electricity consumption and economic growth in the period 1995-2015. In order to identify countries with reverence of development of their renewable energy sector statistical clustering methods are applied. The relations between renewable and non-renewable electricity consumption and economic growth are investigated within the panel VAR framework.

The results reveal that the level of development of the renewable energy sector is crucial for linkages between electricity consumption and economic growth. In a group of countries with a relatively high level of development of the renewable energy sector, renewable electricity consumption and economic growth are mutually dependent. Additionally, the increase in economic growth leads to a short-term increase in renewable electricity consumption. At the same time, the rise in renewable electricity consumption is helpful in increasing economic activity. In the remaining countries renewable electricity consumption and economic growth are independent.

Keywords: *renewable electricity consumption, economic growth, EU countries, panel VAR*
JEL Classification:C3, Q4
DOI: 10.14659/SEMF.2018.01.36

1 Introduction

Most studies investigating single countries confirm the impact of electricity consumption on GDP (Gurgul and Lach, 2012; Marques et al., 2018), or GDP on electricity consumption (Baranzini et al., 2013). Recently observed flourishing of renewable energy attracts the attention of researchers who analyse the renewable energy-growth nexus. Rafindadi and Ozturk (2017) study the relation in Germany, which is a leading renewable energy user in Europe. They find that renewable energy consumption in Germany consolidates the country's economic growth prospects and boosts German economic growth. However, the results obtained for other (European) countries or groups of countries are not so unambiguous (an

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

extensive review of literature devoted to relationships between energy, environment and economic growth can be found in Tiba and Omri's, 2017). In Menegaki's (2011) analysis of European countries no relation is found between economic growth and renewable energy consumption in Europe. As she explains, it could result from the uneven and still insufficient exploitation of renewable energy sources across Europe. Similarly, the analysis of linkages between electricity consumption and GDP for many countries does not yield consistent results (Wolde-Rufael, 2014), although the growth hypothesis is confirmed for selected countries (Costa-Campi et al., 2018). However, to the best of our knowledge, none of the papers has so far addressed the direction of causality in the electricity consumption-growth nexus in relation to the level of development of the renewable energy sector in the EU member countries.

The objective of the paper is to assess the impact of the level of development of the renewable energy sector on the relationship between renewable and non-renewable electricity consumption and economic growth in the EU countries.Since the trend of the electrification of economy (Marques et al., 2018) is so prominent, we decide to base our study on electricity consumption instead of energy consumption, which is a prevailing tendency in literature. It is in line with Otzurk's (2010) survey, in which he states that electricity consumption "is a limiting factor to economic growth". Moreover, we decide to divide the overall electricity consumption into renewable and non-renewable electricity consumption, as the first component corresponds to 'new' renewables. This way we attempt to assess the relations between 'new' renewables and economic growth. Being aware of the limitations listed in Menegaki (2011), we decide to divide the EU countries into two groups according to the level of development of their renewable energy sectors. In order to distinguish the groups, we use aclustering method and the data provided by EurObserv'ER's(2015) barometers, which include employment, investment, and turnover in the renewable energy sector in the EU countries. Obviously, the larger the share of renewable energy sector in economy is, the more significant its interrelations with economy should be. In general, the study is dedicated to verificationof four hypotheses (the growth hypothesis, the conservation hypothesis, the feedback hypothesis, and the neutrality hypothesis) for the energy consumption-growth nexus, which have important implications for energy policy.

The analysis of the relationship between renewableand non-renewable electricity consumption and economic growth is based on the annual panel data from the period 1995 – 2015. However, the actual level of development of the renewable energy sector in the EU countriesis measuredbetween 2010 and 2015. The data are obtained from the EurObserv'ER's barometers (2011-2016).The analysis consists of two main stages.The aim of the first stage is

to identify countries with differentlevels of development of their renewable energy sectors. Clustering is conducted by k-means methods, assuming two groups of countries. In the second stage, dynamic panel modelsare applied to investigate causality between renewable and non-renewableelectricity consumption and economic growth.The analysis is conducted within the groups of countries obtained in the first step. To uncover the links between renewable and non-renewable electricity consumption and economic growth,the panel VAR model is implemented.

The paper contributes to the existing literature in three main aspects. First, our analysis is (to the best of our knowledge)the only attempt devoted to studying the relationship between renewable and non-renewable electricity consumption and economic growth in the EU countries. Second, as noticed by Menegaki(2011),the EU countries are highly diversified, thus the empirical strategy we employ takes into account a potential impact of development of the renewable energy sector on the relationships we analyse. Any relations between renewable electricity consumption and economic growth observed for the group of countries with a relatively high level ofdevelopment of the renewable energy sector could become important premises for other countries. Third, we apply a universal econometric framework, i.e. the panel VAR model, which allows not only for observing Granger causality, frequently used in literature, but also for studying the impulse response function.

2 Methodology

Clustering is conducted by k-means, which is one of the most frequently used methods. The k-means procedure consists of several steps in which objects are relocated into classes (whose number is given). The procedure is interrupted and the groups are established if there are no further relocations of any object. As the procedure is sensitive to the initial choice of seed clusters, it is repeated 500 times with different choices of seeds in order to obtain optimal results.

In order to investigate the relations panel VAR framework is applied. A reduced form of panel *VAR* of order *p* with a panel specific fixed-effect model can be represented by the system of linear equations:

$$Y_{it} = A_1 Y_{it-1} + A_2 Y_{it-2} + ... + A_p Y_{it-p} + u_i + e_{it}, \text{ for } i \in \{1,2,..N\}, t \in \{1,2,..,T\} \quad (1)$$

where: Y_{it} is a $(1xk)$ vector of variables $(\Delta \ln GDP_{it}, \Delta \ln REL_{ti}, \Delta \ln NREL_{it}, \Delta \ln K_{it}, \Delta \ln L_{it})$, u_i is fixed effects and e_{it} is idiosyncratic errors, while $A_1, A_2, ..., A_p$ are (kxk) matrixes of model parameters to be estimated. It is assumed that innovations have the following

characteristics: $E[e_{it}]=0, E[e'_{it}\,e_{it}]=\Sigma$ and $E[e'_{it}\,e_{is}]=0,\ t>s.$ The last assumption indicates a model without dynamic interdependencies. Finally, it is assumed that VAR is stable, which means that all characteristic roots lie outside the unit circle (Lütkepohl, 2005). The advantage of the panel VAR model with fixed effects is that all unobservable time invariant factors at a country level can be captured. It is important, as factors which describe a renewable potential of a particular country (wind intensity, solar radiation), and characteristics critical to economic growth (country size, population), are used in the analysis. The fixed effect indicators impose a challenge when estimation is considered, since they are correlated with lagged regressors. As a consequence, the mean-differencing procedure commonly applied to eliminate fixed effects would create a bias. To avoid the problem, the Helmert procedure (Arellano and Bover, 1995) is applied, which preserves the orthogonality between transformed variables and lagged regressors, so lagged regressors are used as instruments and the coefficients are estimated by the system GMM.

Two applications of panel VAR are used in the study. The first one is Granger causality, based on the Wald test, applied to relations between electricity (renewable and non-renewable) consumption and economic growth. The second one is the ortogonalized impulse-response function, which shows the reaction of one variable to the innovations in another variable in the system. Since the actual variance–covariance matrix of the errors is diagonal, it is necessary to decompose the matrix and make it orthogonal. In the analysis we adopt the Choleski decomposition, which is equivalent to transforming the system in 'recursive'VAR for identification purposes. As a result, it is assumedthat variables which come earlier in the ordering affect the remaining variables with lags and contemporaneously, while the variables that come later affect only the variables with lags.

3 Data and empirical results

The assessment of the impact of development of the renewable energy sector on the relationship between renewable and non-renewable electricity consumption and economic growth is based on the annual panel data. The analysis covers the period 1995-2015 and takes into account 26 EU countries. The actual level of development of the renewable energy sector in the EU countries is measured between 2010 and 2015[4] using three variables: the employment in renewable energysectorsharein total employment (EMP), the turnover in the renewable energy sector sharein GDP (TURN), and the investments in the renewable

[4]No data from the earlier period are available.

energysector sharein GDP (INVEST) in each EU country. These data are obtained from the EurObserv'ER's barometers (2011-2016). In order to study the relationships between renewable and non-renewable electricity consumption and economic growth we use:GDP - real gross domestic product per capita in constant 2010 US dollars, REL and NREL -gross electricity generation from renewable (or non-renewable) sources in TWh per capita, respectively, K - real gross capital formation per capita in constant 2010 US dollars and L - the labour force participation rate (% of total population aged 15+). All variables are in natural logarithms.

The groups are identified by comparing three variables: EMP, TURN, and INVEST. All variables are initially standardized. The *k*-means method is employed to find two clusters in the EU countries related to the share of the renewable energy sector.Thefirst group (GROUP 1) include following countries: Austria, Bulgaria, Denmark, Germany, Estonia, Finland, Latvia, Romania, Swedenand the second group (GROUP 2) include: Belgium, Croatia, the Czech Republic, France, Greece, Hungary, Ireland, Italy, Lithuania, Luxembourg, the Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, the United Kingdom (see also Papież et al., 2017).The first group (GROUP 1) includes countries with the relatively well-developed renewable energy sectors. In comparison to other countries, they have both larger turnover and larger employment in this sector. These countries not only have higher average values of EMP (0.92% in comparison to 0.38%), but are also much more diversified with respect to EMP (standard deviation equals 0.43% for the first group and 0.13% for the second group). The situation is similar when TURN is considered. The first group has decidedly higher values of TURN, but their diversity is slightly lower than in case of employment. The highest value of TURN is observed in Denmark (4.01%), while in Germany it is three times lower (1.25%). Smaller differences between the groups are observed in INVEST. In both groups there are countries with no INVEST, e.g. Latvia (GROUP 1) and Croatia (GROUP 2). High values of investment in RES are noted in Bulgaria (1.00%). The first group is rather heterogeneous with reference to investment, with clear outlier values for Bulgaria and Romania (in which high INVEST accompanies very low EMP) and Denmark (with the highest EMP and TURN and relatively low INVEST).

The analysis of the relationships between renewable and non-renewable electricity consumption and economic growth is conducted for both groups (GROUP 1 and GROUP 2).Before the final panel VAR models are estimated, the dynamic properties of the data are analysed. First, the data are examined for cross-sectional dependence using Pesaran's (2004) cross-sectional dependence (CD) test. As the result confirms cross-sectional dependence, the

tests applied next incorporate this assumption. The second generation panel unit root tests (Pesaran, 2007) are applied to uncover the stochastic properties of the series used.The results of Westerlund's (2007) panel cointegration testsobtained for all four statistics for all specifications(GROUP 1, GROUP 2) reveal that the null cannot be rejected, which is clear evidence of the lack of the cointegration relationship between renewable and non-renewable electricity consumption and economic growth.Similar conclusions regarding economic growth, and renewableand non-renewable energy consumption for 28developedcountries are drawn by Aflonso et al. (2017).

Since the data are nonstationary, and there is no cointegration between them, the final step of the analysis uses the panel VAR estimated for first differences of the series. The modelis used to infer Granger causality relationships between renewable and non-renewable electricity consumption and economic growth.The estimation of the panel vector autoregression (VAR) model, i.e. Eq. (1) usesthe GMM estimator. Based on the selection criteria, we fit a second-order panel VAR model with 2 lags.Granger causality results obtained through the Wald testare presented in Table1, which includes findings obtained for the first group (GROUP 1), and the second group (GROUP 2), respectively.

Table 1. Panel Granger causality test results.

Dependent variable	**Source of causation (independent variables)– Chi2-statistics**		
	ΔlnGDP	**ΔlnREL**	**ΔlnNREL**
GROUP 1			
ΔlnGDP	-	**7.053****	0.639
ΔlnREL	**7.418****	-	9.910***
ΔlnNREL	1.716	4.116	-
GROUP 2			
ΔlnGDP	-	0.877	0.501
ΔlnREL	4.666	-	0.683
ΔlnNREL	2.626	1.431	-

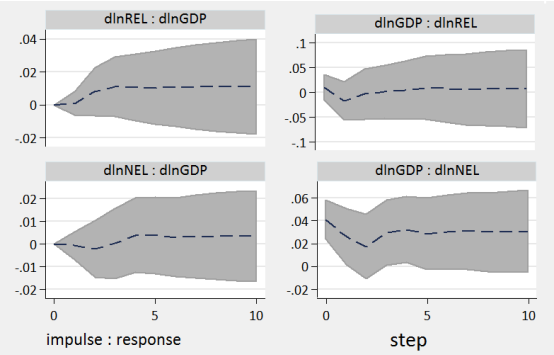
Notes: The null hypothesis: no causal relationship between variables. ***, ** indicate statistical significance at 1 and 5 per cent level, respectively.

The results reported in Table 1 confirm the neutrality hypothesis between both renewable or non-renewable electricity consumption and economic growth in the countries with a relatively low level of development of the renewable energy sector (GROUP 2). In these

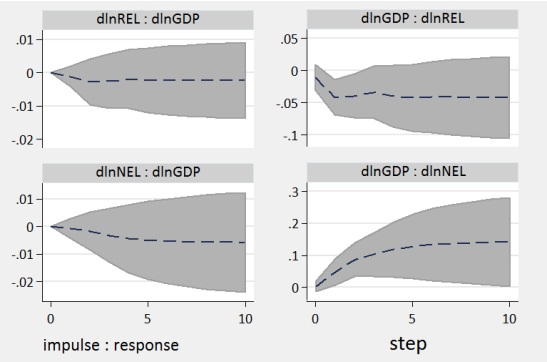
countries changes in electricity consumptiondo not affect economic growth adversely, and vice versa:changesin economic activity do not lead to changes in electricity consumption. In contrast, the results obtained for the countries with a relatively high level of development of the renewable energy sector (GROUP 1) confirm the neutrality hypothesis only between non-renewable electricity consumption and economic growth. The results also indicate thebidirectional causality relationship between renewable electricity consumption and economic growth. These findings support the feedback hypothesis in these countries, which states that renewable electricity consumption and economic growth are mutually dependent.Thus, the results demonstrate that development of the renewable energy sector, i.e. employment growth and the increase in turnover and investment in this sector, has an impact on economic growth. Economy in these countriesis called 'renewable energy dependent', which means that growth of renewable electricity consumption is an important factor in their economic development.

Next, in order to identify the reaction of electricity consumption toeconomic growth and vice versa, the impulseresponse functions (IRF) are found using the Cholesky orthogonalization procedure. The impulseresponse functions are obtained within a band representing a 95% confidence interval estimated by Monte Carlo simulations (1000 iterations). Fig. 1 illustrate the impulseresponse functions of economic growth and renewable and non-renewableelectricity consumptionto shocks (one standard deviation) to these variables. The first row of Fig. 1 shows the response of renewable electricity consumption to one standard deviation shock to economic growth and vice versa.The reaction of renewable electricity consumption to a shock to economic growth is positive but insignificant for the countries with a relatively high level of development of the renewable energy sector (GROUP 1).The increase in economic growth seems to have a positive impact on growth of renewable electricity consumption. In contrast, these reactions for the countries with a relatively low level of development of the renewable energy sector (GROUP 2) is negative but still insignificant. The reaction of economic growth to a shock to renewable electricity consumption is positive but insignificant for the countries with a relatively high level of development of the renewable energy sector (GROUP 1) as well as for all countries. This means that the advancement of the renewable electricity sector and the rise in renewable electricity consumption leads to the increase in their economic activity. But for the countries belonging to the second group the effect of a shock to renewable electricity consumption to economic growth exerts a persistentand negative impact on future economic growth.The second row of Fig. 1 demonstrates the reaction of non-renewable electricity consumption to a

shock to economic growth is positive and statistically significant, for all cases (i.e. for GROUP 1, GROUP 2). This indicates that the increase in economic growth positively affectsnon-renewable electricity consumption for all EU countries, although for countries with a relatively high level of development of the renewable energy sector (GROUP 1) this positive reaction is statistically significant only in the first year, and loses significance later. Also the Fig. 1 illustrates the negative and insignificant reaction of economic growth to a shock to non-renewable electricity consumption for countries with a relatively low level of development of the renewable energy sector (GROUP 2).



a) GROUP 1



b) GROUP 2

Fig. 1. Impulse response functions for the countries belonging to GROUP 1 or GROUP 2.

Conclusions

The two stages of the empirical strategy employed in the study prove to be effective, as the results obtained for particular groups are considerably different, although reasonable. The relations between renewable and non-renewable electricity consumption and economic growth in countries with a high level of renewable energy development are different than in the remaining countries. The greater the share of the renewable energy sector in economy (although limited), the more noticeable the relations between renewable electricity consumption and economic growth. On the other hand, the renewable energy sector in the remaining countries is too small to play a significant role in economy. In particular, our findings can be summarised in two main points.First, the feedback hypothesis between renewable electricity consumption and economic growth is confirmedfor the group of countries with a relatively high level of development of the renewable energy sector. Additionally, the results of the impulse response function suggestthat the rice in economic growth leads to a short-term increase in renewable electricity consumption and the rise in renewable electricity consumption boosts economy.

Second, in the group of countries with a low level of development of the renewable energy sector, the neutrality hypothesis between renewable or non-renewable electricity consumption and economic growth is confirmed.In general, the results obtained in the study point at an important role played by development of the renewable energy sector in establishing the relations between electricity consumption and economic growth, which can be directly translated into several relevant policy implications. The analysis suggests that, in order to find significant relations between renewable electricity consumption and economic growth, the level of development of renewable energy sector has to be large enough, which, is still not the casefor most EU countries. Another important conclusion stemming from the impulse response analysis (even though the responses described usuallyare not statistically significant)states that, in order to expect a positive impact of the renewable sector to economy, sufficient investments into renewable energy have to be made. If the renewable energy sector is small, it is unlikely to exert a largeimpact on economy. But if the renewable energy sector is larger (at least as large as in the countries from the first group), it can be reasonably expected to stimulate economy. This non-linearity lends support to the idea of enlarging the renewable energy sector.

Acknowledgments

This paper was supported by funds from the National Science Centre (NCN), Poland through grant No 2015/17/B/HS4/02685.

References

Afonso, T. L., Marques, A. C. & Fuinhas, J. A. (2017). Strategies to make renewable energy sources compatible with economic growth. *Energy Strategy Reviews*, *18*, 121-126.

Arellano, M. & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29-51.

Baranzini, A., Weber, S., Bareit, M. & Mathys, N. A. (2013). The causal relationship between energy use and economic growthin Switzerland. *Energy Economics*, *36*, 464-470.

Costa-Campi, M. T., García-Quevedo, J. & Trujillo-Baute, E. (2018). Electricity regulation and economic growth. *Energy Policy*, *113*, 232-238.

EurObserv'ER, 2011:2016. The state of renewable energies in Europe. 16th EurObserv'ER Report, 15, 1-137. Available from: http://www.eurobserv-er.org/pdf/annual-overview-2016-en/ [accessed 01.08.17].

Gurgul, H. & Lach, Ł. (2012). The electricity consumption versus economic growth of the Polish economy. *Energy Economics*, *34*(2), 500-510.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Marques, A. C., Fuinhas, J. A. & Neves, S. A. (2018). Ordinary and Special Regimes of electricity generation in Spain: How they interact with economic activity. *Renewable and Sustainable Energy Reviews*, *81*, 1226-1240.

Menegaki, A. N. (2011). Growth and renewable energy in Europe: a random effect model with evidence for neutrality hypothesis. *Energy Economics*, *33*(2), 257-263.

Ozturk, I. (2010). A literature survey on energy–growth nexus. *Energy policy*, *38*(1), 340-349.

Papież, M., Śmiech, S. & Frodyma, K. (2017). The impact of development of the renewable energy sector in the EU on the energy – growth nexus. In: Papież, M. & Śmiech S. (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, 270-279.

Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels.http://papers.ssrn.com/sol3/papers.cfm?abstract_id=572504, access: 3.02.2018.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*(2), 265-312.

Rafindadi, A. A. & Ozturk, I. (2017). Impacts of renewable energy consumption on the German economic growth: Evidence from combined cointegration test. *Renewable and Sustainable Energy Reviews*, *75*, 1130-1141.

Tiba, S. & Omri, A. (2017). Literature survey on the relationships between energy, environment and economic growth. *Renewable and Sustainable Energy Reviews*, *69*, 1129-1146.

Westerlund, J. (2007). Testing for error correction in panel data. *Oxford Bulletin of Economics and Statistics*, *69*(6), 709-748.

Wolde-Rufael, Y. (2014). Electricity consumption and economic growth in transition countries: A revisit using bootstrap panel Granger causality analysis. *Energy Economics*, *44*, 325-330.

Evaluation of investment support from Common Agricultural Policy with propensity score matching method

Aleksandra Pawłowska[1]

Abstract

The study refers to the microeconomic producer theory as a framework and the expected positive relationship between investment support and labour productivity. The aim of the paper was to evaluate the impact of subsidies from Common Agricultural Policy on an increase in labour productivity on Polish crop farms. The applied research tool was propensity score matching method, based on so-called counterfactual results, i.e. potential results possible to be achieved, if the status of treating the given object was different than observed. The study used data from the Farm Accountancy Data Network for individual Polish farms for 2008-2015.

The results show that the positive effect of investment subsidies occurred only in 2011. Back then, the farms which in 2010 received the analysed support, were characterised, on average, by 39 percentage points higher annual increase in the labour productivity compared to the control group. In the remaining years, in turn, the impact of investment support on the increase in labour productivity was negative. However, given relatively high standard errors, the differences between farms which received and did not receive analysed payments were not statistically significant.

Keywords: *farms, labour productivity, agriculture policy, propensity score matching*
JEL Classification: D24, Q12, Q18
DOI: 10.14659/SEMF.2018.01.37

1 Introduction

Improving the competitiveness of agriculture is one of the main challenges of the Common Agricultural Policy (CAP). This objective should be achieved, by, *inter alia*, an increase in the labour productivity, as the effect of investments carried out in farms. The initiatives aimed at modernisation of farm could be therefore supported by specific policy instruments. In case of Poland, measures supporting investments were, for example, "Modernisation of agriculture holdings" and "Setting up of young farmers" under the Rural Development Programme (RDP) for 2007-2013.

The positive relationship between subsidies on investments and the increase in labour productivity can be demonstrated based on the microeconomic producer theory. The foundations of investment are savings collected by the (agricultural) producer. If the investment needs are greater than the possibilities determined by savings, support under the

[1] Corresponding author: Institute of Agricultural and Food Economics – National Research Institute, Mathematics Application in Agricultural Economics Department, Świętokrzyska 20, 00-002 Warszawa, Poland, e-mail: Aleksandra.Pawlowska@ierigz.waw.pl.

CAP could be a catalyst for investment in fixed factors of production, which, in turn, leads to the increase of capital-to-labour ratio. Assuming that at the level of a single farm the labour input is constant, it allows obtaining higher production and consequently leads to increased labour productivity (Rembisz et al., 2014). In accordance with Fig. 1, farms which received investment support in 2007-2015 achieved, on average, higher labour productivity compared to farms which did not receive analysed support, which is to a certain extent confirmed by the mentioned positive relationship between investment (subsidies on investment) and labour productivity.

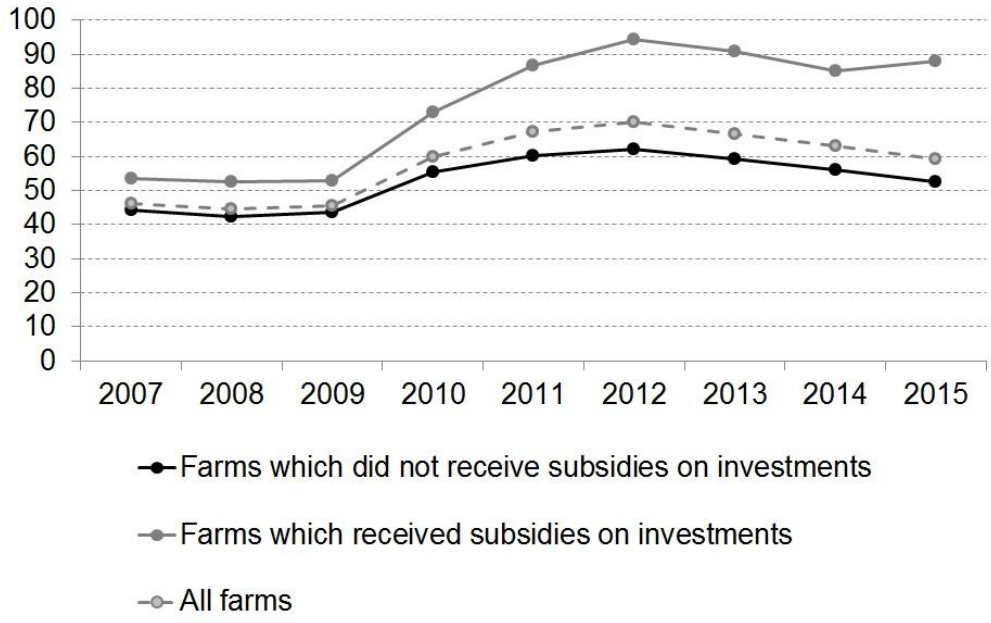


Fig. 1. Average labour productivity vs subsidies on investments in Polish farms in 2007-2015. Source: (Pawłowska and Bocian, 2017).

However, so far, the relationship between subsidies on investments and labour productivity was extensively examined primarily in the context of evaluation of the implemented agriculture policy instruments. The results presented in literature do not give a clear answer as to the positive or negative impact of this specific payments on the increase in the efficiency of the production factors on farms (Nilsson, 2017). The conclusions reached by Zhu and Lansink (2010) and Latruffe (2010) suggest that subsidies have a negative impact on efficiency or productivity. These studies, however, considered only the total amount of subsidies, so they cannot provide evidence on the impact of specific CAP measures. Similar conclusions about the negative impact of investment support on TFP (total factor productivity) were demonstrated by Mary (2013), but in contrast with previous studies it was

shown that in case of French crop farms the effect of investment subsidies was not statistically significant.

The aim of the paper is to evaluate the impact of subsidies on investment from CAP on the increase in labour productivity on Polish crop farms using propensity score matching. The study contributes to the literature on the productivity effects of investment support under RDP for 2007-2013 in Polish crop farms.

2 Propensity score matching

Propensity score analysis was used to assess the net effect of investment support on an increase in labour productivity. This class of statistical methods, and generally the causal analysis introduced by Rosenbaum and Rubin (1983), is based on counterfactual framework.

Let $D = 1$ denote the receipt of treatment (treatment group), $D = 0$ denote non-receipt (control group) and Y_i indicate the measured outcome variable. Each *i*-th subject under evaluation would have two potential outcomes, i.e. Y_{0i} and Y_{1i}, corresponding respectively to the potential outcome in the untreated and treated states. The observed outcome variable is defined as (Guo and Fraser, 2015):

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i} \tag{1}$$

The fundamental problem of causal inference is that we observe only Y_{1i} for treatment group and Y_{0i} for control group (Holland, 1986). The potential outcome that is not observed is the so-called counterfactual result.

The main aim of propensity score analysis is to estimate the counterfactual by evaluating the difference in mean outcomes between treatment and control group. For each *i*-th unit the propensity score $b_{PS}(\mathbf{x}_i)$ can be estimated from logistic regression of the treatment condition D_i on the covariate vector $\mathbf{x}_i$ (Pan and Bai, 2015):

$$ln\left(\frac{b_{PS}(\mathbf{x}_i)}{1 - b_{PS}(\mathbf{x}_i)}\right) = \boldsymbol{\beta}\mathbf{x}_i \tag{2}$$

Rosenbaum and Rubin (1983), defining propensity score as a balancing score, introduced two assumptions about the strong ignorability in treatment assignment (see Pan and Bai, 2015; Leite, 2017):

$$(Y_{0i}, Y_{1i}) \perp D_i | \mathbf{x}_i \tag{3}$$

$$0 < P(D_i = 1|\mathbf{x}_i) < 1 \quad (4)$$

In accordance with the first assumption (formula (3)), treatment assignment D_i and outcomes Y_{0i}, Y_{1i} are conditionally independent, given $\mathbf{x}_i$. In the second one (formula (4)) it is assumed a common support between treatment and control group.

After estimation of the propensity score and evaluation of matching quality, the outcome variable could be analysed. Assuming the absence of the self-selection phenomenon, at the population level one of the treatment effects to be calculated is average treatment effect for the treated (ATT), in accordance with the formula (see Sekhon, 2011; Strawiński, 2014):

$$ATT = E(Y_{1i}|D_i = 1) - E(Y_{0i}|D_i = 1) \quad (5)$$

3 Data

The study used data from the Farm Accountancy Data Network (FADN) for individual Polish farms for 2008-2015. To ensure the homogeneity of analysed farms, only the farms specializing in field crops was considered.

Given that the aim of the study was to estimate the effect of investment support form CAP, $D = 1$ and $D = 0$ denoted, respectively, farms which received and did not receive the subsidies on investments. The potential outcome Y_i was the annual increase in labour productivity, defined as gross value added per annual work unit.

The logit models were used to estimate the impact of all possible combinations of the selected 14 variables on the dichotomous variable that express the fact of receiving (or not) investment support. The observed characteristics of farms, included in the propensity score model, were: economic class size, education of farmer, age of farmer, total utilised agricultural area, total agricultural area out of production, total livestock units, farm use, total external factors, total assets, total liabilities, change in net worth, average farm capital, gross investment on fixed assets and cash flow (see Floriańczyk et al., 2017).

Following the suggestion by Heckman, Ichimura and Todd (1997), the set of such variables was selected, for which the classification accuracy was the highest. However, the main objective of propensity score analysis is obtaining balanced characteristics to ensure the similar distribution of observed covariates for treated and untreated subjects. Hence, if it was not possible for the model with the highest prediction accuracy, for further analysis the author selected the logit model with lower accuracy rate, but ensuring balanced covariates.

In the propensity score matching, the matching ratio 1:1 with replacement was used. For propensity score matching the genetic search algorithm was used, allowing to find automatically the optimal covariate balance (Sekhon, 2011).

Examining the treatment effect of investment support on the increase in labour productivity in Polish crop farms, it was assumed, that the observed covariates from the year *t* affected receiving analysed subsidies in the year *t+1*, the result of which was an increase in the labour productivity in year *t+2*.

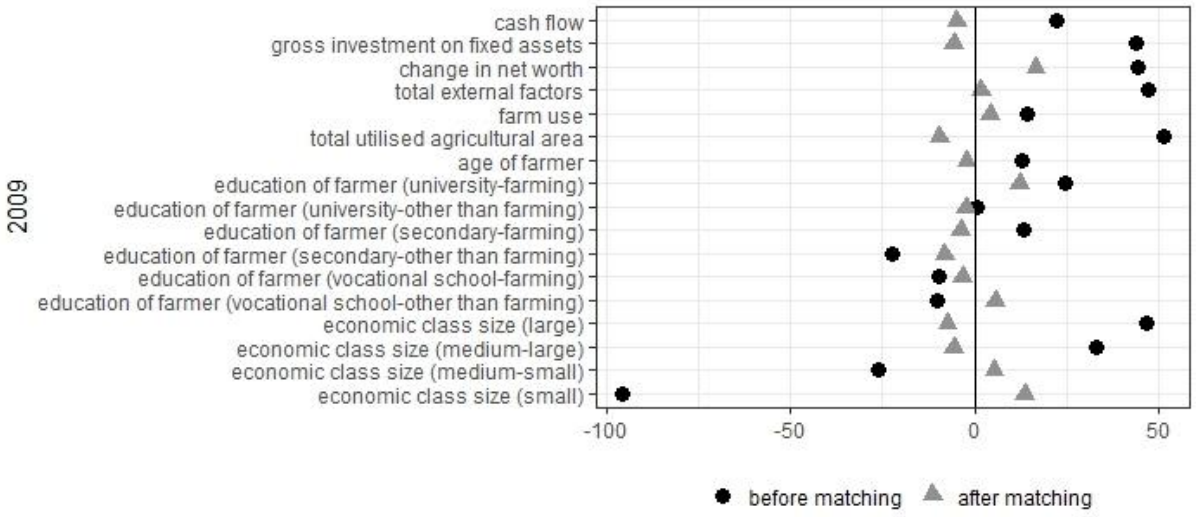
4 Results

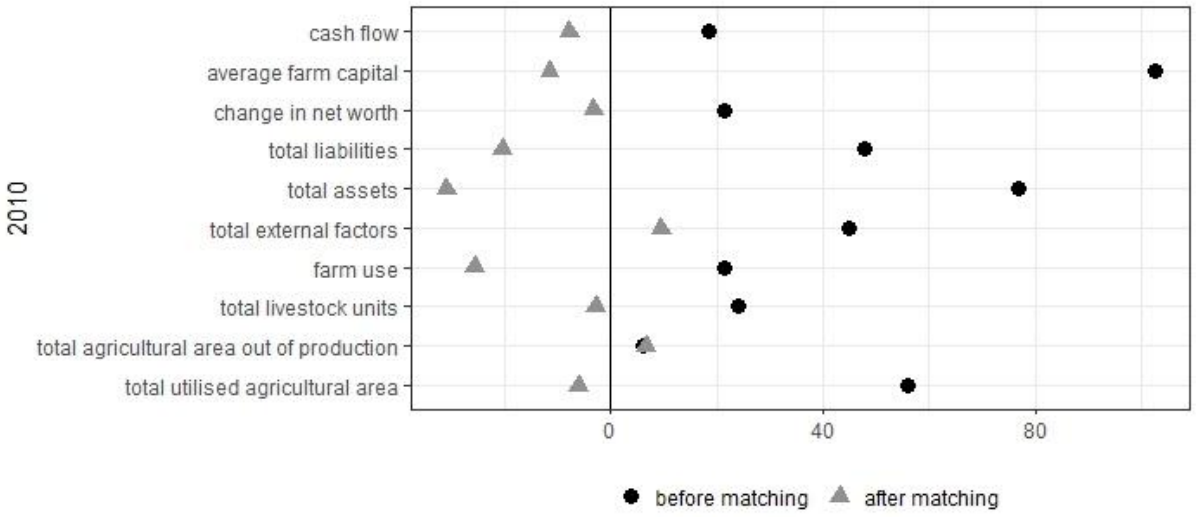
According to Table 1, the characteristics which mostly impacted in 2009-2014 on the probability of farm being assigned to the treatment group (receiving the investment support), were: farm use (positive impact), total external factors (positive impact), total utilised agricultural area (negative impact), change in net worth (positive impact) and average farm capital (positive impact). For each estimated logit model the accuracy rate was at least 0.89.

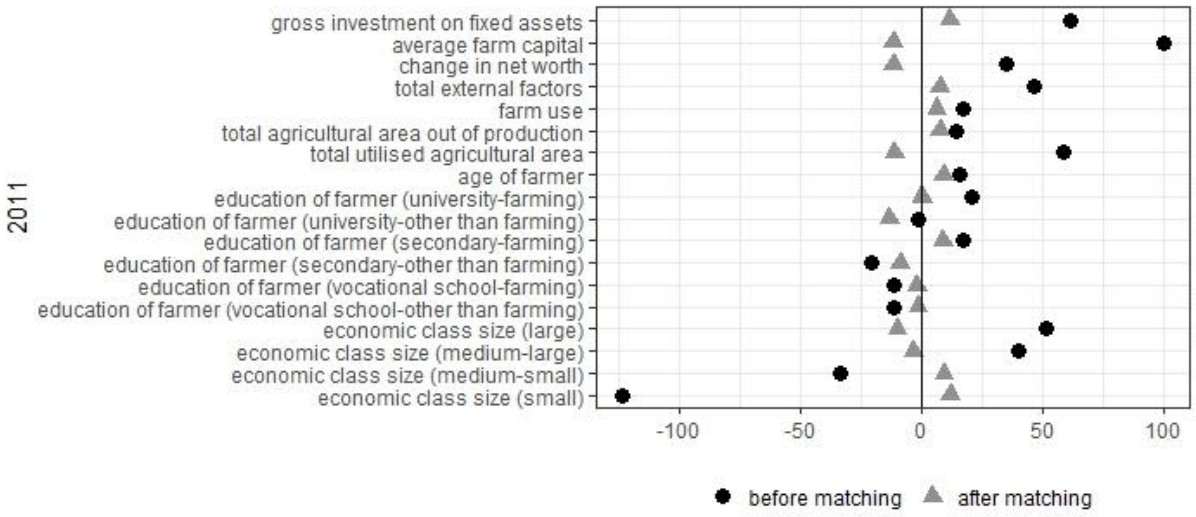
Table 1. Balanced variables included in propensity score model.

variable	2009	2010	2011	2012	2013	2014
economic class size	yes	no	yes	yes	no	yes
education of farmer	yes	no	yes	no	yes	no
age of farmer	yes	no	yes	no	no	no
total utilised agricultural area	yes	yes	yes	no	yes	yes
total agricultural area out of production	no	yes	yes	no	no	no
total livestock units	no	yes	no	yes	no	yes
farm use	yes	yes	yes	yes	yes	yes
total external factors	yes	yes	yes	yes	yes	yes
total assets	no	yes	no	no	yes	yes
total liabilities	no	yes	no	yes	yes	no
change in net worth	yes	yes	yes	no	yes	yes
average farm capital	no	yes	yes	yes	yes	yes
gross investment on fixed assets	yes	no	yes	yes	yes	no
cash flow	yes	yes	no	yes	yes	no
classification accuracy for model	0.91	0.9	0.89	0.93	0.91	0.89

Source: own elaboration based on the FADN data.


2009
cash flow
gross investment on fixed assets
change in net worth
total external factors
farm use
total utilised agricultural area
age of farmer
education of farmer (university-farming)
education of farmer (university-other than farming)
education of farmer (secondary-farming)
education of farmer (secondary-other than farming)
education of farmer (vocational school-farming)
education of farmer (vocational school-other than farming)
economic class size (large)
economic class size (medium-large)
economic class size (medium-small)
economic class size (small)
-100
-50
0
50
before matching
after matching


2010
cash flow
average farm capital
change in net worth
total liabilities
total assets
total external factors
farm use
total livestock units
total agricultural area out of production
total utilised agricultural area
0
40
80
before matching
after matching


2011
gross investment on fixed assets
average farm capital
change in net worth
total external factors
farm use
total agricultural area out of production
total utilised agricultural area
age of farmer
education of farmer (university-farming)
education of farmer (university-other than farming)
education of farmer (secondary-farming)
education of farmer (secondary-other than farming)
education of farmer (vocational school-farming)
education of farmer (vocational school-other than farming)
economic class size (large)
economic class size (medium-large)
economic class size (medium-small)
economic class size (small)
-100
-50
0
50
100
before matching
after matching

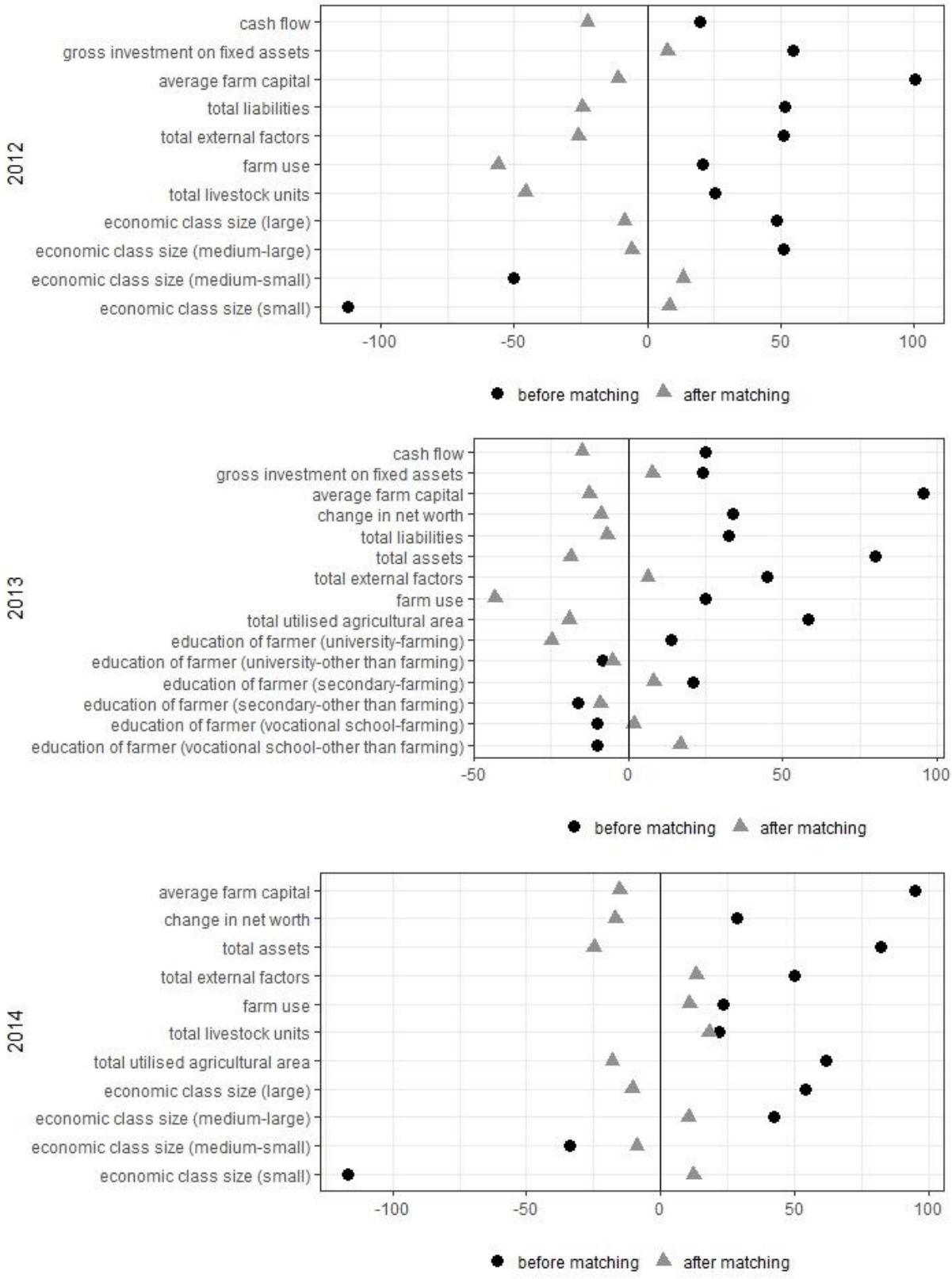


Fig. 2. Standardized differences in means.

Source: own elaboration based on the FADN data.

Fig. 2 shows the standardized mean differences for each explanatory variable in logit models.[2] It confirms that, comparing to unmatched samples, the matching procedure ensured significantly more balanced treatment and control groups.

The crop farms, which in 2009-2014 received support for investment, recorded both negative (in 2010 and 2012-2015) and positive (in 2011) effect of those subsidies on the increase in gross value added per annual work unit. According to Table 2, the negative treatment effect of subsidies on investment was dominant, when compared to the units which did not receive those subsidies. For example, the crop farms which in 2009 did not receive the investment support, recorded, on average, by 14 percentage points higher annual increase in labour productivity compared to the beneficiaries of the programme. Given relatively high standard errors for almost every estimates of ATT, it should be noted that the negative effects were not statistically significant, what confirm the results presented by Mary (2013)

Table 2. Average treatment effect for the treated.

year	2010	2011	2012	2013	2014	2015
ATT	-0.143	0.391	-0.099	-0.067	-0.056	-34.56
standard error	0.681	0.175	0.096	0.188	0.106	46.29
p-value	0.83	0.03	0.31	0.72	0.6	0.46
number of observations	940	987	1017	1080	1106	1152
number of treated observations	118	127	127	138	137	142

Source: own elaboration based on the FADN data.

The only significant difference between treatment and control group occurred in 2011. Then, the crop farms which in 2010 received the analysed support, were characterised, on average, by 39 percentage points higher annual increase in the labour productivity compared to the control group.

Conclusions

The study refers to the microeconomic producer theory as a framework and the expected positive relationship between investments, but also subsidies on investments and labour productivity, as a basis of farmers' income from work. It results from the assumption that for a single farm the employment of the labour factor is constant, so the increase in the use of

[2] Due to page restrictions, the histograms of each covariate by treatment groups before and after matching cannot be presented.

(physical) capital factor should imply an increase in the technical equipment of labour and, consequently, lead to increased labour productivity. These processes rely on investments made by farms (producers). The increase in labour productivity could be also supported by implementation of relevant policy instruments.

The objective of the study was to carry out the quantitative assessment of the treatment effect of investment support on the increase in labour productivity in Polish crop farms in 2010-2015. The results suggest that the positive effect of investment subsidies occurred only in 2011. Back then, the farms which in 2010 received the analysed support, were characterised, on average, by 39 percentage points higher annual increase in the labour productivity compared to the control group. In the remaining years, in turn, the impact of investment support on the increase in labour productivity was negative. However, given relatively high standard errors, the differences between farms which received and did not receive analysed payments were not statistically significant. These conclusions confirm the results presented by Mary (2013) that targeted subsidies have no significant impact on productivity.

Acknowledgement

This publication was prepared as a contribution to the research conducted within the Multi-Annual Programme 2015-2019 entitled "The Polish and the EU agricultures 2020+. Challenges, chances, threats, proposals" in IAFE-NRI.

References

Floriańczyk, Z., Osuch, D. & Płonka, R. (2017). *2015 Standard Results of Polish FADN agricultural holdings*. IERIGŻ-PIB, Warsaw.

Guo, S. & Fraser, M. W. (2015). *Propensity Score Analysis. Statistical methods and applications. Second Edition*. SAGE Publication, Thousand Oaks.

Heckman, J. J., Ichimura, H. & Todd, P. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *The Review of Economic Studies, 64(4)*, 605-654.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81(396)*, 945-960.

Latruffe, L. 2010. Competitiveness, Productivity and Efficiency in the Agricultural and Agri-Food Sectors. OECD Food. *Agriculture and Fisheries Working Papers 30*. OECD Publishing, Paris.

Leite, W. (2017). *Practical Propensity Score Methods Using R.* SAGE Publications, Thousand Oaks.

Mary, S. 2013. Assessing the impacts of Pillar 1 and 2 subsidies on TFP in French crop farms. *Journal of Agricultural Economics*, *64(1)*, 133-144.

Nilsson, P. 2017. Productivity effects of CAP investment support: Evidence from Sweden using matched panel data. *Land Use Policy, 66*, 172-182.

Pan, W. & Bai, H. (eds.). 2015. *Propensity Score Analysis. Fundamentals and development*. The Guilford Press, New York.

Pawłowska, A. & Bocian, M. 2017. Estymacja wpływu polityki rolnej na wydajność pracy z wykorzystaniem propensity score matching. *Publikacje Programu Wieloletniego 2015-2019, 50.* IERIGŻ-PIB, Warszawa.

Rembisz, W., Sielska, A. & Bezat-Jarzębowska, A. 2014. Agricultural policy and the *Multiannual Program Reports 2011-2014, 102.1*decisions of agricultural producers as to income and investment.. IERIGŻ-PIB, Warsaw.

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70(1)*, 41-55.

Sekhon, J. S. (2011). Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R. *Journal of Statistical Software*, *42(7)*, 1-52.

Strawiński, P. (2014). *Propensity score matching. Własności małopróbkowe*. Wydawnictwo Uniwersytetu Warszawskiego, Warszawa.

Zhu, X. & Lansink, A. O. 2010. Impact of CAP subsidies on technical efficiency of crop farms in Germany, the Netherlands and Sweden. *Journal of Agricultural Economics*, *61*, 545-564.

The composition of agricultural macroregions in Eastern Poland - an empirical example of the Aggregation Problem

Michał Bernard Pietrzak[1]

Abstract

The focus of the work is on the problem of the Modifiable Areal Unit Problem (MAUP). The Aggregation Problem being one of the aspects of the MAUP issue will be investigated in the paper. The research objective of this paper will be to consider the Aggregation Problem based on an empirical example, regarding an alternative way of determining agricultural macroregions for the eastern part of Poland. The implementation of the research objective will consist in determining an adequate composition of territorial units for the purpose of analysing economic phenomena related to the situation of agriculture in Poland. Due to the assumed research objective, the author proposed a two-step procedure, the application of which will allow the Aggregation Problem to be solved positively. According to the proposed procedure, the boundaries of agricultural macro-regions were determined based on the analysis of the agrarian structure at the district level (NUTS4). The analysis of the agrarian structure carried out in the paper allowed the determination of homogeneous agricultural macroregions in terms of the culture and development of agriculture. The research is funded by the National Science Centre, Poland under the research project no. 2015/17/B/HS4/01004.

Keywords: *spatial econometrics, Aggregation Problem, agrarian structure, Modifiable Areal Unit Problem*
JEL Classification:C10, C15, C21
DOI: 10.14659/SEMF.2018.01.38

1 Introduction

One of the essential topics discussed within spatial econometrics is the issue of Modifiable Areal Unit Problem (MAUP), which is a crucial issue that is given much consideration within spatial econometrics. (see: Anselin, 1988; Suchecki, 2010; Jaworska et al., 2014; Pietrzak and Ziemkiewicz, 2017). The MAUP is about differentiated results of spatial analyses while modifying the aggregation scale or a composition of territorial units.Within the MAUP issue, Aggregation Problem is considered. The subject literature contains studies of the Aggregation Problem, where the possibility of receiving different research results is indicated depending on the adopted composition of territorial units for the same level of aggregation (see: Openshaw and Taylor, 1979; Openshaw 1984, Jaworska et al., 2014).

The research objective of this paper is to consider the Aggregation Problem based on an empirical example. The selected empirical example concerns an alternative method of

[1]Corresponding author: Nicolaus Copernicus University, Department of Econometrics and Statistics, ul. Gagarina 13a, 87-100 Toruń, Poland, pietrzak@umk.pl.

determining agricultural macroregions for the eastern part of the territory of Poland. The implementation of the objective will consist in determining an adequate composition of territorial units for analysis of economic phenomena related to the situation of agriculture in Poland. The essence of the Aggregation Problem is that the adoption of different compositions of territorial units acting as agricultural macro-regions will result in obtaining different outcomes of research done on agriculture. Therefore, a positive solution to the Aggregation Problem will consist in finding an adequate composition, which will be the only one to allow a correct analysis of the properties of phenomena occurring in agriculture as part of the research problem undertaken.

It should be emphasized that in the case of analyses carried out on agriculture in Poland, the Aggregation Problem was successfully solved and in 2000 the adequate composition of SGM agricultural macroregions was established. The SGM composition is currently used in Poland for administrative purposes related to the Farm Accountancy Data Network (see: Skarżyńska et al., 2005). Therefore, the implementation of the research objective of the paper in the form of an alternative way of determining agricultural macroregions will allow us to emphasize the fact that the Aggregation Problem is a problem of an adequate delimitation of the composition of territorial units. Although the proposed method will be significantly different from the method used to determine agricultural SGM macroregions, the macroregion boundaries will be identical. This results from the fact that the composition of agricultural macroregions was determined in both cases as part of the research problem concerning the establishment of homogenous areas in terms of the culture and development of agriculture.

Moreover, isolating agricultural macroregions only for the eastern part of Poland allows indicating that the Aggregation Problem does not need to refer only to the adoption of a complete composition of territorial units for the entire country. The research carried out in this paper demonstrates the possibility of solving the Aggregation Problem for any area.

2 Methodology and data

Considering the Aggregation Problem within the spatial econometrics stems from the character of the composition of territorial units, whose boundaries of regions can be determined arbitrarily (see: Openshaw, 1984). It should be emphasized, however, that the arbitrariness of determining the boundaries of the composition is only apparent. Characteristics, dependencies for the analysed economic phenomena relate to specific areas, therefore, their boundaries cannot be determined freely. Literature studies on the Aggregation Problem point to the variability of results depending on the arbitrarily adopted compositions

of territorial units for the same level of aggregation (see: Openshaw and Taylor, 1979; Openshaw, 1984). Considering the Aggregation Problem in this way is fruitless in terms of cognition, since for different compositions of territorial units various results must be obtained. This is due to the fact that within the altered boundaries of compositions of territorial units, the properties and relationships between phenomena are mixed up, therefore, almost every result can be obtained. From the perspective of the problems undertaken in the study, the described variability of results should be obvious to researchers.

According to the author, the burden of the Aggregation Problem should be transferred to the adoption of the adequate composition of territorial units based on the earlier identification of the properties of the phenomena studied and the dependencies occurring between them. This means that the key task of the researcher when considering the Aggregation Problem should be the correct identification of spatial variability of phenomena within the research problem undertaken, and then the adoption of the adequate composition of territorial units. However, the selection of the composition should be determined by the nature and variability of the phenomena studied.

Therefore, a positive solution of the Aggregation Problem will consist in determining the adequate composition of territorial units. This means that exclusively the researcher's knowledge and scientific experience can allow a proper determination of a composition. In addition, it should be emphasized that the positive solution of the Aggregation Problem can be achieved only as part of an empirical research problem.

In accordance with the assumed research objective of the paper, the Aggregation Problem will be presented in the light of the research problem which is the determination of agricultural macroregions for the eastern part of the territory of Poland. The established agricultural macroregions should be homogeneous internally, in terms of the culture and development of agriculture (see: Skarżyńska et al., 2005). The designated boundaries of macroregions will form a composition of territorial units that will be adequate for research conducted on agriculture in Poland.

In the first step of the research, the composition of SGM macroregions used administratively in Poland will be presented. The formal adoption by Poland of the composition of SGM macroregions in Poland was related to its European Union accession (see: Skarżyńska et al., 2005). Accessing EU structures imposed on Poland an obligation to establish its Farm Accountancy Data Network (FADN). Taking into account the eastern part of Poland, two SGM macroregions can be distinguished. The Mazowsze and Podlasie Region, which includes the following provinces: *lubelskie*, *łódzkie*, *mazowieckie* and *podlaski*e, and

the Małopolska and Pogórze Region, which, in turn, includes the following provinces: *małopolskie*, *podkarpacki*e, *śląski*e and *świętokrzyskie*. The two above-mentioned agricultural SGM macroregions and provinces forming them are shown in Fig. 1. The establishment of the SGM macroregion composition for Poland in 2000 should be considered as an example of a positive solution of the Aggregation Problem. The research on agricultural issues performed based on the SGM macroregions should lead to the obtainment of correct results.



Fig. 1. Agricultural SGM macroregions in the eastern part of Poland.

The next step of the research will be to propose a procedure allowing positive solutions of the Aggregation Problem. In the first stage of the procedure, it is vital to check whether at present there is a composition of territorial units which would be appropriate for conducting analysis of the properties of the phenomena examined and the relationships between them. If such a composition exists, then it should be applied in analyses, which at the same time will end the procedure of a positive solution of the Aggregation Problem.

If there is no composition of territorial units that can be used in the research undertaken, or if the existing composition is not the adequate one, then it is necessary to move on to the second stage of the procedure. In the second stage of the procedure, the boundaries of the new composition of territorial units should be defined. The boundaries of the new composition must be determined in such a way so that they should allow a proper description of the properties of the analysed phenomena as part of the research problem undertaken. The

methods of multiple-criteria analysis[2], including linear ordering methods (see: Cheba and Szopik-Depczyńska, 2017; Kuc, 2017), cluster analysis (Małkowska and Głuszak, 2016), applications of Structural Equation Modeling (Pilelienė and Grigaliūnaitė, 2017), methods of convergence analysis (see: Próchniak and Witkowski, 2016; Furková and Chocholatá, 2017; Wójcik, 2017). Data Envelopment Analysis (see: Balcerzak et al., 2017) ormethods of locataion quotient (see: Suchecki, 2010; Koľveková and Palaščáková, 2017) are considered as appropriate tools for determining boundaries of macroregions.Implementation of the second stage concludes the procedure of a positive solution of the Aggregation Problem[3].

3 Discussion of research findings

Although the problem of defining the boundaries of SGM macroregions has been solved positively, the paper will attempt to re-identify the composition of agricultural macroregions. Another determination of agricultural macroregions is aimed at presenting an alternative solution of the Aggregation Problem. Instead of the taxonomic analysis carried out in case of the SGM macroregions, the analysis of the agrarian structure was carried out in the paper. The designated macroregions should be homogeneous in terms of the culture and development of agriculture. The agrarian structure of farms is a good tool to measure the level of agricultural development, since it constitutes a basic element of rational management in agriculture. The term ‘agrarian structure’ is defined as the distribution of farms by area (see: Walczak and Pietrzak, 2016). The area of arable land was calculated as the total area of agricultural farms minus land in forest use and wasteland. In order to determine the spatial variability of the agrarian structure, the concentration of arable land at the district level was determined (NUTS 4). The concentration value for each district was determined using Gini coefficient. The following area groups (1-5 hectares), (5-10 hectares), (10-20 hectares), (20-50 hectares) and (50 hectares and more) were selected to calculate the index value. The data used come from the Universal Agricultural Census conducted in 2002.

The spatial diversification of the agrarian structure in the eastern part of Poland is shown in Fig. 2 (the graph on right side), where districts were divided by concentration of agricultural land into three classes. Districts were assigned to classes based on the natural break method. The analysis of the results presented on right side of Fig. 2 allows the existence

[2]The SGM macroregions were determined on the basis of taxonomic measure of development (TMD)(see: Skarżyńska et al., 2005).

[3] It should be emphasized that over years, there may occur a change in the spatial diversity of the studied phenomena, which means that the Aggregation Problem will need to be resolved again.

of the spatial differentiation of the agrarian structure in the eastern part of Poland to be verified. This confirms the need to create macroregions that would differ in their level of the culture and development of agriculture.

The analysis of spatial diversity of the agrarian structure makes it possible to state that in the *warmińsko-mazurskie* province all its districts were assigned to the third class with the highest concentration of agricultural land. In the *śląskie* province, the dominant ones are (in a similar number) districts assigned to the third and second classes. In the *łódzkie, mazowieckie, podlaskie* and *lubelskie* provinces, districts assigned to the second class prevail. However, in the case of the *małopolskie, świętokrzyskie* and *podkarpackie* provinces, there can be noted a numerical advantage of districts assigned to the first class with the lowest level of concentration of agricultural land.

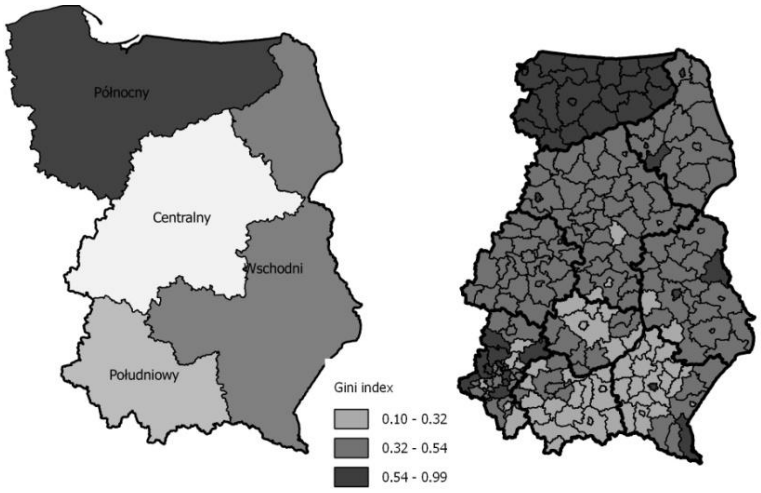


Fig. 2. Part of the composition of territorial units NUTS 1 (regions) and the concentration of agricultural land at the level of districts (NUTS4).

The *warmińsko-mazurskie* province is characterized by the highest level of agricultural development, though it is bases mainly on the cultivation of cereals. The characteristics of the *warmińsko-mazurskie, mazowieckie, podlaskie,* and *lubelskie* provinces are medium-sized farms and, unfortunately, poor quality of soils. In these provinces there is a large share of permanent grassland, which results in a high cattle density. One can also notice there a high share of cereals, however, the low level of fertilization means that the yields obtained are low. All this contributes to the low intensity of agricultural production in the region. The *małopolskie, podkarpackie, śląskie* and *świętokrzyskie* provinces are distinguished by the largest agrarian fragmentation in the country with the smallest average area of farms. The low intensity of production is also affected by a low consumption of mineral fertilizers and

purchases of concentrated feed. However, the cast of animals is the highest in the country. Based on these considerations, relatively homogeneous groups can be formed, where the first group will be created by the *warmińsko-mazurskie* province alone. The second group will be formed by the *łódzkie, mazowieckie, podlaskie,* and *lubelskie* provinces. The third group, in turn, will be formed by *małopolskie, podkarpackie, świętokrzyskie,* and *śląskie* provinces[4].

In accordance with the proposed procedure for determining the appropriate composition of territorial units, in the first stage an attempt will be made to select an already existing composition. For this purpose, it will be assumed that territorial units are available only under the NUTS classification and, based on them, agricultural macro-regions will be built. Since the number of SGM macroregions should not be too high, the NUTS 1 composition (regions) will be adopted as the initial composition of agricultural macroregions. Fig. 2 (the graph on left side), shows the separated eastern part of the NUTS 1 composition of territorial units. The comparison of the NUTS 1 composition of territorial units with the suggested groups of provinces indicates that this composition is inadequate for conducting research on the development of agriculture in eastern Poland.

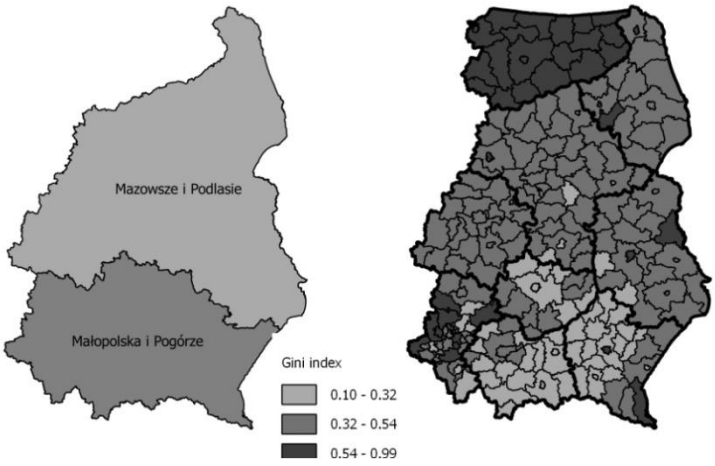


Fig.3. Part of the composition of SGM territorial units and the concentration of agricultural land at the district level (NUTS 4).

Therefore, in the second stage of the procedure agricultural macroregions will be determined based on the analysis of the agrarian structure. The basic spatial unit used to create macroregions will be provinces. The *warmińsko-mazurskie* province will not be included, since it creates the first group with too small space on its own. The second group

[4] The *śląskie* province was included in the third group due to the spatial cohesion criterion.

and the third group make up areas with the size large enough to form macroregions. Therefore, the provinces of the first group will create the first agricultural macroregion, and the provinces from the second group will form another macroregion. It turns out that agricultural macroregions determined on the basis of the agrarian structure analysis coincide with the administrative composition of SGM macroregions (see Fig. 3). Determining the appropriate boundaries of agricultural macroregions based on the analysis of the agrarian structure shows that for any research problem there may be different, alternative ways to solve the Aggregation Problem.

Conclusions

The research objective of the study was to consider the Aggregation Problem based on an empirical example, regarding an alternative way of determining agricultural macroregions for the eastern part of Poland. A positive solution of the Aggregation Problem consists in determining the composition of territorial units, which provides the opportunity to correctly analyse phenomena within the research problem. Therefore, the paper proposed a two-stage procedure, the use of which allows a positive solution of the Aggregation Problem.

According to the proposed procedure, the boundaries of agricultural macroregions were determined for the part of the eastern territory of Poland based on the analysis of the agrarian structure. The analysis of the agrarian structure at the level of districts (NUTS 4) carried out in the paper allowed us to determine homogenous agricultural macroregions in terms of the culture and development of agriculture. In this way, a positive solution of the Aggregation Problem was presented in the form of determining an adequate composition of territorial units for analysis of economic phenomena related to the situation of agriculture in Poland.

The analysis conducted showed large disparities in the agrarian structure in the eastern part of Poland, which is reflected in significant disproportions in the level of key variables in agricultural activity. Therefore, the state agricultural policy should be diversified and stimulate various directions of agricultural development depending on the agricultural macroregions. A method for determining the appropriate systems of territorial units included in the draft proposal can be the basis for delimitation issues.

Acknowledgements

The research is funded by National Science Centre, Poland under the research project no. 2015/17/B/HS4/01004.

References

Anselin, L. (1988). *Spatial Econometrics: Method and Models*. Kluwer Academic Publishers, Netherlands.

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

Cheba, K. & Szopik-Depczyńska, K. (2017). Multidimensional comparative analysis of the competitive capacity of the European Union countries and geographical regions. *Oeconomia Copernicana, 8*(4), 487–504.

Furková, A. & Chocholatá, M. (2017). Interregional R and D spillovers and regional convergence: a spatial econometric evidence from the EU regions. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(1), 9-24.

Koľveková, G. & Palaščáková, D. (2017). Analysis of Bratislava and Žilina as urban areas in Western Slovakia in the context of associations among employment and industries. *Oeconomia Copernicana, 8*(4), 547–562.

Kuc, M. (2017). Social convergence in Nordic countries at regional level. *Equilibrium-Quarterly Journal of Economics and Economic Policy*, *12*(1), 25-41.

Jaworska, R., Łaszkiewicz, E. & Modranka, E. (2014). Dane przestrzenne-podstawowe zagadnienia. In: Suchecka, J. (eds.), *Statystyka przestrzenna, metody analizy struktur przestrzennych*. C.H. Beck, Warszawa.

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

Openshaw, S. & Taylor, P. J. (1979). A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In: Wrigley, N. (eds.), *Statistical methods in the spatial sciences*. Pion, London.

Openshaw, S. (1984). *The Modifiable Areal Unit Problem*. GeoBooks, CATMOG 38, Norwich.

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

Pilelienė, L. & Grigaliūnaitė, V. (2017). Colour temperature in advertising and its impact on consumer purchase intentions. *Oeconomia Copernicana, 8*(4), 667–679.

Próchniak, M. & Witkowski, B. (2016). On the use of panel stationarity tests in convergence analysis: empirical evidence for the EU countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 11*(1), 77-96.

Skarżyńska, A., Goraj, L. & Ziętek, I. (2005). *Metodologia SGM „2002" dla typologii gospodarstw rolnych w Polsce*. IERiGŻ-PIB, Warszawa.

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

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

Wójcik, P. (2017). Was Poland the next Spain? Parallel analysis of regional convergence patterns after accession to the European Union. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(4), 593–611.

Comparing populations based on squared Euclidean distance between sets of variables

Dominika Polko-Zając[1]

Abstract

In economics as well as in other researches there is often a need to detect the differences between multidimensional populations. This paper concerns the problem of comparing such populations using distance between two analysed sets of objects being characterized by many variables. The study use squared Euclidean distance measure but method can be used with any kind of distance measure between objects. In order to identify differences permutation tests were used. The method is illustrated by analysing economics data sets. Included empirical example contains data from Central Statistical Office of Poland. All calculations were done in R program.

Keywords: *multidimensional data, distance approach, permutation tests, Monte Carlo study*
JEL Classification: C12, C15, C30
DOI: 10.14659/SEMF.2018.01.39

1 Introduction

Multidimensional methods are important part of statistical methods. The purpose of the research is usually to detect the similarity or existing differences between the examined objects being characterized by many diagnostic variables. This stems mainly from the fact that in many areas of empirical research they relate to phenomena of complex, multidimensional structure.

Statistical methods of multidimensional analysis e.g. cluster analysis are used to divide one set of data into homogeneous groups of objects. The level of assessment of the similarity/distance of objects allows to include them in the same group of objects or to conclude that there are no similarities between the objects of the study. Cluster analysis allows to organize objects when their structure is unknown (classification within groups). This structure should only be discovered while having multidimensional data about objects. The division of the set of objects is based on a certain distance measure between the examined objects $d(x_i, x_j)$, where $i, j = 1,...,n.$. The calculated distance matrices containing distances for each of the object pairs allow to evaluate the differences between these objects. Distance

[1] Corresponding author: University of Economics in Katowice, Department of Statistics, Econometrics and Mathematics, ul. 1 Maja 50, 40-287 Katowice, Poland, e-mail: dominika.polko@ue.katowice.pl

measurements are characterized by the fact that the increase in value means an increase in the diversity between the examined objects.

The article considers a different approach. Samples (groups) were randomly taken from *k* populations of any continuous distributions. The purpose of the study was to test the differences between the *k* populations using for that the distance between objects in the considered samples. Proposed method can be applied even when the sample size is small.

2 Dissimilarity (distance) measures

Comparing the examined objects in data sets, e.g. countries, regions, etc., measures which determine the similarity or differences between pairs of these objects are calculated. The measures used most often are divided into measures of proximity and measures of distance. Literature provides numerous suggestions of measures applicable in multidimensional methods. Applying different measures of similarity of objects is not a problem when all variables describing objects are measured on a scale of one type. Tables with numerous distance measures for variables measured on an interval, order and nominal scale are presented by Walesiak and Gatnar (2009). When the quantitative variables are analysed (interval), the following distance measures are distinguished to measure the distance between points $x=(x_1,x_2...x_n)$ and $y=(y_1,y_2...y_n)$ in the *n*–dimensional space, e.g.:

- Manhattan

$$d(x,y)=\sum_{i=1}^{n}|x_i-y_i|, \tag{1}$$

- Euclidean

$$d(x,y)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}, \tag{2}$$

- Squared Euclidean

$$d(x,y)=\sum_{i=1}^{n}(x_i-y_i)^2, \tag{3}$$

- Chebychev

$$d(x,y)=\max_{1\le i\le n}|x_i-y_i|, \tag{4}$$

- Minkowski

$$d(x,y)=\sqrt[p]{\sum_{i=1}^{n}|x_i-y_i|^p}. \tag{5}$$

Method described in the next paragraphs can be used with any kind of distance measure between objects but further considerations (example) use only squared Euclidean distance measure.

3 Distances between groups

Applying multidimensional analysis, various methods are used to calculate distances between groups. These methods are used to optimize the pre–made division into groups of objects. There are five commonly used hierarchical clustering methods (Everitt et al., 2011):

- average between groups linkage, calculated as the average distance between all pairs of objects belonging to groups A and B,

$$d(A,B)=\frac{\sum_{i=1}^{n_A}\sum_{j=1}^{n_B} d\left(O_{A_i},O_{B_j}\right)}{n_A \cdot n_B}, \tag{6}$$

- average within groups linkage, determined as the arithmetic mean of the distance between all possible pairs of objects belonging to groups A and B,

$$d(A,B)=\frac{\sum_{i=2}^{n_A}\sum_{p=1}^{i} d\left(O_{A_i},O_{A_p}\right)+\sum_{j=2}^{n_B}\sum_{q=1}^{j} d\left(O_{B_j},O_{B_q}\right)+\sum_{i=1}^{n_A}\sum_{j=1}^{n_B} d\left(O_{A_i},O_{B_j}\right)}{\frac{n_A\left(n_A-1\right)}{2}+\frac{n_B\left(n_B-1\right)}{2}+n_A \cdot n_B}, \tag{7}$$

- centroid method, calculated as the distance between the centroids of objects of groups A and B,

$$d(A,B)=d\left(\overline{X}_A,\overline{X}_B\right), \tag{8}$$

  where: $\overline{X}_A,\overline{X}_B$ are midpoints of objects of groups A and B,

- median method, determined as the median distance between objects belonging to groups A and B

$$d(A,B)=\underset{i,j}{median}\left\{d\left(O_{A_i},O_{B_j}\right)\right\}, i=1,...,n_A,\ j=1,...,n_B, \tag{9}$$

- Ward's method, where distance between groups defined as sum of squares within groups, after fusion, summed over all variables,
- methods based on the vicinity of objects, e.g. the nearest neighbourhood (single linkage), i.e. the distance between the nearest objects belonging to groups A and B respectively, or the furthest neighbourhood (complete linkage), i.e. the distance between the most distant objects belonging to groups A and B respectively.

For all methods of calculating the distance between groups the distance matrix is calculated using the suitable measures presented in the second part of this article. The methods discussed above excluding Centroid method, Median method and Ward's method can be used with any kind of similarity or distance measure between objects. Above mentioned three methods use squared Euclidean distances.

The goal of those methods is to obtain the most concentrated values in groups, so often in multidimensional analyses, methods based on determining the average distance between objects within groups are used. In the method presented in the article, the goal is not to obtain homogeneous groups but testing existing differences between them. The analysis used selected measures of inter–group distances or their modifications to determine test statistics during the verification of the hypothesis about the similarity of the studied populations based on sets of variables.

4 Testing differences based on distances between groups

Let's assume that there are two sets of objects $A=\{O_1,O_2,...,O_{n_A}\}$ and $B=\{O_1,O_2,...,O_{n_B}\}$ of sizes n_A and n_B respectively. Each object is described using p diagnostic variables. Comparing populations based on two sets of variables the null hypothesis can be stated as follows: "samples were taken from multidimensional populations having the same distributions", which suggests the equality of two distributions. The alternative hypothesis is formulated: "samples were taken from multidimensional populations having different distributions". In a two–sample testing problem formally these hypotheses can be written as follows

$$H_0 : F = G, \tag{10}$$

and the alternative

$$H_1 : F \neq G \tag{11}$$

where:

F, G – continuous but unknown multidimensional distributions of populations.

To test the null hypothesis the permutation test can be used. A test statistic is computed using two sets of independent p–dimensional observations. Permutation tests in general take a test statistic T used for a parametric test, or one derived intuitively (Baker, 1995). Unfortunately

many classical statistical methods do not have their counterparts for multidimensional data. To test null hypothesis against alternative hypothesis the following test statistic can be used

$$T^{(1)} = \frac{\sum_{i=1}^{n_A}\sum_{j=1}^{n_B} d\left(O_{A_i}, O_{B_j}\right)}{n_A \cdot n_B}, \tag{12}$$

where,

$d\left(O_{A_i}, O_{B_j}\right)$ distances between *i*-th object belonging to group A and *j*-th object belonging to group B in multidimensional space R^p, i.e. distance between $O_{A_i} = \left(x_{i1}, x_{i2}, ..., x_{ip}\right)$ and $O_{B_j} = \left(y_{j1}, y_{2j}, ..., y_{jp}\right)$, where $i, j = 1, ..., n$.

It is also possible to test the differences between populations using as a test statistic the distance between the centroids of the population

$$T^{(2)} = d\left(\overline{X}_A, \overline{X}_B\right), \tag{13}$$

where: $\overline{X}_A, \overline{X}_B$ are midpoints of objects of groups A and B.

The distance between centroids, or midpoints in a multidimensional space, can also be expressed directly from the distances between items in each group. The formula to find the centroid distance were presented by Apostol and Mnatsakanian (2003)

$$T^{(3)} = \frac{1}{n_A \cdot n_B}\sum_{i=1}^{n_A}\sum_{j=1}^{n_B} d^2\left(O_{A_i}, O_{B_j}\right) - \frac{1}{n_A^2}\sum_{i=2}^{n_A}\sum_{p=1}^{i} d^2\left(O_{A_i}, O_{A_p}\right) - \frac{1}{n_B^2}\sum_{j=2}^{n_B}\sum_{q=1}^{j} d^2\left(O_{B_j}, O_{B_q}\right), \tag{14}$$

where,

$\sum_{i=2}^{n_A}\sum_{p=1}^{i} d^2\left(O_{A_i}, O_{A_p}\right)$ is the sum of squared distances between objects in group A,

$\sum_{j=2}^{n_B}\sum_{q=1}^{j} d^2\left(O_{B_j}, O_{B_q}\right)$ is the sum of squared distances between objects in group B,

$\sum_{i=1}^{n_A}\sum_{j=1}^{n_B} d^2\left(O_{A_i}, O_{B_j}\right)$ is the sum of squared distances between objects in group A and those in group B.

In order to identify differences between sets of variables permutation tests were used. Tests based on permutations of observations were introduced by R. A. Fisher in 1930's (Welch, 1990). Because of the need to perform complex calculations method was widely used only in recent decades, when computing capabilities of computers increased. Currently, the problem of using the permutation tests in statistical analysis is popular among researchers. The most

important references are Good (1994), Good (2005), Good (2006), Pesarin (2001), Basso et al. (2009), Pesarin and Salmaso (2010) and Kończak (2016).

Most of multidimensional two–sample tests perform poorly for high dimensional data and many of them are not applicable when the dimension of the data exceeds the sample size (Biswas and Ghosh, 2014). Permutation tests do not require additional assumptions about the form of the distribution in the population; are suitable for small sample sizes and are robust to outliers. The goal of the test is to verify hypothesis at certain level of significance to discover a differences between data sets. After the value of the statistic T_0 had been calculated, N permutations of variables were performed and values T_i ($i = 1, 2, \ldots, N$) were determined. The decision concerning a verified hypothesis is made on the basis of *ASL* (*achieved significance level*) value (Efron and Tibshirani, 1993):

$$ASL = P_{H_0}\{T \geq T_0\}. \tag{15}$$

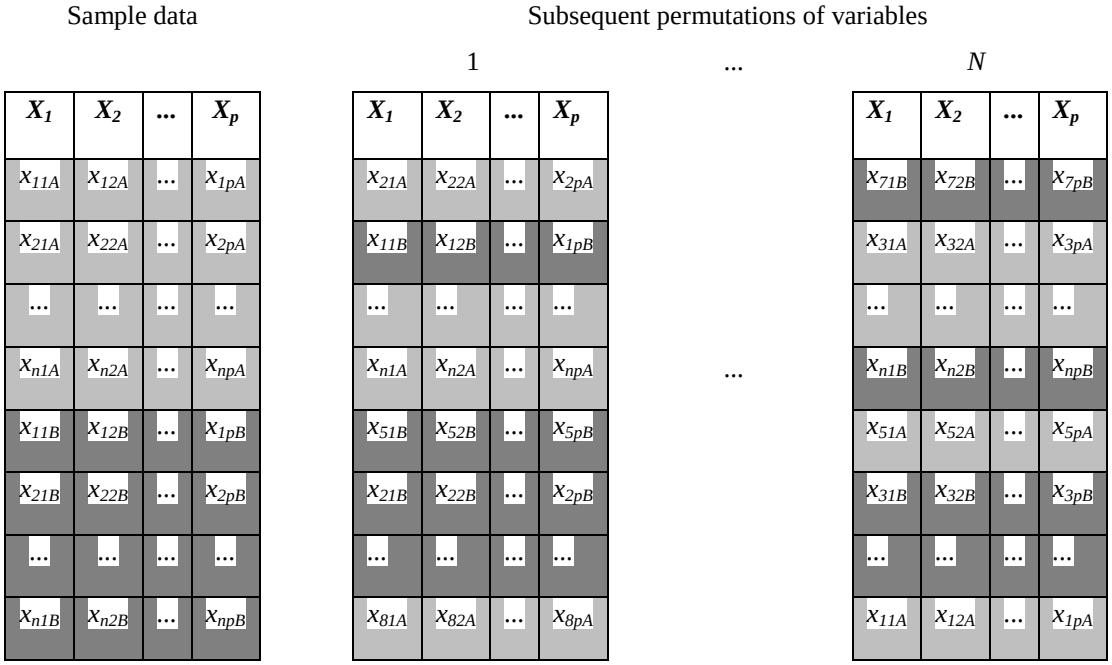


Fig. 1. The scheme of permutation variables.

On the basis of the random variable of large size (it is recommended in most cases the number of permutation to be greater than 1000) taken from the set of all possible permutations of the data set the *ASL* is determined using formula (Kończak, 2016)

$$\hat{ASL} = \frac{card\{i : T_i \geq T_0\}}{N}. \tag{16}$$

The smaller the value of *ASL*, the stronger evidence against H_0. Formally we choose a significance level α and reject H_0 if *ASL* is less than α.

The steps in the permutation test conducted to determine the significance of the difference between two sets of variables are as follows:

1. Assume the level of significance α;
2. Calculate the value of the statistics T_0 for the sample data;
3. Proceed permutations of data. A random permutation of data can be obtained by recreating the original data and destroying existing structure of variables (see. Fig. 1). Calculate test statistic values T_i for each permutation of data.
4. Create empirical distribution of T_i, where (i = 1, 2, … , N) and locate calculated value of T_0 on this distribution and estimate *ASL* value.

The simulation study was performed using R program (R Core Team, 2016). The author also presents another proposal for comparing multidimensional populations based on two sets of variables using permutation tests (Polko–Zając, 2017).

5 Empirical example

To illustrate the possibilities of application the method for the analysis of economic data, data from the Local Data Bank of the Central Statistical Office were used. The study concerned a comparison of the situation on the labour market. For this purpose, data for Polish voivodeships were used for two compared periods: 2004 and 2016 (Table 1). The data set contained five diagnostic variables:

X_1 – unemployment rate (in %),

X_2 – the percentage of the long–term unemployed, i.e. those unemployed for over a year in the total number of unemployed,

X_3 – the percentage of unemployed among people under 25 in the total number of unemployed,

X_4 – the percentage of unemployed people with no experience or length of employment up to 1 year in the total number of unemployed,

X_5 – the percentage of unemployed people with higher education in the total number of unemployed.

Table 1. Data of the situation on the labour market in 2004 and in 2016.

Voivodeship	Year 2004					Year 2016				
	X_1	X_2	X_3	X_4	X_5	X_1	X_2	X_3	X_4	X_5
Dolnośląskie	22.4	49.6	21.4	34.6	4.4	7.2	36.7	10.5	26.4	12.0
Kujawsko-Pomorskie	23.6	52.9	25.8	34.3	3.3	12.0	41.9	14.4	31.7	8.7
Lubelskie	17.8	54.7	28.0	49.9	7.4	10.3	45.4	16.0	46.1	15.4
Lubuskie	25.6	47.3	21.8	31.6	3.4	8.6	33.1	12.4	28.2	10.2
Łódzkie	19.5	54.8	22.4	36.3	5.4	8.5	43.1	11.2	29.3	11.8
Małopolskie	15.0	51.2	28.8	39.3	5.4	6.6	39.9	15.8	33.3	16.3
Mazowieckie	14.7	56.0	22.7	41.7	5.6	7.0	45.4	12.3	34.0	15.4
Opolskie	20.0	50.0	21.6	29.6	4.3	9.0	37.3	13.0	28.4	11.9
Podkarpackie	19.1	55.3	26.3	42.2	5.5	11.5	45.2	15.1	37.1	15.0
Podlaskie	16.1	50.2	25.6	43.2	6.5	10.3	46.0	14.6	38.5	14.5
Pomorskie	21.4	53.3	23.4	30.6	4.3	7.1	35.6	14.2	29.6	12.9
Śląskie	16.9	50.1	24.2	41.0	5.3	6.6	37.9	11.3	32.9	13.7
Świętokrzyskie	22.0	53.0	25.4	41.0	7.3	10.8	37.7	15.1	37.1	16.3
Warmińsko-Mazurskie	29.2	52.3	23.4	36.6	3.4	14.2	38.9	13.5	30.9	9.4
Wielkopolskie	15.9	48.4	27.4	34.1	4.3	4.9	35.0	14.6	27.5	13.0
Zachodniopomorskie	27.5	51.5	21.0	40.6	4.1	10.9	36.5	11.9	33.2	10.8

Source: Central Statistical Office of Poland (GUS).

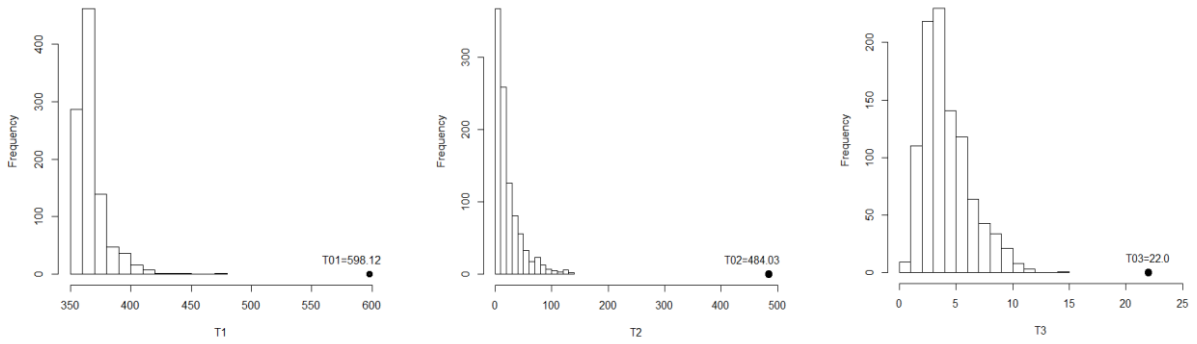


Fig. 2. Empirical distributions of statistics $T^{(1)}$, $T^{(2)}$ and $T^{(3)}$.

To test null hypothesis (10) the permutation test was used. Significance level $\alpha = 0.05$ was assumed and N=1000 permutations of variables were performed. As test statistics (12) – (14) were used. Statistics values were determined using the squared Euclidean distance measure (3). Values of test statistics calculated for the sample data are: T_{01}=598.12, T_{02}=484.03 and T_{03}=22.0. Empirical distributions of statistics were presented on Fig. 2. *ASL* values calculated with empirical distributions of statistics equal to zero so they are lower than assumed significance level. Verified hypothesis H_0 should be rejected in favour of alternative hypothesis, which means that there is the significant difference between situation on the labour market in 2004 and in 2016.

Conclusion

Many parametric and nonparametric methods are available for the multidimensional two–sample testing problem. Article deals with a permutation distance approach to multidimensional problem of hypothesis testing. The limitation of commonly used classical statistical methods makes the simulation methods being used in a variety of data analyses. The paper presents a method for testing distance between two analysed sets of objects characterized by many variables. In order to identify differences between populations permutation test was proposed. Introduced test based on comparing groups obtained by permuting the group's membership. The procedure using the permutation test is used to estimate the distribution of the test statistic. The proposed method is illustrated by an empirical example from real application for economics data sets.

References

Apostol, T. M. & Mnatsakanian, M. A. (2003). Sums of squares of distances in m-space. *Math. Assoc. Am. Monthly, 110*, 516.

Baker, R. D. (1995). 2 permutation tests of equality of variances. *Statistics and Computing*, *5*, 289–296.

Basso, D., Pesarin, F., Salmaso, L. & Solari, A. (2009). *Permutation Tests for Stochastic Ordering and ANOVA*. Heidelberg: Springer Science + Business Media.

Biswas, M. & Ghosh, A. K. (2014). A nonparametric two–sample test applicable to high dimensional data. *Journal of Multivariate Analysis*, *123*, 160–171.

Efron, B. & Tibshirani, R. (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall.

Everitt, B. S., Landau, S., Leese, M. & Stahl, D. (2011). *Cluster Analysis*. *5th edition*. John Wiley & Sons.

Good, P. I. (1994). *Permutation Tests: A Practical Guide for Testing Hypotheses*. New York: Springer–Verlag.

Good, P. (2005). *Permutation, Parametric and Bootstrap Tests of Hypotheses*. New York: Springer Science Business Media.

Good, P. (2006). *Resampling Methods. A Practical Guide to Data Analysis*. Boston–Basel–Berlin: Birkhauser.

Kończak, G. (2016). *Testy permutacyjne. Teoria i zastosowania*. Katowice: Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.

Pesarin, F. (2001). *Multivariate Permutation Tests with Applications in Biostatistics*. Chichester: John Wiley & Sons, Ltd.

Pesarin, F. & Salmaso, L. P. (2010). *Permutation Tests for Complex Data. Theory, Applications and Software*. Chichester: John Wiley & Sons, Ltd.

Polko–Zając, D. (2017). On comparing populations based on two sets of variables. In: Papież, M. & Śmiech S. (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, 349-358. ISBN: 978–83–65173–85–0 (HTML).

R Core Team (2016). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. URL https://www.R-project.org/.

Walesiak, M. & Gatnar, E. (2009). *Statystyczna analiza danych z wykorzystaniem programu R*. Warszawa: Wydawnictwo Naukowe PWN.

Welch, W. J. (1990). Construction of Permutation Tests. *Journal of the American Statistical Association. Theory and Methods, 85*(411), 693–698.

Entrepreneurship conditions in Poland at NUTS 3 level. Application of taxonomic measure of development based on median vector Weber

Elżbieta Rogalska[1]

Abstract

The main aim of the article is to analyse factors influencing entrepreneurship conditions in Poland at NUTS 3 level. The entrepreneurship conditions are considered as a multiple-criteria phenomenon. Thus, it is analysed based on 5 criteria. In this context the choice of variables describing entrepreneurship conditions at the NUTS 3 level was the biggest limitation of the research. Due to availability of data it was possible to conduct dynamic research for the years 2010-2015. In the research taxonomic measure of development was assessed with application of TOPSIS method based on median vector Weber. The obtained values of taxonomic measure of development enabled to rank the NUTS 3 regions starting with the once characterised with the best conditions for entrepreneurs to the once with the worst conditions and to analyse the stability of the obtained results in time. The analysis indicates that relatively stable disparities at regional level in regard to entrepreneurial conditions can be considered as a significant problem for regional policy in Poland.

Keywords: *entrepreneurship, multiple-criteria analysis, taxonomic measure of development, TOPSIS, median vector Weber, NUTS 3, Poland*

JEL Classification: C38, L26, P25

DOI: 10.14659/SEMF.2018.01.40

1 Introduction

After successful transformation all Central European economies face a challenge of avoiding middle income trap. Many international studies indicate that regional sustainability and good conditions for entrepreneurship have crucial role in obtaining that aim (Żelazny and Pietrucha, 2017; Simionescu et al., 2017). These factors are especially important for Poland, which on the one hand, is the biggest country in the region, thus, it is an economy with big potential for taking advantage of economies of scale. But on the other hand, Poland is commonly considered as the country facing the problem of regional divergence and significant regional disparities (Wójcik, 2017; Kisiała and Suszyńska, 2017; Bartkowiak-Bakun, 2017). The main aim of the article is to analyse factors influencing entrepreneurship conditions in Poland at NUTS 3 level. The conducted literature review indicates that the entrepreneurship conditions should be analysed with application of multiple-criteria analysis tools. Therefore, in the research TOPSIS method based on median vector Weber was applied.

[1] Corresponding author: University of Wamia and Mazury, Department of Microeconomics, ul. Plac Cieszyński 1, 10-719 Olsztyn, ela.n.rogalska@gmail.com.

The research was conducted for the years 2010-2015, where the main limitation for this period was the availability of data at regional level.

2 Methodology

In the case of current research taxonomic measure of development (TMD) based on TOPSIS method was applied, where the object is compared to pattern and anti-pattern of development (Balcerzak and Pietrzak, 2016; 2017). In order to be able to use the method, the phenomenon under research should be divided into economic aspects, which can be described with available diagnostic variables. The diagnostic variables are usually selected after two stages: a) preliminary selection of variables based on the experience of a researcher; b) evaluation of the diagnostic variables with application of formal taxonomic criteria. The variables should be characterised with high level of variation, high information value, which means that the variables should reach high values with relatively great difficulty and relatively low level of correlation (Balcerzak, 2016).

After obtaining the final set of diagnostic variables TMD can be assessed. For this purpose the TOPSIS method with application of median Weber (Cheba and Szopik-Depczyńska, 2017) can be used:

1. The final diagnostic variables should be normalized with application of formula 1 and 2 (Lira et al., 2002).

$$z_{ij} = \frac{x_{ij} - \theta_j}{1{,}4826 \cdot s_j} \tag{1}$$

$$s_j = \underset{i=1,2,\ldots n}{med} \left| x_{ij} - \theta_j \right| \tag{2}$$

where: $\theta = (\theta_1, \theta_2, ..., \theta_m)$ is the Weber median, s_j is the absolute median deviation, (i=1,2, ,...n) –number of the object, $(j = 1,2,...m)$ – number of the diagnostic variable.

2. Selection of pattern z_j^+ and anti-pattern z_j^- of economic development based on maximum value of the variable z_j^+ for the pattern and minimum value of the variable z_j^- for the anti-pattern in the case of stimulants and based on minimum value of the variable z_j^+ for the pattern and maximum value of the variable z_j^- for the anti-pattern in the case of des-stimulants. For dynamic research the constant pattern and anti-pattern of economic development must be taken, which is necessary for obtaining comparable results in time (Pietrzak and Balcerzak, 2016).

3. Assessment of distance from the pattern (equation 3) and anti-pattern (equation 4) with application of absolute median deviation:

$$d_i^+ = \underset{j=1,2,\ldots,m}{med} \left| z_{ij} - z_j^+ \right| \tag{3}$$

and

$$d_i^- = \underset{j=1,2,\ldots,m}{med} \left| z_{ij} - z_j^- \right| \tag{4}$$

4. Estimation of TMD with application of equation 5:

$$TMD_i = \frac{d_i^-}{d_i^- + d_i^+} \tag{5}$$

In the case of dynamic research, the stability of the obtained rankings can be analysed with application Kendall rank correlation coefficient, which can provide additional information on the potential tendencies in the case of analysed phenomenon.

3 Empirical research

The entrepreneurship conditions are formed by many long and short term factors, which can be related to institutional order of given economy and current economic policy (Bednarz et al., 2017; Pietrzak et al., 2017a). The most commonly pointed determinants of entrepreneurship conditions are the formal regulations influencing barriers for entering given markets and increasing scale of activates of enterprises, which influences competitive environment (Kruk and Waśniewska, 2017), and effectiveness of financial sector or availability of financing for enterprises(Kljucnikov and Belas, 2016; Ivanová, 2017; Balcerzak et al., 2017; Meluzin, et al., 2017; Pietrzak et al., 2017b).

As a result, the entrepreneurship conditions should be analysed with application of multiple-criteria tools. In the case of regional research – especially at lower aggregation level such as NUTS 3 region analysis, which was proposed in current article – the most important limitation for multivariate analysis is an availability of data that describes selected aspects of given phenomenon. This factor can be also attributed to current research. Therefore, in the analysis the final set of diagnostic variable given in Table 1 was applied. All the diagnostic variables were classified as stimulants. The data for the period 2010-2015 was provided by Central Statistical Office of Poland (Local Data Bank). In the first step, the standardization of the variables was carried out jointly for the entire data set for the years 2010-2015. Then in the research the methodology described in previous section was applied. The final results are given in Table 2.

The obtained rankings and the values of TMD confirm significant disparities in regard to entrepreneurial conditions at regional level. The highest positions in rankingswere obtained by the NUTS 3 dominated by the biggest municipal centres. In the case of the lowest positions one can find peripheral regions mostly located in Eastern Poland.

Table 1. The set of diagnostic variables.

Variable	Description of the variable
X_1	Number of entities included in the REGON registration per 10 thousand inhabitants
X_2	Share of commercial law companies in the number of economic entities
X_3	Share of companies with foreign capital in the total number of commercial law companies
X_4	Gross value of fixed assets in enterprises per capita
X_5	Capital expenditures in enterprises per capita

Table 2. Ranking of NUTS 3 regions in regard to entrepreneurship conditions.

NUTS 3 Region	2010		2011		2012		2013		2014		2015	
	TMD	Rank	TMD	Rank	TMD	Rank	TMD	Rank	TMD	Rank	TMD	Rank
m. Warszawa	0.893	1	0.898	64	0.881	1	0.911	1	0.868	1	0.869	1
m. Poznań	0.491	2	0.51	25	0.555	2	0.505	2	0.534	2	0.543	2
m. Wrocław	0.445	4	0.472	16	0.479	3	0.495	3	0.491	3	0.477	3
trójmiejski	0.471	3	0.472	5	0.466	4	0.465	4	0.473	4	0.473	4
m. Kraków	0.41	5	0.414	58	0.406	6	0.436	5	0.457	5	0.449	5
m.Szczecin	0.376	7	0.372	22	0.347	8	0.394	6	0.375	6	0.356	6
gliwicki	0.304	10	0.334	50	0.336	9	0.344	9	0.333	9	0.348	7
warszawskizachodni	0.374	8	0.345	62	0.347	7	0.37	8	0.364	7	0.333	8
legnicko-głogowski	0.289	12	0.31	19	0.277	11	0.304	11	0.31	10	0.327	9
opolski	0.238	16	0.244	14	0.216	18	0.236	18	0.263	15	0.323	10
katowicki	0.376	6	0.393	49	0.41	5	0.379	7	0.337	8	0.313	11
poznański	0.262	14	0.298	26	0.263	12	0.253	16	0.282	14	0.303	12
tyski	0.305	9	0.296	46	0.295	10	0.325	10	0.301	11	0.302	13

m. Łódź	0.259	15	0.272	71	0.246	14	0.274	12	0.282	13	0.291	14
wrocławski	0.291	11	0.25	17	0.26	13	0.257	14	0.288	12	0.272	15
szczeciński	0.208	21	0.21	21	0.24	15	0.253	15	0.24	16	0.23	16
bydgosko-toruński	0.271	13	0.227	13	0.207	19	0.219	21	0.221	18	0.229	17
piotrkowski	0.22	19	0.247	70	0.206	20	0.231	20	0.228	17	0.226	18
gorzowski	0.222	18	0.251	32	0.231	17	0.259	13	0.204	20	0.219	19
jeleniogórski	0.168	27	0.178	20	0.154	28	0.182	26	0.2	21	0.214	20
świecki	0.146	34	0.166	9	0.11	46	0.1	54	0.11	50	0.21	21
zielonogórski	0.205	22	0.194	31	0.193	23	0.238	17	0.195	22	0.21	22
bielski	0.227	17	0.22	53	0.235	16	0.235	19	0.207	19	0.207	23
sosnowiecki	0.162	29	0.189	47	0.166	26	0.205	23	0.186	24	0.197	24
tarnobrzeski	0.132	38	0.167	38	0.147	30	0.167	29	0.175	27	0.187	25
słupski	0.125	41	0.118	7	0.132	36	0.182	25	0.135	38	0.184	26
koniński	0.113	46	0.129	29	0.146	32	0.101	52	0.153	31	0.182	27
warszawskiwschodni	0.17	26	0.169	63	0.16	27	0.147	33	0.133	40	0.18	28
płocki	0.208	20	0.183	61	0.192	24	0.206	22	0.19	23	0.177	29
rybnicki	0.155	32	0.169	48	0.199	21	0.197	24	0.183	25	0.175	30
lubelski	0.167	28	0.177	43	0.198	22	0.179	27	0.176	26	0.171	31
koszaliński	0.179	24	0.137	24	0.149	29	0.148	32	0.159	30	0.168	32
rzeszowski	0.117	43	0.128	39	0.128	39	0.133	40	0.139	35	0.167	33
leszczyński	0.158	30	0.16	28	0.146	31	0.134	39	0.166	28	0.166	34
starogardzki	0.172	25	0.156	6	0.136	35	0.155	31	0.163	29	0.151	35
krakowski	0.086	57	0.101	59	0.114	45	0.127	43	0.143	33	0.148	36
oświęcimski	0.102	52	0.142	56	0.123	41	0.135	37	0.137	36	0.142	37
białostocki	0.117	44	0.125	37	0.121	42	0.128	42	0.131	41	0.139	38
olsztyński	0.146	35	0.144	1	0.138	34	0.141	34	0.133	39	0.139	39
częstochowski	0.179	23	0.137	51	0.129	37	0.136	35	0.144	32	0.138	40
łódzki	0.126	40	0.12	72	0.129	38	0.135	36	0.129	43	0.135	41
gdański	0.156	31	0.14	8	0.143	33	0.159	30	0.137	37	0.133	42
kaliski	0.114	45	0.119	30	0.125	40	0.135	38	0.13	42	0.13	43

wałbrzyski	0.131	39	0.168	18	0.166	25	0.168	28	0.122	47	0.128	44
pilski	0.113	47	0.109	27	0.109	48	0.109	49	0.111	49	0.127	45
bytomski	0.123	42	0.117	52	0.106	50	0.123	45	0.128	44	0.127	46
kielecki	0.138	36	0.137	34	0.12	44	0.129	41	0.125	45	0.124	47
skierniewicki	0.1	53	0.103	68	0.109	49	0.123	46	0.14	34	0.122	48
szczecinecko-pyrzycki	0.089	56	0.085	23	0.083	58	0.116	48	0.084	55	0.118	49
radomski	0.092	54	0.09	65	0.089	56	0.117	47	0.108	51	0.111	50
inowrocławski	0.102	51	0.111	10	0.1	51	0.106	51	0.121	48	0.111	51
nyski	0.136	37	0.129	15	0.091	54	0.108	50	0.098	53	0.106	52
elbląski	0.104	48	0.12	3	0.121	43	0.127	44	0.108	52	0.103	53
chojnicki	0.078	59	0.062	4	0.064	65	0.072	62	0.084	56	0.099	54
włocławski	0.147	33	0.094	11	0.109	47	0.101	53	0.125	46	0.097	55
suwalski	0.063	65	0.094	35	0.096	52	0.067	64	0.083	57	0.095	56
tarnowski	0.089	55	0.098	55	0.089	55	0.09	55	0.089	54	0.092	57
łomżyński	0.073	62	0.086	36	0.083	59	0.078	59	0.078	60	0.091	58
grudziądzki	0.102	50	0.1	12	0.085	57	0.09	56	0.082	58	0.081	59
ciechanowski	0.071	64	0.07	67	0.071	61	0.068	63	0.07	62	0.075	60
siedlecki	0.075	60	0.07	60	0.071	62	0.078	58	0.08	59	0.075	61
nowotarski	0.058	67	0.045	54	0.034	70	0.028	72	0.033	71	0.072	62
puławski	0.061	66	0.072	42	0.064	64	0.076	60	0.072	61	0.072	63
sieradzki	0.073	61	0.069	69	0.062	67	0.066	66	0.067	63	0.071	64
ełcki	0.052	69	0.068	2	0.066	63	0.046	70	0.057	66	0.067	65
ostrołęcki	0.056	68	0.06	66	0.055	69	0.066	65	0.063	64	0.066	66
nowosądecki	0.045	70	0.053	57	0.056	68	0.063	67	0.056	67	0.059	67
krośnieński	0.104	49	0.117	41	0.093	53	0.083	57	0.054	68	0.051	68
sandomiersko-jędrzejowski	0.085	58	0.103	33	0.071	60	0.063	68	0.063	65	0.045	69
- bialski	0.034	72	0.016	45	0.017	72	0.072	61	0.039	69	0.038	70
chełmsko-zamojski	0.043	71	0.036	44	0.03	71	0.03	71	0.027	72	0.028	71
przemyski	0.071	63	0.076	40	0.062	66	0.053	69	0.036	70	0.025	72

In the final step of the research the stability of the obtained ranking with application of Kendall rank correlation coefficient was verified. The result are given in Table 3.

Table 3. Kendall rank correlation coefficients for the obtained ranking in the year 2010-2015.

Years	2010	2011	2012	2013	2014	2015
2010	1	0.857	0.84	0.809	0.808	0.780
2011	0.857	1	0.882	0.840	0.826	0.815
2012	0.84	0.882	1	0.878	0.867	0.833
2013	0.809	0.84	0.878	1	0.861	0.826
2014	0.808	0.826	0.867	0.861	1	0.885
2015	0.780	0.815	0.833	0.826	0.885	1

The critical value from the normal distribution for the 5% significance level is equal to 1.960. The test statistics for the lowest value (0.780) of Kendall rank correlation coefficient is 9.693, which indicates statistical significance of all parameters presented in Table 3. Intuitively, the Kendall correlation between two ranks will be higher (close to 1), when they are similar. It can be seen that the ranking form the year 2010 becomes less similar to the rankings obtained in the next years, which shows a systematic tendency of the analysed phenomenon. It can be concluded that there are some changes in rankings of the entrepreneurship conditions of analysed regions, though the changes can be considered as relatively slow.

Conclusions

Good conditions for entrepreneurship are currently considered as one of the most important intangible factor influencing growth both at national and regional level. It is especially important in such countries as Poland that should create conditions for closing its development gap in relation to developed countries of the European Union and in the same time create good conditions for regional sustainability. As a result, in current paper the research concerning conditions for entrepreneurship at the NUTS 3 level was conducted. In the research the dynamic approach was taken. The subject of the research was considered as the multiple-criteria phenomenon, therefore TOPSIS method based on median vector Weber was used.

The conducted research confirms significant disparities in Poland at regional level in regard to entrepreneurial conditions. The disparities are also relatively stable, which confirms

that the phenomenon of unbalanced – therefore, unsustainable regional structure of economy – should be considered as a significant problem for regional policy in Poland.

The proposed research can be characterised with the following limitations. First of all, the period of the research is relatively short. The second most important critics for the provided study can relate to the selection of diagnostic variables used in the research. However, the most important determinants for both mentioned limitations are the consequence of the data availability at the NUTS 3 level. In spite of these factors the obtained results are consistent with other research in the field.

References

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

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

Balcerzak, A. P. & Pietrzak, M. B. (2016). Quality of Institutions for Knowledge-based Economy within New Institutional Economics Framework. Multiple Criteria Decision Analysis for European Countries in the Years 2000–2013. *Economics & Sociology*, *9*(4), 66-81.

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

Bartkowiak-Bakun, N. (2017). The diversity of socioeconomic development of rural areas in Poland in The Western Borderland and the problem of post-state farm localities. *Oeconomia Copernicana*, *8*(3), 417–431.

Bednarz, J., Bieliński, T., Nikodemska-Wolowik, A. M. & Otukoya, A. (2017). Sources of the Competitive Advantage of Family Enterprises: An International Approach Focusing on China, Nigeria and Poland. *Entrepreneurial Business and Economics Review*, *5*(2), 123-142.

Cheba, K. & Szopik-Depczyńska, K. (2017). Multidimensional comparative analysis of the competitive capacity of the European Union countries and geographical regions. *Oeconomia Copernicana*,*8*(4), 487–504.

Ivanová, E. (2017). Barriers to the development of SMEs in the Slovak Republic. *Oeconomia Copernicana*, *8*(2), 255–272.

Kisiała, W. & Suszyńska, K. (2017). Economic growth and disparities: an empirical analysis for the Central and Eastern European countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(4).

Kljucnikov, A. & Belas, J. (2016). Approaches of Czech entrepreneurs to debt financing and management of credit risk. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 11*(2), 343-365.

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

Lira, J., Wagner, W. & Wysocki, F. (2002). Mediana w zagadnieniach porządkowania liniowego obiektów wielocechowych. In: Paradysz, J. (ed.), *Statystyka regionalna w służbie samorządu lokalnego i biznesu.* Poznań: Internetowa Oficyna Wydawnicza, Centrum Statystyki Regionalnej, Akademia Ekonomiczna w Poznaniu, 87-99.

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

Pietrzak, M. B. & Balcerzak, A. P. (2016).Assessment of Socio-Economic Sustainability in New European Union Members States in the years 2004-2012. In: Papież, M. & Śmiech S. (eds.). *The 10th Professor AleksanderZelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena*. *Conference Proceedings*. Cracow: Foundation of the Cracow University of Economics, 120-129.

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

Pietrzak, M. B., Fałdziński, M., Balcerzak, A. P., Meluzín, T. & Zinecker, M. (2017b). Short-term Shocks and Long-term Relationships of Interdependencies Among Central European Capital Markets. *Economics & Sociology, 10*(1), 61-77.

Simionescu, M., Lazányi, K., Sopková, G., Dobeš, K. & Balcerzak, A. P. (2017). Determinants of Economic Growth in V4 Countries and Romania. *Journal of Competitiveness, 9*(1), 103-113.

Wójcik, P. (2017). Was Poland the next Spain? Parallel analysis of regional convergence patterns after accession to the European Union. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(4).

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

Application of the AG mean index for CPI substitution bias reduction

Elżbieta Roszko-Wójtowicz[1], Jacek Białek[2]

Abstract

The Consumer Price Index (CPI) approximates changes in the costs of household consumption assuming the constant utility (COLI, Cost of Living Index). In practice, the Laspeyres price index is used to measure the CPI despite the fact that many economists consider the superlative indices to be the best approximation of COLI. The Fisher index is one of the superlative indices and, additionally, it satisfies most of tests from the axiomatic price index theory. Nevertheless, the Fisher price index makes use of current-period expenditure data and its usefulness in CPI measurement is limited. Lent and Dorfman (2009) show that a weighted average of the arithmetic and geometric Laspeyres indices (the so-called AG mean index) used as a proxy for the Fisher formula can provide a simple alternative to the Lloyd-Moulton index. To use the AG mean index in practice we have to approximate the right parameter being in the index's body. Theoretically, the parameter should not change rapidly over time since it denotes the elasticity of substitution. In the paper we apply the AG mean index for the Fisher index approximation using CPI data from the United Kingdom and Bulgaria. The main aim of the paper is to examine fluctuations in the estimated parameter and its dependence on the level of data aggregation.

Keywords: *CPI, COLI, the Fisher index, the Laspeyres index, the AG Mean index*
JEL Classification: C43
DOI: 10.14659/SEMF.2018.01.41

1 Introduction

The Consumer Price Index (CPI) is commonly used as a basic measure of inflation. The index approximates changes in the costs of household consumption assuming the constant utility (COLI, Cost of Living Index). In practice, the Laspeyres price index is used to measure the CPI (White, 1999; Clements and Izan, 1987), although the Laspeyres formula does not take into account changes in the structure of consumption, which occur as a result of price changes in the given time interval. Many economists consider the superlative indices (like the Fisher index or the Törnqvist index) to be the best approximation of COLI (Von der Lippe, 2007). The difference between the Laspeyres index and the superlative index should approximate the value of the commodity substitution bias. However, there are some other ways to reduce that bias, like using the Lloyd–Moulton price index (see Lloyd, 1975; Moulton, 1996; Shapiro and

[1] Corresponding author: University of Lodz, Department of Economic and Social Statistics, ul. Rewolucji 1905 r. nr 41, 90-214 Lodz, email: eroszko33@gmail.com.
[2] University of Lodz, Department of Statistical Methods, ul. Rewolucji 1905 r. nr 41, 90-214 Lodz, email: jbialek@uni.lodz.pl.

Wilcox, 1997; Białek, 2014), the AG Mean index (Lent and Dorfman, 2009; Białek, 2017) or the Lowe and Young indices (Armknetch and Silver, 2012).

In the paper we focus on application of the AG mean index for CPI substitution bias reduction. To use the AG mean index in practice one needs to approximate the right parameter being in the index's body. Theoretically, the parameter should not change rapidly over time since it denotes the elasticity of substitution. In the paper we apply the AG mean index for the Fisher index approximation using CPI data from the United Kingdom and Bulgaria[3]. The main aim of the paper is to examine fluctuations in the estimated parameter and its dependence on the level of data aggregation.

2 Cost of Living Index (COLI) vs Consumer Price Index (CPI)

Let $E(P,\bar{u}) = \min_{Q}\{P^T Q | U(Q) \geq \bar{u}\}$ be the expenditure function of a representative consumer which is dual to the utility function $U(Q)$. In other words, it is the minimum expenditure necessary to achieve a reference level of utility $\bar{u}$ at the vector of prices P. Then, the Konüs cost of living price index is defined as

$$P_K = \frac{E(P^t,\bar{u})}{E(P^s,\bar{u})}, \tag{1}$$

where t denotes the current period, s denotes the base period, and in general, the vector of N considered prices at any moment τ is given by $P^\tau = [p_1^\tau, p_2^\tau, ..., p_N^\tau]^T$. I_K is the true cost of living index in which the commodity Q changes as the vector of prices facing the consumer changes. The CPI, in contrast, measures the change in the cost of purchasing a fixed basket of goods at a fixed sample of outlets over a time interval, *i.e.* $Q^s = [q_1^s, q_2^s, ..., q_N^s]^T = Q^t$. The CPI is a Laspeyres-type index being the weighted arithmetic mean of price relatives

$$P_{La} = \frac{\sum_{i=1}^{N} q_i^s p_i^t}{\sum_{i=1}^{N} q_i^s p_i^s} = \sum_{i=1}^{N} w_i^s \frac{p_i^t}{p_i^s}, \tag{2}$$

so we assume here the constant consumption vector on the base period level. It can be shown (Diewert, 1993) that

[3] Currently there are no differences between the CPI and HICP (Harmonized Index of Consumer Prices) indices in the case of these countries. In our research we use HICP data from Eurostat from COICOP-4 digit and COICOP-3 digit levels of aggregation.

$$P_K = \frac{E(P^t, U(Q^s))}{E(P^s, U(Q^s))} \leq P_{La}, \tag{3}$$

so $P_{La} - P_K$ is the extent of the commodity substitution bias.

3 Proxies for the Fisher price index

In the so-called economic price index approach many authors use superlative price indices to approximate the P_K index (White, 1999). The most popular superlative index is the Fisher index P_F (Fisher, 1922), which can be written as the geometric mean of the Laspeyres (P_{La}) and Paasche (P_{Pa}) price indices, i.e.

$$P_F = \sqrt{P_{La} P_{Pa}}\,. \tag{4}$$

where

$$P_{Pa} = \frac{\sum_{i=1}^{N} q_i^t p_i^t}{\sum_{i=1}^{N} q_i^t p_i^s}. \tag{5}$$

Nevertheless, the Fisher price index makes use of current-period expenditure data and its usefulness in CPI measurement is limited. Using a Constant Elasticity of Substitution (CES) framework, a superlative Fisher price index can be approximated once the elasticity of substitution (σ) is estimated. The Lloyd–Moulton price index (Lloyd, 1975; Moulton, 1996; Shapiro and Wilcox, 1997) does not make use of current-period expenditure data, and it follows the formula

$$P_{LM}(\sigma) = \{[\sum_{i=1}^{N} w_i^s (\frac{p_i^t}{p_i^s})^{1-\sigma}]^{\frac{1}{1-\sigma}}, \tag{6}$$

where σ is some real parameter and w_i^s denotes the expenditure share of commodity i in the base period s. Unfortunately, there is no analytical solution for the equation $P_F = P_{LM}(\sigma)$ with respect to σ and thus this parameter can be calculated only numerically (Feenstra and Reinsdorf, 2007; Biggeri and Ferrari, 2010; Greenlees, 2011; Armknecht and Silver, 2012; Białek, 2017).

Lent and Dorfman (2009) prove that a weighted average of the Laspeyres index and the geometric Laspeyres index can approximate the Lloyd-Moulton index and thus it also approximates the superlative target index. In particular, the above-mentioned weighted average (called the AG Mean index) provides a close approximation to the Fisher price index, namely

$$P_F \approx P_{AG} = \sigma \prod_{i=1}^{N} \left(\frac{p_i^t}{p_i^s}\right)^{w_i^s} + (1-\sigma) \sum_{i=1}^{N} w_i^s \left(\frac{p_i^t}{p_i^s}\right). \quad (7)$$

Solving the equation $P_F = P_{AG}$ with respect to σ we obtain (Armknecht and Silver, 2012)

$$\hat{\sigma} = \frac{P_F - P_{La}}{P_{GLa} - P_{La}}, \quad (8)$$

where P_{GLa} denotes the geometric Laspeyres price index (Von der Lippe, 2007)

$$P_{GLa} = \prod_{i=1}^{N} \left(\frac{p_i^t}{p_i^s}\right)^{w_i^s}. \quad (9)$$

Since the parameter should not change rapidly over time, we can estimate it using historical data. Apparently, we should have a good tool for approximation of the current value of the Fisher price index. In the empirical study we verify the utility of the AG Mean index in CPI substitution bias reduction using CPI (HICP) data from the United Kingdom and Bulgaria.

4 Empirical study

In the following section we apply the AG Mean index for the Fisher index approximation using CPI data from the United Kingdom and Bulgaria. Currently there are no differences between the CPI and HICP (Harmonized Index of Consumer Prices) indices in the case of these countries. Thus, we use yearly data from Eurostat from COICOP-4 digit and COICOP-3 digit levels of aggregation, and we calculate main price indices and also the parameter of the elasticity of substitution (8) for years 2011-2016.

Table 1 and Table 2 present results for Bulgaria, whereas Tab. 3 and Tab. 4 concentrate on the United Kingdom.

Table 1. Price indices and the estimated parameter of the elasticity of substitution.

Index	**Bulgaria (COICOP 3-digit)**					
formula	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
P_{La}	1.03645	1.02033	1.00396	0.98295	0.98841	0.98592
P_{GLa}	1.03546	1.01960	1.00365	0.98256	0.98738	0.98540
P_{Pa}	1.03393	1.02381	1.00392	0.98389	0.99047	0.98730
P_F	1.03519	1.02207	1.00394	0.98342	0.98944	0.98661
$\hat{\sigma}$	1.26854	-2.39129	0.06734	-1.22779	-0.99247	-1.32732

Notes: Bulgaria, years: 2011-2016, data from COICOP 3-digit level of aggregation.

Table 2. Price indices and the estimated parameter of the elasticity of substitution.

Index formula	Bulgaria (COICOP 4-digit)					
	2011	**2012**	**2013**	**2014**	**2015**	**2016**
P_{La}	1.04385	1.02504	1.00469	0.98087	0.98481	0.98419
P_{GLa}	1.04205	1.02411	1.00419	0.98019	0.98304	0.98294
P_{Pa}	1.03839	1.02772	1.00536	0.98119	0.98797	0.98581
P_F	1.04111	1.02638	1.00503	0.98103	0.98639	0.98500
$\hat{\sigma}$	1.51943	-1.44693	-0.66349	-0.23119	-0.89077	-0.65484

Notes: Bulgaria, years: 2011-2016, data from COICOP 4-digit level of aggregation.

Table 3. Price indices and the estimated parameter of the elasticity of substitution.

Index formula	United Kingdom (COICOP 3-digit)					
	2011	**2012**	**2013**	**2014**	**2015**	**2016**
P_{La}	1.04425	1.02648	1.02237	1.01129	0.99802	1.00485
P_{GLa}	1.04339	1.02609	1.02207	1.01104	0.99759	1.00456
P_{Pa}	1.04417	1.02732	1.02227	1.01222	0.99839	1.00581
P_F	1.04421	1.02690	1.02232	1.01175	0.99821	1.00533
$\hat{\sigma}$	0.04953	-1.07063	0.17058	-1.90674	-0.42942	-1.67218

Notes: United Kingdom, years: 2011-2016, data from COICOP 3-digit level of aggregation.

Fig. 1 compares $\hat{\sigma}$ estimates calculated for Bulgaria and the United Kingdom for each year of observation and depending on the level of aggregation (COICOP). Our results show that the application of the AG Mean index for CPI substitution bias reduction (or, equivalently, for the Fisher price index approximation) is limited in practice.

Table 4. Price indices and the estimated parameter of the elasticity of substitution.

Index formula	United Kingdom (COICOP 4-digit)					
	2011	**2012**	**2013**	**2014**	**2015**	**2016**
P_{La}	1.04591	1.02164	1.02001	1.00813	0.98914	0.99841
P_{GLa}	1.04473	1.02085	1.01954	1.00777	0.98817	0.99797
P_{Pa}	1.04557	1.02311	1.01990	1.00858	0.98978	0.99926
P_F	1.04574	1.02237	1.01996	1.00835	0.98946	0.99884
$\hat{\sigma}$	0.14339	-0.93094	0.10857	-0.63530	-0.33362	-0.98098

Notes: United Kingdom, years: 2011-2016, data from COICOP 4-digit level of aggregation.

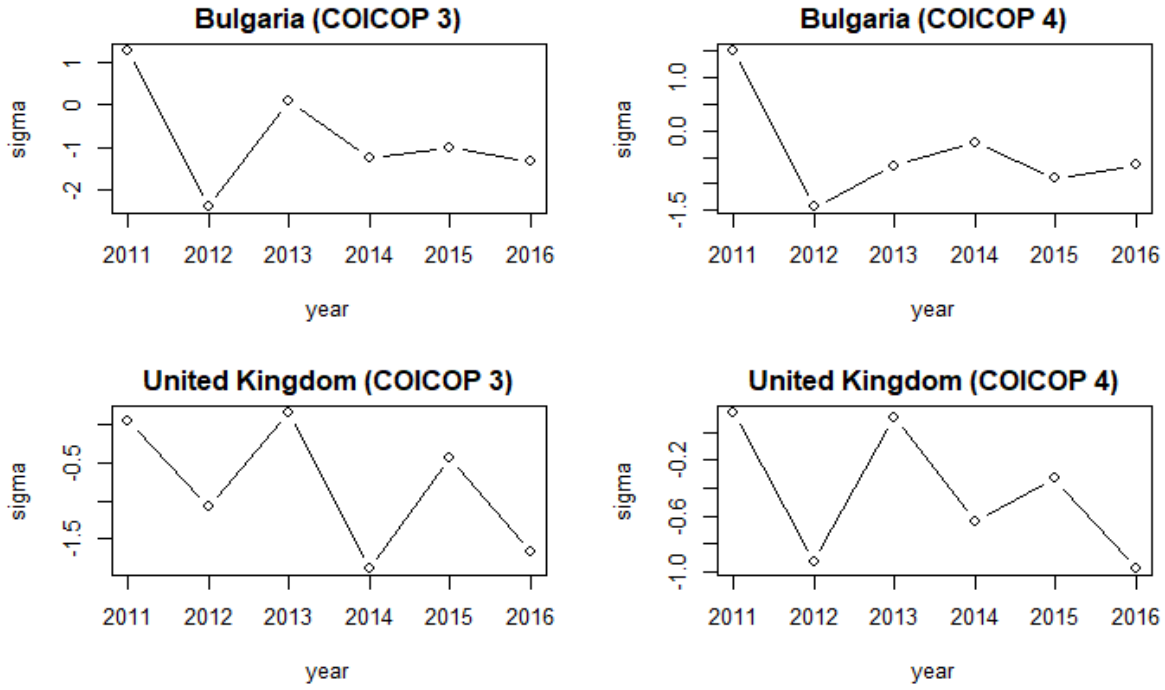


Fig. 1. Estimates $\hat{\sigma}$ for Bulgaria and the United Kingdom for years 2011-2016.

Conclusions

In our opinion the usefulness of the AG Mean index as a proxy for the ideal Fisher price index is very limited in practice. From a theoretical point of view, the AG Mean index seems to be a great alternative to the Laspeyres price index in inflation (CPI) measurement, since it approximates the Fisher index (considered the best approximation of Cost of Living Index) and it does not need the current-period expenditure data. Nevertheless, estimation of the parameter σ, which reflects the elasticity of substitution, has some practical drawbacks.

Firstly, its estimator given by (8) do need the current-period expenditure data. Obviously, we could change the method of estimation of σ using some econometric approach or numerical methods for the Lloyd-Moulton index, but the usefulness of any estimations is still bounded since this parameter strongly fluctuates in time (see Fig. 1). Thus, the computed value of this parameter for the year $t-1$ as a rule is completely different from its value calculated for the current year t (see Fig. 1 and Table 1 – 4). Secondly, the considered parameter seems to strongly depend on the level of aggregation (compare Table 1 with Table 2 and Table 3 with Table 4). The main conclusion from these observations is that within the framework considered in this paper it is impossible to provide a single, general value of the parameter σ or to propose some narrow interval of its possible values. Thus, the application of the AG Mean index for CPI substitution bias reduction is doubtful in practice.

Acknowledgements

Authors of the paper would like to thank the National Science Centre in Poland for financing this publication (grant no. 2017/25/B/HS4/00387).

References

Armknecht, P. & Silver, M. (2012). Post-Laspeyres: The Case for a New Formula for Compiling Consumer Price Indexes. *International Monetary Fund (IMF) Working Paper*.

Białek, J. (2014). Obciążenie wskaźnika cen towarów i usług konsumpcyjnych wynikające z substytucji dóbr. *Wiadomości Statystyczne, 10*, 30-38.

Białek, J. (2017). Approximation of the Fisher price index by using Lowe, Young and AG Mean indices. *Communications in Statistics – Simulation and Computation, 46*(8), 6454-6467.

Biggeri, L. & Ferrari, G. (eds.) (2010). *Price Indexes in Time and Space*. Springer-Verlag.

Clements, K. W. & Izan, H. Y. (1987). The Measurement of Inflation: A Stochastic Approach. *Journal of Business and Economic Statistics, 5*, 339-350.

Diewert, W. E. (1993). The economic theory of index numbers: a survey. In: Diewert, W. E. & Nakamura, A. O. (eds.), *Essays in index number theory, 1*, 177-221.

Feenstra, R. C. & Reinsdorf, M. B. (2007). Should Exact Index Numbers Have Standard Errors? Theory and Application to Asian Growth. In: Berndt, E. R. & Hulten, C. R. (eds.), *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches, 10*, 483 – 513.

Fisher, I. (1922). *The Making of Index Numbers.* Houghton Mifflin.

Greenlees, J. (2011). Improving the preliminary values of the chained CPI-U. *Journal of Economic and Social Measurement, 36,* 1–18.

Lent, J & Dorfman, A. H. (2009). Using a Weighted Average of Base Period Price Indexes to Approximate a Superlative Index. *Journal of Official Statistics, 25* (1), 139-149.

Lloyd, P. J. (1975). Substitution Effects and Biases in Nontrue Price Indices. *The American Economic Review, 65*(3), 301–313.

Moulton, B. R. (1996). *Constant Elasticity Cost-of-Living Index in Share-Relative Form*. Washington DC: Bureau of Labour Statistics.

Shapiro, M. D. & Wilcox, D. W. (1997). Alternative Strategies for Aggregating Prices in the CPI. *Federal Reserve Bank of St. Louis Review, 79*(3), 113–125.

White, A. G. (1999). Measurement Biases in Consumer Price Indexes. *International Statistical Review, 67*(3), 301-325.

Von der Lippe, P. (2007). *Index Theory and Price Statistics*. Peter Lang.

A note on the private housing rental market in Poland

Michał Rubaszek[1]

Abstract

Housing rental market share in Poland is low. I explore the reasons behind this underdevelopment with a survey conducted among SGH Warsaw School of Economics students. It turns out that tenure preferences of respondents are tilted towards owning, which can be attributed predominantly to psychological factors. Using a life-cycle model I evaluate the effect of the reforms aimed at improving the functioning of the rental market by increasing the quality of rental services, reducing the risk of investment in rental housing and removing fiscal incentives for owning. The results indicate that the reforms, if introduced simultaneously, would significantly increase the rental market share.

Keywords: *Housing rental market, survey data, counterfactual simulations*
JEL Classification: D91, E21, R21
DOI: 10.14659/SEMF.2018.01.42

1 Introduction

A safe place to live is among basic needs of households. The most popular form of satisfying housing needs is ownership. In this case the house serves a dual purpose: it provides utility and is an investment vehicle that allows for storing value. The alternative housing tenure is renting. It allows to separate the dual role of housing: the tenant derives utility from housing services and the landlord obtains profits from housing investment. The literature indicates that the tenure structure of the rental market has an important impact on the macroeconomic outcome. Some authors show that the availability of rental housing diminish the variability of the housing sector and reduce the risk of a house price bubble (Arce and Lopez-Salido, 2011; Cuerpo et al., 2014; Rubaszek and Rubio,2017), whereas other authors argue that a well-functioning rental market enhances residential mobility and limits long-run unemployment (Blanchflower and Oswald, 2013; Lisi, 2016). On the contrary, there is a strand of literature that claims that home owners are more likely to invest in local community and social capital (DiPasquale and Glaeser, 1999).

Given the above discussion in the literature, in this study I argue that housing rental market underdevelopment in Poland should be regarded as a structural weakness. The scale of this underdevelopment is well illustrated by Eurostat data on the housing tenure structure. The detailed data for Poland in 2015 are as follows: "market price" tenants (4.5%), "reduced

[1] Corresponding author: SGH Warsaw School of Economics, Econometrics Institute, al. Niepodległości 162, 02-554 Warsaw, Poland, michal.rubaszek@sgh.waw.pl.

price" tenant (11.8%), owners without mortgage (72.8%) and owners with mortgage (10.9%). These figures compare to the share of "market price" tenants amounting to 15.4% in Italy, 19.8% in France, 39.9% in Germany and 49.2% in Switzerland. In the further part of the article I explain why renting is unpopular in Poland using the results of a survey on housing tenure preferences conducted among 195 students of the SGH Warsaw School of Economics. Then I will briefly discuss what kind of policies would help in developing the private rental market, using the results of model simulations described in detail in Rubaszek (2017).

2 The reasons behind rental market underdevelopment in Poland

The high proportion of owners and a marginal share of tenants observed in Poland has several key reasons. As indicated by Lux and Sunega (2014) a very important factor behind a declining trend in the rental market share was the transfer of public rental housing into private hands, which took the form of a massive sale to sitting tenants. This is well illustrated by the Eurostat data, according to which the share of public rental in Poland decreased from 34.9% in 2007 to just 11.8% in 2015. The second factor is related to the development of the mortgage market. A steady decrease of nominal interest rates, better access to FX denominated loans and two programs enhancing house purchases on credit ("Rodzina na Swoim" and "Mieszkanie dla Młodych") have led to an increase in the proportion of owners with a mortgage from 2.9% in 2007 to 10.9% in 2015. The next reason is related to ineffective regulations, excessive protection of "bad tenants" within the regular rent contract for instance, and the lack of consistent housing policy to support the rental market. This is aptly summarized by Priemus and Mandic (2000), who claim (as indicated by the title of their article) that in the countries of the region both private and public rental market at the beginning of the twenty-first century was "no man's land." The last factor behind high ownership rate is that Poles strongly prefer ownership to renting due to psychological factors. As indicated by Rubaszek and Czerniak (2017) owning is perceived as the only option to "feel at home".

The survey presented in this paper builds upon my recent work with Adam Czerniak (Rubaszek and Czerniak, 2017), in which we describe the results of the survey conducted by IPSOS in June 2016 among a representative group of 1005 Poles. In this article I extend these results by conducting the same survey among 195 students from SGH Warsaw School of Economics (112 respondents from Jan. 2017 and 83 respondents from Jan. 2018).

Table 1. Characteristics of the respondents.

Question	Fraction of respondents
Place of residence during childhood	
Warsaw	31.8%
Other city	68.2%
Housing tenure status	
Owner	36.4%
Tenant	39.0%
SGH campus	4.6%
living with parents	20.0%
Duration of stay in the current dwelling	
below 1 year	18.5%
from 1 to 5 years	49.7%
above 5 years	31.8%
Plans to change the dwelling	
yes, before 5 years	81.0%
yes, after 5 years	8.2%
No	6.7%
The most likely choice in case of moving	
Buying	28.7%
Renting	65.2%
A sentence closer to your opinion	
Buying makes more sense as it is a good investment	72.3%
Renting is better as it enables financial liquidity	27.7%

I believe that analyzing the answers of WSE students provide value added in comparison to the results based on the representative sample for at least three reasons. First, understanding the perception of young adults towards housing tenures is important as the availability of stable rental housing has a tremendous impact on their strategies how to start a life of their own. This relates not only to the satisfaction from utilized housing, but also to professional career development or family formation. Second, WSE students are more mobile and more often leave in rented apartments in comparison to the rest of population, hence their opinion on the functioning of the rental market is more informative than the opinion of respondents who have always been homeowners. Third, WSE students are relatively fluent in economics,

hence should be better prepared to evaluate the relative merits of owning and renting. However, given that the sample is limited to young adults with economic educational attainment, the results of the survey should be treated with some reservation.

Let us focus on the results of the survey. Table 1 describes selected characteristics of the participating students. It turns out that about one third of them is from Warsaw and two thirds from another city or country. About half of them live in a rented apartment or in the university campus, whereas the other half cohabit with parents or is a homeowner. The majority of the respondents plan to change the place of residence in the short term horizon, where the most likely choice will be renting. This choice is however at odds with the general belief about the pros and cons of both tenure alternatives. About three quarters of students believes that *buying a house is a good investment decision over the life cycle*, whereas only one quarter of them selected that *renting makes more sense because it enables flexibility and financial liquidity and is a better deal than owning.*

Table 2. Preferred housing tenure.

	buy	**no opinion**	**rent**
economic factors			
mortgage vs. rent costs	40.5%	16.4%	43.1%
house price vs. rent fluctuations	54.4%	25.1%	20.5%
psychological factors			
social status	70.3%	23.5%	6.2%
freedom and independence	44.6%	12.9%	42.5%
comfort	80.5%	10.8%	8.7%
family	78.0%	13.3%	8.7%

The next set of questions was related to a detailed reasons behind housing tenure choices. Following the literature, I have focused on both financial and psychological factors (Coolen et al., 2002; Ben-Shahar, 2007; Bourassa and Hoesli, 2010). The upper panel of Table 2 shows that the level of mortgage installments in considered to be higher than the level of rents for 40.5% of the respondents, whereas 43.1% of them is of the opposite opinion. The distribution of answers is more tilted towards owning if the risk of rent increases is compared to the risk of house price decline. In this case 54.4% of respondents prefers to buy and only 20.5% to rent. This might be influenced by the common believe that buying a house is a good

investment over the life cycle. The bottom panel of Table 2 shows that due to psychological factors renting might lose as a viable alternative to ownership as a long-term solution to satisfy housing needs. The big majority of the respondents think that owning raises social status, improves comfort and satisfaction from housing services as well as provides a good and secure place for a family. On the contrary, for a large group of students renting allows for mobility, which has a positive impact on freedom and independence.

In the final set of questions I checked which factors, in the opinion of SGH students, are decreasing the attractiveness of being a tenant. The results in Table 3 show that constraints in arranging apartment and inspections by landlords diminish the comfort of living in a rented apartment. This might explain why psychological factors are tilted so much towards owning. In turn, the risk of rent increase as well as high level of rents are important in explaining why so many students believe that buying a house is a better financial decision than renting it. Finally, the popularity of renting is diminished by inefficient regulations related to tenant protection. In particular, the restrictive eviction regulations within regular rental contracts cause that most of the newly-signed contracts are temporary, with usual duration of one year. Consequently, renting is treated as a short-term method of satisfying housing needs rather than as a long-term solution.

Table 3. Factors decreasing the attractiveness of renting.

	yes	**no opinion**	**no**
constrains in arranging apartment	68.2%	21.5%	10.3%
no protection against rent increases	63.6%	22.1%	13.4%
often inspections by landlords	50.8%	30.8%	18.4%
high rents	48.2%	39.5%	12.3%
no protection against eviction	44.6%	37.4%	18.0%
inadequate offer of houses to let	45.1%	38.0%	16.9%

3 Model simulations

So far I have shown that the share of private market tenants in Poland is low and presented a series of arguments explaining the reasons behind this outcome. In this section I will present a more formal discussion on the impact of economic and psychological factors on rental market share. For this purpose I discuss the results of counterfactual simulations conducted with a life-cycle, heterogenous agent model that I describe in details in Rubaszek (2017). In short,

the model economy is inhabited by many households, which differ by age, income, financial and housing assets. They derive utility from consumption of non-housing goods and housing services, which might be satisfied by owning or renting. The model includes several features important for the functioning of the housing market such as: (i.) taxes and subsidies, (ii.) mortgage interest rate spread and minimum down-payment constraint, (iii.) disutility of renting and (iv.) higher depreciation rate of rented dwellings than the owned ones. I show that these features have important impact on tenure decisions by calculating how housing market reform changes the share of the private rental market share.

The main reasons why in under the benchmark parametrization of the model households prefer to own rather to rent are as follows. First, the utility derived from living in rented house is lower than in the same house that is owned. The utility loss is assumed to be 15% and can be attributed to psychological factors described in the previous section. Second, the depreciation rate of rented apartments is higher than that for the owned ones (by 1 pp. per year). I interpret this difference as a reward for the risk of letting an apartment to a "bad tenant", who cannot be evicted, combined with potential higher utilization rate of housing stock by tenants in comparison to owners (Sinai and Souleles, 2005). Third, there are taxes on income from renting (8.5%) and no taxes on imputed rents. Additionally I assume that the government subsidises interest payments on mortgages (a proxy for programs "Rodzina na Swoim" and "Mieszkanie dla Młodych"). The last two factors raise the relative cost of renting in comparison to owning: each year tenants have to pay about 2% of the house value more than owners for living in the same dwelling.[2] There are also reasons why households rent: transaction costs, the spread on the mortgage rate and maximum loan-to-value restriction. If a house is bought with a mortgage then the economic advantage to own diminishes from 2% to 0.4%, whereas the transaction costs cause that the size of the owned house might be inadequate to households needs.

I apply the proposed model to evaluate the results of selected changes in the functioning of the rental market. In particular, I focus on the following reforms:

1. Increasing the quality of renting services. In the model economy the disutility of renting is lowered from 15% to merely 5%, i.e. levels observed in European countries with well-functioning rental markets (Diaz-Serrano, 2009).

[2]According to the data collected by Global Property Guide gross rental yields in Poland are relatively high and amount to 5.5% per year, see http://www.globalpropertyguide.com/Europe/rent-yields.

2. Increasing the protection of landlords against bad tenants or introducing policies subsidizing build-to-let investment. In the model economy the depreciation rate of rented houses declines by 1pp, i.e. to the level for the owned houses.
3. Removing the tax on income from renting and mortgage rate subsidies.

Table 4. The effects of the rental market reform.

	no reform	**1**	**2**	**3**	**full**
fraction of tenants (%)	9.6	12.7	17.4	14.2	35.4
average age of first house purchase	28.0	29.1	30.6	30.1	37.9
fraction of households with debt (%)	20.0	17.8	16.4	16.0	4.6
mortgage debt to GDP (%)	40.2	37.2	36.2	34.4	12.5

Table 4 presents the aggregate effects of the reforms. The first row of the table shows that in the benchmark economy everyone tries to buy a house as quick as possible and only credit constrained households cause that some of the youngest households decide to rent (9.6% of all households). A single reform helps in increasing the rental market share, but the gains are not large: 3.1 pp., 7.8 pp and 4.6 pp for reforms 1, 2 and 3, respectively. However, given that there are interaction effects in the model, if the three reforms are introduced together, the share of tenants in the economy rises by enormous 25.8 pp to 35.4%. This shows that to make renting an attractive alternative to owning policymakers need to remove or alleviate all barriers that make renting unattractive. The next rows of the table indicate that the full reform also switches the average moment of the first house purchase by almost 10 years. Moreover, in the reformed economy most of house purchases is financed from savings rather than by mortgages. As a result the mortgage debt to GDP ratio decreases from 40.2% to merely 12.5%.This makes the financial sector more stable and the economy less susceptible to financial shocks. It should be noted, however, that in the model it is assumed that landlords finance rental housing purchases from their savings. This describes well the situation in Germany, where the majority of the rental housing is owned by investment and pension funds. However, it is also possible that landlords finance rental housing purchase with “buy-to-let” mortgages, which are very popular in the UK for instance. In this case the effect of the rental market reform on the stability of the financial sector is much lower.

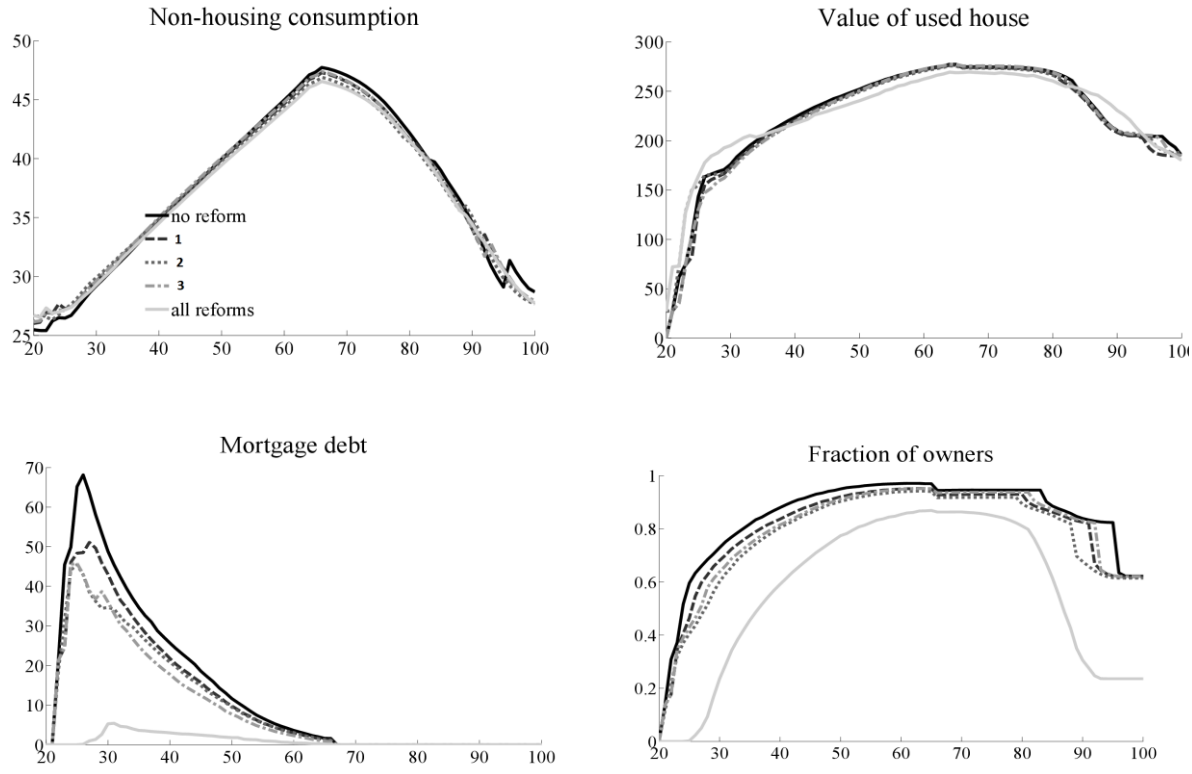


Fig. 1. The effects of rental market reform by age cohorts.

The effects of the rental market reform on the decisions taken during the lifecycle are presented in Fig. 1. The upper panels of the figure show that the average paths of spending on consumption and housing services are little affected by the reform. In contrast, the bottom panels demonstrate that the impact of the full reform on the life-time path of mortgage debt and tenure structure is sizeable. The reform is strongly limiting the demand of households to take a mortgage in the early stage of life as they now satisfy their demand for housing services by renting.

Conclusions

The share of the rental housing market in Poland is low. Using the survey conducted among SGH Warsaw School of Economics students I have showed that housing tenurepreferences are tilted towards owning,predominantly due to psychological reasons. The respondents perceive ownership not only as a cheaper form of satisfying housing needs, but also as the only way to provide a safe place for the family and to really “feel at home”. The survey also identifies the most important barriers to demand for and supply of rental housing. Among the former, inefficient institutions and the lack of professional renting services turned out to be of

crucial importance. Given the above diagnosis, I have used a life-cycle model to quantify the effects of several rental market reforms. Simulation results indicate that a complete reform would shift the rental share from below 10% in the benchmark economy to about 35% in the reformed one. The additional result of the reform is a more stable financial sector, as the household debt to GDP ratio decreases substantially.

Based on the results of this study one might formulate a set of recommendations for housing policy. First of all, lowering the relative cost of renting in comparison to owning is very important to develop private rental market. This can be achieved by implementing regulations limiting the risk associated with investing in rental housing, eliminating fiscal measures promoting ownership or even introducing rent subsidies. Second, to develop the rental market housing policy should also take into account non-financial factors, which might decrease the satisfaction from being a tenant. In the longer horizons, stimulating the professionalization of rental services and smart regulations should contribute to changing psychological attitudes towards renting.

Acknowledgements

This project was financed by the National Science Centre grant No. 2014/15/B/HS4/01382.

References

Arce, O. & Lopez-Salido, D. (2011). Housing bubbles. *American Economic Journal: Macroeconomics*, *3*(1), 212–41.

Ben-Shahar, D. (2007). Tenure choice in the housing market: Psychological versus economicfactors. *Environment and Behavior*, *39(*6), 841–858.

Blanchflower, D. G. & Oswald, A.J. (2013). *Does High Home-Ownership Impair the LaborMarket?* NBER Working Papers 19079, National Bureau of Economic Research, Inc.

Bourassa, S. & Hoesli, M. (2010). Why Do the Swiss Rent? *The Journal of Real Estate Financeand Economics*, *40*(3), 286–309.

Coolen, H., Boelhouwer, P. & Driel, K.V. (2002). Values and goals as determinants of intended tenure choice. *Journal of Housing and the Built Environment*, 17(3), 215–236.

Cuerpo, C., Pontuch, P. & Kalantaryan, S. (2014). Rental market regulation in the EuropeanUnion.*European Economy Economic Papers 515*, European Commission.

Diaz-Serrano, L. (2009). Disentangling the housing satisfaction puzzle: Does homeownership eally matter? *Journal of Economic Psychology*, *30*(5), 745 – 755.

DiPasquale, D. & Glaeser, E. L. (1999). Incentives and social capital: Are homeowners better citizens? *Journal of Urban Economics*, *45*(2), 354–384.

Lisi, G. (2016). Mortgage Market, Housing Tenure Choice and Unemployment. *The Journal of Real Estate Finance and Economics*, *53*(4), 472–493.

Lux, M. & Sunega, P. (2014). Public Housing in the Post-Socialist States of Central and Eastern Europe: Decline and an Open Future.*Housing Studies29*(4), 501-519.

Priemus, H. & Mandic, S. (2000). Rental housing in Central and Eastern Europe as no man's land. *Journal of Housing and the Built Environment*, *15*(3), 205–215.

Rubaszek, M.(2017). *Reforming housing rental market in a life-cycle model*,WP2017-028, WSE Collegium of Economic Analysis.

Rubaszek, M. & Czerniak, A. (2017). Preferencje Polaków dotyczące struktury własnościowej mieszkań: opis wyników ankiety, *Bank i Kredyt48*(2), 197-234.

Rubaszek, M. & Rubio, M. (2017). *Does rental housing market stabilize the economy? Amicro and macro perspective*. Discussion Papers 2017/06, University of Nottingham, Centre for Finance, Credit and Macroeconomics (CFCM).

Sinai, T. & Souleles, N.S. (2005). Owner-occupied housing as a hedge against rent risk. *Quarterly Journal of Economics120*(2), 763-789.

An Application of conjoint analysis to study determinants of Polish direct investment located in European countries

Marcin Salamaga[1]

Abstract

The purpose of this article is to identify main factors determining the choice of the destination of Polish FDI (foreign direct investments) in Europe. The research includes cost factors, market factors, efficiency factors, law factors, social factors, and political factors that may be important in the search for the beneficiaries of FDI. The basic research method used in the paper is the conjoint analysis. Its application allows for estimating the respondents' utilities, which enables calculating the relative importance of each variable representing the determinants of the choice of Polish FDI destinations. In addition, estimated part-worth utilities have enabled the segmentation of enterprises according to similar preferences of the selection of FDI location factors. Data from a survey conducted among companies that invest abroad or plan this form of investment has been used in the analysis.

Keywords: *foreign direct investment, conjoint analysis, analysis of variance, k-means method*
JEL Classification: F21, C38
DOI: 10.14659/SEMF.2018.01.43

1 Introduction

Poland has been perceived as an attractive area for foreign direct investments (FDI) for a long time. At present, also Polish companies increasingly seek investment opportunities in foreign markets. It is natural that due to the geographic location, European Union member states are the main recipients of Polish FDI. Legal conditions developed in the EU that guarantee free flow of people, goods, services and capital as well the geopolitical location of Poland are conducive to this situation. On the other hand, numerous Polish companies also look for opportunities for locating foreign direct investments outside EU. One of various reasons for potential attractiveness of such a destination may be the willingness to reduce costs of manufacturing or to seek new markets, and other conditions typical of individual countries which are beneficiaries of FDI. Cost factors, market factors, commercial factors and factors that create the investment climate are relatively frequently mentioned among numerous reasons for FDI enumerated in literature (Lukas, 1993; Rymarczyk, 2004; Salamaga, 2015). There are diverse economic theories that explain the reasons for the locations of FDI, for example, the theory of location factors (Lukas, 1993; Rymarczyk, 2004), economic

[1]Cracow University of Economics, Department of Statistics, 31-510 Kraków, ul. Rakowicka 27, salamaga@uek.krakow.pl

development (Dunning, 1981), comparative benefits (Kojima, 1980), dynamic comparative advantages (Ozawa, 1992), eclectic theory of international production (Dunning, 1988), and others. It was the theory of location factors that inspired the author of the article to carry out empirical research among Polish businesses in order to identify factors of key importance to the selection of FDI location. The analysis was based on data retrieved from a survey conducted among Polish companies that invest abroad or plan foreign direct investments. The conjoint analysis was the basic research tool used in the paper. With the method, partial usability of respondents was estimated and the validity level of variables that represent reasons for undertaking FDI by Polish companies was calculated. Furthermore, the estimated partial utilities enabled the segmentation of the surveyed companies according to similar preferences when it comes to the choice of reasons for undertaking FDI. The conjoint analysis in each of the separated segments enables a detailed description of reasons for undertaking FDI with consideration given to the validity of factors which are decisive for the choice of destinations of FDI. The results presented in the article also constitute a review of current preferences of Polish businesses in the scope of European destinations of FDI.

2 Research method

The conjoint procedure which, in principle, is a method for the classification and analysis of data used to measure preferences of respondents has been applied in research (Walesiak and Bąk, 2000). Respondents evaluate the set of profiles described with the use of selected explanatory variables (attributes). This enables one to gain knowledge on total preferences of respondents with regard to the distinguished profiles. The next stage of the conjoint analysis involves the calculation of part-worth utilities of attribute levels based on the decomposition of total utilities carried out in response to results of preference evaluations (Green and Wind, 1975). The usability of surveyed companies has been calculated with the use of the parameters of the following regression function (Walesiak and Bąk, 2000).

$$\hat{Y}_k = b_{0k} + \sum_{p=1}^{n} b_{pk} X_{pk} + e \,. \tag{1}$$

where:

Y_k – variable expressing the evaluation of preferences of k-th respondent who represents the surveyed company, b_{0k}, b_{1k}, b_{2k},..., b_{nk} – parameters of the regression model, X_1, X_2,...,X_n – artificial variables which represent attribute levels, e – random term of the model.

Respondents in surveyed companies have allocated ranks to individual profiles, so the values of variable *Y* are measured on an ordinal scale. Explanatory variables in the research were of qualitative nature and took two categories. That is why artificial variables are implemented in model (1) and they express the influence of each level of variable on the assessment attributed to profiles by respondents. The construction of artificial variables with the use of quasi-experimental coding and the method for calculating partial utility of respondents are presented in Table 1 (Walesiak and Bąk, 2000).

Table 1. Quasi-experimental coding of explanatory variable of two states.

Explanatory variable	Artificial variable Xp	Part-worth utilities
Level I	1	$U_{j1}^{k} = b_{pk}$
Level II	-1	$U_{j2}^{k} = -b_{pk}$

Part-worth utilities also allowed for calculating relative validity of each explanatory variable for *k*-th respondent, according to the following formula (Hair et al., 1995).

$$W_j^k = \frac{\max\limits_{l_j}\left\{U_{jl_j^i}^k\right\} - \min\limits_{l_j}\left\{U_{jl_j^i}^k\right\}}{\sum\limits_{j=1}^{m}\left(\max\limits_{l_j}\left\{U_{jl_j^i}^k\right\} - \min\limits_{l_j}\left\{U_{jl_j^i}^k\right\}\right)} \cdot 100\% \,. \tag{2}$$

The following formula, in turn, was used to determine average validity of *j*-th variable

$$W_j = \frac{1}{K}\sum_{k=1}^{K} W_j^k \,. \tag{3}$$

3 Research results

In the survey concerning the reasons for undertaking foreign direct investments, respondents were representatives of Polish companies who made foreign direct investments or had potential to use this form of investment. The survey was carried out in January 2016 and had approximately 750 subjects - enterprises operating in diverse economy branches. Following final selection of the responses sent in the survey, 408 completed questionnaires were taken into account in the research. Respondents assessed factors decisive for the export of foreign direct investments to destinations in Europe in accordance with their preferences. Six groups of factors were finally selected from various factors (Jun and Singh, 1996; Salamaga, 2015; Bayraktar-Saglam and Boke, 2017): cost factors, social factors, efficiency factors, law factors, market and political factors (cf. Table 2). The choice of these factors is not accidental. As long-term investments, foreign direct investments cause that, apart from purely economic

factors, an investor must also consider the stability of political and legal conditions in final destinations of FDI, as potential future social and political turbulences in the country of FDI destination may result in considerable losses on the part of an investing company. The analysis of non-economic factors is particularly justified, especially considering an unstable global economic and political situation in recent years.

Table 2. Determinants of a decision to choose FDI.

Variable No.	Factors	Variable level
1	cost	labour costs [A] prices of natural resources [B]
2	social	qualifications of staff [A] level of society [B]
3	efficiency	cooperation with local companies [A] modernization of production methods [B]
4	law	customs and fiscal policy [A] economic and administrative legislation [B]
5	market	new markets [A] seeking market niches [B]
6	political	degree of state's interference with economy [A] political stability [B]

Based on six distinguished variables and corresponding levels, a set of hypothetical variants for the choice of profiles has been created. Ten variants were proposed for 64 profiles and they were subject to evaluation by representatives of the surveyed enterprise. The specification of selected profiles is given in Table 3.

The surveyed respondents ranked the mentioned profiles according to their preferences: from the least preferred one (rank 1) to the most preferred one (rank 10). Based on the evaluation of profiles by respondents, parameters of model (1) were estimated with the least squares method. The factors presented in Table 3 were implemented in the model with the use of artificial variables in accordance with the quasi-experimental coding procedure described earlier.

Table 3. Profiles of the choice of destination of FDI.

Profiles	Factors determining preferences for the choice of destination					
	1	2	3	4	5	6
1	A	A	B	B	A	A
2	A	B	B	B	A	B
3	A	A	A	A	B	B
4	A	B	A	B	A	A
5	B	A	B	A	A	B
6	B	A	A	B	A	A
7	B	B	A	B	A	B
8	B	A	B	A	B	A
9	A	A	B	B	B	B
10	A	A	A	B	A	B

The estimation of the parameters of model (1) for each respondent allowed for calculating part-worth utilities in accordance with formulas given in Table 1. The results of respondents' part-worth utilities and average part-worth utilities in the conjoint analysis are given in Table 4.

Table 4. Results for part-worth utilities in the conjoint analysis.

Factors	Variable level	Respondent No.					Average
		1	2	....	407	408	
cost	labour costs	3.050	1.119	.....	1.322	0.937	1.941
	prices of natural resources	-3.050	-1.119	.....	-1.322	-0.937	-1.941
social	qualifications of staff	-0.050	-0.056	.....	1.918	1.188	1.505
	level of society	0.050	0.056	.....	-1.918	-1.188	-1.505
efficiency	cooperation with local companies	0.200	1.725	.....	0.496	0.750	1.254

	modernization of production methods	-0.200	-1.725	.....	-0.496	-0.750	-1.254
law	customs and fiscal policy	1.300	1.150	.....	2.442	2.000	1.901
	economic and administrative legislation	-1.300	-1.150	.....	-2.442	-2.000	-1.901
market	new markets	0.750	0.969	.....	2.380	2.938	1.382
	seeking market niches	-0.750	-0.969	.....	-2.380	-2.938	-1.382
political	degree of state's interference with economy	0.350	1.456	.....	0.899	0.813	0.989
	political stability	-0.350	-1.456	.....	-0.899	-0.813	-0.989

Based on part-worth utilities, relative validity of attributes was calculated in accordance with formula (2) and average relative validity of variables was calculated in accordance with formula (3) (cf. Table 5).

Table 5. Relative validity of attributes (in %) in the conjoint analysis.

Factors	**Respondent No.**					**Average**
	1	2	....	407	408	
cost	53.51	17.28	....	13.98	10.87	22.05
social	0.88	0.87	....	20.28	13.77	15.21
efficiency	3.51	26.64	....	5.24	8.70	13.42
law	22.81	17.76	....	25.82	23.19	21.06
market	13.16	14.96	....	25.17	34.06	15.21
political	6.14	22.49	....	9.51	9.42	13.05

The results shown in the last column of Table 5 suggest that cost and law factors are of the greatest importance to Polish companies when it comes to factors decisive for the location of FDI. Political and efficiency factors are of relatively low significance when making decisions on FDI.

Part-worth utilities of respondents specified in Table 4 reflect their responses to individual profiles evaluated (Walesiak and Bąk, 2000). This allows for using these utilities to segment the surveyed enterprises in accordance with similar factors decisive for the export of FDI. The method of *k*-means is used where four homogeneous clusters of enterprises are separated.
The created segments contain the following numbers of companies (as per the order of concentration): 84, 106, 79, 137. The conjoint analysis has been conducted separately in each segment distinguished and the average value of variables per segments is presented in Table 6.

Table 6. Average value of variables by segments of surveyed companies.

Factors	**Segment No.**			
	I	II	III	IV
	Average importance of variables (in %)			
cost	27.21	16.40	10.54	19.80
social	13.12	10.21	17.62	15.12
efficiency	5.14	14.20	13.21	29.85
law	21.45	27.89	17.21	21.40
market	16.12	16.30	24.24	8.13
political	16.96	15.00	17.18	5.70

In the first group of companies cost factors are of the greatest importance for launching FDI. They are followed by law factors whereas efficiency factors are of secondary importance. In the second segment, law and cost factors are of the greatest importance. Social factors are the least important there. The third segment of enterprises, in turn, is distinguished by a considerable importance of market and social factors, and the least importance of cost factors. In the fourth segment of enterprises, efficiency factors are the most important and political factors are the least important. The essence of the research lies in the linking of the distinguished segments of enterprises that export FDI and the choice of specific destinations

in Europe. Table 7 demonstrates the structure of the indicated chosen destinations in accordance with four segments of enterprises.

Table 7. The structure of selected destinations indicated in accordance with segments of businesses (in %).

Destination	**Segment No.**			
	I	II	III	IV
Cyprus	5.5	22.4	3.2	6.1
Czech Republic	13.5	7.3	10.1	5.8
France	7.9	5.1	4	11.5
Spain	6.2	9.2	10.5	8.2
Luxembourg	4.3	17.1	2.5	6.3
Germany	7.7	6.5	14.3	14.3
Norway	1.2	9.9	6.3	7.2
Russia	14	3.4	16.5	2.1
Switzerland	6	4.9	1.8	5.1
Ukraine	20.3	4.2	13.1	4.8
Great Britain	8.2	7.3	5.6	16.8
Other	5.2	2.7	12.1	11.8
Total	100	100	100	100

Based on Table 7, it may be stated that enterprises belonging to the first segment are most inclined to point at Ukraine, and further the Czech Republic as preferred destinations for FDI. For companies in the second segment, Cyprus and Luxembourg are optimal destinations of FDI whereas the majority of enterprises in the third sector choose Russia as the destination of their FDI. In the fourth segment, companies prone to invest in the United Kingdom are predominant. To sum up, it may be stated that if cost or market factors are important for FDI exporters, they tend to choose destinations in Eastern Europe most often. When law factors are most important, Cyprus or Luxembourg are the most frequent choices. In the situation when efficiency factors are decisive, exporters will be most inclined to point at destinations in Western Europe.

Conclusions

The export of FDI is associated with specific risk, e.g. political, market, FX and other risks. These risks must be taken into consideration by an enterprise which is to function in foreign markets in the long term. Therefore, the reasoned choice of FDI destinations is important here as it allows, at least partly, for restricting certain risks associated with international business activity of a company. Cost and law factors turned out to be most important for the selection of destination of FDI in European countries in the analysed research. A relatively low level of validity is attributed to political factors, which may be striking, for example, in the context of the current unstable situation in Eastern Europe. Priority treatment of law factors may suggest interest of some companies in FDI in the form of financial flows (capital in transit). It should be emphasised that legal and tax regulations of selected host countries of FDI encourage registration of subsidiaries, and this is conducive to "tax optimization". The combination of conjoint analysis results with the method for grouping *k*-means has allowed for separating homogeneous segments of companies and detailed presentation of their preferred factors decisive for the launch of FDI. The dynamically changing geopolitical and market situation in Europe and worldwide will certainly generate changes both in the scope of preferred destinations of FDI and in the scope of the assessment of determinants of the choice of these destinations. That is why such research should be repeated. This enables updating the combinations of determinants for the choice of FDI and verifying their validity level on an ongoing basis.

Acknowledgements

Publication was financed from funds allocated to Faculty of Management (Cracow University of Economics) within grants to maintain research capacity.

References

Bayraktar-Sağlam, B. & Böke, S.S. (2017), Labor Costs and Foreign Direct Investment: A Panel VAR Approach. Economies *5*(4), 1-23.

Dunning, J. H. (1981). Explaining the International Direct Investment. Position of Countries: Towards a Dynamic or Developmental Approach. *Weltwirtschaftliches Archiv Review of World Economics, 117*(1), 30-64.

Dunning, J. H. (1988). The Eclectic Paradigm of International Production: A Restatement and Some Possible Extensions. *Journal of International Business Studies*, *19*(13), 1-31.

Dunning, J. H. (2004), Determinants of Foreign Direct Investment: Globalization-Induced Changes and the Role of Policies. In: Tungodden, B., Stern, N. & Kolstad, I. (eds.), *Toward Pro-Poor Policies. Aid, Institutions and Globalization*. World Bank, Washington, 279-290.

Green P.E. & Wind Y. (1975). New Way to Measure Consumers' Judgments. *Harvard Business Review*, *53*, 107-117.

Hair J.F., Black W.C., Babin, B. J., Anderson, R. E. & Tatham R.L. (1995). *Multivariate data analysis*. Englewood Cliffs, Prentice-Hal.

Jun K.W. & Singh H. (1996). The Determinants of Foreign Direct Investment in Developing Countries. *Transnational Corporations*, *5*(2), 67-105.

Kojima K. (1980). Direct Foreign Investment. A Japanese Model of Multinational Business Operations. *Weltwirtschaftliches Archiv Review of World Economics, 116* (2), 388-390.

Lukas R. (1993). On the Determinants of Direct Foreign Investment: Evidence from East and Southeast Asia. *World Development, 21*(3), 391-406.

Ozaw, T. (1992). Foreign Direct Investment and Economic Development. *Transnational Corporation, 1*(1), 27-54.

Rymarczyk, J. (2004). *Internacjonalizacja i globalizacja przedsiębiorstwa*. PWE, Warszawa.

Salamaga, M. (2015). Wielowymiarowa analiza statystyczna motywów podejmowania bezpośrednich inwestycji zagranicznych przez polskie firmy. *Przegląd Statystyczny*, *62*(4), 415-433.

Walesiak, M., & Bąk A. (2000). *Conjoint analysis w badaniach marketingowych*. Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.

Anatomy of entrepreneurial activity and challenges of the new century

Katerina Shestaeva[1], Elena Sibirskaya[2], Paweł Lula[3], Evgenia Progonova[4]

Abstract

Recent research of entrepreneurial activity is undoubtedly among the most dynamically developing areas of socio-economic studies and this is well maintained by the number of published papers, researchers attending panel discussions on entrepreneurial activity, and the growing lists of international peer review journals and conferences. However, along with the entrepreneurship legitimization problems, so far, prospects for development of entrepreneurial activity research remain uncertain. Changes in methods and structure of this research mostly relate to the fact that Russia is undergoing a large-scale development of statistical methodologies to ensure consistency of the country's statistics with international standards, and those of OECD as such, improve National Accounts System along with demographic data and National Healthcare statistics, design methodology for basic tables "costs-output", and for statistical surveys of workforce, quality of life and households. In order to align with the business-logic and obtain timely and reliable statistical data on entrepreneurship, we find the conducted work critically important to meet current challenging issues. From this perspective, the research of anatomy of entrepreneurial activity can become a key element of the economic development evaluation and address the challenges of the modern society.

Keywords: statistics, entrepreneurship, entrepreneurial activity, economy, challenges.

JEL Classification: M21; O1

DOI: 10.14659/SEMF.2018.01.44

1 Introduction.

Entrepreneurial activity has proved to be a critical index to measure economic efficiency; indeed, businessmen form a social group that plays a crucial role in building economic well-being of a society. The very term of entrepreneurship implies exploration, search for optimal conditions, and maximizing resources, all of which are vital in our dynamically developing 21st century, and as a result, this area of studies has gained increased relevance in the modern world. This research aims at evaluating changes of entrepreneurial activity in Russia in 2008-2016, spreading the knowledge about indicators of entrepreneurial activity in the country and

[1]Corresponding author: Plekhanov Russian University of Economics, Situation Centre, address: Russia, Moscow, 117997, Stremyanny pereulok, 36, e-mail: shestaeva@gmail.com
[2]Plekhanov Russian University of Economics, Department of Statistics, address: Russia, Moscow, 117997, Stremyanny pereulok, 36, e-mail: Sibirskaya.EV@rea.ru
[3]Cracow University of Economics, Department of Computational Systems, address: Republic of Poland, Cracow, 27 Rakowicka St., 31-510, e-mail: pawel.lula@post.pl
[4]Higher School of Economics, National Research University, Department of English for Economic and Mathematical Studies, address: Russia, Moscow, 101000, Myasnitskaya Str., 20, e-mail: yprogonova@gmail.com

defining attitude to entrepreneurs and entrepreneurial activity in the society.

2 Methodology (materials used, description of the subject, methods and techniques of research).

Today, the best approach to define entrepreneurial activity in compliance with statistics is presented by S. Voinova and I. Savelieva "Evaluation of entrepreneurial activity: identification and types of economic activities" (Voinova and Savelieva, 2012), the authors were the first to tackle entrepreneurial activity as a specific economic indicator (indicators of entrepreneurial activity, grouped by types of economic activity and Russian Federation regions). According to the authors, the index of entrepreneurial activity evaluates entrepreneurship and intensity of its performance. Moreover, analysis of other definitions of entrepreneurial activity (Smelov et al., 2017; Sibirskaya et al., 2015; Stroeva et al., 2015; Harrisson et al., 2012; Joel and Bogers, 2014; London et al., 2010; Moss, 2005), leads to the following inferences: entrepreneurial activity can be measured, it is related to the population, and it fulfills its own functions, shows intensity of entrepreneurship, and drives the economy.

This research approaches entrepreneurial activity based on the information by Russia's information database - Statistical Register of Economic Entities of Federal State Statistics Service of the Russian Federation (Rosstat). The register is a database of economic entities, established and registered on the territory and under the law of the Russian Federation. In 2014, according to the recommendations of Statistical Office of the European Union (Eurostat) and Organization for Economic Co-operation and Development, Rosstat developed the official statistical methodology to produce indicators of business demography purposed at studying entrepreneurial activity as the number of high-growth enterprises and the number of their employees in the framework of demographic events (birth and death of enterprises).

Once the methodology was approved, Federal State Statistics Service released the initial business demography indicators (pilot project for 7 regions), from 2017 general information on Russian Federation is regularly updated with growth indicators evaluating the number of employees especially highlighted (high-growth enterprises; high-growth enterprises, including "mice enterprises"; enterprises with high growth potential; enterprises with high growth potential, including "gazelles"[5] enterprises) and growth evaluation by turnover in the same

[5]Enterprises with high growth potential are the enterprises, which maintain employees' growth or turnover rate at 10% a year within three years. High growth enterprises are the enterprises with an average annual growth higher than 20% within three years. "Gazelles" are a subgroup of high growth enterprises of five years and younger, i.e. all enterprises of four or

group. This paper applies the method of statistical observation of 2008-2016 business demography data. The tool for researching entrepreneurial activity in the world economy is Global Entrepreneurship Monitor (GEM 2015 Global Entrepreneurship Monitor (2016). The leading scientists of the UK, the USA, Finland and Ireland launched this project for country-specific studies of entrepreneurship and information exchange in 1997, and in 2016, it involved 66 countries with 69% of the world's population and 85% of the world's GDP. Russia has been carrying out these studies since 2006.

Two methods of collecting information were utilized: 1) surveys of adult working-age population with specially designed questionnaires (Adult Population Surveys — APS); 2) expert surveys (National Expert Surveys — NES), i.e. interviews of entrepreneurs and entrepreneurship experts.

3 **Findings.**

The indicators such as the overall number of enterprises and the number of sole proprietors can describe the country's level of entrepreneurial activity, since the individual entrepreneurship provides a more comprehensive understanding of the level of entrepreneurial activity among population. Table 1 shows the trends in the number of enterprises in the Russian Federation by Rosstat data. Table 1 shows that the number of enterprises in the Russian Federation decreased in 2007-2012, with 2008 plunge stemming from the world economic and financial crisis. The year of 2013 also demonstrates a drop in the number of organizations in the Russian Federation. However, Rosstat data in Table 2 shows a slight increase in the number of sole proprietors in 2016.

Table 2 also shows a fall in the entrepreneurial activity in Russia in 2008-2011, while in 2012 there was an increase in the number of individual entrepreneurs, 2013-2014 also witnessed a decrease of the indicator. In 2016, the situation levelled off and then showed an upward trend. The key indicators of entrepreneurial activity in the Russian Federation are the above-mentioned coefficients of birth and death of businesses per 1000 organizations. As can be seen from Fig.1, the average birth rate is 102 with that of death at 70. However, 2016, first for the period of the study, witnessed the death index higher than the birth figure (by 53 companies). Prospects for birth are lower than in 2015 when 102 companies were registered

five years with an average annual growth of 20% and up within three years, should be identified as "Gazelles". An average number of employees and turnover measure the growth. In addition, to avoid the undercount of a large number of high growth enterprises, they are identified within the group of enterprises with five to ten employees at the beginning of the growth period. These enterprises are called "mice".

per each 1000 existing organizations (Simonova et al., 2016).

Table 1. Changes in the number of enterprises and organizations (business entities) in the Russian Federation in 2008-2016 (Rosstat, 2018).

Years	Indicator, unit	Growth ratechained, %	Growth rate, basic, %	Rate of increase chained, %	Rate of increaseb asic, %	Absolute value of one percent of the increase
2008	94341.00	-	-	-	-	-
2009	93707.00	99.33	81.49	-0.67	-18.51	943.41
2010	92007.00	98.19	80.01	-1.81	-19.99	937.07
2011	90745.00	98.63	78.92	-1.37	-21.08	920.07
2012	89868.00	99.03	78.15	-0.97	-21.85	907.45
2013	92242.00	102.64	80.22	2.64	-19.78	898.68
2014	86471.00	93.74	75.20	-6.26	-24.80	922.42
2015	84222.00	97.40	73.24	-2.60	-26.76	864.71
2016	83333.00	98.94	72.47	-1.06	-27.53	842.22

Table 2. Operating sole proprietors in the Russian Federation in 2008-2016[6] (Rosstat, 2018).

Years	Indicator, ‘000 people	Growth rate chained, %	Growthr ate, basic, %	Rate of increase chained, %	Rate of increaseb asic, %	Absolute value of one percent of the increase
2008	2742.00	-	-	-	-	-
2009	2663.90	97.15	97.15	-2.85	-2.85	27.42
2011	2505.10	94.04	91.36	-5.96	-8.64	26.64
2012	2602.30	103.88	94.91	3.88	-5.09	25.05
2013	2499.00	96.03	91.14	-3.97	-8.86	26.02
2014	2413.80	96.59	88.03	-3.41	-11.97	24.99
2016	2523.60	104.55	92.04	4.55	-7.96	24.14

In 2016,the business demography data reflected a sharp increase in the number of liquidated companies compared to the newly registered ones. This new trend emerged for the first time from 2009, and, therefore, was not observed even in the recession of 2008-2009. However, it should not be interpreted as a negative trend only. The table below shows a selective analysis of entrepreneurial activity among “Gazelles”, data collected in a pilot project according to ROSSTAT information. “Gazelles” appear to be an illustrious example

[6] Data for 2010 and 2015 are not available; a statistical survey was not conducted.

of a healthy and mass-scale Russian business that today has affected the country's economy both in quantitative, gauged by GDP and employment growth, and qualitative, with launched and transferred innovations, factors.

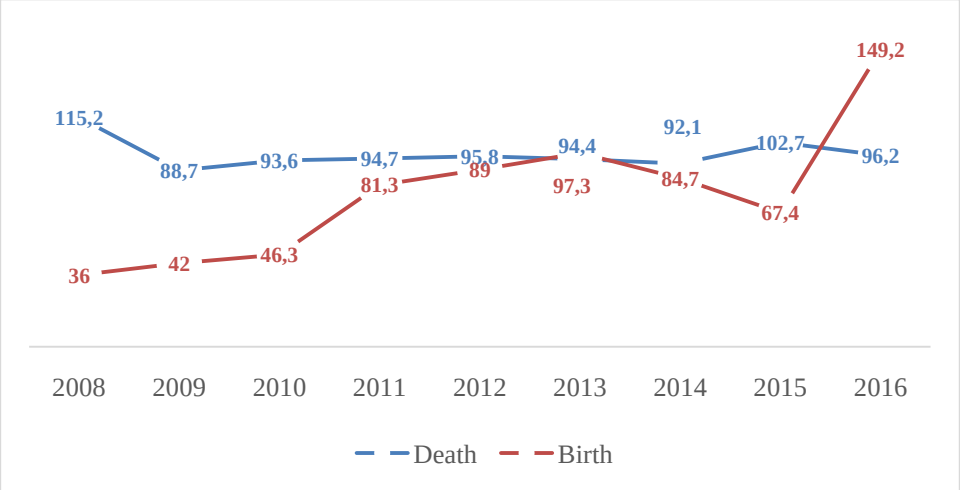


Fig.1. Trends in birth and death rates. (National report. Global Entrepreneurship Monitor. Russia 2016/2017, 2018).

Table 3. "Gazelles" enterprises (evaluation by turnover) in 2016 (Rosstat, 2018).

Region of the Russian Federation	Gazelles
Astrakhanskaya oblast	56.00
Belgorodskaya oblast	44.00
Vologodskaya oblast	54.00
Novosibirskaya oblast	923.00
Permskiy kray	60.00
Republic of Tatarstan	289.00
Sverdlovskaya oblast	610.00

Analysis of Table 3 reveals that in 2014 Novosibirskaya oblast boasted the top position among 7 regions of the Russian Federation by the number of "gazelles" enterprises (923 companies) with wholesale and retail trade as the most popular type of business (30% of all gazelles), while education was defined as the least popular. Fig. 2-5 show life style and interaction types of enterprises by the pilot project data as a study of the anatomy of entrepreneurial activity. However, the model of Global Entrepreneurship Research suggests

dividing indicators of entrepreneurial activity by individual and national characteristics showing how entrepreneurs are perceived and valued in a society (Shirokova et al., 2014).

In 2016 estimation of opportunities for opening a business was indicated as 17.9%. while perception of the business environment in the place of residence as favorable dropped if compared to 2014. when more than a quarter of the respondents estimated it positively. also slightly more than a quarter (28.4%) believed that their knowledge and experience were sufficient to start their own business. This indicator remained stable during the period of studies fluctuating from 22.7 in 2010 to 33.2 in 2011.



Fig.2. Trends in individual entrepreneurship characteristics[7]. (National report. Global Entrepreneurship Monitor. Russia 2016/2017.2018).

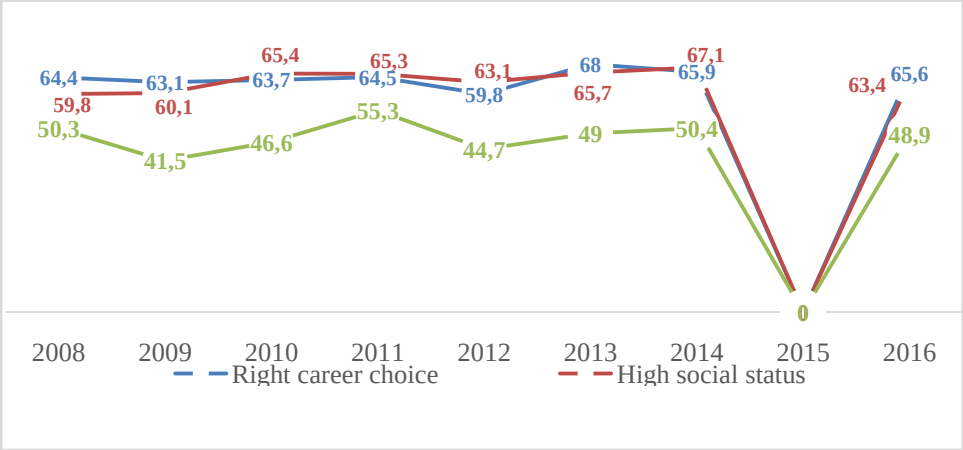


Fig.3. Trends in individual entrepreneurship characteristics. (National report. Global Entrepreneurship Monitor. Russia 2016/2017.2018).

[7] No research was conducted in 2015 and official data are not available(Fig. 3-6).

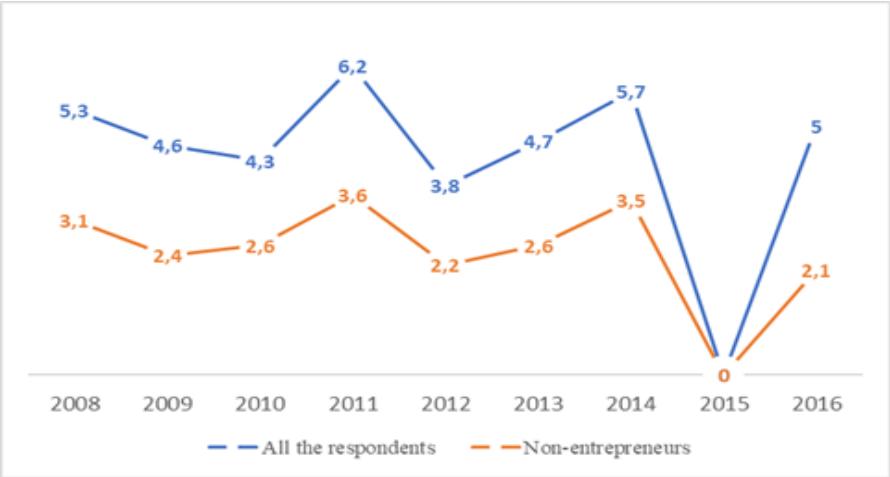


Fig.4. Trends in national entrepreneurship characteristics. (National report. Global Entrepreneurship Monitor. Russia 2016/2017. 2018).

Compared to 2014. 2016 saw a decrease of respondents deterred from starting their business by the fear of failure (from 59.1 to 55.3). However. the largest part of the surveyed noted that it was the fear of failure that kept them from going into entrepreneurial activity. During the period of studies. the majority of Russians highly valued an entrepreneur's status and deemed entrepreneurial career appealing. In 2016 65.6 and 63.4% of Russian citizens agreed with this opinion. which is 1.5 and 2.5% fewer than in 2014.

Also fewer than half of the respondents stated that mass media fairly often released informative publications about entrepreneurs who were able to start their business from scratch and became leaders in their industry. if referred to numbers. the highest figure of "Media attention to entrepreneurship" was recorded in 2011 (55.3) and the lowest in 2009 (41.5). As for the indicator of "entrepreneurial intentions". Fig. 5 illustrates the trends among all the interviewees and a group of non-entrepreneurs.

The carried out study demonstrates that there was a decrease in the birth rate of enterprises in the Russian Federation with a fluctuation of the death rate indicator and its slight upward trend. i.e. the key business demography indicators experienced a downward trend showing a decline in the entrepreneurial activity in the country.

The above-mentioned indicators are very important for cross-country comparisons for Global Entrepreneurship Monitor (GEM 2015 Global Entrepreneurship Monitor. 2016). which is used as an international information database. and. at present. it has become one of the most reliable and influential sources of information on entrepreneurial activity (Table 4).

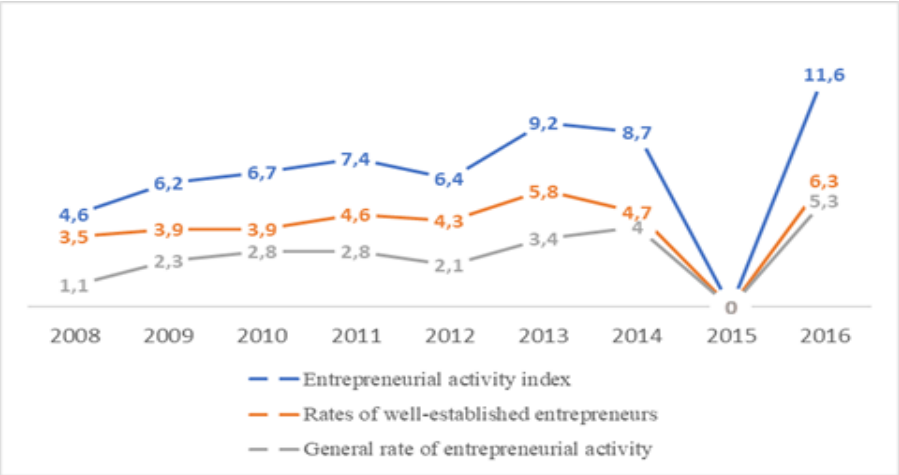


Fig. 5. Trends in entrepreneurial intentions. (National report. Global Entrepreneurship Monitor. Russia 2016/2017.2018).

Table 4. GEM countries by types of economy. (GEM 2015 Global Entrepreneurship Monitor. 2016).

Economy type	Economy factors	Countries
Factor-driven economy	Prices. traditional economy factors	Angola. Bolivia. Bosnia and Herzegovina. Columbia. Ecuador. Egypt. India. Iran
Efficiency-driven economy	Efficient production. labour intensity. education system aligned with economy. new markets and technologies	Argentina. Brazil. Chile. Croatia. Dominican Republic. Hungary. Jamaica. Latvia. Republic of Macedonia. Mexico. Peru. Romania. **Russia**. Serbia. South Africa. Turkey. Uruguay.**Kazakhstan**
Innovation-driven economy	Innovations. high-tech manufacturing and markets	Belgium. Denmark. Finland. **France. Germany.** Greece. Iceland. Ireland. **Israel.** Italy. Japan. Korea. Holland. Norway. Slovenia. Spain. **Switzerland**. **the UK**. **the USA**

Conclusion.

According to the "great challenges" logic and high expectations from the development of science, technology and innovations, we expect entrepreneurs not only to adequately address existing problems but also rather to boost quality of life and humanity development beyond

the limits of the present day. Comprehension of the anatomy of entrepreneurial activity by scientific and research community from the perspective of business demography as a structure (high-growth enterprises; high-growth enterprises. including "mice enterprises"; enterprises with high growth potential; enterprises with high growth potential. including "gazelles" enterprises), a composition (indicators of entrepreneurial activity grouped by the types of economic activity and Russian Federation regions), life style (indicators split by individual and national characteristics) and a type of interaction (how entrepreneurs are perceived and valued),enables to beneficially utilize this knowledge and enhance the quality of life.

References

GEM 2015 Global Entrepreneurship Monitor (2016) // http://datasebrae.com.br/documentos2/pesquisas/GEM%20%20Global%20Entrepreneurship%20Monitor/Apresentacao%20GEM%202015_PPT_COMPLETO.pdf.

Harrisson D., Chaari, N. & Comeau-Vallée, M. (2012). Intersectoral alliance and social innovation: When corporations meet civil society. *Annals of Public and Cooperative Economics, 83,* 1–24.

West, J. & Bogers, M. (2014). Leveraging External Sources of Innovation: A Review of Research on Open Innovation. *Journal of Product Innovation Management,31*,814 – 831.

London, T., Anupindi R. & Sheth S. (2010). Creating mutual value: Lessons learned from ventures serving base of the pyramid producers. *Journal of Business Research,63*, 582–594.

Moss, L. (2005). Encouraging small cultural enterprises in Russia: Gateways and barriers in Russian society. *The International Journal of Entrepreneurship and Innovation,6,* 193-200.

Rosstat, (2018) //http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/enterprise/reform/

Shirokova, G.V., Tsukanova, T. V. & Bogatyreva, K. A. (2014). *GUESSS National report* (Global research of entrepreneurial spirit among students) in Russia in 2013/2014 - http://www.ruaee.ru/about/documents/GUESSS_2013_Rus.pdf.

Sibirskaya, E., Stroeva, O. & Simonova, E. (2015). The characteristic of the institutional and organizational environment of small innovative and big business cooperation. *Procedia Economics and Finance*, *27*, 507 – 515.

Simonova, L. M. & Ovsyankina, M. V. (2016). Demography of organizations as a navigator in managerial solutions. *Vestnik of Tyumen State University. Socio-economic and legal*

studies,2, 209–218.

Smelov, P. A., Epshteyn, N., Egorova, E. A., Shestaeva, K. A. & Izvarina, V.A. (2017). Labor migration of the population of the Russian Federation. *4th International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM* 2017. 77-384.

Stroeva, O., Sibirskaya, E. & Shmanev, S. (2015). Character regional innovative policy. *Procedia Economics and Finance,27*,516 – 521.

Voinova, S. I. & Savelieva, I. P. (2012). Evaluation of entrepreneurial activity: the notion and types of economic activity. *Current problems of science and education,1*, 244.

National report. Global Entrepreneurship Monitor. Russia. 2016/2017 (2018) http://gsom.spbu.ru/files/docs/gem_russia_2016-2017.pdf

Fiscal and monetary effects in Ukraine: SVAR approach

Victor Shevchuk[1], Roman Kopych[2], Marianna Golynska[3]

Abstract

This paper investigates the impact of fiscal and monetary shocks on the output gap, producer price inflation (PPI) and the current account in Ukraine when applying SVAR model. On the basis of quarterly data for the 2000–2016 period, it has been found that the budget surplus is expansionary and anti-inflationary, with an improvement in the current account. The tightening of monetary policy, as measured by a decrease in the money supply with respect to the equilibrium trend, is associated with an improvement in the current account and a negative effect upon the output gap, while being neutral with respect to PPI. Our findings are in line with the predictions of the dependent economy model when considering the money-based expectations of the exchange rate. The evidence does indicate that a higher PPI makes a positive contribution to the current account, while there is a negative effect upon output. There is a favorable temporary effect of the current account upon the output gap, with a weak reverse causality running from the latter to the former. Monetary policy has standard anti-inflationary response, while being pro-cyclical. Also, it is worth noting that budget deficits are associated with an increase in the money supply. However, fiscal policy is independent of all endogenous variables.

Keywords**:** *budget balance, money supply, current account, output gap, inflation*
JEL Classification**:** C5, E5, H6
DOI: 10.14659/SEMF.2018.01.45

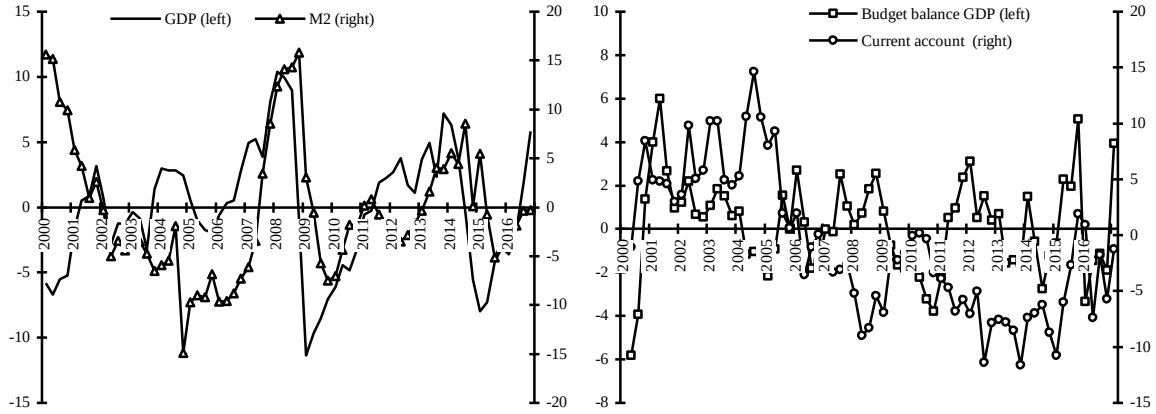
1 Introduction

Over the recent decade, the Ukraine's economy has experienced two harsh financial crises of 2008–2009 and 2014–2015, with the plunge of output below its trend to 11% and 8%, respectively (Fig. 1). Considering a possible link between the financial turmoil and macroeconomic policies, there appears to be a correlation between the money supply and the output gap. While excessive money supply can be blamed for a significant current account worsening in 2007–2008 and again in 2012–2013, the budget balance does not reveal any connections to the twin-deficit hypothesis.

Most international monetary and macro models predict that expansionary money supply shocks lead to a temporary increase in output (if any), worsening of the current account and a

[1] Corresponding author: Cracow University of Technology, Institute of Economy, Sociology and Philosophy, Warszawska 24, 31-155 Kraków, Poland, e-mail: vshevchuk@pk.edu.pl.
[2] Lviv National University named by Ivan Franko, Ukraine, Universytetska 1, 79-001 Lviv, Ukraine, e-mail: kopych_r@bigmir.net.
[3] Lviv University of Trade and Economics, Tuhan-Baranovskoho 5, 79-005 Lviv, Ukraine, e-mail: marianna.yasinska@ukr.net.

higher price level (at least in the long run). Macroeconomic effects of the fiscal stimuli are similar, though with different timing and amplitude. A number of studies have empirically supported these standard predictions for both money supply (Canova and Menz, 2011; Favara and Giordani, 2009) and fiscal policy (Ilzetzki et al., 2013). However, it has not been ruled out that money supply is ineffective with respect to output, even in the short run (Uhlig, 2005), or it has restricted impact on economies with financial market constraints (Rojas-Suarez, 1992). Similarly, the so-called Non-Keynesian fiscal policy effects imply output expansion following budget deficit cut.



a) GDP and money supply (deviation from trend, %)

b) budget and current account balances (% of GDP)

Fig. 1. Ukraine: selected macroeconomic indicators, 2000–2016.

Note: GDP and money aggregate M2 are de-trended with the Hodrick–Prescott filter.

Source: Ukraine's Ministry of Finance, IMF *International Financial Statistics*.

The aim of this study is to examine the budget balance and money supply shocks upon producer price inflation (PPI), the current account and output gap in Ukraine. We implement structural vector auto-regression (SVAR) approach for modeling the inter-dependencies between fiscal and monetary policies and main macroeconomic indicators.

The rest of the paper is organized as follows. Section 2 provides a brief review of the related literature. Section 3 describes data and methodology. Section 4 discusses empirical results and Section 5 provides conclusions.

2 Related literature

Although in the majority of modern monetary models the stock of money has disappeared, it is hard to accept such approach not only because of the quantity theory orthodoxy, the balance sheet effects or liquidity constraints, but also because of numerous evidence in favor of real money supply effects on both industrial and developing countries (Canova and Menz, 2011). However, earlier studies concerning post-Soviet countries provide mixed evidence of the real effects of monetary policy (Starr, 2005). While earlier studies had found smaller (but similar) responses of CEE countries to the monetary policy shock than those found for the Western European countries (Anzuini and Levy, 2007), later ones report that the price responses imply the possibility of even stronger effects of monetary policy on prices than in the euro area countries, especially for lags which are longer than one year (Jarocinski, 2010).

The relationship between money and macroeconomic indicators is likely to be country-specific. For example, it has been found that the balance of payments shocks is important in price level movements in Hungary, while nominal shocks are dominant in affecting prices in Poland; on the other hand, monetary shocks affect output in the short run in Hungary, while supply shocks dominate output movements in Poland (Dibooglu and Kutan, 2005). Standard features of monetary policy are found for the Czech Republic (Borys et al., 2009). Although the money-prices relationship is supposed to weaken once inflation lowers, money supply is likely to play a role in explaining inflation dynamics in Ukraine due to inflationary inertia, dollarization and unstable economic environment.

The effect of monetary policy upon the current account depends on the relative strength of familiar income absorption and expenditure switching effects. For several European countries, it has been revealed that the expenditure-switching effect is stronger, meaning that improvement in the current account is brought about by monetary contraction (Kim, 2001).

Most of the studies for the CEE countries report more conventional Keynesian features of fiscal policy (Ambriŝko et al., 2013; Kabashi, 2017). However, earlier findings are in favor of the inverse Non-Keynesian relationship between the budget deficit and output (Segura-Ubiergo et al., 2006). As proposed by Rojas-Suarez (1992), expansionary effects of both budget surplus and money supply contraction could be explained by the dependent-economy model with financial constraints under the assumption of the money-based expectations of a nominal exchange rate.

3 Data and statistical methodology

The data set is quarterly for the sample period of 2000–2016 and has been primarily collected from IMF International Financial Statistics (IFS) and Ukraine's State Committee of Statistics. As shown in Fig. 1, the cyclical component of real output (index, 1994=100) and money supply (money aggregate M2, million of hryvnas) is extracted by the Hodrick–Prescott filter. Both budget and current account balances are measured as % of GDP. To capture price effects producer price inflation (PPI) is used (%). All data were seasonally adjusted using the Census X12 procedure, except for PPI. Both ADF and PP stationarity tests indicate that all the macroeconomic variables are stationary at the 5% significance level (not reported).

Assuming that the VAR model of Ukraine's economy is represented by a structural-form equation in the form of $A_0 X_t = A(L)X_{t-1} + B\varepsilon_t$, the reduced-form is as follows:

$$X_t = A_0^{-1}A(L)X_{t-1} + A_0^{-1}B\varepsilon_t = C(L)X_{t-1} + u_t \tag{1}$$

where X_t is the $n\times 1$ vector of the endogenous variables, $A(L)$ is a polynomial variance-covariance matrix, A_0 is a non-singular matrix normalized to have ones on the diagonal and summarizes the contemporaneous relationships between the variables in the model contained in the vector X_t, $C(L)$ is a matrix representing the relationship between lagged endogenous variables, L is the lag operator, ε_t is a $n\times 1$ vector of normally distributed, serially uncorrelated and mutually orthogonal white noise disturbances, u_t is a $n\times 1$ vector of VAR residuals with a zero mean and that are serially uncorrelated but could be contemporaneously correlated with each other.

Assuming that the reduced-form VAR disturbances are related to the structural disturbances as $A_0 u_t = B\varepsilon_t$, the specification of our SVAR is as follows (in terms of the contemporaneous innovations):

$$bd = u_1, \tag{2}$$

$$m = a_1 bd + a_2 ca + u_2, \tag{3}$$

$$p = b_1 bd + b_2 m + u_3, \tag{4}$$

$$ca = c_1 y + c_2 p + u_4, \tag{5}$$

$$y = d_1 bd + d_2 m + d_3 p + d_4 ca + u_5, \tag{6}$$

where *bd* is the budget balance (% of GDP), *m* is the money supply (%), *ca* is the current account (% of GDP), *p* is the PPI (%), *y* is the output gap (%).

All variables in equations (2)–(6) represent the first stage VAR residuals. As implementing new fiscal measures in response to specific macroeconomic developments, typically taking longer than three months, it is assumed that quarterly variables allow setting the discretionary contemporaneous responses of the budget balance to changes in other endogenous variables to zero (equation (2)). De-trended money supply is influenced by both budget and current account balances, reflecting realities of a *de facto* fixed exchange rate regime, as practised in Ukraine over the period of 2000–2013 (equation (3)). Producer price dynamics is affected by fiscal and monetary policies (equation (4)). The current account is a function of output and PPI shocks (equation (5)). Finally, domestic business cycle is influenced by all other endogenous variables in the current period (equation (6)).

Among exogenous variables, our SVAR includes a nominal effective exchange rate (index, 2010=100), present and lagged foreign direct investments (% of GDP), world metal and crude oil prices (index, 2005=100) and business cycle in Russia (%), while dummy variable is used to control the financial turmoil of 2008–2009 and 2014–2016. In estimation, we use two lags of each endogenous variables, as implied by most of lag length criteria.

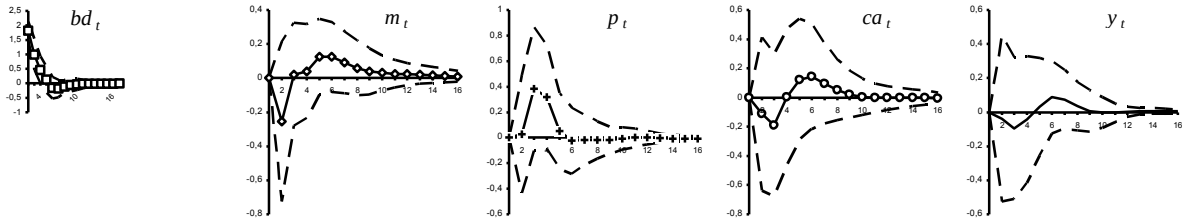
4 Estimation results

The effects of macroeconomic shocks are analyzed through impulse-response functions and the forecast error variance decompositions (FEVDs). Fig. 2 presents the impulse-response functions for endogenous shocks. Table 1 shows the portion of the FEVD for endogenous variables.

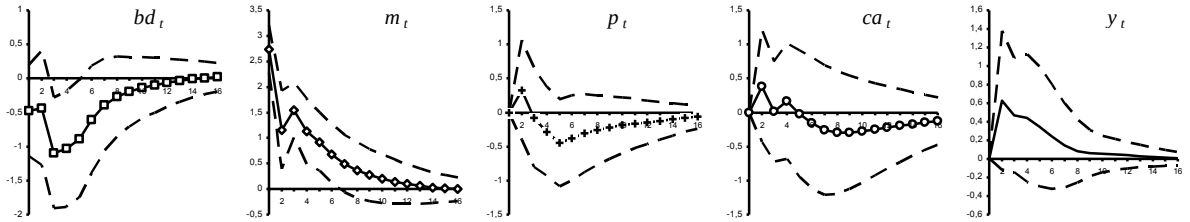
As stated in Fig. 2, an improvement in the budget balance is followed by a short-term anti-inflationary impact combined with an improvement in the current account, as suggested by the twin deficits hypothesis and a persistent pro-growth effect prevails. It means that Ukraine is characterized by the Non-Keynesian pattern of fiscal policy. Somewhat surprisingly, the budget surplus is associated with a prolonged decrease in the money supply. On the other hand, there is an expansionary reaction of fiscal policy to monetary tightening on impact, with a gradual reversal in 4 to 6 quarters. A positive shock to prices leads to an immediate improvement in the budget balance, while it takes time to establish the restricted monetary stance. Fiscal policy does not react to changes in the current account and business cycle, while monetary policy is pro-cyclical. Also, there is evidence of the current account monetary impact which is consistent with a fixed exchange rate regime.

In contrast to fiscal policy, a restricted monetary stance has no significant anti-inflationary effect, while a positive money supply shock leads to an immediate worsening of the current

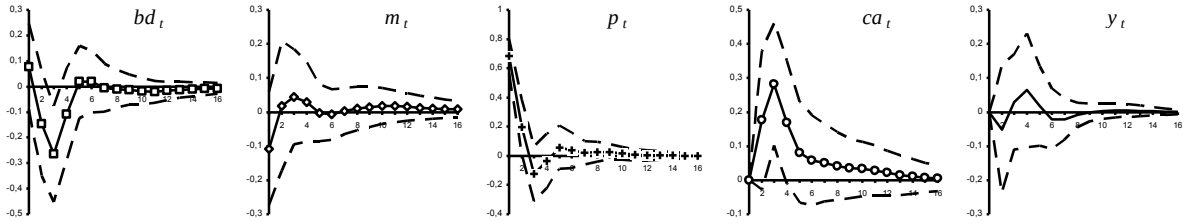
account that is corrected in 3 quarters. However, there is a significant negative impact on output, up to 8 quarters after the shock. Within the framework of a dependent economy model (Rojas-Suarez, 1992), it supports the assumption of the money supply-based expectations of the exchange rate. If there is a high exchange rate elasticity of substitution in demand for traded and non-traded goods combined with a strong link between the supply of both goods



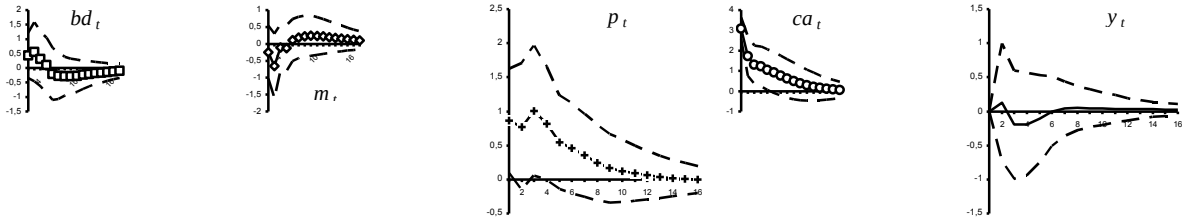
a) response of the budget balance



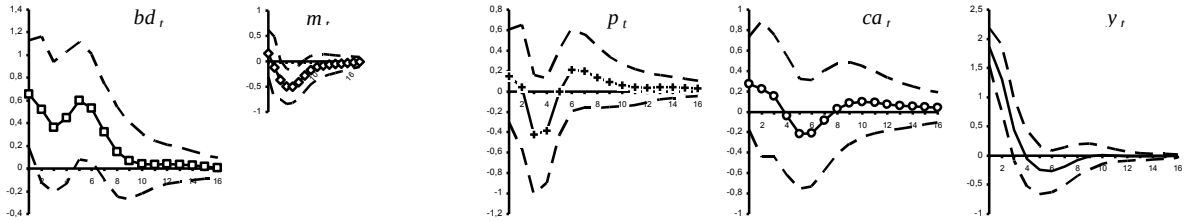
b) response of the money supply



c) response of producer price inflation



d) response of the current account



e) response of output

Fig. 2. Impulse response functions of endogenous variables.

Note: Solid lines are the point estimates of the impulse-response mean. Dashed lines are the point estimates ± 2 standard deviations.

Table 1. Forecast error variance decomposition.

Responses to innovations in	**Forecast horizons**		
	4	**8**	**16**
bd to innovations in			
bd	92	90	90
m	1	2	2
p	5	5	5
ca	1	2	2
y	0	1	1
m to innovations in			
bd	16	20	20
m	76	70	68
p	1	4	4
ca	1	2	3
y	5	5	5
p to innovations in			
bd	14	13	14
m	2	2	2
p	66	65	64
ca	17	19	19
y	1	1	1
ca to innovations in			
bd	3	4	5

m	2	2	3
p	15	15	15
ca	79	78	77
y	0	0	0
ca to innovations in			
bd	14	20	20
m	6	11	11
p	5	5	5
ca	2	3	3
y	73	61	61

and real money supply, it is possible to obtain a decrease in output against the backdrop of a weak (if any) inflationary impact.

After an inflationary shock, there is a significant improvement in the current account combined with a restricted impact upon the output gap, reaching a peak within a year. On the other hand, PPI accelerates in response to the current account shock, however there is no reaction to the output gap. As improvement in the current account is likely to have an expansionary effect upon output, the former seems to be neutral with respect to the latter.

Analysis of the FEVD indicates that shocks to fiscal policy account for 20% of the fluctuations in the output and 14% of the fluctuations in PPI (Table 1). However, there is no significant fraction of the budget balance justifying the current account developments. Nominal and real effects of money supply shock are less significant in both respects. However, fiscal and monetary policy together account for a significant (above 30%) proportion of output forecasts, while the combined value of this indicator for the PPI and the current account is 16% and 8%, respectively. The budget balance determines up to 20% of the variation in the money supply. Other endogenous variables do not play any role in the variation of money supply. It is also the case of fiscal policy.

Among other results, the PPI is driven mostly by the changes in budget balance (14%) and the current account (19%). In turn, inflationary developments are an influential factor of the variation in the current account (15%). The output gap is of marginal dependence on PPI and current account shocks, with their combined fraction in variance decomposition at just 8%.
Comparing with the relevant empirical studies of the CEE countries, our results are consistent with earlier findings of the Non-Keynesian fiscal policy effects, as established by Segura-Ubiergo et al. (2006), while there is no support for a strong relationship between the money

supply and prices, as it is found by Jarocinski (2010). Following Kim (2000), a short-lived inverse relationship between the money supply shock and the current account can be explained by the relative strength of the expenditure-switching effect compared to the income absorption effect. As there is a significant negative impact of the money supply on output in Ukraine, it is running counter to standard expansionary effects of money aggregates for the CEE countries, with the Czech Republic and Hungary being an example (Borys et al., 2005; Dibooglu and Kutan, 2005).

Conclusions

The main outcome of the empirical analysis is that the budget surplus is expansionary, anti-inflationary and helpful for the current account adjustment. A decrease in the money supply seems to be expansionary as well, with a simultaneous short-lived improvement in the current account and neutrality with respect to PPI. Such fiscal and monetary policy effects are in line with the predictions of the dependent economy model under an assumption of the money-based expectations of the exchange rate. Monetary policy has standard anti-inflationary response, but it is pro-cyclical with respect to GDP, while fiscal policy is independent of all endogenous variables. Our findings are of particular policy relevance in the context of ongoing debate on the merits of fiscal consolidation and monetary tightening as viable options for a sustainable macroeconomic stabilization in Ukraine, as well as considering important consequences of a recent switch to a 'pure' floating exchange rate regime since February 2014.

References

Ambriško, R., Babecky, J., Rysanek, J. & Valenta, V. (2015). Assessing the Impact of Fiscal Measures on the Czech Economy. *Economic Modelling*, *44*(C), 350-357.

Anzuini, A. & Levy, A. (2007). Monetary policy shocks in the new EU members: a VAR approach. *Applied Economics*, *39*(9), 1147-1161.

Borys, M. M., Horváth, R. & Franta, M. (2009). The effects of monetary policy in the Czech Republic: an empirical study. *Empirica*, *36*(4), 419-443.

Canova, F. & Menz, T. (2011). Does Money Matter in Shaping Domestic Business Cycles? An International Investigation. *Journal of Money, Credit and Banking*, *43*(4), 577–607.

Dibooglu, S. & Kutan, A. (2005). Sources of inflation and output movements in Poland and Hungary: Policy implications for accession to the economic and monetary union. *Journal of Macroeconomics*, *27*(1), 107-131.

Favara, G. & Giordani, P. (2009). Reconsidering the Role of Money for Output, Prices and Interest Rates. *Journal of Monetary Economics*, *56*(3), 419-430.

Ilzetzki, E., Mendoza, E. & Végh, C. (2013). How big (small?) are fiscal multipliers? *Journal of Monetary Economics*, *60*(2), 239-254.

Jarocinski, M. (2010). Responses to monetary policy shocks in the east and the west of Europe: a comparison. *Journal of Applied Econometrics*, *25*(5), 833-868.

Kabashi, R. (2017). Macroeconomic effects of fiscal policy in the European Union, with particular reference to transition countries. *Public Sector Economics*, *41*(11), 39-69.

Kim, S. (2001). Effects of monetary policy shocks on the trade balance in small open European countries. *Economics Letters*, *71*(2), 197-203.

Lo, M. & Granato, J. (2008). What explains recent changes in international monetary policy attitudes toward inflation? Evidence from developed countries. *Economics Letters*, *100*(3), 411-414.

Rojas-Suarez, L. (1992). Limited Capital Markets And The Real Effects Of Monetary Stabilization Policies Under Alternative Exchange Rate Regimes. *Journal of International Money and Finance*, *11*(6), 594-613.

Segura-Ubiergo, A., Simone, A. & Gupta, S. (2006). New Evidence on Fiscal Adjustment and Growth in Transition Economies. *IMF Working Paper*, 06/244, Washington: International Monetary Fund.

Starr, M. (2005). Does money matter in the CIS? Effects of monetary policy on output and prices. *Journal of Comparative Economics*, *33*(3), 441-461.

Uhlig, H. (2005). What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics*, *52*(2), 381-419.

Armed conflicts and multicriterial evaluation of the development of the country in a long-term perspective

Agata Sielska[1]

Abstract

Armed conflicts have significant impact on economics, although their indirect influence is difficult to assess. Indirect effects of the conflict can occur both before and after the conflict and result from changes in the allocation of resources, which in turn results i.a. from a shortage of investment.

The influence of conflicts is not always assumed to be negative, as they may contribute to the improvement of effectiveness. In the paper methods of constructing multi-criteria rankings were used to assess the situation of selected countries in the period 1810-1980. In the next step, evaluation results were analysed for the relationship between the occurrence of conflicts and situation of country. Due to the change in the nature of armed conflicts in the analyzed period three sub-periods were specified: before 1910, 1910-1940 and after 1940. Results depend on the assumptions, i.e. the methods used to construct rankings and the assumptions regarding the relevance of the criteria.

Keywords: *Multicriteria rankings, development, welfare, armed conflicts*
JEL Classification: I310, C440, F510
DOI: 10.14659/SEMF.2018.01.46

1 Introduction

Economic agents react to changes in the environment. Therefore war modifies the allocation of resources as capital and labor move to purposes related to conflict which can lead to the shortage of investment. Impact of civil war on economic growth was discussed among others by (Kang, Meernik, 2005). Although many indicators used for analysis of the impact of war base on direct effects like damage costs, indirect influence is difficult to assess (cf. Lacina and Gleditsch, 2005). Those effects occur in many spheres related to the development and quality of life, like life expectancy (Plümper and Neumayer, 2006) or mortality (Li and Wen 2005). Impact of war on trade is one of the most discussed topics (Barbieri and Levy, 1999; Feldman and Sadeh, 2018). Economists also take into consideration changes in supply of basic goods and services like education (Lai and Thyne, 2007) or healthcare (Plümper and Neumayer, 2006; Lai and Tyne, 2007). This results not only in lowered quality of life and life satisfaction, but has serious economic consequences as it reduces the quality of human capital and contributes towards costs increase. On the other hand, those changes may result in the

[1] Corresponding author: Warsaw School of Economics, Department of Applied Economics, Al. Niepodległości 162, 02-554 Warsaw, e-mail: asiels@sgh.waw.pl.

improvement of efficiency (Kang and Meernik, 2005). All of the effects mentioned above may occur both before and after conflict. Not only wars were included among conflicts analysed in the paper due to the fact that even smaller conflicts or tensions can cause wide-spread results or they can escalate (Barbieri and Levy, 1999).

2 Methods and data

According to (Gal et al., 1999) one of the conditions justifying the use of multicriteria rankings is the situation when criteria are greatly different which makes expressing them on one common scale a very difficult task. Rankings allow to overcome problems resulting from unclear compensation or preference thresholds. In case of the assessment of the level of development the abovementioned features make them very useful and efficient tools.

In the paper following methods were used to construct multicriteria rankings: WSA (the method also referred to as Simple Additive Weighting, cf. Geldermann and Rentz, 2000), further denoted in tables by W, PROMETHEE II (Brans et al., 1986), further denoted in tables by P, and TOPSIS (Hwang and Yoon, 1981), further denoted in tables by T. For PROMETHEE II all criteria were considered being of the usual (I) type. Each method is based on different approach towards ranking. They are used for ranking economic entities like countries and regions, often in combination with other approaches.

All three methods require an analyst to specify the set of weights. In case of decision-aiding problems weights are either directly given by the decision-maker or derived via an interview. In the paper criteria were analyzed in accordance with two sets of weights. The first set, hereinafter referred to as E variant is based on the assumption that all criteria are equally important. Therefore for each criterion its weight w_i is computed as follows:

$$w_i = \frac{1}{N} \tag{1}$$

where: i=1,...,N, and N denotes the number of criteria.

In the second variant hereinafter referred to as NE weights were derived assuming that criteria should be complementary not substitutional to one another. It can be supposed that if values of a criterion are not clearly differentiated between alternatives it should not be considered as important as it does not provide much additional information. Similarly, it may be assumed that correlated variables transfer the same information. Therefore for each criterion weight w_i is computed in accordance with the procedure (Sielska, 2010) similar to the CRITIC method (Diakoulaki et al., 1995). In the first step coefficients of variation of

criteria values are computed and normalized (2) and secondly relations between criteria are considered (3-4).

$$v_i = \frac{S_i}{\overline{f}_i} \tag{2}$$

$$w_i^1 = \frac{|v_i|}{\sum_{i=1}^{N}|v_i|} \tag{3}$$

$$w_i^2 = \frac{\sum_{j=1}^{N}\left|r_{ij}^P\right|}{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|r_{ij}^P\right|} \tag{4}$$

where: S_i – standard deviation of the i-th criterion, f_i – average of the i-th criterion; $i=1,..,N$, r^P_{ij} – coefficient of correlation between criteria i and j.

Final weights for NE variant are computed as follows:

$$w_i = \left(\frac{w_i^1}{w_i^2}\right) \Big/ \left(\sum_{i=1}^{N}\frac{w_i^1}{w_i^2}\right) \tag{5}$$

Data used in the paper come from (van Zanden et al., 2014). Variables approximate the level of development by taking into account 3 different spheres: economic performance, population health and the quality of environment. Due to the varying availability of data sets of criteria differ between periods. However, in case of single missing observations, missing data was supplemented either with the average of neighboring values or with values determined using a trend function. The following criteria were maximized: GDP per capita expressed in 1990 PPP dollars (period 1820-1980); mean height (1820-1980); democracy index (1850-1980); life expectancy at birth (1870-1980); mean species abundance (1820-1980); polity2 index (1850-1980) and real wage of construction workers (1820-1980). The set of minimized criteria includes: homicide rates for 100 000 people (1850-1980); SO2 emission per capita (1850-1980); income inequality measured by Gini index (1820-1980); CO2 emission per capita (1820-1980). Data on occurrence of conflicts come from Clioinfra database (https://www.clio-infra.eu). The following countries were analyzed: Great Britain (GBR), Netherlands (NLD), France (FRA), Germany (DEU), Italy (ITA) and Spain (ESP).

3 Results and discussion

The first place in the ranking (the Netherlands) turned out to be very stable. Results obtained using PROMETHEE method were similar for both variants of weights (see Tables 1-2). Letters in the indices in tables denote the method used for construction of the ranking.

Table 1. Development ranks (equal weights).

period	GBR	NLD	FRA	DEU	ITA	ESP
1820-1830	$2^W6^T2^P$	$1^W2^T1^P$	$3^W5^T6^P$	$4^W3^T5^P$	$6^W4^T3^P$	$5^W1^T4^P$
1830-1840	$1^W4^T1^P$	$2^W2^T2^P$	$4^W6^T6^P$	$3^W3^T3^P$	$6^W5^T4^P$	$5^W1^T5^P$
1840-1850	$1^W4^T1^P$	$2^W3^T2^P$	$5^W5^T5^P$	$3^W2^T3^P$	$6^W6^T6^P$	$4^W1^T4^P$
1850-1860	$1^W6^T1^P$	$2^W5^T2^P$	$3^W3^T3^P$	$4^W4^T6^P$	$5^W1^T5^P$	$6^W2^T4^P$
1860-1870	$1^W6^T1^P$	$2^W5^T3^P$	$4^W3^T2^P$	$3^W4^T3^P$	$5^W1^T6^P$	$6^W2^T5^P$
1870-1880	$2^W1^T1^P$	$1^W2^T3^P$	$2^W3^T2^P$	$4^W4^T4^P$	$5^W5^T6^P$	$6^W6^T5^P$
1880-1890	$2^W1^T1^P$	$1^W2^T3^P$	$3^W3^T2^P$	$4^W4^T4^P$	$6^W6^T6^P$	$5^W5^T5^P$
1890-1900	$2^W1^T1^P$	$1^W2^T2^P$	$3^W4^T3^P$	$4^W3^T4^P$	$6^W5^T6^P$	$5^W6^T5^P$
1900-1910	$2^W1^T1^P$	$1^W2^T1^P$	$3^W4^T3^P$	$4^W3^T4^P$	$5^W5^T5^P$	$6^W6^T6^P$
1910-1920	$2^W1^T2^P$	$1^W2^T1^P$	$3^W4^T4^P$	$4^W3^T3^P$	$5^W5^T5^P$	$6^W6^T6^P$
1920-1930	$2^W1^T2^P$	$1^W2^T1^P$	$3^W3^T3^P$	$4^W4^T4^P$	$6^W5^T6^P$	$5^W5^T5^P$
1930-1940	$2^W2^T2^P$	$1^W1^T1^P$	$3^W3^T3^P$	$5^W4^T4^P$	$6^W5^T4^P$	$4^W5^T6^P$
1940-1950	$2^W1^T2^P$	$1^W3^T1^P$	$3^W4^T3^P$	$5^W2^T4^P$	$6^W5^T5^P$	$4^W6^T6^P$
1950-1960	$3^W1^T2^P$	$1^W2^T1^P$	$5^W3^T4^P$	$4^W4^T5^P$	$6^W5^T3^P$	$2^W6^T6^P$
1960-1970	$3^W1^T2^P$	$2^W2^T1^P$	$4^W4^T4^P$	$4^W3^T5^P$	$6^W5^T6^P$	$1^W6^T3^P$
1970-1980	$3^W3^T4^P$	$2^W1^T1^P$	$3^W2^T3^P$	$5^W3^T6^P$	$6^W5^T5^P$	$1^W6^T2^P$

In the next stage of the analysis, the relationship between the occurrence of conflicts and the rank and evaluation result for a given country was examined. The examination is based on correlation coefficients and analysis of variance. Therefore, the results do not refer to causal relationships, only the coexistence of phenomena. A significance level of 0.05 was assumed.

In the first step, analysis of variance was carried out. Binary variables refer to the occurrence of conflicts. Results are reported in Table 3. Variable InterX means that the country participated in an international conflict X decades earlier and variable IntraX refers to the occurrence of intranational conflict X decades earlier. In the analysis both evaluations and ranks were taken into account. * denotes the significance at 0.05 level. First element in pair denotes the significance of differences in ranks, second element denotes the significance of differences in evaluations. For E variant it can be seen that evaluation results and ranks depend on the occurrence of internal conflicts for all methods except TOPSIS.

Table 2. Development ranks (unequal weights).

period	GBR	NLD	FRA	DEU	ITA	ESP
1820-1830	$6^W6^T5^P$	$1^W1^T1^P$	$2^W2^T4^P$	$5^W4^T6^P$	$4^W5^T3^P$	$3^W3^T2^P$
1830-1840	$6^W6^T5^P$	$1^W1^T1^P$	$3^W2^T6^P$	$5^W4^T4^P$	$4^W5^T2^P$	$2^W3^T3^P$
1840-1850	$6^W6^T5^P$	$1^W1^T1^P$	$3^W3^T6^P$	$4^W4^T4^P$	$5^W5^T2^P$	$2^W2^T3^P$
1850-1860	$6^W5^T5^P$	$2^W2^T4^P$	$1^W1^T3^P$	$3^W4^T6^P$	$5^W3^T2^P$	$4^W6^T1^P$
1860-1870	$2^W2^T2^P$	$4^W4^T5^P$	$1^W1^T1^P$	$5^W5^T6^P$	$6^W6^T4^P$	$3^W3^T2^P$
1870-1880	$2^W2^T2^P$	$4^W4^T4^P$	$1^W1^T1^P$	$5^W6^T6^P$	$6^W5^T5^P$	$3^W3^T3^P$
1880-1890	$3^W2^T2^P$	$5^W5^T4^P$	$1^W1^T1^P$	$4^W4^T5^P$	$6^W6^T6^P$	$2^W3^T3^P$
1890-1900	$3^W2^T2^P$	$4^W4^T3^P$	$1^W1^T1^P$	$5^W5^T6^P$	$6^W6^T5^P$	$2^W3^T4^P$
1900-1910	$4^W4^T3^P$	$2^W3^T2^P$	$1^W1^T1^P$	$5^W6^T6^P$	$6^W5^T4^P$	$3^W2^T5^P$
1910-1920	$4^W4^T3^P$	$1^W1^T1^P$	$2^W2^T2^P$	$3^W3^T4^P$	$6^W6^T5^P$	$5^W5^T6^P$
1920-1930	$3^W4^T2^P$	$1^W1^T1^P$	$2^W2^T3^P$	$5^W5^T6^P$	$6^W6^T5^P$	$4^W3^T4^P$
1930-1940	$2^W2^T2^P$	$1^W1^T1^P$	$3^W4^T4^P$	$6^W6^T6^P$	$4^W3^T3^P$	$5^W5^T5^P$
1940-1950	$4^W4^T3^P$	$1^W1^T1^P$	$2^W2^T4^P$	$6^W6^T6^P$	$3^W3^T2^P$	$5^W5^T5^P$
1950-1960	$2^W3^T3^P$	$1^W1^T1^P$	$4^W4^T4^P$	$5^W5^T6^P$	$3^W2^T2^P$	$6^W6^T5^P$
1960-1970	$3^W4^T3^P$	$1^W1^T1^P$	$4^W5^T5^P$	$6^W6^T6^P$	$2^W3^T2^P$	$5^W2^T4^P$
1970-1980	$3^W5^T3^P$	$1^W2^T1^P$	$4^W3^T5^P$	$6^W6^T6^P$	$5^W4^T4^P$	$2^W1^T2^P$

In the second step of analysis, correlations between the intensity of conflicts and evaluation results were examined. Intensity of conflict was defined as the number of years during which a conflict took place in a given decade (variables IInterX and IintraX, where X denotes a lag in decades). Correlation coefficients are presented in Table 4. * denotes the significance at 0.05 level. Results show that in the case of equal importance of criteria, relationships were not strong but in line with predictions, i.e. negative for the intensity of internal conflicts, positive for the intensity of external ones. The only exception was the TOPSIS method, in case of which no statistically significant relationship between intensity of internal conflicts and evaluations can be seen. In the case of different importance of criteria significant relationships were found only for external conflicts that took place from 2 to 6 decades earlier.

Table 3. F-test results for the differences in ranks, evaluations and occurrence of conflicts.

Conflict	E W	E T	E P	NE W	NE T	NE P
Intra0	-,*	-,-	*,*	-,-	-,-	-,-
Intra1	*,*	-,-	*,*	-,-	-,-	-,-
Intra2	*,*	-,-	-,*	-,-	-,-	-,-
Intra3	*,*	-,-	*,*	-,-	-,-	-,-
Intra4	*,*	-,-	*,*	*,-	-,-	-,-
Intra5	*,*	-,-	-,*	-,-	-,-	-,-
Intra6	*,*	-,-	-,*	-,-	-,-	-,-
Inter0	*,-	-,-	*,-	-,-	-,-	-,-
Inter1	-,-	-,-	-,-	-,-	-,-	-,-
Inter2	-,-	-,-	-,-	-,-	-,-	-,-
Inter3	-,-	-,-	-,-	-,-	-,-	-,-
Inter4	-,-	-,-	-,-	-,-	-,-	-,-
Inter5	-,-	*,-	*,-	-,-	-,-	-,-
Inter6	-,-	*,*	*,*	-,-	-,-	*,-

Table 4. Correlation coefficients for evaluations and intensity of conflicts.

Conflict	E W	E T	E P	NE W	NE T	NE P
IIntra0	-0,200	-0,097	-0,218*	-0,042	-0,050	-0,050
IIntra1	-0,265*	-0,133	-0,277*	-0,098	-0,071	-0,071
IIntra2	-0,379*	-0,128	-0,386*	-0,173	-0,103	-0,14
IIntra3	-0,397*	-0,081	-0,394*	-0,113	-0,012	-0,136
IIntra4	-0,384*	-0,031	-0,351*	-0,142	-0,028	-0,128
IIntra5	-0,349*	0,018	-0,312*	-0,160	-0,063	-0,110
IIntra6	-0,339*	-0,005	-0,305*	-0,181	-0,098	-0,092
IInter0	0,280*	0,172	0,297*	0,109	0,089	0,194
IInter1	0,283*	0,202*	0,281*	0,183	0,167	0,165
IInter2	0,323*	0,245*	0,336*	0,276*	0,258*	0,204*
IInter3	0,325*	0,279*	0,32*	0,336*	0,312*	0,272*
IInter4	0,262*	0,214*	0,288*	0,264*	0,235*	0,251*
IInter5	0,324*	0,322*	0,391*	0,283*	0,261*	0,336*
IInter6	0,350*	0,387*	0,411*	0,311*	0,305*	0,326*

Before 1910 the relationships between international conflicts and evaluations are positive, statistically significant but not strong (Table 5). In case of internal conflicts and equal criteria weights the relationships are often negative, weak and not statistically significant. ANOVA shows the significant impact of international conflicts lagged by maximum 3 decades.

In the period 1910-1940 relationships of evaluations and internal conflicts in case of equal weights of criteria are mostly negative and significant (the only exceptions are results obtained for TOPSIS method) while relationships obtained for international conflicts lagged by 3-5 decades are positive. ANOVA shows significant impact of internal conflicts (Table 6).

After 1940 differences in directions of influence of internal and international conflicts are unclear. Chosen internal conflicts and international conflicts lagged by 3 decades have strong negative impact on the evaluations. Similarly, the influence of international conflicts lagged by 6 decades is strong and positive. Also the ANOVA results show the significant influence of some internal and international conflicts (Table 7).

Table 5. Correlation coefficients for evaluations and intensity of conflicts before 1910.

Conflict	**E W**	**E T**	**E P**	**NE W**	**NE T**	**NE P**
IIntra0	-0.184	-0.016	-0.201	0.018	0.004	0.067
IIntra1	-0.202	0.018	-0.250	0.043	0.069	0.069
IIntra2	-0.345*	-0.001	-0.354*	-0.046	0.024	0.024
IIntra3	-0.333*	-0.080	-0.321*	0.006	0.104	0.000
IIntra4	-0.228	-0.079	-0.231	0.153	0.254•	0.077
IIntra5	-0.123	-0.055	-0.055	0.201	0.231	0.217
IIntra6	-0.125	0.000	-0.040	0.122	0.148	0.285*•
IInter0	0.667*•	0.440*	0.720*•	0.232	0.170	0.400*
IInter1	0.587*•	0.470*	0.596*•	0.346*	0.313*	0.326*
IInter2	0.518*•	0.411*•	0.518*•	0.368*•	0.342*	0.260
IInter3	0.443*•	0.430*	0.425*•	0.389*•	0.345*	0.327*
IInter4	0.277*	0.350*	0.330*	0.222	0.190	0.2400
IInter5	0.349*	0.471*•	0.505*•	0.202	0.156	0.383*
IInter6	0.284*	0.502*•	0.406*	0.195	0.231	0.273*

Table 6. Correlation coefficients for evaluations and intensity of conflicts 1910-1940.

Conflict	E W	E T	E P	NE W	NE T	NE P
IIntra0	-0.315	-0.045	-0.301	-0.071	-0.006	-0.151
IIntra1	-0.418*•	-0.100	-0.369•	-0.33	-0.327	-0.247
IIntra2	-0.453*•	-0.146	-0.457*•	-0.276	-0.224	-0.287
IIntra3	-0.465*•	-0.096	-0.452*•	-0.170	-0.082	-0.220
IIntra4	-0.577*•	-0.192	-0.51*•	-0.396•	-0.337	-0.347•
IIntra5	-0.681*•	-0.302	-0.638*•	-0.52*	-0.458*	-0.461*•
IIntra6	-0.570*•	-0.313•	-0.535*•	-0.417*•	-0.382•	-0.342•
IInter0	-0.236•	-0.368•	-0.315•	-0.101	-0.076	-0.210
IInter1	0.012	-0.109•	0.009	0.024	-0.008	0.012
IInter2	0.264	0.233	0.311	0.292	0.279	0.283
IInter3	0.635*	0.577*	0.678*	0.607*	0.575*	0.638*
IInter4	0.651*	0.550*	0.644*	0.681*	0.653*	0.651*
IInter5	0.437*	0.371	0.402	0.519*	0.517*	0.432*
IInter6	0.310	0.308	0.310	0.399	0.388	0.334

Table 7. Correlation coefficients for evaluations and intensity of conflicts after 1940.

Conflict	E W	E T	E P	NE W	NE T	NE P
IIntra0	-0.147	-0.278	-0.247•	-0.223•	-0.309	-0.335•
IIntra1	-0.442•	-0.693*•	-0.39•	-0.694*•	-0.725*•	-0.631*•
IIntra2	-0.323	-0.541*	-0.301	-0.498*	-0.495*	-0.436
IIntra3	-0.454•	-0.182	-0.473*•	-0.376	-0.306	-0.341
IIntra4	-0.343	0.147	-0.278	-0.226	-0.032	-0.097
IIntra5	-0.096	0.427	-0.071	0.035	0.290	0.064
IIntra6	-0.352	0.126	-0.364	-0.216	-0.008	-0.218
IInter0	-0.046	0.014	-0.010	0.128	0.098	0.143
IInter1	-0.249•	-0.275	-0.309•	-0.124•	-0.202	-0.206
IInter2	-0.207	-0.267	-0.203	-0.096	-0.208	-0.139
IInter3	-0.613*•	-0.498*•	-0.679*•	-0.416•	-0.398•	-0.602*•
IInter4	-0.417•	-0.374	-0.420•	-0.374	-0.363	-0.373•
IInter5	0.095	0.118	0.048•	0.107	0.256	0.005
IInter6	0.709*	0.454	0.667*	0.609*	0.560*	0.521*

Conclusion

Armed conflicts influence not only the material sphere of the economics but also other factors that decide of the country's development level like population health, life expectancy or quality of the environment. Results of the analysis provided in the paper suggest that there is a relationship between evaluations of the country's development and internal conflicts that occurred on its territory. There exists a negative relationship between evaluations and the intensity of internal conflicts and positive relationship in case of external ones. However, detailed conclusions depend on assumptions concerning importance of criteria and specifics of the ranking method. Relationships mentioned above evolved in 3 analyzed periods: before 1910, 1910-1940 and after 1940.

Acknowledgments

Paper is a part 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.

References

Barbieri, K. & Levy, J.S. (1999). Sleeping with the Enemy: The Impact of War on Trade. *Journal of Peace Research*, *36*(4), Special Issue on Trade and Conflict, 463-479.

Brans, J.P., Mareschal, B. & Vincke, P. (1986). How to select and how to rank projects: The PROMETHEE method for MCDM. *European Journal of Operational Research*, *24*(2), 228-238.

Diakoulaki, D., Mavrotas, G, & Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: the CRITIC method. *Computers and Operations Research, 22*(7), 763-770.

Feldman, N. & Sadeh, T. (2018). War and Third-Party Trade. *Journal of Conflict Resolution, 62*(1), 119-142.

Gal, T., Stewart, T. J. & Hanne, T. (1999). *Multicriteria decision making: advances in MCDM Models, Algorithms, Theory, and Applications*. Boston: Kluwer Academic.

Geldermann, J. & Rentz, O. (2000, March, 30-31). *Bridging the gap between american and european MADM-approaches?*. Paper presented at 51st Meeting of the European Working Group Multicriteria Aid for Decisions, Madrid.

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.

Lacina, B. & Gleditsch, N.P. (2005). Monitoring Trends in Global Combat: A New Data-set of Battle Deaths. *European Journal of Population*, *21*(2/3), 145-66.

Lai, B. & Thyne, C. (2007). The Effect of Civil War on Education, 1980-97. *Journal of Peace Research*, *44*(3), 277-292.

Li, Q. & Wen, M. (2005). The Immediate and Lingering Effects of Armed Conflict on Adult Mortality: A Time-Series Cross-National Analysis. *Journal of Peace Research*, *42*(4), Special Issue on the Demography of Conflict and Violence (Jul., 2005), 471-492.

Plümper, T. & Neumayer, E. (2006). The Unequal Burden of War: The Effect of Armed Conflict on the Gender Gap in Life Expectancy. *International Organization*, *60*(3), 723-754.

Sielska, A. (2010). O pewnej propozycji zastosowania rankingów wielokryterialnych w analizie portfelowej. In T. Trzaskalik (Ed.), *Modelowanie preferencji a ryzyko '09* (Prace naukowe, pp. 85-99). Katowice: Akademia Ekonomiczna w Katowicach.

Zanden, J.L. & Baten, J. (2014). *How was life?: global well-being since 1820*. Paris: OECD Publ.

Clio Infra | Reconstructing Global Inequality. Retrieved July 20, 2017, from https://www.clio-infra.eu/

International technology transfer and smart growth of EU countries in 2010-2015

Iwona Skrodzka[1]

Abstract

Smart growth is based on knowledge and innovation. The notion of smart growth, its factors and measuring methods are new categories which emerge from the concept of EU's strategic development objectives. Technology transfer, meanwhile, is a multidimensional process, whose effects include both the implementation and diffusion of technologies in new economic environments. It is regarded as one of the key factors behind diminishing technology gaps and a driver of innovation-based growth. International transfer of technology involves those technologies which have been devised in a country different from that where they are implemented. The purpose of the present paper is to examine the role that international technology transfer plays in the development of EU countries. Owing to the fact that neither of the two categories is measurable, the study uses a soft-modelling method which allows for measuring and analysis of the relationships among the latent variables.

Keywords: *international technology transfer, smart growth, European Union, soft modelling*
JEL Classification: C59, O33
DOI: 10.14659/SEMF.2018.01.47

1 Introduction

Modern economies operate in the conditions of globalization, considerable developmental disparities, and technology gaps. Transfer of technology (TT), as a multidimensional process consisting in implementation and diffusion of technologies in new economic environments is considered to be a key factor which helps narrow development gaps and which determines development based on innovation (Ciborowski, 2015).

Technology can be defined as the ability to apply knowledge for solving practical problems and achieving utilitarian goals, or as a collection of methods and procedures allowing its users to obtain certain resources and transform them into useful products (Kubielas, 2009). A new technology can be an outcome of own R&D efforts, production experience, knowledge derived from relevant literature published elsewhere in the world, purchase of patents, recruitment of human capital, collaboration among enterprises and higher education institutions, takeovers of companies, joint ventures, purchase of licences, know-how, or research contracts (Freeman, 1992). The significance of technology for economic

[1] Corresponding author: University of Bialystok, Department of Econometrics and Statistics, Warszawska 63, 15-062 Bialystok, e-mail: i.skrodzka@uwb.edu.pl.

growth and development is emphasised in both theoretical and empirical models. Historical analyses also indicate that TT, combined with accumulation of domestic technologies, lies at the core of accelerated economic growth (Hoekman et al., 2005; Saggi, 2002).

TT is not a new phenomenon. Authors stress, however, that it is still difficult to agree on a definition of the term, among other things, because of the complex nature of the process and the variety of components that must be taken into account (Ciborowski, 2016). TT can be defined in two ways: narrowly and broadly. As an example of a narrow approach can serve the UNCTAD definition, according to which the process of technology transfer is a mechanism as a result of which, through agreements concluded between parties (a supplier and a recipient), diffusion of technology takes place (UNCTAD, 2005). This interpretation is practically limited to indicating the participants of the process and the necessity of co-operation between them. Meanwhile, the broad approach to defining technology transfer makes it possible to consider the process of creating technology and knowledge, its conveyance to the place where it is implemented, as well as the eventual acceptance and implementation of the technology by the final user (Ciborowski, 2016).

Not only defining of TT, but also its measurement is far from easy. This is caused, among other things by the difficulty in making a clear-cut distinction between its particular stages and effects. Transfer of technology is a phenomenon that has many aspects and dimensions. Its constituent elements frequently overlap, and their consequences are ambiguous. In the majority of studies, authors focus on analysing individual, selected components of technology transfer or, alternatively, on a part of the process which can be distinguished and measured quantitatively. The lack of a universal and comprehensive measure makes macroeconomic analysis and international comparisons rather difficult (Ciborowski, 2016).

International technology transfer (ITT) refers to technologies devised in a country other than that in which they are implemented. There are numerous channels through which technology may be transferred across international boundaries. One major channel is trade in goods, especially capital goods and technological inputs. A second is foreign direct investment (FDI), which may be expected generally to transfer technological information that is newer or more productive than that of local firms. A third is technology licensing, which may be done either within firms or between unrelated firms at arm's-length. Licenses typically involve the purchase of production or distribution rights (protected by some intellectual property right) and the technical information and know-how required to make effective the exercise of those rights. In this regard patents, trade secrets, copyrights, and trademarks serve as direct means of information transfer (Maskus, 2004).

The purpose of this study is to examine the role which ITT plays in the processes of smart growth in the EU countries. Smart growth is based on knowledge and innovation. The notion of smart growth, its factors and measuring methods are new categories which emerge from the concept of EU's strategic development objectives (EU Commission, 2010). Although the concept of smart growth is relatively new, it has already been discussed by other authors, but studies concerning the issue are so far not very numerous. Authors unanimously emphasize that more in-depth research, both of theoretical and empirical nature, is required.

In this paper it was assumed that the level of smart growth can be defined by two aspects:
- by knowledge that is created and developed within the R&D sector,
- by the level of innovation in the economy, which is reflected by the innovative activity of individuals and enterprises.

This paper proposes the following research hypothesis: ITT has a positive impact on the level of smart growth of EU countries. Because of the multi-dimensional and intangible character of both phenomena, a soft-modelling method was applied. The obtained results have allowed the author to verify a research hypothesis, identify the channels of technology transfer that are the most significant for ITT processes in EU states, identify the most crucial aspects of smart growth, and order the countries according to the size of ITT and level of smart growth. The study encompasses the years 2010-2015, the selection of which was determined by the availability of statistical data. Two soft models are estimated: one for the beginning of the studied period (2010), the other for its end (2015).

2 Soft modelling method

The soft modelling method (in the literature also referred to as PLS Path Modeling) was developed by H. Wold (1980). The method makes it possible to investigate relations between variables which are not directly observable (latent variables). The values of such variables cannot be measured in a straightforward manner because of the lack of a widely accepted definition or a uniform method of their measurement.

The soft model consists of two sub-models: an internal one (structural model) and an external one (measurement model). The internal sub-model depicts the relationships between the latent variables on the basis of the assumed theoretical description. The external sub-model defines latent variables by means of observable variables (indicators). Indicators allow for direct observation of latent variables and are selected according to the assumed theory or the intuition of the researcher. A latent variable can either be defined (with the use of indicators) inductively: the approach is based on the assumption that the indicators make up

latent variables (formative indicators), or deductively: when it is assumed that indicators reflect the respective theoretical notions (reflective indicators). Under the deductive approach, the latent variable, as a theoretical notion, is a point of departure for a search of empirical data (the variable is primary to a given indicator). In the inductive approach, it is the indicators that are primary to the latent variable which they comprise. Both the approaches use latent variables that are estimated as the weighted sums of their indicators. However, depending on the definition, indicators should be characterized by different statistical properties – no correlation in the case of inductive definition and high correlation in the deductive one (Rogowski, 1990).

The estimation of the parameters of the soft model is performed by means of the partial least squares method (PLS method). The description of the method can be found in: Boardman et al. (1981), Lohmoller (1988), Westland (2005). The quality of the model is assessed with the use of determination coefficients (R^2), established for each equation. The significance of the parameters is checked by means of the standard deviations calculated with the Tukey's cut method ('2s' rule: a parameter significantly differs from zero if double standard deviation does not exceed the value of the estimator of this parameter). Besides, in the case of the external submodel, the estimators of factor loadings can be treated as the degree in which the indicators match the latent variable that they define. The prognostic property of model can be evaluated by means of the Stone-Geisser test [3], which measures the accuracy of the forecast obtained as a result of the model's application as compared with a trivial forecast. The test statistics take values from the range $<-\infty,1>$. In the ideal model, the value of the test equals 1 (the forecasts are perfectly accurate in comparison with trivial forecasts). When the value of the test equals zero, the quality of the model's forecast and the trivial forecast tend to be virtually identical. Negative values indicate a low quality of the model (its weak predictive usefulness compared with a trivial forecast).

Using the partial least squares method, it is possible to obtain the estimated values of latent variables, which can be regarded as the values of synthetic measures. They can be employed for linear ordering of the examined objects. These values depend not only on the external relationships, but also on the relationships between the latent variables which are assumed for the internal model. This means that the cognitive process hinges not merely on the definition of a given notion, but also on its theoretical description (Rogowski, 1990).

3 Model specification

The model used in the present paper to reach its aim of determining the influence of ITTon the level of smart growth contains the following equation

$$SG_t = \alpha_1 ITT_t + \alpha_0 + \xi \tag{1}$$

where *SG* – the level of smart growth,*ITT* – international technology transfer, α_0, α_1 – structural parameters of the model, ξ – random parameter,*t* – year 2010 or 2015.

The latent variables *ITT* and *SG* are defined by means of observable variables on the basis of the deductive approach, i.e. the latent variable, as a theoretical concept, serves as a starting point to identify empirical data. The statistical data come from the Eurostat and World Bank databases. The indicators for the model were selected based on criteria of substantive and statistical nature. Using the available domestic and international literature, primary sets of indicators of the variables *ITT* and *SG* were developed. The methodologies used comprised, among others, 'Knowledge Assessment Methodology' (Chen and Dahlman, 2005) and 'European Innovation Scoreboard Methodology' (EU Commission, 2017). The developed database was checked in terms of missing data. Data shortages were overcome by using naive prognosis, consisting in replacing a lacking value with an adjacent one. Two countries, Greece and United Kingdom, were excluded from the estimation (due to significant data shortages).

From the statistical point of view, the following considerations were taken into account: variability of indicator values (coefficient of variation above 10%) and analysis of the quality of the estimated model (ex post analysis).

The set of indicators reflecting *ITT*:

*ITT*01 – Foreign direct investment, net inflows as % of GDP (%).

*ITT*02 – High-tech import as % of total (%).

*ITT*03 – Product and/or process innovative enterprises, engaged in any type of innovation co-operation with a partner in EU countries, EFTA or EU candidates countries (% of total).

*ITT*04 – Product and/or process innovative enterprises, engaged in any type of innovation co-operation with a partner in United States (% of total).

*ITT*05 – Product and/or process innovative enterprises, engaged in any type of innovation co-operation with a partner in China or India (% of total).

The set of indicators reflecting *SG*:

*KNOW*01 – Researchers as percentage of total employment (%).

KNOW02 – Researchers in business enterprise sector as percentage of total employment (%).

*KNOW*03 – Graduates in tertiary education, in science, mathematics, computing, engineering, manufacturing, construction per 1000 of population aged 20-29 (person).

*KNOW*04 – Graduates at doctoral level, in science, mathematics, computing, engineering, manufacturing, construction per 1000 of population aged 25-34 (person).

*KNOW*05 – Scientific and technical journal articles per 1 million inhabitants (number).

*INNO*01 – Patent applications to the EPO per 1 million inhabitants (number).

*INNO*02 – Exports of high technology products as a share of total exports (%).

*INNO*03 – Product and/or process innovative enterprises as percentage of total (%).

*INNO*04 – Organization and/or marketing innovative enterprises as percentage of total (%).

*INNO*05 – Total turnover of innovative enterprises as percentage of GDP (%).

*INNO*06 – Charges for the use of intellectual property (receipts) as percentage of GDP (%).

The data on three of the indicators *ITT*01, *KNOW*05, *INNO*06 were obtained from the World Bank database, while all the others from the Eurostat.All the indicators qualified for the model were stimulants of latent variables.

4 Results

Table 1 presents the estimations of factor loadings (parameters of the external model). Because of the adopted deductive approach, estimations of weights are not given.

As was expected, the estimations of the factor loadings for all the indicators were positive. However, not all the indicators proved to be statistically significant. In 2010, six indicators, *ITT*01, *ITT*02, *KNOW*03, *INNO*02, *INNO*04, and *INNO*06 had no statistical significance, while in 2015, only one of them – *KNOW*03.

Both in 2010 and 2015, indicators *ITT*04 and *ITT*05 were strongly or very strongly correlated with the ITT variable[2]. As for the *SG* variable, the correlated indicators included: *KNOW*01, *KNOW*02, *KNOW*05, and *INNO*01. The *ITT* variable was insignificantly (2010) or moderately significantly (2015) reflected by indicators *ITT*01 and *ITT*02. The *SG* variable was insignificantly or weakly (depending on the year) associated with the following indicators: *KNOW*03, *INNO*02, *INNO*04, *INNO*05, and *INNO*06.

[2]The following interpretation of factor loading ρ were adopted: $|\rho| < 0.2$ – no correlation; $0.2 \leq |\rho| < 0.4$ – weak correlation; $0{,}4 \leq |\rho| < 0{,}7$ – moderate correlation; $0.7 \leq |\rho| < 0.9$ – strong correlation; $|\rho| \geq 0.9$ very strong correlation.

Table 1. Estimations of factor loadings in soft model 2010 and soft model 2015.

Symbol	Soft model 2010		Soft model 2015	
of indicator	Factor loading	Standard dev.	Factor loading	Standard dev.
*ITT*01	0.251	0.277	0.402	0.062
*ITT*02	0.205	0.286	0.588	0.048
*ITT*03	0.723	0.252	0.602	0.042
*ITT*04	0.964	0.270	0.929	0.018
*ITT*05	0.957	0.280	0.872	0.033
*KNOW*01	0.834	0.341	0.856	0.043
*KNOW*02	0.956	0.233	0.939	0.048
*KNOW*03	0.066	0.298	0.234	0.144
*KNOW*04	0.609	0.389	0.710	0.028
*KNOW*05	0.866	0.291	0.878	0.025
*INNO*01	0.893	0.290	0.859	0.057
*INNO*02	0.282	0.279	0.393	0.056
*INNO*03	0.628	0.210	0.548	0.114
*INNO*04	0.018	0.210	0.338	0.130
*INNO*05	0.336	0.169	0.494	0.078
*INNO*06	0.072	0.320	0.563	0.069

$$\underset{(0.101)}{\hat{SG}_{2010} = 0.641 \cdot ITT_{2010}} + \underset{(0.855)}{0.587}, \quad R^2 = 0.41, \tag{2}$$

$$\underset{(0.017)}{\hat{SG}_{2015} = 0.667 \cdot ITT_{2015}} + \underset{(0.181)}{1.129}, \quad R^2 = 0.44. \tag{3}$$

Positive values of the estimated parameters of the ITT_t variables are consistent with expectations.Also, both parameters are significantly different from zero (in accordance with the '2s' rule). The values of the coefficient of determination R^2 are not high, suggesting that the independent variables ITT_t determine the variability of the dependent variables SG_tonly to a limited extent. The overall values of the S-G test are positive and stand at 0.087 for the 2010 model and 0.123 for the 2015 model. This proves that the prognostic capacity of the model estimated on the basis of 2015 data is better.

The obtained results indicate a positive, statistically different from zero, dependence between ITT and the level of smart growth in the studied countries in 2010 as well as in 2015. It can be concluded, therefore, that those countries where ITT was more intense were also

characterised by higher levels of smart growth. ITT turns out to be a key factor of innovation growth in the entire studied group. The differences concern only its structure: transfer occurs via other channels and is embodied in different forms.

Table 2. Rankings of EU countries according to levels of international technology transfer and smart growth in 2010 and 2015.

Country	ITT_{2010}	SG_{2010}	ITT_{2015}		SG_{2015}	
Austria	10	7	11	↓	7	•
Belgium	9	8	10	↓	10	↓
Bulgaria	22	25	22	•	25	•
Croatia	19	21	19	•	22	↓
Cyprus	4	22	9	↓	21	↑
Czech Republic	14	14	13	↑	12	↑
Denmark	5	2	3	↑	4	↓
Estonia	17	12	6	↑	17	↓
Finland	1	1	5	↓	3	↓
France	8	11	14	↓	9	↑
Germany	21	4	20	↑	6	↓
Hungary	18	16	15	↑	16	•
Ireland	7	9	1	↑	5	↑
Italy	25	19	25	•	18	↑
Latvia	13	23	17	↓	23	•
Lithuania	15	20	21	↓	20	•
Luxemburg	3	5	7	↓	8	↓
Malta	11	15	16	↓	14	↑
Netherlands	16	6	8	↑	2	↑
Poland	**20**	**24**	**18**	↑	**24**	•
Portugal	24	13	26	↓	13	•
Romania	23	26	24	↓	26	•
Slovakia	12	18	4	↑	19	↓
Slovenia	6	10	12	↓	11	↓
Spain	26	17	23	↑	15	↑
Sweden	2	3	2	•	1	↑

Basing on the estimation of the value of latent variables, rankings of the analysed countries according to ITT size and smart growth level were compiled. The rankings are presented in Table 2.

When comparing the rankings, one notices considerable changes in the ordering of the countries in terms of international technology transfer (Spearman's rank correlation coefficient: 0.82) and relatively small changes in the innovation ranking (Spearman's rank correlation coefficient: 0.96). The following countries moved up spectacularly in the ITT rankings: Estonia (17th in 2010, 6th in 2015), Netherlands (16th in 2010, 8th in 2015), Slovakia (12th in 2010, 4th in 2015), Ireland (7th in 2010, 1st in 2015). Five countries recorded significant drops: Slovenia (6th in 2010, 12th in 2015), France (8th in 2010, 14th in 2015), Lithuania (15th in 2010, 21th in 2015), Cyprus (4th in 2010, 9th in 2015) and Malta (11th in 2010, 16th in 2015).

Poland took one of the last positions among European countries (24th position) in both the SG_{2010} and SG_{2015} rankings. Increasing international technology transfer, in particular a wider range of cooperation with innovative companies from USA, China and India may in the future affect the currently unfavourable situation of Poland in terms of the level of smart growth.

Conclusions

The conducted research has demonstrated that in the years 2010 and 2015, ITT was a significant determinant of smart growth in EU countries. Co-operation with partners from the USA, China and India proved to be the most important ITT channel. The significance of foreign direct investments and high-tech imports was either insignificant or limited. This can be due to the fact that the models took into consideration both highly-developed and less developed countries. The problem, therefore, requires further investigation, particularly in order to examine the impact of ITT on smart growth of highly-developed economies as compared with its influence on less developed ones.

The obtained results also make it possible to conclude that smart growth of EU countries was mainlybased on scientific and R&D activity (researchers, patents, papers in science journals) and that the influence of innovating companies is becoming increasingly important.

The formulated conclusions can be used in practice by government institutions, for example for planning the economy policy as well as innovation policy of countries.

References

Boardman, A. E., Hui, B. S. & Wold, H. (1981). The partial least squares-fix point method of estimating interdependent systems with latent variables. *Communications in Statistics - Theory and Methods, 10(7),* 613-639.

Chen, D. H. C. & Dahlman, C. J. (2005). *The Knowledge Economy, the KAM Methodology and World Bank Operations*. World Bank Institute.

Ciborowski, R. (2016*). Międzynarodowy transfer technologii a innowacyjność krajów Europy Środkowo-Wschodniej*. Bialystok: PTE.

EU Commission (2017). *Innovation Union Scoreboard 2017.*

EU Commission (2010). *EUROPE 2020. A European strategy for smart, sustainable and inclusive growth.*

Freeman, C. (1992). Formal Scientific and Technical Institutions in the NSI. In: Lundvall, B. A. (ed.), *National System of Innovation*, 173-192.

Hoekman, B. M., Maskus, K. E. & Saggi K. (2005). Transfer of technology to developing countries: Unilateral and multilateral policy options. *World Development, 33(10),* 1587-1602.

Kubielas, S. (2009). *Innowacje i luka technologiczna w gospodarce globalnej opartej na wiedzy.* Warsaw: Wydawnictwo UW.

Maskus, K. E. (2004). Encouraging International Technology Transfer.*UNCTAD-ICTSD Project on IPRs and Sustainable Development, 7*.

Lohmoller, J. B. (1988). The PLS Program System: Latent Variables Path Analysis with Partial Least Squares Estimation. *Multivariate Behavioral Research, 23(1),* 125-127.

Rogowski, J. (1990). *Modele miękkie. Teoria i zastosowanie w badaniach ekonomicznych.* Bialystok: Wydawnictwo Filii UW w Białymstoku.

Saggi, K. (2002). Trade, foreign direct investment, and international technology transfer: a survey. *World Bank Research Observer, 17 (2),* 191-235.

UNCTAD (2005). International Investment Agreements: Key Issues, 3.

Westland, J. C. (2015).Partial Least Squares Path Analysis. In: Westland, J. C., *Structural Equation Models,* 23-46.

Wold, H. (1980). Soft modelling: Intermediate between Traditional Model Building and Data Analysis. *Mathematical Statistics, 6(1),* 333-346.

Are uncertainty measures powerful predictors of real economic activity in the Euro Area?

Sławomir Śmiech[1], Aleksandra Mucha[2]

Abstract

In the paper we propose and collect measures that are related to different aspects of economic uncertainty and examine their predictive power for real economic activity in the Euro Area. The set of uncertainty measures investigated includes: the European News Index, Economic Policy Uncertainty index, uncertainty indices related to the global financial and commodity markets, and the Euro Area industry uncertainty measures calculated using business surveys provided by the European Commission. Macroeconomic activity indicators used in the study describe production, investment and unemployment. The analysis covers quarterly data from 2000 to 2014. The study is based on the rolling scheme and different specifications of VAR models. Two main conclusions can be drawn. First, uncertainty measures are, to a great extent, independent of one another. Second, some uncertainty indices perform well in forecasting economic activity indicators. It seems, however, that the best strategy is to apply forecast averaging techniques and to combine information provided by all uncertainty indices.

Keywords: *uncertainty measures, the Euro Area activity, forecasting*
JEL Classification:C53, E27, D81
DOI:10.14659/SEMF.2018.01.48

1 Introduction

Economic uncertainty is an important economic category in both theoretical analyses (Ellsberg, 1961; Dequech, 1999) and empirical studies (Bloom, 2009; Baker et al., 2016; Kang et al., 2016). It seems obvious that high uncertainty affects the decisions made by consumers, who restrict their purchases, and entrepreneurs, who limit hiring and investing, which leads to a decrease in production and employment.

The consensus about the definition of economic uncertainty has not been reached so far: in the Knightian tradition it is "non-quantitative" (Knight, 1921), while in more recent literature attempts are made at measuring uncertainty, and the most popular approaches include: economic and political uncertainty (measured with the EPU index) (Baker et al., 2016); financial uncertainty (measured with the VIX index and the Financial Stress Index STRESS), uncertainty related to credit risk or uncertainty related to expectations of entrepreneurs

[1]Corresponding author: Cracow University of Economics, Department of Statistics, Rakowicka 27, 31-510 Kraków, Poland, e-mail: smiechs@uek.krakow.pl.
[2]Cracow University of Economics, Rakowicka 27, 31-510 Kraków, Poland, email: aleksandra.mucha94@gmail.com

(Bachmann et al., 2013). Although all these measures are expected to quantify the same economic phenomenon, it is not clear to what extent they are independent of one another and how accurate they are in explaining actual economic activity.

The paper has two objectives: first, to collect and compare uncertainty measures which are related to European economy, and, second, to analyse their predictive power with reference to economic activity in Europe. The set of uncertainty measures includes seven indices, which reflect different aspects of uncertainty: (1) economic policy uncertainty measured with the European News Index (ENI) and the Policy Uncertainty Index (EPI); (2) industrial sector uncertainty measured with the Euro Area Industrial uncertainty indices (INDU, INDU_f); (3) uncertainty in financial markets measured with the common volatility of stock exchange indices in the Euro Area (FIN_EA) and the index of common volatility of stock exchange indices (FIN); (4) uncertainty in the global commodity markets measured with the index of common volatility of prices on these markets (COMU). Economic activity in Europe is measured with: industrial production (PROD), employment (EMPL) and gross fixed capital formation (INVEST). All macroeconomic variables are seasonally adjusted.

The study covers quarterly data spanning the period from January 2000 to January 2014[3], which yields 65 observations. The data are obtained from the Eurostat database, European Commission website, Economic Policy Uncertainty website, Word Bank, and Thompson Reuters DataStream.

The empirical strategy consists of two steps. In the first one, the uncertainty indices are collected and analysed. Two measures (EPI, ENI) are taken from the Economic Policy Uncertainty website, and the reaming uncertainty indices are estimated. As the measures describe different aspects of uncertainty, their mutual relations are investigated by analysing correlations and Granger causality between the measures. In the second step, the predictive power of the uncertainty indicators regarding the US are compared to predicted economic activity in Europe. The comparison is performed by estimating two-dimensional VAR models with a number of lags from 1 to 4 and different settings of a deterministic trend. The models include one variable relating to economic activity (PROD, INVEST, UNEMP) and one of the uncertainty measures (EPU, ENI, INDU, INDU_f, FIN_EA, FIN, COMU). Each of the models (21 models in total) is used to determine the forecasts for one period ahead in the rolling forecasting scheme. The rolling windows include observations of 31 consecutive quarters. The first forecasts are obtained for Q4 2007, which means that the largest recent

[3]There are two reasons to limit the dataset until 2014: first, the EPU index avulability, and second, the Euro Zone countries seem to be recovering since the end of 2013.

slowdown of economic activity in Europe, related to the global financial crisis, is taken into account. The last forecasts are obtained for the first quarter of 2014. The forecasts obtained from the VAR models are compared with the forecasts obtained with benchmark models, which include AR(1), AR(4) and naive forecasts. Accuracy of forecasts is assessed using RMSE and MAPE. Forecasts of trend changes are also analysed by comparing forecast errors with differences of original variables.

The contributionof the paper is twofold. First, new measures of uncertainty which are related to different aspects of uncertainty in Europe are proposed, and a detailed comparison of these measures is made. Second, different aspects of uncertainty in prediction of economic activity in the Euro Area are analysed. Both the mutual relations between different aspect of uncertainty and the utility of uncertainty indices in predicting the main macroeconomic indicators are of great importance for both theoretical and practical considerations.

The study consists of 4 sections. The introduction is followed by a section presenting uncertainty measures and another one reporting the empirical results, while conclusions are drown in the last section.

2 Uncertainty measures

A1. *Euro Area industrial uncertainty (INDU)*

In order to measure industrial uncertainty in the Euro Area, we follow two approaches proposed by Bachmann et al. (2013) using the survey data for the Euro Area provided by the European Commission[4]. The first approach is based on the assumption that heterogeneity of economic sentiment surveys reflects the dispersion of expectations that can be used as the proxy for macroeconomic uncertainty. The greater the heterogeneity of the survey responses, the larger uncertainty. On the other hand, high homogeneity of managers' expectations reflects high confidence and low uncertainty. As the survey results include the answer to the question "Q5: *Production expectations for the months ahead*", we calculate standard deviation of the aggregate answers (the difference between answers "increase" and "decrease") for different industrial sub-sectors:

$$INDU_t = sd\left(Q5_{t,i}\right), i = 1,..23, t = 1,..217,$$

where: $Q5_{t,i}$ indicates answers to Q5 question obtained for the i^{th}industrial subsector in the period *t*.

[4]https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/business-and-consumer-surveys_en

The second approach assumes that uncertainty can by related to the forecast error of the expectations. Thus, we compare answers to two questions: Q5 (previously used) obtained in the period *t* and Q1: "*Production trend observed in recent months*" in the period *t*+3. High uncertainty appears if production expectations are not in line with production observed. In order to measure uncertainty in the period *t*, answers regarding different industrial subsectors are used in the following way:

$$INDFU_t = mean\,|(Q5_{t,i} - Q1_{t+3,i})|, i = 1,..23, t = 1,..217.$$

A2. *Policy uncertainty indices*

The European News Index (ENI) and the Economic Policy Uncertainty index measure the European policy-related economic uncertainty based on newspaper articles and are obtained from the website of Economic Policy Uncertainty. Data collected from 10 newspapers in 5 European countries: France, Germany, Italy, Spain, and the United Kingdom, are developed by Baker et al. (2016).[5]

A3. *Global commodity market uncertainty index (COMU)*

The global commodity market uncertainty index (COMU) is estimated for 24 commodities[6], which represent the following commodity market sectors: agriculture, livestock, energy, industrial metals and precious metals. Commodities which have been selected are believed to be both sufficiently significant to the world economy and are tradable through futures contracts. With the exception of several metals contracts (aluminium, copper, lead, nickel and zinc) whichtrade on the London Metals Exchange ("LME"), and the contract for Brent crude oil and gas oil, the remaining commodities are the subject of trade on different U.S. exchanges. The details regarding the underlying futures contracts and the exchanges in which they are traded are available on request. The data are obtained from Thomson Reuters DataStream.

In order to measure uncertainty related to the global commodity market, we apply methodology proposed in Kang et al. (2016). The proposition assumes that uncertainty can be estimated as a common component of volatility of returns on financial instruments. Calculating the global commodity market uncertainty requires three steps. First, commodity returns are calculated as:

[5]See http://www.policyuncertainty.com/europe_monthly.html.
[6] Details are available from the authors on request.

$$r_{c,t} = ln\frac{y_{c,t}}{y_{c,t-1}}$$

where $y_{c,t}$ denotes the average monthly price for a given commodity *c*, in period $t= 1,..,T$. Next, volatility proxy for each commodity *c* is calculated as:

$$V_{c,t} = \left(r_{c,t} - \bar{r}_{c,t}\right)^2,$$

where $\bar{r}_{c,t}$ is the sample mean of $r_{c,t}$.

Given the data matrix (24xT dimension) with $V_{c,t}$ for 24 commodities in the final step, the principal component for the correlation matrix is estimated. The first principal component is used as the global commodity market uncertainty proxy.

A4. *Financial market uncertainty index (FIN, FIN_EA)*

The global financial market uncertainty index (FIN) is estimated for the main indices of 10 largest stock exchanges in the world in terms of market capitalization. The data are obtained from Thomson Reuters Datastream.

The global financial market uncertainty index is calculated in the same way as the global commodity market uncertainty index, which means that common volatility of returns is calculated using the principal component analysis.

The Euro Area financial market uncertainty index (FIN_EA) is based on the same idea, but uses equity indices of 10 largest stock exchanges in the Euro Zone countries.

3 Empirical results

The first step of the study is dedicated to the analysis of mutual relations between the uncertainty measures. Fig. 1 presents standardized series of uncertainty measures[7]. The results presented reveal a rather moderate similarity between the indices. Most shocks identified, similar to Jurado et al. (2015)[8], are *specific* to the particular aspect of uncertainty. In general, there is no overlap between shocks. The only exception is the outbreak of the global financial crisis in which different uncertainty shocks occur simultaneously. Correlations between the uncertainty indices are positive, yet moderate for most pairs (see Fig. 1 the right panel). Correlations larger than 0.5 are obtained only for (NEWS, EPU), (FIN, FIN_EA) and (FIN, INDU_f) pairs. Granger causality for each pair of uncertainty indices is examined as well. *p*-values of the tests are presented in Table 1. There is no Granger causality between most pairs,

[7]The range of the series is limited to 2014 as EPI index is available up to this date.

[8]An uncertainty shock is an event in which a given uncertainty index deviates from its mean level by more than 1.65 standard deviation.

which means that past values of one uncertainty index cannot reduce forecast variance of another uncertainty index. There are, however, some exceptions. FIN Granger causes both industrial uncertainty measures. INDU_f is influenced (in the Granger sense) by almost all other uncertainty aspects.

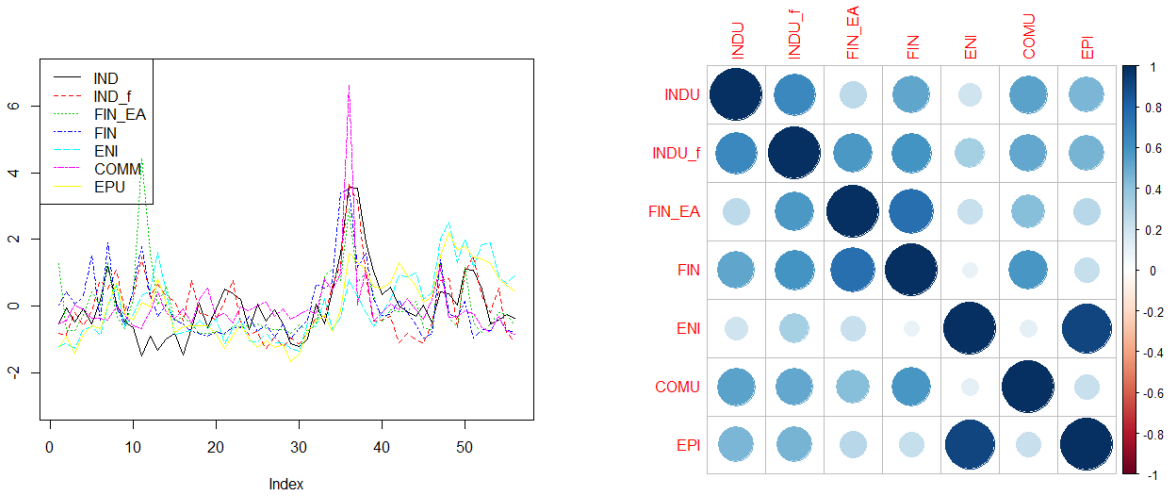


Fig. 1. Standardized uncertainty indices (left) and the correlation between them (right).

Table 1. Granger causality test for uncertainty measures (p-values).

Cause\Effect	EPU	FIN_EA	ENI	INFU	INFU_f	FIN	COMU
EPU		0.485	0.305	0.237	**0.003**	**0.055**	0.965
FIN_EA	0.608		0.501	0.056	**0.014**	**0.038**	**0.000**
ENI	0.250	0.825		0.187		0.262	0.872
INDU	0.732	0.171	0.896		**0.013**	0.163	0.378
INDU_f	0.201	0.931	0.059	0.139		0.565	0.791
FIN	0.367	0.509	0.446	0.469	**0.044**		0.769
COMU	0.617	0.217	0.446	0.509	**0.002**	**0.017**	

Note: bold numbers indicate statistically significant relations.

In the next step, the predictive power of uncertainty measures for forecasting economic activity in the Euro Area is examined. Forecasts for the next quarter are generated using VAR(1) and VAR(4) models with or without trend component within the rolling forecasting scheme. The first forecasts are obtained for the fourth quarter of 2007, the last for the fourth quarter of 2013. The smallest errors are obtained for the most parsimonious specification i.e.

VAR(1) without trend[9]. Owing to the limited space, only the results obtained with this specification are discussed here. The results obtained for the VAR models are compared with the set of benchmark models, which includes: AR (1), AR (4) and the naive model. Fig.2 shows a series of forecast errors for each macroeconomic variable (the panel on the left), and the comparison of forecast errors with changes in the direction (difference) of the predicted variable. Two remarks should be made here. First, a high variety of forecast errors is observed. For almost all analysed periods there are models that produce both negative and positive forecast errors. The largest forecast errors are obtained for the periods between 2008 and 2009, which results from the global financial crises (a structural change is observed for uncertainty as well as macroeconomic variables). Second, there is no correlation between forecast errors and the change of activity in the Euro Area when all the models are considered. Both above observations suggest thata forecast combination could serve as a potentially effective tool for predicting macroeconomic variables.

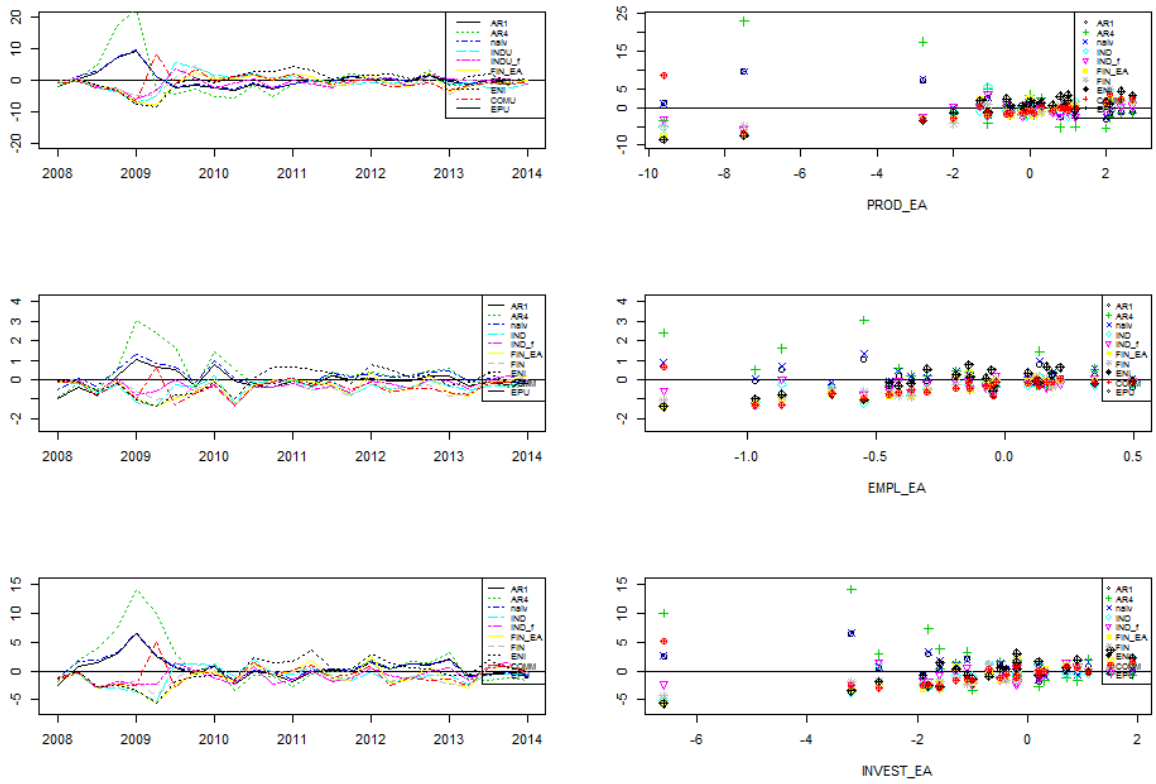


Fig.2 Forecast errors and the relations between the difference of macroeconomic variables and forecast errors.

[9]The results of the remaining specifications are available from the authors on request.

Forecast error measures for all models are presented in Table 2.The last two columns show the average rank of the models regarding RMSE and MAPE. The last row presents the results obtained for the combination of forecasts obtained from seven VAR models. Three observations can be made here. First, different uncertainty measures provide different forecasting accuracy. The smallest forecast errors are obtained for VAR models that include INDU_f and EPU indices. At the opposite end there are models that use ENI index. Second, the benefits of using uncertainty measures to forecast economic activity in the Euro Area are not obvious. AR(1) model turns out to be the best (in comparison to VAR models) in predicting employment. This model performs well when MAPE is taken as the criterion for the two reaming variables. Third, forecast combinations (the last row in Table 2) obtained as the average of forecasts from VAR models generate forecasts with the smallest errors (for all variables when MAPE is taken into account).

Table 2. Forecast error measures obtained for benchmark models and models incorporating uncertainty measures.

	PROD		**EMPL**		**INVEST**			
	RMSE	**MAPE**	**RMSE**	**MAPE**	**RMSE**	**MAPE**	**Rank 1**	**Rank 2**
AR1	2.78	1.7	**0.44**	**0.24**	1.83	1.19	4.7	3.00
AR4	6.35	3.55	0.96	0.46	4.14	2.53	11.00	11.00
naiv	2.77	1.63	0.46	0.24	1.87	1.25	6.33	3.00
INDU	2.68	1.99	0.54	0.29	1.84	1.38	6.00	7.00
INDU_f	***1.97***	*1.42*	0.46	*0.25*	1.47	1.22	2.33	3.00
FIN_EA	2.66	1.88	0.6	0.33	2.05	1.58	7.67	8.67
FIN	2.22	1.69	0.56	0.31	1.69	1.35	4.67	5.67
ENI	2.91	2.05	0.56	0.32	2.05	1.53	8.67	8.67
COMU	2.67	1.82	0.59	0.33	1.86	1.43	7.00	7.67
EPU	2.81	2.2	**0.44**	*0.25*	1.67	1.31	4.33	6.33
FC	2.02	**1.39**	0.45	**0.23**	**1.43**	**1.09**	**2.00**	**1.00**

Notes: bold numbers indicate the models that yield the most accurate forecasts; numbers in italics indicate the best model among models using the uncertainty measures; Rank 1 (Rank 2) show an average rank obtained for models regarding RMSE (MAPE) for three macroeconomic indicators.

Conclusion

The study proposes measures of uncertainty, next examines their mutual relations, and finally assesses their predictive power in forecasting real activity in the Euro Area between 2000 and 2014. The results obtained reveal that individual uncertainty measures are - to a great extent - independent and seem to be related to different aspects of economic uncertainty. The predictive power of uncertainty measures is varied as well. Uncertainty in the industrial sector turns out to be the best predictor of real production and investment in the Euro Area. The economic policy uncertainty index is the best predictor of employment. Nevertheless, the information provided by single measures of uncertainty does not seem to be enough to beat all benchmark models.

Considerable independence of the uncertainty measures makes the forecast combination useful. The forecasts calculated as the average of forecasts obtained from all models using uncertainty measures turn out to be the most effective in predicting the real activity in the Euro Area.

Acknowledgements

The authors gratefully acknowledge financial support from the National Science Centre in Poland (grant no 2017/25/B/HS4/01058).

References

Bachmann, R., Elstner, S. & Sims, E. R. (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics*, *5*(2), 217-49.

Baker, S. R., Bloom, N. & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, *131*(4), 1593-1636.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, *77*(3), 623-685.

Dequech, D. (1999). Expectations and confidence under uncertainty. *Journal of Post Keynesian Economics*, *21*(3), 415-430.

Diebold, F. X. & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal*, *119*(534), 158-171.

Diebold, F. X. & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, *28*(1), 57-66.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, *75*(4), 643-669.

Jurado, K., Ludvigson, S. C. & Ng, S. (2015). Measuring uncertainty. *American Economic Review*, *105*(3), 1177-1216.

Kang, W., Ratti, R. A. & Vespignani, J. L. (2016). *Global uncertainty and the global economy: Decomposing the impact of uncertainty shocks.*

Knight, F. H. (1921). *Risk, uncertainty and profit.* New York: Hart, Schaffner and Marx.

Comparison of income poverty and social exclusionin the EU in 2008 and 2016

Erik Šoltés[1], Tatiana Šoltésová[2]

Abstract

Poverty, material deprivation and joblessness are serious problems to which the European Union still has to pay close attention since, according to the European Commission, meeting the Europe 2020 strategy goals in the area of poverty and social exclusion seems improbable. The aim of the article is to map a spatial distribution of income poverty and social exclusion from point of view of three-dimensional concept including poverty, material deprivation and joblessness in EU-28 in 2016 (the most recent available data from EU-SILC survey and selected statistics provided by Eurostat). For that purpose, multivariate statistical methods were used, such as correlation analysis, factor analysis and cluster analysis. Results gained for the year 2016 are compared to the reference year 2008 (the most recent data available when the target for Europe strategy 2020 was adopted (in 2010)). The paper puts emphasis on the visualisation of results obtained by statistical methods, therefore, the analyses were carried out by means of SAS JMP.

Keywords: *poverty and social exclusion, income poverty, material deprivation, joblessness, cluster analysis*
JEL Classification: I32, C38, E24
DOI: 10.14659/SEMF.2018.01.49

1 Introduction

Combating against poverty and social exclusion is one of the headline targets of Europe 2020 strategy. This strategy for smart, sustainable and inclusive growth is approaching its final year so we decided to compare the conditions of income poverty and social exclusion in the EU member countries in 2016 (the most recent available data from EU-SILC survey) and in 2008. In order to assess poverty and social exclusion, Europe 2020 strategy uses 3-dimensional concept which take into account three dimensions: income poverty, material deprivation and labour market exclusion. These three negative social phenomena influence one another. In recent 10 years more studies have appeared that evaluate a one dimension of poverty and social exclusion in relation to other dimensions rather than in isolation. From scientific works that analyse relation between poverty and material deprivation or even deal with consistent poverty, we were inspired by (Guio and Maquet,2006; Labudová et al.,2010; Nolan and

[1] University of Economics in Bratislava, Faculty of Economic Informatics, Department of Statistics, Dolnozemskácesta 1, 852 35 Bratislava, Slovakia, e-mail: erik.soltes@euba.sk
[2] University of Economics in Bratislava, Faculty of Economic Informatics, Department of Mathematics and Actuarial Science, Dolnozemskácesta 1, 852 35 Bratislava, Slovakia, e-mail: tatiana.soltesova@euba.sk

Whelan, 2010; Želinský, 2010).The influence of labour market exclusion or low work intensity of households on poverty was proven for example in papers (Guagnano et al., 2013; Mysíková et al., 2015; Kis and Gábos, 2016).Ayllón and Gábos (2015) and Řezankováand Želinský(2014) confirmed the impact of very low work intensity and joblessness of households on material deprivation in Central and Eastern Europe and the Czech Republic, respectively.

The article maps and compares the conditions of income poverty and social exclusion in the member countries of EU in 2008 (the year 2008 is the reference year for strategy Europe 2020) and in 2016. As the partial indicators (at-risk-of poverty rate, severe material deprivation rate, very low work intensity rate) of the aggregate indicator AROPE (at risk of poverty or social exclusion) map "only" the occurrence of income poverty and social exclusion but not the depth of those negative phenomena, we decided to also use some indicators characterising the severity of poverty and social exclusion in the EU to create a more objective and more complex picture. Each dimension of poverty and social exclusion was captured in the paper by means of 4 indicators. For the dimension of *Income poverty and income inequality* we used the following indicators: the *at-risk-of poverty rate after social transfers* (AROP), the *relative median at-risk-of-poverty rate gap* (PG), the *income quintile share ratio* or*S80/S20 ratio* (S80-S20), the *persistent at-risk-of poverty rate* (Persistent_P). The dimension of *Material deprivation* was represented by the following indicators: the *material deprivation rate* (MD), the *severe material deprivation rate* (SMD), the *mean number of deprivation items among the deprived* (Depth_MD), the *severe housing deprivation rate* (Housing_D).The dimension of *Exclusion from labour market* was represented by the following indicators: the *unemployment rate* (UR), the *long-term unemployment rate, % of active population aged 15-74* (Long_term_U), the *jobless households rate* (Jobless_H), the *very low work intensity rate* (VLWI).

2 Analysis of source variable dependence and data preparation for cluster analysis

As it was mentioned in the introduction of the article there are many scientific studies that confirmed significant relationships between dimensions of poverty and social exclusion. For this reason, it is not surprising that most of observed indicators, especially those ones that belong to the same dimension, are mutually dependent (Fig. 1).

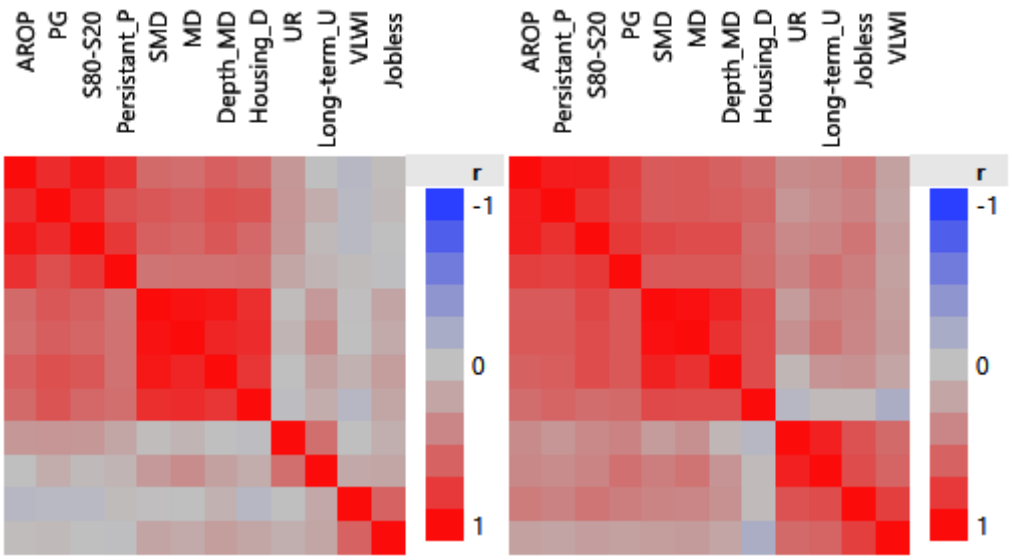


Fig. 1. Correlation maps of source indicators for 2008 (on the left) and 2016 (on the right).

Source: Eurostat, self-processed in SAS JMP.

In both analysed years we can notice weaker relationship between unemployment rates on the one side and very low work intensity rate or joblessness of households on the other side than between the unemployment rates themselves or between indicators of labour market exclusion of households. It is particularly visible for the year 2008. On the one hand, the analysed indicators characterize poverty and social exclusion from various perspectives, on the other hand, those perspectives more or less overlap as the significant correlations among the monitored indicators testify. As a result, for the purpose of the cluster analysis, the set of original indicators had to be redesigned into a set of new, mutually independent variables. Factor analysis was implemented to serve that purpose. We attempted to create such factors that would be determined by those source indicators which would facilitate their interpretation. Simultaneously, we wanted to decrease the number of dimensions, i.e. to achieve a reduced number of factors compared to the original indicators while those factors would still carry at least 85% of information provided by the original indicators.

To assess the suitability of source indicators for the factor analysis, we applied the Kaiser-Meyer-Olkin measure (Stankovičová and Vojtková, 2007). The KMO statistics (Table 1) showed excellent suitability of the source variables for factor analysis.

If we applied the Kaiser's rule for eigenvalues in correlation matrices which states that only factors with eigenvalues greater than average eigenvalue should be used (the average eigenvalue of a correlation matrix is 1) then we would consider 4 factors in 2008 and 2 factors

in 2016 (Fig. 2). In order to obtain comparable results, we decided to set the number of factors to 4.

Table 1. Values of Kaiser-Meyer-Olkin measure for source indicators.

Kaiser's Measure of Sampling Adequacy: Overall MSA = 0.88382018											
AROP	**PG**	**S80-S20**	**Persistent_P**	**SMD**	**MD**	**Depth_MD**	**Housing_D**	**UR**	**Long-term_U**	**VLWI**	**Jobless_H**
0.885	0.925	0.822	0.949	0.855	0.870	0.830	0.923	0.857	0.872	0.926	0.913

Source: Eurostat, self-processed in SAS EG.

Number	**2008**			**2016**		
	Eigen-value	**Percent**	**Cum Percent**	**Eigen-value**	**Percent**	**Cum Percent**
1	**6.0453**	**50.377**	**50.377**	**6.8598**	**57.165**	**57.165**
2	**1.8684**	**15.570**	**65.948**	**2.5375**	**21.146**	**78.311**
3	**1.5201**	**12.668**	**78.615**	**0.9342**	**7.785**	**86.096**
4	**1.1729**	**9.774**	**88.389**	**0.6416**	**5.346**	**91.442**
5	0.5213	4.345	92.734	0.2781	2.317	93.759
6	0.2889	2.408	95.141	0.2633	2.194	95.953
7	0.2186	1.821	96.963	0.1945	1.621	97.574
8	0.1818	1.515	98.478	0.1332	1.110	98.684
9	0.0749	0.625	99.102	0.0876	0.730	99.414
10	0.0617	0.514	99.616	0.0429	0.357	99.771
11	0.0384	0.320	99.937	0.0234	0.195	99.966
12	0.0076	0.063	100.000	0.0041	0.034	100.000

Fig. 2. Eigenvalues of the correlation matrices (PCA method) for 2008 and 2016.

Source: Eurostat, self-processed in SAS JMP.

After obbiquartimax rotation we obtained factor loadings shown in Table 2. Based on those factor loadings, we found out that the 1st factor had strong positive correlation with the indicators of material deprivation, the 2nd factor demonstrated strong positive correlation with the indicators of income poverty and income inequalities, the 3rd factor showed strong positive correlation with the indicators of labour market exclusion of inhabitants (unemployment rate and long-term unemployment rate) and the 4th factor was characterized by mostly labour market exclusion of households and had moderate positive correlation with the very low work intensity rate and the jobless households rate.

Table 2. Factor loadings after obbiquartimax rotation for 2008 and 2016.

	2008				2016			
	Factor				Factor			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
AROP	-0.033	**0.997**	0.046	0.017	0.056	**0.969**	-0.017	-0.011
PG	0.194	**0.584**	0.167	0.085	0.251	**0.736**	0.074	0.043
S80-S20	0.179	**0.772**	0.010	0.154	0.133	**0.895**	0.013	0.022
Persistant_P	0.039	**0.904**	0.018	0.001	0.101	**0.775**	0.080	-0.073
SMD	**0.926**	0.043	0.093	0.102	**0.933**	0.107	0.015	0.061
MD	**0.901**	0.045	0.272	-0.044	**0.943**	0.052	0.152	-0.064
Depth_MD	**0.806**	0.127	-0.233	0.293	**0.776**	0.264	-0.093	0.254
Housing_D	**0.677**	0.256	-0.062	-0.270	**0.782**	0.179	-0.029	-0.064
UR	-0.105	0.104	**0.912**	0.095	-0.231	0.299	**0.717**	-0.015
Long-term_U	0.182	-0.037	**0.780**	0.212	0.276	-0.259	**0.830**	0.120
VLWI	0.001	-0.029	0.118	**0.707**	-0.109	-0.010	0.007	**0.651**
Jobless	0.002	0.192	0.174	**0.702**	0.082	0.007	0.026	**0.582**

3 Cluster analysis of EU member countries in terms of income poverty and social exclusion in 2008 and 2016

The factor analysis resulted in 4 mutually independent factors, each representing one dimension of poverty and social exclusion. These factors were appropriate for the cluster analysis with the aim to create clusters of EU member countries where the countries falling into a common cluster would be most similar in terms of poverty and social exclusion while the countries in different clusters would be significantly different. Using Ward's method (Hebák et al., 2005) which due to its excellent results belongs among the most popular hierarchical procedures (Loster and Pavelka, 2013), we obtained a dendrogram in Fig. 3. The dendrogram is supplemented by colour maps of the 4 factors. The colour map in the 1st column refers to the 1st factor representing the material deprivation dimension, the colour map of the 2nd factor representing the dimension of income poverty and income inequality is shown in the 2nd column, and in the 3rd and 4th column we can find the colour map of the 3rdand 4thfactor characterising labour market exclusion of inhabitants and households, respectively.

In 2008, *Cluster 1* includes Belgium, Germany, Ireland and France and is characterized by the highest labour market exclusion of households. Comparable poor values of indicators from this dimension were recorded in Bulgaria and Italy, as well. Cluster 1 achieved above-average good results in material deprivation.

Cluster 2 is created by three countries out of V4 countries, specifically Hungary, Poland and Slovakia. For the countries of this cluster was typical high threat of material deprivation. In 2008 the highest material deprivation across EU-27 was in Bulgaria and Romania followed by countries of Cluster 2. Slovakia recorded high unemployment rate and long-term unemployment rate (3rd factor). On the other hand, Slovakia and Hungary as well as the Czech Republic (from Cluster 5) achieved the best results within 1st dimension (2nd factor) – income poverty and income inequality.

Cluster 3 includes Baltic States (Estonia, Lithuania and Latvia), most countries of Southern Europe (Greece, Italy, Spain, Portugal) and the United Kingdom. These countries manifested above-average risk of income poverty and income inequality. Latvia even had significantly the worst situation in this dimension. Although countries of Southern Europe did not create a separate cluster, we can see that already in 2008 these countries were subject to a larger labour market exclusion than the rest of Cluster 3. While Portugal, Spain and Greece were threatened by high unemployment rates, Italy had trouble with high occurrence of households with very low work intensity and jobless households.

Cluster 4 is very specific due to extremely high social exclusion for reason of material deprivation. In addition, Bulgaria and Romania, which belong to this cluster, had to face a considerable income poverty and income inequality in 2008.

Overall, *Cluster 5* achieved the best results in area poverty and social exclusion. To Cluster 5 were merged up to 9 countries (the Czech Republic, Slovenia, Luxembourg, Malta, Denmark, the Netherlands, Austria, Finland and Sweden) so this cluster naturally shows some heterogeneity. Although most of countries had positively low values of factors as well as original indicators, in Slovenia and the Czech Republic we can observe slightly higher material deprivation but relatively very low income poverty and income inequality. A specific situation was in Cyprus which created a separate cluster (*Cluster 6*) with the lowest exclusion of households from labour market. Moreover, Cyprus had also good condition in other dimensions of poverty and social exclusion. Several states (mainly from Southern Europe) have failed to recover from the economic crisis yet and this was also revealed by cluster analysis for 2016.

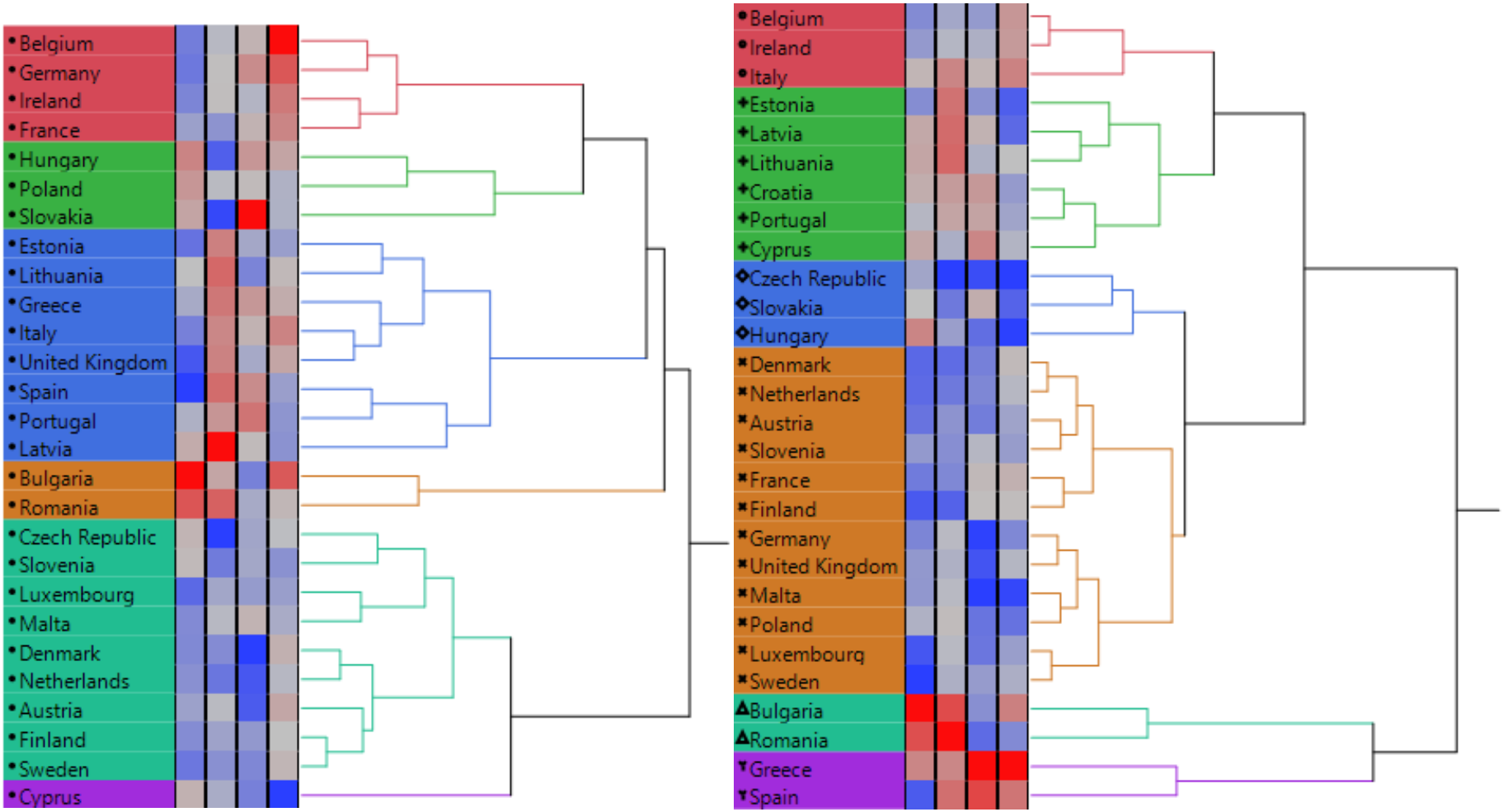


Fig. 3. Dendrogram of EU country clusters according to poverty and social exclusion factors in 2008 (on the left) and 2016 (on the right).

In 2016, *Cluster 1*consists of Belgium, Ireland and Italy and is characterized by relatively high exclusion of households from labour market.

Cluster 2 similarly like Cluster 3 from the year 2008 includes Baltic States and most countries of Southern Europe. But the only common country of Southern Europe for these two clusters is Portugal though. To Cluster 2 also belong Croatia and Cyprus. This cluster recorded relatively high income poverty and income inequality, especially in Baltic States.

Cluster 3 is very similar to Cluster 2 from the year 2008 and consists of three out of V4 countries but this time the cluster consists of Hungary, Slovakia and the Czech Republic. This cluster achieves the lowest exclusion of households from labour market. The leader of the group is the Czech Republic which does not have such problems with material deprivation as Hungary and neither problem with unemployment of population as Slovakia.

The most populous among the other clusters is *Cluster 4* that consists of 12 EU-28 member countries (Denmark, the Netherlands, Austria, Slovenia, France, Finland, Germany, the United Kingdom, Malta, Poland, Luxembourg and Sweden). This cluster together with Cluster 3 manifests the lowest risk of poverty and social exclusion. On a basis of the dendrogram in Fig. 3 Cluster 4 could be divided into 2 sub-clusters. The first sub-cluster includes first 6 abovementioned countries that are characterized by a bit higher degree of labour market exclusion but lower income poverty and income inequality than the second group of six countries.

The remaining 2 clusters were much more affected by poverty and social exclusion than others. Bulgaria and Romania which created *Cluster 5*, have in 2016 equally like in 2008 a significantly worst condition in area of material deprivation despite progress they have made since 2008. Moreover, these 2 countries recorded negative trend in area of income poverty and their inhabitants have to face the largest risk of poverty, persistent poverty and the highest income inequalities.

Cluster 6 includes Greece and Spain. In 2016, in both countries we can observe the largest labour market exclusion of population as well as households. Furthermore, both countries have experienced a deteriorating situation in the area of income poverty and income inequality. Greece also reached poor results in indicators of material deprivation and followed only by Bulgaria and Romania. In contrast with Greece, in 2016 Spain recorded a relatively satisfactory incidence and depth of material deprivation. Differences in the dimension of material deprivation between Spain and Greece caused that Cluster 6 was created last of all the clusters.

Conclusions

The paper evaluates and compares poverty and social exclusion in EU member countries based on statistical analyses of selected indicators in 2008 and 2016. Multidimensional statistical methods were used for that purpose, such as correlation analysis, factor analysis and cluster analysis.The correlation analysis confirmed the strong dependence among indicators included in each dimension. On the basis of the results of the factor analysis we compiled 4 relevant factors of poverty and social exclusion. These independent factors were created from the set of 12 original indicators. Our analysis showed that the 1st factor characterizes material deprivation, the 2nd one represents income poverty and income inequality, 3rd factor reflects labour market exclusion of population and 4th factor characterizes labour market exclusion of households. If we look at 3-dimensional concept which Eurostat uses for monitoring of progress in fighting against poverty and social exclusion, we find out that the first 2 factors obtained by our analysis coverthe first 2 dimensions and the 3rd dimension is divided into 2 factors that evaluate labour market exclusion separately for population and for households.

The cluster analysis highlighted differences in the area of poverty and social exclusion within European Union in 2016 and revealed some changes that have occurred since 2008.According to European Commission (2017) it has been around four years since the EU economy started its slow though consistent recovery from economic crisisdue to which the employment level in the EU now exceeds the 2008 peak (although the impact of this is yet to be fully reflected in all social indicators). Our analysis confirmed that the impact of economic crisis and recovery after crisis were different in various states of EU.

Acknowledgements

Creation of this paper was supported by a grant agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic VEGA. Project: VEGA no. 1/0548/16*The progress of the SR in achieving of the strategy Europe 2020 in the area of poverty and social exclusion*.

References

Ayllón, S. & Gábos, A. (2015). *The interrelationships between the Europe 2020 social inclusion indicators* (pp. 1-28, Working paper No. 15/01). Antwerp: Herman Deleeck Centre for Social Policy – University of Antwerp.

European Commission. (2017). *Social protection committee annual report 2017. Review of the social protection performance monitor and developments in social protection*

policies(pp. 1-491, Rep.). Luxembourg, European Union: Publications Office of the European Union.

Guagnano, G., Santarelli, E. & Santini, I. (2013). *Subjective poverty in Europe: The role of household socioeconomic characteristics and social capital* (pp. 1-27, Working paper No. 113). SAPIENZA University of Rome.

Guio, A. & Maquet, I. E. (2006, November). "Material deprivation and poor housing" What can be learned from the EU-SILC 2004 data? How can EU-SILC be improved in this matter? In:*Comparative EU Statistics on Income and Living conditions: issues and Challenges*. Retrieved December 20, 2017, from https://www.stat.fi/eusilc/guio_maquet.pdf

Hebák, P., Hustopecký, J., Pecáková, I., Průša, M., Řezanková, H., Svobodová, A. & Vlach, P. (2005). *Vícerozmernéstatistickémetody (3)*. Praha: Informatorium.

Kis, B. A. & Gábos, A. (2016). Consistent poverty across the EU. *Corvinus Journal of Sociology and Social Policy*, *7*(2), 3-27.

Labudová, V., Vojtková, M. & Linda, B. (2010). Application of multidimensional methods to measure poverty. *E & M Ekonomie a management*, *13*(1), 6-22.

Loster, T. & Pavelka, T. (2013). Evaluating of the results of clustering in practical economic tasks. In:*The 7th International Days of Statistics and Economics* (pp. 804-818). Prague: University of Economics in Prague.

Mysíková, M., Večerník, J. & Želinský, T. (2015). Impact of the Low Work Intensity on Poverty in the Czech Republic and the Slovak Republic. *Ekonomickýčasopis, 63*(6), 555-575.

Nolan, B. & Whelan, C. T. (2010). Using non-monetary deprivation indicators to analyze poverty and social exclusion: Lessons from Europe?*J Policy Anal Manage*, *29*(2), 305-325.

Řezanková, H. & Želinský, T. (2014). Factors of material deprivation rate in the Czech Republic by household type.*Ekonomický časopis, 62*(4), 394-410.

Stankovičová, I., & Vojtková, M. (2007). *Viacrozmerné štatistické metódy s aplikáciami*. Bratislava: Iura Edition.

Želinský, T. (2010). Analysis of poverty in Slovakia based on the concept of relative deprivation. *Politickáekonomie, 58*(4), 542-565.

Evaluation of the socio-economic situation of European Union countries, taking into account accuracy of statistical data

Małgorzata Stec[1], Małgorzata Wosiek[2]

Abstract

The aim of this paper is to evaluate the socio-economic situation of EU countries in 2016, taking into consideration accuracy of statistical data. The study used ten variables defining the socio-economic situation of EU countries. Linear ordering of EU countries was made using the zeroed unitarisation method. An assessment of the impact of uncertainty in the measurement of diagnostic variables on the value of a synthetic measure was also carried out. For this purpose, a procedure using the Monte Carlo method was proposed. The results indicate that the accuracy of statistical data may influence the results of the linear ordering of EU countries.

Keywords: European Union countries, synthetic measure, Monte Carlo method
***JEL Classification*:**C15, O11, O57
DOI: 10.14659/SEMF.2018.01.50

1 Introduction

Credible and high quality statistical data should provide a basis for the objective empirical research, as this decides about the research's final results. Person who uses statistical data, when taking them from official sources, considers them to be accurate. However, the way of gathering statistical information will not be error-free. It must be highlighted that a reduction of a number of errors in data published in yearbooks is achieved by applying complicated analyses and corrective calculations, which are employed by institutions gathering such data (e.g. Eurostat and statistical offices of each country), but obtained statistical data cannot be deemed accurate. In the economic literature, the analysis concerning influence of errors, which stem from inaccuracy of statistical data, is ignored, which may result in drawing incorrect conclusions in regard to the researched topic.The aim of this paper is to evaluate a socio-economic situation of EU countries, including the accuracy of statistical data. 10 variables determining socio-economic situation of EU countries in year 2016 were employed. The linear ordering of EU countries was done by the zero edunitarisation method. Moreover, an influence of uncertainty of measurement of diagnostic variables on the values of synthetic

[1] Corresponding author: University of Rzeszów, Faculty of Economics, Department of Quantitative Methods and Economic Informatics, M. Ćwiklińskiej 2 St., 35-601 Rzeszów, Poland, e-mail: malgorzata.a.stec@gmail.com.
[2] University of Rzeszów, Faculty of Economics, Department of Microeconomics, M. Ćwiklińskiej 2 St., 35-601 Rzeszów, Poland, e-mail: mwosiek@ur.edu.pl.

measure was investigated. For this, a procedure using the Monte Carlo method was proposed.

2 Theoretical basis for statistical data accuracy

The quality of statistical data is defined by three characteristics (Domański and Pruska, 2000; Kordos, 1988):

- usefulness of data in regard to users' needs (this postulate is fulfilled when a user solving specific topics, plans and realizes special research),
- validity (results are less useful with passing of time),
- accuracy (expressed by similarity of statistical information to real values, that is, the value which would be obtained if for all units of a group under investigation, the data was gathered and processed without errors).

Measurement error is a dissimilarity between determined value (measured) and real value. Question arises, this being the case when the published data can differ from the real data, whether it is worth using them for research or not? For such question, the answer can only be positive, however, it is important to ensure that such statistical data is as accurate as possible.

In technical science the problem regarding evaluation of the results' accuracy was settled in 1995 in the following document Guide to the Expression of Uncertainty in Measurement (2008). It was determined that the components of a measured result are: measured value and a bracket of uncertainty around this value. The formal definition of the term 'uncertainty of measurement' is as follows: uncertainty (of measurement) parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measured. 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).As Diettrich (1991) notes, all measurements are subject to error because no quantity can be known exactly, hence, any measurement has a probability of lying within a certain range. No measurement is perfect. The idea of an uncertainty in measurements is nevertheless something that has to be accepted as far as possible allowed for.

The accuracy of statistical data can be, as in technical science, identified with an error or uncertainty of measurement. In case of statistical data, the real value of measured quantity is very often unknown. In such situation, the Uncertainty of Measurement theory is employed.

In the Fig. 1, an interpretation of uncertainty of statistical parameter is presented. Read from a yearbook or taken from a data base, a value is deemed as a nominal value of a variable X. It is marked with X_n symbol for a multi-dimensional analysis, where n denotes a number of

the following variable taken from a set of diagnostic variables. The essence of the presented approach is an assumption that there is no certainty whether this nominal value is the real value. It is presumed, however, that with the assumed distribution of probability the real value can be found in a range indicated by this distribution. Both distribution of probability and range of uncertainty of diagnostic variables were calculated on the basis of researcher's knowledge regarding methods of collecting this kind of data.

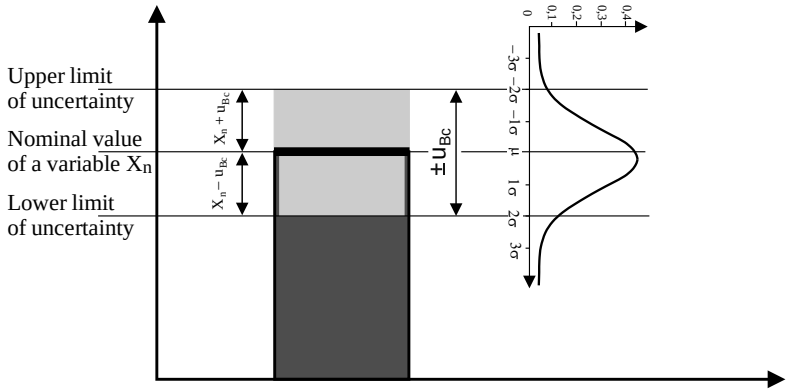


Fig. 1. The essence of setting the uncertainty of statistical value.

In yearbooks, no information regarding the calculated values of uncertainty affecting a specific statistical value is published. Due to this situation, the authoress performed such calculations on a basis of the available knowledge regarding methods of obtaining data used to determine a specific value.

Nominal value for each $X_n$variable employed to calculate a synthetic measure was taken from the Eurostat database. Calculated relative value of uncertainty $u_{Bc}$was converted to an upper and a lower value of a range created around the nominal value, in which, with regard to an assumed probability (depending on an adopted distribution), the real value is present.

$$X_R = X_n \pm u_{Bc} \tag{1}$$

where: X_R - range where the real value of variable is present,

X_n - nominal value of variable,

u_{Bc} - calculated value of variable's uncertainty.

Depending on the way of obtaining data, a value of uncertainty will differ. The most accurate data is collected by official state registers, which are legally responsible for its updates. Such registers include various records, for instance, census, legal entities, institutions, etc. However, state registers are not fully up-to-date which means that they are susceptible to uncertainty concerning real values included in data. One example can be errors

caused by a lack of regular updates, which results in data being different from the real values. For instance: people going abroad permanently or for a longer period of time not always report this fact for a record. Similar situations also occur for other registers.

Assuming that the subsequent variables, which are a basis for developing a synthetic measure, are affected by uncertainty (their real values are unknown, only their estimations), it is necessary to analyse whether these ranges are not too vast to "blur" the difference between the subjects of research. Fig. 2 in a graphical way presents the essence of comparison of the values of the variables of two objects (countries), for which the ranges of uncertainty "overlap". Such case will occur if for a variable being a stimulant, the upper limit of uncertainty for a subject that has a lower position in a rank has bigger value, than the lower limit of uncertainty of a subject that is higher in a rank (Stec, 2017).Overlapping of ranges of uncertainty may take place if values of variables for two or more objects (countries) barely differ and calculated uncertainties are relatively big. With little diversification of subjects' values of variables the following situation may arise, in which a few objects can be characterised by similar values of a specific variable, which hinders the interpretation of real differences between these objects. Analogous situation relates to synthetic measures, too.

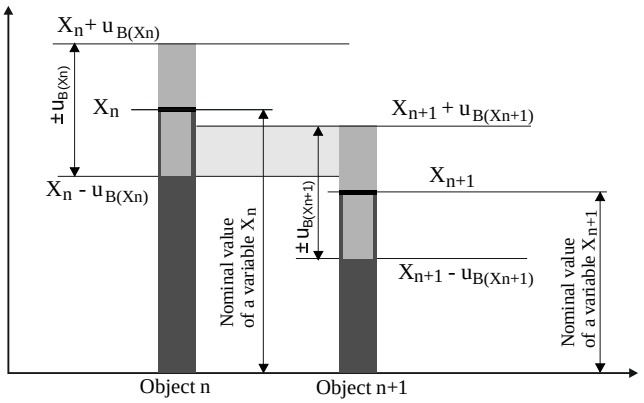


Fig. 2. A case of overlapping ranges of uncertainty between two subjects.

3 Methods applied

In this paper, the zeroed unitarisation method was used in order to calculate a value of synthetic measure for all countries in terms of socio-economic situation in year 2016. (Kukuła, 2014).

A normalization of the variable values was conducted using the following formulas:

$$z_{ij} = \frac{x_{ij} - \min_i\{x_{ij}\}}{R_j} \qquad \text{for stimulating factors} \qquad (2)$$

$$z_{ij} = \frac{\max_i\{x_{ij}\} - x_{ij}}{R_j} \qquad \text{for non-stimulating factors} \qquad (3)$$

where:z_{ij}- normalized value of j-th variable for the i-th object,x_{ij},- value of j-thvariable for the i-th object,R_j- range for the j-th variable.

The synthetic measure was calculated as an arithmetic mean of the normalized value of variables:

$$MS_i = \sum_{j=1}^{m} z_{ij} \qquad (4)$$

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:

- 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 normalization,
- on the basis of the normalized 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.

[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 Diagnostic variables employed in the research

The evaluation of a socio-economic situation for 28 EU countries was done with an employment of 10 diagnostic variables (Table 1)[4]. 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.

Table 1. Diagnostic variables determining socio-economic situation of all EU countries in year 2016.

Diagnostic variables		Calculated value of uncertainty	Number of collisions
X1.	Crude rate of natural change of population per 1000 persons (S)	0.01%	1
X2.	Infant deaths rate per 1000 live births (D)	0.01%	10
X3.	Employed persons per 1000 population (S)	0.1%	5
X4.	Unemployment rate (based on LFS) in % (D)	0.5%	3
X5.	Gross value added by kinds of activity services in% (S)	1.0%	19
X6.	Exports of goods and services in % of GDP (S)	1.0%	5
X7.	Investment rate in % (S)	1.0%	16
X8.	Research and development expenditure (% of GDP) (S)	1.0%	7
X9.	At-risk-of-poverty rate in % (D)	1.0%	11
X10.	Students of higher education institutions per 10 thous. population (S)	0.01%	1

Depending on the way of obtaining statistical data, different values of uncertainty were adopted. In case of three diagnostic variables, the uncertainty was assumed at the level of 0,01%, for five variables 1%. For two cases, uncertainty was assumed as follows - at the level of 0,1% and 0,5%.In case of a diagnostic variable X5, there were 19 colliding situations; that is those in which the ranges of uncertainty of these variables partially overlap. It should be

[4]Examples of similar research can be found in: Barro, 1991;Del Campo et al., 2008;Ertur and Koch,2006;Grzebyk and Stec, 2015; Qizilbash, 2001; Rakauskienė and Kozlovskij, 2014; Stec et al., 2014.

notes that with 28 objects (countries), the maximum number of collisions (overlapping of neighbouring objects) totals 27.

5 Empirical results

Table 2 presents the values of synthetic measures for EU countries calculated by the zeroedunitarisation method with the values of uncertainty range for this measure. The nominal values of the respective synthetic measures are subject to analysis, which were put in order in the traditional way. Ranges of uncertainty constitute additional information enabling to verify the created rank of objects. Standard deviation is an uncertainty measure for each country's synthetic measure. To make the obtained result more trustworthy, standard deviation σ_i was multiplied by coefficient 1,96. This determined value, added to the nominal value of the MS_imeasure, created the upper limit of the uncertainty range and when subtracted - the lower limit of uncertainty ($MS - 2\sigma; MS + 2\sigma$).

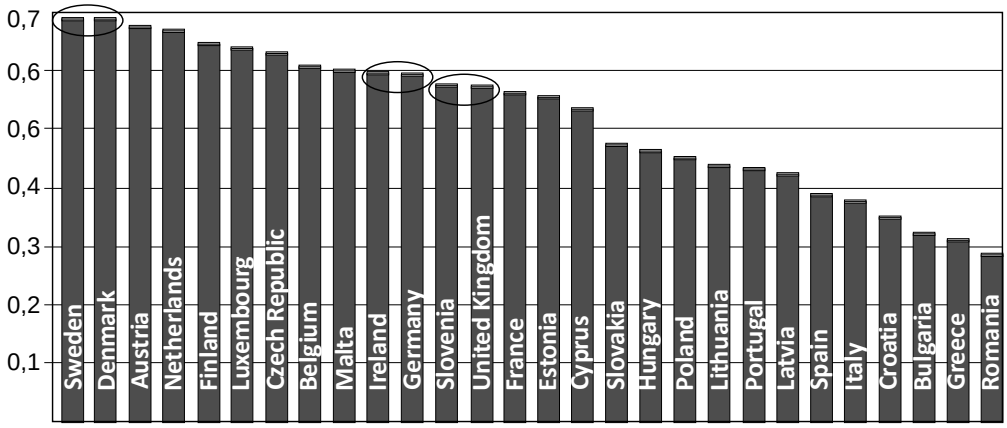


Fig. 3. Rank of EU countries in the area of a socio-economic situation with reference to the uncertainty of diagnostic variables in year 2016.

The uncertainty of the synthetic measure expressed in a relative form ranges from 0,13% for Luxembourg to 0,55% for Greece.The value of the synthetic measure MS_i is a basis for ordering the EU countries in terms of a socio-economic situation. In 2016, as regards the socio-economic situation, the leading positions in the rank of EU countries were taken by: Sweden, Denmark, Austria and Netherlands. Last positions were taken by: Romania, Greece, Bulgaria and Croatia. The comparison of upper and lower limit values of neighbouring uncertainty ranges in the rank of objects allowed for verification whether the differences between nominal values of synthetic measures of these objects are not so small enough that there are no grounds to differentiate their positions (Fig.3).

Table 2. Value of synthetic measure for EU countries calculated by the zeroedunitarisation method with the values of uncertainty ranges for this measure*.

No	Countries	Value of synthetic measure	2σ	$MS - 2\sigma$	$MS + 2\sigma$
1	Sweden	0.6877	0.0017	0.6860	0.6894
2	Denmark	0.6861	0.0019	0.6843	0.6880
3	Austria	0.6730	0.0018	0.6712	0.6748
4	Netherlands	0.6660	0.0019	0.6641	0.6679
5	Finland	0.6448	0.0017	0.6431	0.6465
6	Luxembourg	0.6363	0.0008	0.6355	0.6371
7	Czech Republic	0.6290	0.0016	0.6274	0.6306
8	Belgium	0.6066	0.0019	0.6047	0.6085
9	Malta	0.5991	0.0019	0.5972	0.6010
10	Ireland	0.5945	0.0011	0.5934	0.5956
11	Germany	0.5930	0.0017	0.5913	0.5947
12	Slovenia	0.5730	0.0016	0.5714	0.5746
13	United Kingdom	0.5719	0.0018	0.5701	0.5738
14	France	0.5599	0.0019	0.5580	0.5618
15	Estonia	0.5548	0.0016	0.5532	0.5564
16	Cyprus	0.5338	0.0020	0.5318	0.5358
17	Slovakia	0.4726	0.0016	0.4711	0.4742
18	Hungary	0.4619	0.0015	0.4604	0.4634
19	Poland	0.4494	0.0015	0.4479	0.4509
20	Lithuania	0.4358	0.0014	0.4344	0.4373
21	Portugal	0.4324	0.0016	0.4308	0.4340
22	Latvia	0.4226	0.0016	0.4210	0.4242
23	Spain	0.3870	0.0016	0.3854	0.3886
24	Italy	0.3762	0.0016	0.3746	0.3778
25	Croatia	0.3479	0.0015	0.3464	0.3494
26	Bulgaria	0.3203	0.0015	0.3189	0.3218
27	Greece	0.3100	0.0017	0.3083	0.3117
28	Romania	0.2856	0.0015	0.2841	0.2872

*by underlining we mean overlapping ranges

The analysis of the obtained results allows for confirmation that the differences of synthetic measures are too small to unequivocally acknowledge the positions of the following countries: Sweden and Denmark, Ireland and Germany, Slovenia and United Kingdom. There is, therefore, a risk of making a mistake that a country classified lower in ranks has, after all, a higher value of synthetic measure than a country classified higher in ranks. It was therefore concluded that there are no grounds for diversification of the positions of these counters.

Conclusions

The following conclusions can be drawn based on the carried out research:

- The evaluation of a socio-economic situation for 28 EU countries in 2016 was done with use of 10 diagnostic variables. The zeroed unitarisationmethod was employed for the empirical research.
- The results confirm the diversity of EU countries in terms of a socio-economic situation. In 2016 Sweden, Denmark, Austria and the Netherlands were the leaders. The lowest level in this context represent the following countries: Romania, Greece, Bulgaria and Croatia.
- Taking into consideration the uncertainty of the values of synthetic measures, it may influence the final conclusions drawn from the research.
- In case of comparison research, in which the results of linear order of the objects have a considerable meaning, it seems purposeful to include the influence of uncertainty of the values of diagnostic variables on the calculation of the synthetic measure.

References

Balazs, A. (2008). International Vocabulary of Metrology-Basic and General Concepts and Associated Terms. *Chemistry International*, *25*.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, *106*(2), 407-443.

Del Campo, C., Monteiro, C. M. & Soares, J. O. (2008). The European regional policy and the socio-economic diversity of European regions: A multivariate analysis. *European Journal of Operational Research*, *187*(2), 600-612.

Dietrich, C. F. (1991). *Uncertainty, calibration and probability: the statistics of scientific and industrial measurement*. CRC Press, 3.

Domański, C. & Pruska, K. (2000). *Nieklasyczne metody statystyczne*. Polskie Wydawnictwo Ekonomiczne, 65.

Ertur, C. & Koch, W. (2006). Regional disparities in the European Union and the enlargement process: an exploratory spatial data analysis, 1995–2000. *Annals of Regional Science*, *40*(4), 723–765.

Grzebyk, M. & Stec, M. (2015). Sustainable development in EU countries: concept and rating of levels of development. *Sustainable Development*, *23*(2), 110-123.

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*. CeDeWu, 221.

Kordos, J. (1988). *Jakość danych statystycznych*. Państwowe Wydawnictwo Ekonomiczne.

Kukuła, K. (2014). Budowa rankingu województw ze względu na wyposażenie techniczne rolnictwa w Polsce. *WiadomościStatystyczne*, *7*, 62-76.

Liu, J. S. (2008). *Monte Carlo strategies in scientific computing*. Springer Science & Business Media.

Niemiro, W. (2013). *Symulacje stochastyczne i metody Monte Carlo*. University of Warsaw, Warszawa.

Qizilbash, M. (2001). Sustainable development: Concepts and rankings. *Journal of Development Studies*, *37*(3), 134-161.

Rakauskienė, O. G. & Kozlovskij, V. (2014). Overview of EU economic cohesion process performance for new member states. *Procedia-Social and BehavioralSciences*, *110*, 284-292.

Stec, M. (2017).*Taksonomiczna analiza poziomu rozwoju społeczno-gospodarczego województw Polski. Studium przypadku-województwo podkarpackie*. Wydawnictwo Uniwersytetu Rzeszowskiego, 87-99.

Stec, M., Filip, P., Grzebyk, M. & Pierścieniak, A. (2014). Socio-economic development in the EU member states—concept and classification. *Eng. Econ*, *25*(5), 504-512.

Walesiak, M. & Gatnar, E. (Eds.). (2009). *Statystyczna analiza danych z wykorzystaniem programu R*. Wydawnictwo Naukowe PWN.

On stationarity of changes in the trends of selected refining variables

Jacek Stelmach[1]

Abstract

Forecasting time series representing macroeconomic variables becomes increasingly difficult. These variables, often correlated with each other are influenced with political and social decisions, climatic disasters or warfares. The phenomenon of globalization has increased the impact of these factors, the effect of which is temporary or permanent change of correlation between these variables – making it difficult to make management decision based on forecasts. One of solution could be trend's forecasts. The article presents an analysis of selected, important refinery variables in the last 25 years, in particular stationarity related to the trends.

Keywords: *time series, stationarity, unit roots tests*
JEL Classification: C220
DOI: 10.14659/SEMF.2018.01.51

1 Introduction

Still the most important global refinery raw material is crude oil. In the last years its price fluctuated – with the maximum value of 124 USD/b in 2012 and minimum value of 31 USD/b in the beginning of January 2016. The crucial question is whether oil prices will decline or rise and what will be the impact on other products. The possibility to support management decisions applying statistical methods can be an important competitive advantage. Dynamic environment – changes in correlation between investigated variables - can make it difficult. Then even forecasting of trend`s changes can be useful. However, it requires choosing proper range of investigated sample to prevent spurious correlation, in particular its stationarity. This study investigates the prices and price changes of refinery raw materials and subproducts, taking into account the interval from November 1992 up to November 2017 (25 years), divided into three subperiods according to variance size and examines their stationarity to determine the possibility of building statistical models. Known literature did not examine neither trend's changes nor stationarity in smaller subperiods.

2 International oil market changes

The domination of crude oil began after World War II, since that time the impact of political decisions is visible as sudden changes of crude oil price – see Fig. 1. First visible crises took

[1] Corresponding author: Polwax S.A., jacek.stelmach@polwax.pl.

place in 1973 after embargo of Arab countries (as a reaction on supporting Israel). Next crises happened in 1979 (Iranian Revolution and Iran-Iraq War). Stabilisation in the eighties ended with Gulf War (1990). Very deep price change had place between 1997 and 1998 (Asian financial crisis), extremely high level – occurred in 2010 as a result of "Arab Spring". Together with political decision, the changes in the known oil reserves and production had a visible impact (see Fig. 2) –increase the reserves in North America in 1998, visible decrease in Middle East and very significant (three times) increase in Central and South America in 2008. Therefore the current question is the possibility of further use statistical methods (and statistical models) to forecast price changes in the future.

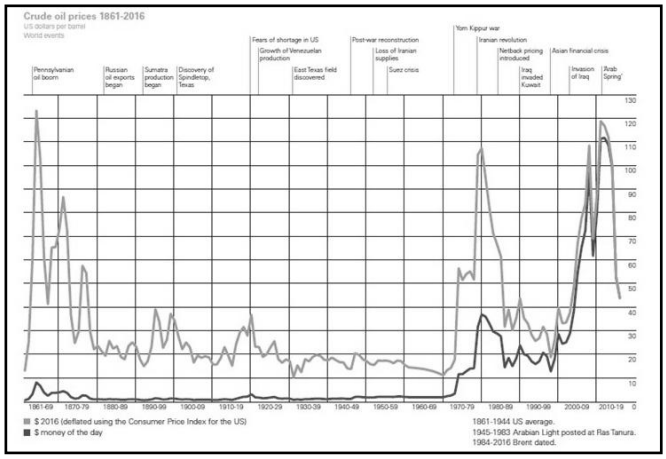


Fig. 1. An impact of political events on crude oil prices (BP Statistical Review of World Energy June 2017).

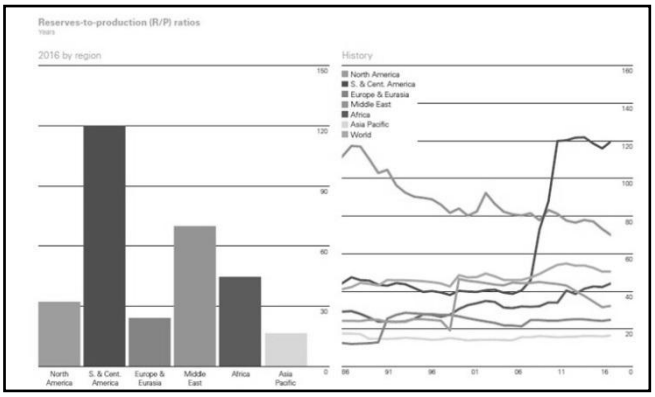


Fig. 2. Reserves to production ratios according to regions in 2016 (left) and historically (right) (BP Statistical Review of World Energy June 2017).

3 Review of econometric models of crude oil price

Known literature let to classify the most popular models in several groups:

- time series models,
- financial models based on the dependencies between spot and future prices,
- structural models, considered as an extension of autoregressive specification.

This paper investigates the possibility of apply time series models with the most popular model as and autoregressive process AR(p):

$$P_i = \varphi_{i-1}P_{i-1} + \cdots + \varphi_{i-p}P_{i-p} + \varepsilon_i \quad (1)$$

where spot price P_i depends on previous values together with uncorrelated error term ε_i.

Mean reverting process with the assumption that a price reverts to the average after a shock can be presented as below:

$$P_{i+1} - P_i = \alpha(P_i^* - P_i) + \varepsilon_i \quad (2)$$

where spot price change depends on long-run equilibrium P_i^* and mean reversion rate α (see Engle and Granger, 1987). Abosedra (2005) proposes another model, observing period January 1991 to December 2001 with monthly unbiased predictor of the future oil price X:

$$P_{i+1} = \alpha + \beta X_{i-1} + \varepsilon_i \, . \quad (3)$$

However, numerous studies (Lalonde at al. 2003; Ye et al., 2005; Pindyck, 1999; Zeng and Swanson, 1998) show rather poor prediction ability. The dynamics of parameters of any model is too significant. Time-series models of other refining variables may behave very similarly. Therefore, both prices and trends of these prices were analysed.

4 The analysis of chosen refinery variables

Taking into consideration that prices of products coming from crude oil depend on crude oil price or on crude oil demand/supply equilibrium – this phenomenon in models should increase the accuracy of forecasts. The analysis covered refinery variables noted monthly in period from November 1992 up to November 2017 (last 25 years), only for variable describing the hydrorefined paraffin, the period is from January 2003 up to November 2017 (www.indexmundi.com, 09.12.2017).

1. Crude oil, dated Brent, monthly price (**CO**) USD/barrel.
2. Australian thermal coal, monthly price (**AC**) USD/metric ton.
3. New York harbour gasoline regular spot FOB (**GA**) USD/gallon.
4. New York harbour heating oil spot FOB (**HO**) USD/gallon.
5. US Gulf Coast Kerosene-type Jet fuel spot FOB (**JF**) USD/gallon.

6. Henry hub natural gas (**NG**) USD/million metric British thermal unit.
7. Mont Belvieu TX propane spot FOB (**PR**) USD/gallon
8. Hydrorefined paraffin ICIS (**HP**) EUR/ton.

However, analysis of such time series indicates a variation of dependencies, see Fig. 3.This phenomenon is better visible analysing an example: a raw material – crude oil (**CO**) and a final product – hydrorefined paraffin (**HP**), see Fig. 4. Up to 2014, there is visible correlation with a shift of 2 months – a time necessary to process raw material into ready product. Last three years, however, due to shutting down number of refineries producing paraffin, the dependence changed. It created new dependencies in which the price of the raw material plays a smaller role. Moreover, analysed time series could be non-stationary, what complicates the prediction. But in such unstable economic world, even a good forecast of changes in trend (like prediction of the sign of the first derivative) can lead to a competitive advantage. Still the stability of model parameters over time and the stationarity of time series should be taken into account as one of the fundamental assumption in modelling econometrics variables.

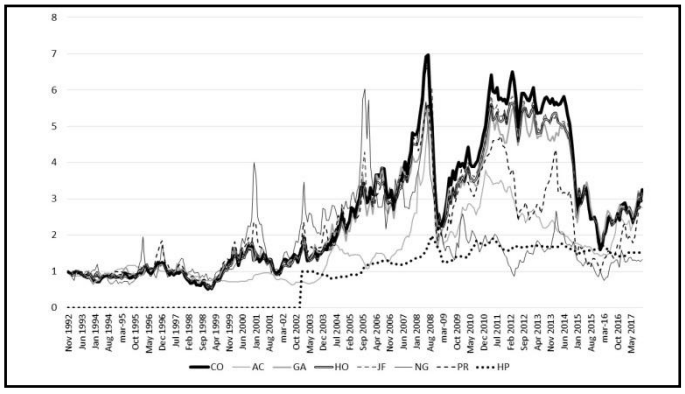


Fig. 3. Plot of analysed time series, values are related to the first observation.

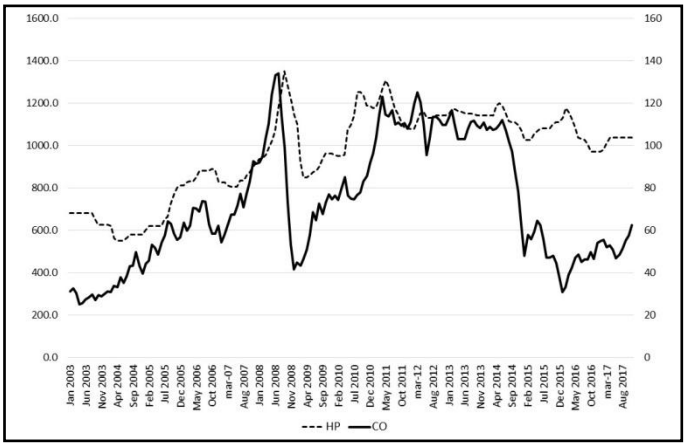


Fig. 4. Plot of crude oil (**CO**) and hydrorefined paraffin (**HP**) time series.

5 The analysis of trends

Trends of investigated variables were estimated decomposing time series with moving average smoothing method with orders: 6, 12, 15 (to check the impact of annual seasonality) and applying Spencer filter (see Brockwell and Davis, 2002).Although the differences between trends estimated by the above methods were small, further investigations were carried for all the moving average methods. The plots of the trends (moving average, order=12) and first derivative (calculated in the form of the first differences) of trends are presented in Fig. 5.

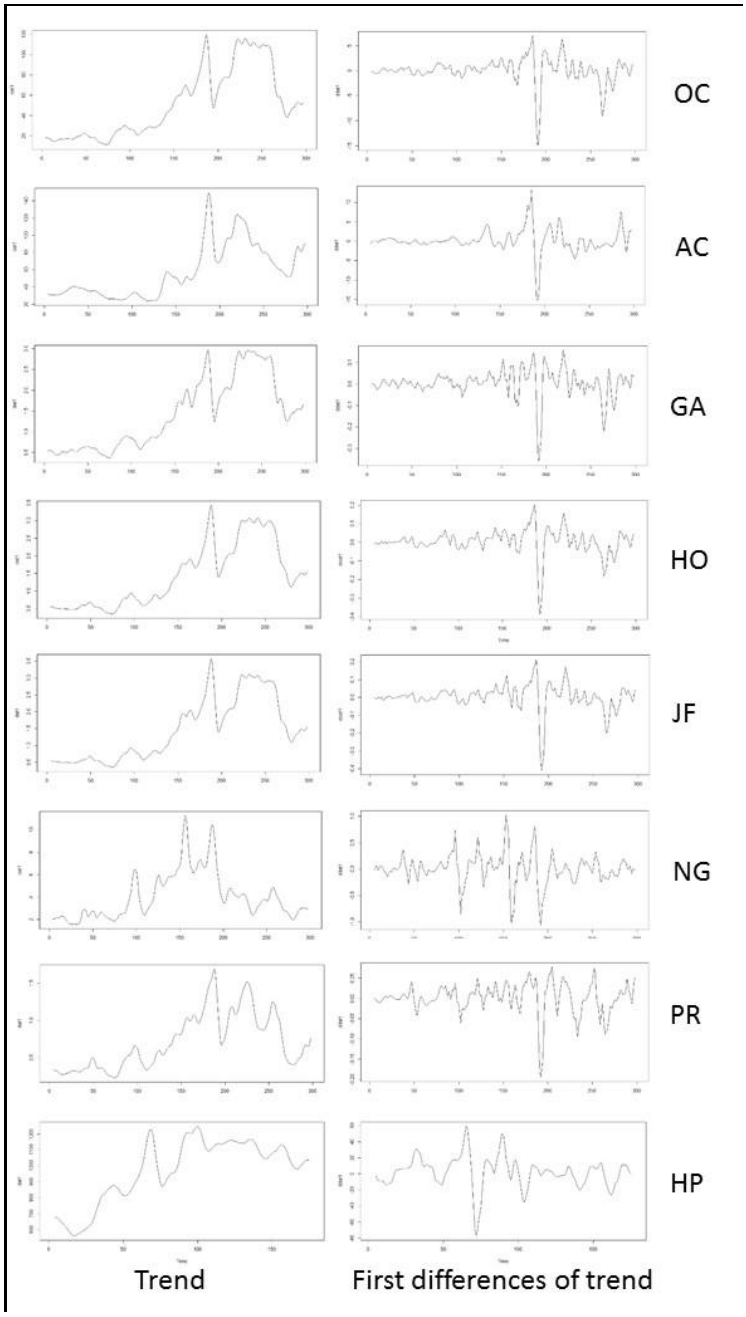


Fig. 5. Plots of trends and their differences of investigated time series.

Analysing these plots, it was decided to divide analysed period into three subperiods:

A. January 1992 – March 2005 (January 2003 – January 2007 for **HP**),
B. April 2005 – January 2011 (February 2007 – January 2011 for **HP**),
C. February 2011 – November 2017,

taking into consideration variation of first differences. First group represents the period before increased variation, second – high changes in unstable environment, third – period of 'stabilisation'. Described studies on stationarity were performed for these three subperiods.

There is a variety of the stationarity tests described in the literature, most commonly used: Dickey-Fuller tests (DF and ADF) (Dickey and Fuller, 1979, 1981), ADF-GLS test (Elliot at al., 1996), Philips-Perron test (Philips and Perron, 1988), KPSS test (Kwiatkowski et al., 1992), and rarely used like: Schmidt-Philips (Schmidt and Philips, 1992) test or Zivot-Andrew test (Zivot and Andrew, 1992). Unfortunately, a power of unit root tests depends on the length of time series. As a result time series can be classified improperly as non-stationary whereas in fact is stationary. Chosen tests are described briefly below. All the tests except KPSS-tests verify a hypothesis of non-stationarity. KPSS-test verifies alternative hypothesis – the stationarity of time series.

5.1 Augmented Dickey-Fuller test

Although probably the most known test is Dickey-Fuller test (Dickey and Fuller, 1979), augmented modification of this test (Dickey and Fuller, 1981) – which was constructed for the cases when non-systematic component is autocorrelated, was chosen. It bases on autoregressive process model (only the limiting distribution of test statistics was tabulated):

$$y_t = \theta_1 y_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta y_{t-1} + \varepsilon_t \, . \tag{4}$$

5.2 ADF-GLS test

The test is a modification of Augmented Dickey-Fuller test (Elliot at al., 1996).Time series is estimated with GLS (generalised least square) method performing the transformation:

$$\tilde{y}_1 = y_1; \tilde{y}_t = y_t - \rho y_{t-1}; x_1 = 1; x_t = 1 - \rho \tag{5}$$

After that constant and trend are removed:

$$y_t^* = y_t - (\hat{\beta}_0 + \hat{\beta}_1 t) \tag{6}$$

and final model of time series represents following formula:

$$\Delta y_t^* = \beta_0 + \varphi_1 y_{t-1}^* + \sum_{i-1}^{p} \gamma_i \Delta y_{t-1}^* + \varepsilon_t. \tag{7}$$

The critical values were calculated based on simulation (Elliot at al., 1996).

5.3 Philips-Perron test

The test is a modification of standard Dickey-Fuller test with non-parametrically modified test statistics and centred time variable (Philips and Perron, 1988). The test statistics *Z* can be reduced to Dickey-Fuller tests statistics if ε_t is not autocorrelated.

5.4 KPSS test

This test verifies the opposite hypothesis that the time series is stationary. The idea developed by Kwiatkowski, Philips, Schmidt and Shin (Kwiatkowski et al., 1992) bases on a sum of a deterministic trend, random walk and stationary random error:

$$y_t = d_t + r_t + \varepsilon_t; r_t = r_{t-1} + u_t \tag{8}$$

where:$d_t = \sum_{i=0}^{p} \beta_i t^i$.

The test base on the hypothesis that the random walk has a variance equal to zero. Critical values were calculated with the simulation (Kwiatkowski et al., 1992).

5.5 Schmidt-Philips test

This test bases on an alternative parametrization (Schmidt and Philips, 1992):

$$y_t = \psi + \xi t + x_t; x_t = \beta x_{t-1} + \varepsilon_t \tag{9}$$

extracted from the score of LM principle under the assumptions that random error $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. The most important advantage of this test is that the meaning of the parameters governing level and trend is independent of whether or not null hypothesis is true. The tabulated critical values were calculate with the simulation (Schmidt and Philips, 1992).

5.6 Zivot-Andrews test

This test is an extension of Philips-Perron test (Zivot and Andrew, 1992) in which it is assumed that breakpoints should be fixed not estimated (like in P-P test). Additionally the new ideas were introduced: the effect on empirical results of fat-tailed and temporally dependent innovations. The tabulated critical values were calculate with the simulation (Zivot and Andrew, 1992).

6 Description of investigations

An experiment was carried out for first differences of trends, estimated as mentioned above with moving average smoothing method with orders: 6, 12, 15 and applying Spencer filter. The investigations were performed with the following tests:

1. Augmented Dickey-Fuller (ADF) test with no trend and drift,
2. Elliott-Rothenberg-Stock DF-GLS test,
3. Elliott-Rothenberg-Stock P-test,
4. KPSS test, no trend,
5. KPSS test, linear trend,
6. Philips-Perron test, Z-alpha statistics,
7. Philips-Perron test, Z-tau statistics,
8. Schmidt-Philips test,
9. Zivot-Andrew test.

The tests were carried out for whole range of samples and three subperiods of chosen time series: A, B and C described below for the same significant level $\alpha=0,05$

7 Results

The results were coded: for all the tests except KPSS-tests – if the test does not reject the hypothesis, the result was set as "0" (otherwise – "1") and for KPSS-tests – if the test does not reject the hypothesis, the result was set as "1" (otherwise – "0"). Coded values, added up across all the tests and across four method of trend's estimation are presented in Table 1.

Table 1. Aggregated results of the tests.

Variable	Whole period	Subperiod A	Subperiod B	Subperiod C
CO	29	7	25	9
AC	31	27	25	12
GA	30	14	26	11
HO	29	15	25	7
JF	29	10	26	8
NG	32	26	22	17
PR	32	25	24	14
HP	32	7	21	25

The results are rather unexpected. Although the tests mostly rejected hypothesis of non-stationarity, for the two subperiodscovering: January 1992 – March 2005 and February 2011 – November 2017, performed tests did not give a clear result. For majority of variables most of the tests (especially for subperiod: February 2011 – November 2017) did not reject the hypothesis of non-stationarity (or reject stationarity – for KPSS-tests) although this

subperiods seems to be rather economically stable. It means that you have to be very careful creating models for forecasting the changes of trends refining variables.

Conclusion

The globalization of today's economy resulting in greater influence of even distant political, social and economic phenomenon interfere seemingly stable economic processes. In contrast to natural phenomena, the dependencies between economic variables constantly change. It makes difficult to build long-term models and limits the accuracy of forecasts. Described study indicates that time series in a seemingly stable environment could be non-stationary. This also applies to changes of trends – on the example of refinery product prices although examined samples coming from rather economically period. Nevertheless, the lower variation of the first differences of trends could allow to set the forecasts useful in making of management decisions taking into account their non-stationarity. This issue will be the subject of further research.

References

Abosedra, S. (2005), Futures versus Univariate Forecast of Crude Oil Prices, *OPEC Review, 29,* 231-241.

Brockwell, P. J. & Davis, R. A. (2002), *Introduction to Time Series and Forecasting.*Springer.

Dickey, D. A. & Fuller, W.A. (1979), Distribution of the estimators for Autoregressive time series with a unit root. *Journal of the American Stat. Association, 74,* 427-431.

Dickey, D. A. & Fuller, W.A. (1981), Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica, 49,* 1057-1072.

Elliott, G., Rothenberg, T. J. & Stock J. H., (1996), Efficient tests for an autoregressive unit root. *Econometrica, 64(4),* 813-836.

Engle, R. F. & Granger, C. W. J. (1987), Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica, 55,* 251-276.

Kwiatkowski, D., Philips, P. C. B., Schmidt, P. & Shin, Y. (1992), Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics, 54,* 159-178.

Lalonde, R., Zhu, Z. & Demers, F. (2003), Forecasting and Analyzing World Commodity Prices. *Bank of Canada, Working Paper 2003-24.*

Philips, P. C. B., & Perron, P. (1988), Testing for a unit root in a time series regression. *Biometrika, 75,* 335-346.

Pindyck, R. S. (1999), The Long-run Evolution of Energy Prices. *The Energy Journal, 10,* 1-27.

Schmidt, P. & Philips, P.C.B. (1992), LM test for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics & Statistics, 54(3),* 257-287.

Ye, M., Zyren, J. & Shore, J. (2005), A monthly Crude Oil Spot Price Forecasting Model Using OECD Petroleum Inventory Levels. *International advances in Economics Research, 8,* 324-334.

Zeng, T. & Swanson, N. R. (1998), Predictive Evaluation of Econometric Forecasting Models in Commodity Future Markets. *Studies in Nonlinear Dynamics and Econometrics, 2,* 159-177.

Zivot, E., & Andrews, D.W.K. (1992), Further evidence on the Great Crash, the Oil-Price Stock, and the unit-root hypothesis. *Journal of Business & Economic Statistics, 10(3),* 251-270.

Spatio-temporal decomposition of the growth of communal budgets expenses on culture in Poland in 2003-2016. Implementation of the SSANOVA model

Adam Mateusz Suchecki[1]

Abstract

Political transformation and decentralisation of public administration in Poland had a significant impact on the cultural sector and forms of its organization and financing (cover these expenses from own income).The main aim of the article is to analyse changes of the expenditure on the culture of selected four groups: libraries, community cultural centers, heritage and monuments protection, and others (in example galleries and Artistic Exhibition Offices) and also the sum of communal budgets in individual provinces in the years 2003- 2016. The analysis used a multidimensional sample of 16x4x14 (space – categories - time). To illustrate the diversity of the communal expenditure on culture in particular years of the analysis was made using the dynamic modification of Knudsen's regression model of share transfers proposed by Berzeg (SSANOVA model). The results of the analysis indicate. The general trend of expenditure growth for culture in the years 2003-2016 is responsible for an average of 61,2% of the rate of changes in spending expenditure on culture on individual voivodships and their categories. Individual structural effects differ significantly depending on the category of expenditure. The volatility of the temporal effects of the national component indicates a clear reduction of expenses - 2011/2010 by more than 3%, an even greater negative change is visible in more than 10% in 2015/2014. Dynamic analysis provides additional information on trends in changes in the areas of economic and social activity under study.

Keywords: *spatial shift-share analysis, cultural economics, public finances, communal budgets*
JEL Classification: Z1, H7, C5
DOI: 10.14659/SEMF.2018.01.52

1 Introduction

In most countries, the provision of cultural services is divided between the public and the private sector. The public sphere is also divided into different administrative levels – the central level, regional and country levels, and local municipalities (Håkonsen and Løyland, 2013). Studies of various aspects of local cultural policy and local cultural spending can be grouped in two recent examples. Most of the studies analyse local cultural expenditures as a whole, i.e. as an aggregate of all cultural expenditures in the municipality (Depalo and Fideli, 2011; Benito, Bastida and Vicente, 2013; Werck, Heyndels and Geyes, 2008). In addition, some studies analyse a specific local or regional cultural institution, such as an opera house

[1] Corresponding author: University of Lodz, Department of Public Finance, 90-214 Lodz, 37 Rewolucji Street 1905, e-mail: adam.suchecki@uni.lodz.pl.

(Schultze and Ursprung, 2000) or a theatre (Getzner, 2004). In this study, the author will combine the two examples mentioned above and begin with by analysing the aggregated expenditures of communal budgets in Poland on culture and national heritage for the economic background, then conduct an analyses of four groups of cultural expenses of communes budgets (global, libraries, community cultural centers, monuments and heritage protection).

The main purpose of the changes in the field of culture after the systemic transformation was the introduction of mechanisms supporting the efficient and fair management of public resources, the introduction of changes in the competences of the public administration relating to the organisation of the financing process for culture, and the introduction of new solutions in the financing, organisation and management of cultural institutions such as the decentralisation of cultural institution management, increasing autonomy of cultural institutions, and the development of legal frameworks for patronage and sponsorship in this area (Wrona, 2011).

One of the most important changes in management processes and financing in self-government units was the transmission of the cultural tasks to the communes as obligatory own tasks. On one hand it gave to it gave to a great scope of the financial autonomy to Polish communes, but on the other hand those changes forced the communes to supply the financial background of all transmitted tasks from the own incomes. This situation caused a great variety of the communal revenues for the culture and national heritage protection correlated with the communal wealth and the way how local politicians treat the cultural goods with respect to socio-economic categories. So the tryouts of quantitative evaluations of this category of communal revenues have been carried out with the consideration of spatial-sectional and time range. This multidimensional approach let to evaluate and estimate the level of the regional development according to the referential area and to analyse the tempo and the structure of changes. One of the methods used to such analysis are methods of so-called the Shift-Share Analysis (SSA) and its modifications enlarged of the time dimension (Knudsen, 2000; Berzeg, 1998). Those methods were widely used in other researches of financing the public sector and regional analysis (Jewczak and Żółtaszek, 2011; Theil and Gosh, 1980).

In this research, the multidimensional data of the Polish communal cultural expenditures by the regions(NUTS 2, voivodships) on the different types of cultural goods and services in the period of fourteen years (2003-2016) has been completed. It allows using more complex methods of the analysis with dynamic aspects.

The collected sample of statistic data consider the information about cultural revenues of Polish communes (by the regions) in the years 2003-2016 diversified into four groups of expenditures: for libraries, for community cultural centres, for heritage and monuments protection, and others (in example galleries and Artistic Exhibition Offices). The statistical information considers the multidimensional data about communal expenses in 16 regions to 4 categories of cultural goods and services in 14 year periods (2003-2016). That structure and the large quantity of statistical samples (16x4x14 observations) allows using the SSANOVA panel model as a tool to analyse the communal cultural expenditures in Poland.

2 Analysis of expense on culture of Polish communal budgets in years 2003-2016

The expenses on culture of Polish communal budgets had been growing since year 2003 till 2014 (Table 1).

Table 1. The total cultural expenditures of communal budgets in years 2003, 2011, 2016 (in thousands PLN) and its percentage in total expenditures (in %).

Regions	2003	%	2011	%	2016	%
Poland	1911065	3	5245835	3.8	5274949	3.3
Dolnośląskie	189685	4	521547	4.0	565099	4.6
Kujawsko-Pomorskie	84684	2	251269	3.3	234337	2.6
Lubelskie	89524	3	288593	5.0	257979	3.6
Lubuskie	47068	2	174989	6.8	129841	2.4
Łódzkie	117137	3	332941	4.4	319139	3.4
Małopolskie	164627	3	428639	4.9	466831	3.9
Mazowieckie	288558	3	791225	3.7	890024	3.4
Opolskie	47996	2	140163	6.0	133599	3.9
Podkarpackie	89300	2	238897	2.6	227789	2.3
Podlaskie	52939	3	150123	2.9	133809	2.5
Pomorskie	109283	3	383461	4.6	358286	3.9
Śląskie	267808	3	538853	3.2	567693	2.8
Świętokrzyskie	42068	3	144005	2.9	116290	2.3
Warmińsko-Mazurskie	62228	3	167994	2.4	157335	1.8
Wielkopolskie	170770	3	439382	3.5	475477	3.1
Zachodniopomorskie	87340	3	253353	4.0	241421	3.1

After that year the slightly deprecation of cultural expenditures has been noticed. In the first year of analysis the amount of cultural expenditures of Polish communes stands at the level of 19,11065 billion of PLN and reached the level of 52,74949 billion PLN in the last estimated year. The highest level of budgetary afford of communal budgets has been reached in the year 2014 (62,43012 billion PLN).

In the years 2003 – 2016 the highest level of cultural expenditures was noticed for the communes in the Mazowieckie and Śląskie regions. The cultural afford of communal budgets in Mazowieckie was highest in the year 2016 and reached 8,90024 billion PLN, where the communes in Śląskie region reached the highest amount of 8,31339 billion PLN in the year 2014.The year to year changes of cultural expenditures of Polish communes has been decreasing since the year 2011. The percentage of the cultural expenses in total budgetary expenses of communes in the analysed periodhave reached the average level of 2 to 4% and was highest in the communes of Dolnośląskie region in 2014.

3 Methodology

For the study of spatio-temporal relations, the tools that consider the specificity of the analysed phenomenon or process are considered to be the main in areas of great social importance. Among them, a group of methods for analysing shift-share deserves particular attention. These tools fall within the scope of spatial statistics and econometrics. They are multidimensional in space-sector-time. This approach allows the assessment and examination of the level of a given area against the background of the reference area. It also allows you to account for the dynamics and structure of changes. Generally, two different models are considered - the Berzeg's model and its dynamic modification - Knudsen's model (Marimor and Ziliboti, 1998).

This study presents and discusses the results of the estimation of three SSA structural and geographic models:

1. ANOVA 2 – full, static decomposition of cultural expenditures growth in whole period 2003-2016 (with two factors without interaction):

$$tx_{ri.}=a_0+\Sigma_i a_i s_i+\Sigma_r b_r w_r+\varepsilon_{rit},$$

enabling the decomposition of the total increase in expenditure on culture (the symbols $s_i$and w_r were marked with zero-one variables indicating individual categories of expenditure on culture and on individual regions):

$$tx_{16/3}=[(x_{2016}-x_{2003})/x_{2003}]$$

on 4 sectoral effects (components) ai (by cultural expenses categories) and on 16 regional effects.

2. Berzeg's model - dynamic decomposition of the cultural expenditures growth, which explains (decomposes) the variable growth rate in subsequent years $tx_t=[(x_t-x_{t-1})/x_t]$ including the division into components by 4 types of expenditure and 16 regions with permanent effect of the temporal dimension (intercept):

$$tx_{rit}=a_0+\Sigma_i a_i.s_i+\Sigma_r b_r w_r+\varepsilon_{rit}.$$

3. Knudsen's model – dynamic decomposition of the cultural expenditure growth, which explains (decomposes) the variable growth rate in subsequent years $tx_t=[(x_t-x_{t-1})/x_t]$, including the division into components by 4 types of expenditures and 16 regions with a variable effect of the temporal dimension, i.e. with different values of the total growth parameter (differentiation of the intercept) in individual years:

$$tx_{rit}=\Sigma_t a_{0t}+\Sigma_i a_i.s_i+\Sigma_r b_r w_r+\varepsilon_{rit}.$$

The results of the estimation of selected models allow analysis as well total growth, as more detailed analysis of the effects of sectoral changes (according to culture expenditure groups) and regional competitiveness (comparative advantage) of regions.

4 The results of the empirical analysis

The general trend of expenditure growth for culture in the years 2003-2016 is responsible for an average of 61,2% of the rate of changes in spending expenditure on culture on individual voivodships and their categories. Individual structural effects differ significantly depending on the category of expenditure (Table 2). A similar assessment of regional growth to the national average is provided by the Lubelskie Region, with the exception here being the assessment for the Mazowieckie Region (5,59) far above the values of the other assessments, among which there are also negative values (Łódzkie, Podlasie, Śląskie). In the case of the assessment of the parameters of the structure of the type of communal expenditure on culture, one should notice a positive increase in the protection of monuments and a negative value (decrease) in the average rate of expenditure on libraries and cultural centers (in the period 2003-2016).

In ANOVA 2 model, from 14 consecutive annual observations, we can create a 13 element decomposed variable (annual growth rate). The possibilities of comparing results confirm the values of ratings from the Berzeg's model (Table 3) multiplied by the number of effective observations (here: 13), which are shown in the fifth column of the scoreboard. However, it could be noted that the results point to other objects reaching high or low values.

In terms of regional growth throughout the 2003/2016/2015 sample period, regions: Wielkopolskie (16,47%), Świętokrzyskie (16,37%) and Lubelskie (13,72%) have the highest values. Also, the domestic factor indicates an increase in expenditure on culture by about 79% compared to about 161% from the static model. Also in the case of assessments of growth components by expenditure groups, significant differences should be noted, especially in the assessment of the pace of changes in expenditure on the monuments and heritage protection: a positive value from the ANOVA2 model and a negative one from the Berzeg's model.

Table 2. Results of estimation of the static model ANOVA 2 decomposition of the growth rate of expenditure on culture by 4 groups of expenses and 16 voivodships in Poland in the period 2003-2016.

	Parameter	**Evaluation**	**In %**	**t-stat**	**p-value**
Domestic Factor	A0	1.6122	161.22	1.11	0.27
Libraries	A1	-1.4017	-140.17	-1.49	0.14
Cultural Centres	A2	-0.5953	-59.53	-0.63	0.53
Heritage Protection	A3	1.9607	196.07	2.08	0.04
Others	A4	-0.0121	Res.		
Dolnośląskie	B1	0.17	-0.06	0.32	-0.29
Kujawsko-Pomorskie	B2	0.17	-0.03	0.10	0.00
Lubelskie	B3	0.22	0.12	0.20	0.04
Lubuskie	B4	0.07	-0.17	0.20	0.01
Łódzkie	B5	0.26	0.12	0.06	0.04
Małopolskie	B6	0.02	-0.13	0.25	0.08
Mazowieckie	B7	0.15	-0.23	0.11	0.02
Opolskie	B8	0.04	-0.19	0.07	0.09
Podkarpackie	B9	0.26	0.01	-0.01	-0.03
Podlaskie	B10	0.04	0.02	0.00	0.08
Pomorskie	B11	-0.03	0.15	0.15	0.00
Śląskie	B12	0.14	-0.05	0.25	-0.11
Świętokrzyskie	B13	0.06	-0.09	0.09	-0.06
Warmińsko-Mazurskie	B14	0.07	-0.08	0.17	0.03
Wielkopolskie	B15	0.13	0.07	0.01	0.10
Zachodniopomorskie	B16	0.05	-0.06	-0.08	0.00

Table 3. Results of estimating the dynamic Berzeg's model the decomposition of the growth rate of expenditure on culture by 4 groups of expenses and 16 regions in Poland in 2004/2003 - 2016/2015.

	Parameter	**Evaluation**	**In %**	**%*13**	**t-stat**	**p-value**
Domestic Factor	A0	0.0605	**6.05**	**78.63**	9.46	0.00
Libraries	A1	-0.0150	-1.50	-19.50	-4.03	0.00
Cultural Centres	A2	-0.0081	-0.81	-10.52	-2.50	0.02
Heritage Protection	A3	-0.0198	-1.98	-25.71	-3.24	0.00
Others	A4	0.0527	5.27	68.45	Res.	
Dolnośląskie	B1	0.0036	0.36	4.73	0.50	0.62
Kujawsko-Pomorskie	B2	-0.0049	-0.49	-6.41	-0.55	0.58
Lubelskie	B3	0.0106	1.06	13.72	1.21	0.23
Lubuskie	B4	-0.0148	-1.48	-19.22	-1.48	0.14
Łódzkie	B5	-0.0026	-0.26	-3.36	-0.31	0.76
Małopolskie	B6	0.0059	0.59	7.68	0.77	0.45
Mazowieckie	B7	-0.0015	-0.15	-1.99	-0.22	0.83
Opolskie	B8	-0.0028	-0.28	-3.63	-0.27	0.79
Podkarpackie	B9	-0.0054	-0.54	-7.00	-0.60	0.55
Podlaskie	B10	0.0007	0.07	0.97	0.07	0.94
Pomorskie	B11	0.0056	0.56	7.25	0.69	0.49
Śląskie	B12	-0.0194	-1.94	-25.28	-2.66	0.01
Świętokrzyskie	B13	0.0126	1.26	16.37	1.18	0.24
Warmińsko-Mazurskie	B14	-0.0052	-0.52	-6.72	-0.52	0.60
Wielkopolskie	B15	0.0127	1.27	16.49	1.63	0.11
Zachodniopomorskie	B16	-0.0099	-0.99	-12.88	Res.	

The results from the Knudsen's model (Table 4) shown above, in addition to the sectoral components (expenditure and regional groups, additionally give the possibility to draw conclusions about trends of changes in the studied areas of economic and social activity.) In particular, attention should be paid to taking into account the variability of temporal effects of the national component. The volatility of the temporal effects of the national component indicates a clear reduction of expenses - 2011/2010 by more than 3%, an even greater negative change is visible in more than 10% in 2015/2014. The values of t-stat and p-value statistics also prove the advantages of the Knudsen model. Most assessments of the

components of changes over time and sectoral components are here statistically significantly different from zero at a 5% level of significance.

Table 4. Results of the estimation of the Knudsen dynamic model of the decomposition of the growth rate of expenditure on culture by 4 groups of expenditures and 16 regions in Poland in 2004/2003 - 2016/2015.

	Parameter	**Evaluation**	**In %**	**%*13**	**t-stat**	**p-value**
Domestic Factor	A0	0.0605	**6.05**	**78.63**	9.46	0.00
Libraries	A1	-0.0150	-1.50	-19.50	-4.03	0.00
Cultural Centres	A2	-0.0081	-0.81	-10.52	-2.50	0.02
Heritage Protection	A3	-0.0198	-1.98	-25.71	-3.24	0.00
Others	A4	0.0527	5.27	68.45	Res.	
Dolnośląskie	B1	0.0036	0.36	4.73	0.50	0.62
Kujawsko-Pomorskie	B2	-0.0049	-0.49	-6.41	-0.55	0.58
Lubelskie	B3	0.0106	1.06	13.72	1.21	0.23
Lubuskie	B4	-0.0148	-1.48	-19.22	-1.48	0.14
Łódzkie	B5	-0.0026	-0.26	-3.36	-0.31	0.76
Małopolskie	B6	0.0059	0.59	7.68	0.77	0.45
Mazowieckie	B7	-0.0015	-0.15	-1.99	-0.22	0.83
Opolskie	B8	-0.0028	-0.28	-3.63	-0.27	0.79
Podkarpackie	B9	-0.0054	-0.54	-7.00	-0.60	0.55
Podlaskie	B10	0.0007	0.07	0.97	0.07	0.94
Pomorskie	B11	0.0056	0.56	7.25	0.69	0.49
Śląskie	B12	-0.0194	-1.94	-25.28	-2.66	0.01
Świętokrzyskie	B13	0.0126	1.26	16.37	1.18	0.24
Warmińsko-Mazurskie	B14	-0.0052	-0.52	-6.72	-0.52	0.60
Wielkopolskie	B15	0.0127	1.27	16.49	1.63	0.11
Zachodniopomorskie	B16	-0.0099	-0.99	-12.88	Res.	

Conclusions

The highest expenditures on culture and heritage protection were taken by the communes in Mazowieckie Region which seems to be the richest ones. It could be also confirmed by the empirical study carried out with the help of the static SSA model - the regional growth assessment for the communes of this region significantly exceeded the values of the

assessments for communes from other regions. The results of the assessment of the parameters of the generic structure of municipalities' expenditure on culture are also interesting, where a positive increase in the average expenditure rate was noted only for expenditure on the protection of monuments and national heritage. This may mean that the municipalities are trying to strengthen their cultural potential in order to increase the tourist attractiveness.

However, using the Berzeg's model, significantly different results were obtained than in the case of the static SSA model. These differences were visible both in terms of the growth of the regional national factor, and in particular in the case of assessments of growth components according to the groups' expenditure on communal budgets for culture.

The most reliable results were obtained by estimating the Knudsen's model, as evidenced by the t-stat and p-value values. The analysis of the results of this estimation revealed that in the case of assessments of components of regional growth, the communes in the Wielkopolskie, Świętokrzyskie, Lubelskie, Małopolskie and Dolnośląskie regions received positive evaluation values. In the case of this method, the downward trend in changes in the value of assessments of the national component was clearly visible. It means, that both Berzeg's and Knudsen's models are reliable tools to analyse the changes in public expenditures on culture, as well as in other genres of public sector (Jewczak and Żółtaszek, 2011).

References

Benito, B., Bastida, F. & Vicente, C. (2013). Municipal elections and cultural expenditure. *Journal of Cultural Economics, vol. 37,* pp. 3-32.

Berzeg, K., (1984) A note on statistical approaches to shift-share analysis. *Journal of Regional Science, vol. 24*, issue 2, pp. 277-285.

Depalo, D. & Fideli, S. (2011). *Heterogenecity in the Cultural Expenditures of Municipalities. Evidence from Italian Data (1998-2006)*. Working Paper no. 139, Sapienza University of Rome.

Getzner, M. (2004). Exploring voter preferences in cultural policy: a case study for Austria.*Empirica, vol. 31*(1), pp. 27-42.

Håkonsen, L. & Løyland, K. (2013). Local government allocation of cultural services, *ACEI working papers series,* p. 2, ACEI, *http://www.culturaleconomics.org/awp/AWP-06-2013.pdf.*

Jewczak, M. & Żółtaszek, A., (2011). Spatial and dynamic analysis of health care expenditures in OECD countries. *Folia Oeconomica, no. 252/2011*, 143-158.

Knudsen, D., (2000). Shift-share analysis: further examination of models for the description of economic change. *Socio-Economic Planning Sciences, vol. 34*, pp. 177-198.

Marimon, R. & Zilibotti, F., (1998). Actual versus virtual employment in Europe. Is Spain different? *European Economic Review, vol. 42*, 123–153.

Shulze, G. & Ursprung, H., (2000). La donna e mobile – or is she? Voter preferences and public support for the performing arts. *Public Choice, vol. 102*, 131-149.

Theil, H. & Gosh, R., (1980). A comparison of Shift-Share and RAS Adjustment. *Regional Science and Urban Economics, no. 10.*

Werck, K., Heyndels, B. & Geys, B., (2008). The impact of 'central places' on spatial spending patterns: evidence from Flemish local government cultural expenditures. *Journal of Cultural Economics, vol. 32*, 35-58.

Wrona, T., (2011). Doświadczenia 20 lat samorządu terytorialnego – stan finansów samorządowych. *Samorząd Terytorialny, no 7-8*, 5.

Cluster analysis of the World Gross-Domestic Product based on the emergent self-organization of a swarm

Michael C. Thrun[1]

Abstract

Cluster analysis is a task of unsupervised classification which seeks high-dimensional structures if natural clusters exist in data. The Databionic swarm (DBS) can adapt itself to structures of natural clusters characterized by distance and density based structures resulting in a topographic map and clustering. It is the first swarm-based technique that shows emergent properties while exploiting concepts of swarm intelligence, self-organization, and game theory. DBS was applied to the World GDP dataset which was constructed by selecting the purchasing power parity converted gross domestic product (GDP) per capita for the years: 1970-2010. The dynamic time warping distances were used to compute the optimal alignment between time series. The number of clusters was derived from, and the quality of the clustering and were verified by the topographic map which is a 3D representation of data structures. A clear cluster structure is also shown in heat map and silhouette plot. The rules deduced from CART show that the clusters are defined by an event occurring 2001. In its aftermath, the world economy was experiencing its first synchronized global recession in a quarter-century. Therefore, the first cluster consists not-affected, mostly African and Asian countries and a second cluster consists of affected countries which are mostly European and American countries.

Keywords: *machine learning, cluster analysis, swarm intelligence, visualization, self-organization*
JEL Classification: O47, F01, C380
DOI: 10.14659/SEMF.2018.01.53

1 An approach to cluster analysis for multivariate time series

The multivariate time series inspected in this work covers repeated measures of the gross domestic product (GDP) of 190 countries published in (Heston et al., 2012). The GDP consists of the total market value of all final goods and services produced in a country. Thus, the GDP is an indicator of economic achievement of a country Mazumdar (2000). Each country's data has to be converted into common currency to compare the GDP between various currencies. An exchange rate is defined through the purchasing power parity PPP at which the currency of a country is converted into that of another country to purchase the same volume of goods and services in both countries (Rogoff, 1996).

The World GDP data set of was extracted from the multivariate time series of (Heston et al., 2012) by selecting the PPP-converted GDP per capita for the years from 1970 to 2010 by

[1] Corresponding author: University of Marburg, Data Bionics Research Group, Hans-Meerwein-Straße 6, D-35032 Marburg, mthrun@mathematik.uni-marburg.de.

Leister (2016). In this work, the World-GDP data set will be investigated in the context of economic similarity between nations by using cluster analysis.

In cluster analysis, the methods rely on some concept of the similarity between pieces of information encoded in the data of interest. However, no accepted definition of clusters exists in the literature (Hennig et al., 2015, p. 705). Additionally, Kleinberg showed for a set of three simple axioms, scale-invariance, consistency, and richness, that there exists no clustering algorithm which can satisfy all three (Kleinberg, 2003). By concentrating on distance and density based structures, this work restricts clusters to "natural" clusters (c.f. Duda et al., 2001) and therefore omits the axiom of richness where all partitions should be achievable. Thus, natural clusters consists of objects which are similar within clusters and dissimilar between clusters. "[Clusters] can be of arbitrary shapes [structures] and sizes in multidimensional pattern space. Each clustering criterion imposes a certain structure on the data, and if the data happen to conform to the requirements of a particular criterion, the true clusters are recovered." (Jain and Dubes, 1988). Here, the Databionic swarm (DBS) is used (Thrun, 2018) to find natural clusters without imposing a particular structure on the data contrary to conventional algorithms. An example of an algorithm imposing structures would be spectral clustering which searches for clusters with "chain-like or other intricate structures" (Duda et al., 2001) (see also Hennig et al., 2015). Spectral clustering lacks "robustness when there is little spatial separation between the clusters" (Handl et al., 2005). Such effects were made visible on simple artificial datasets for conventional algorithms (Thrun, 2018).

2 Distance-based cluster algorithm

The Databionic swarm (DBS) implements a swarm of agents interacting with one another and sensing their environment. DBS can adapt itself to structures of high-dimensional data such as natural clusters characterized by distance and density based structures in the data space (Thrun, 2018).

The DBS algorithm consists three modules. First, the projection method Pswarm, second the visualization technique of a topographic map based on the generalized U-matrix, and third, the clustering approach itself.

Pswarm is a swarm of intelligent agents called DataBots (Ultsch, 2000). It is a parameter-free focusing projection method of a polar swarm that exploits concepts of self-organization and swarm intelligence (Thrun, 2018). During construction of this type of projection, which is called learning phase and requires an annealing scheme, structure analysis shifts from global optimization to local distance preservation (focusing). Intelligent agents of Pswarm operate on

a toroid grid where positions are coded into polar coordinates allowing for a precise definition of their movement, neighborhood function and annealing scheme. The size of the grid and, in contrast to other focusing projection methods (e.g. Ultsch and Lötsch, 2017) the annealing scheme is data-driven and therefore, this method does not require any parameters. During learning, each DataBot moves across the grid or stay in its current position in the search for the most potent scent that means it searches for other agents carrying data with the most similar features to itself with a data-driven decreasing search radius (Thrun, 2018). Contrary to other projections methods and similar to the emergent self-organizing map, the Pswarm projection method does not possess a global objective function which allows the method to apply self-organization and swarm intelligence (Thrun, 2018).

Second, the projected points[2] are transformed to points on a discrete lattice; these points are called the best-matching units (BMUs) $bmu \in B \subset \mathbb{R}^2$ of the high-dimensional data points j. Then the generalized U*-matrix can be applied to the projected points by using a simplified emergent self-organizing map (ESOM) algorithm which is an unsupervised neural network (Thrun, 2018). The result is a topographic map with hypsometric tints (Thrun, Lerch, Lötsch, and Ultsch, 2016). Hypsometric tints are surface colors that represent ranges of elevation (see (Thrun et al., 2016)). Here, contour lines are combined with a specific color scale. The color scale is chosen to display various valleys, ridges, and basins: blue colors indicate small distances (sea level), green and brown colors indicate middle distances (low hills), and white colors indicate vast distances (high mountains covered with snow and ice). Valleys and basins represent clusters, and the watersheds of hills and mountains represent the borders between clusters. In this 3D landscape, the borders of the visualization are cyclically connected with a periodicity (L,C). A central problem in clustering is the correct estimation of the number of clusters. This is addressed by the topographic map which allows assessing the number of clusters (Thrun et al., 2016).

Third, in (Lötsch and Ultsch, 2014) it was shown that a single wall of the AU-matrix represents the actual distance information between two points in the high-dimensional space: the generalized U-matrix is the approximation of the abstract U-matrix (AU-matrix) (Lötsch and Ultsch, 2014). Voronoi cells around each projected point define the abstract U-matrix (AU-matrix) and generate a Delaunay graph. For every BMU all direct connections are weighted using the input-space distances $D(l, j)$, because on each border between two Voronoi cells a height is defined.

[2] Of DataBot positions on the hexagonal grid of Pswarm.

For the distances $D(l,j)$ the dynamic time warping (DTW) distances were calculated using the CRAN package in R "dtw" (Giorgino, 2009). "The DTW distance allows warping of the time axes to align the shapes of the two times series better. The two series can also be of different lengths. The optimal alignment is found by calculating the shortest warping path in the matrix of distances between all pairs of time points under several constraints. The point-wise distance is usually the Euclidean. The DTW is calculated using dynamic programming with time complexity O(n2)" (Mörchen, 2006).

Now, the distances between two points in the high-dimensional space is considered as the distance between two time series. All possible Delaunay paths $p_{j,l}$ between all points are calculated toroidal because the topographic map is toroidal. Then, the minimum of all possible path distances $p_{j,l}$ between a pair of points $\{j,l\} \in O$ in the output space is calculated as the shortest path $G(l,j,\mathcal{D})$ using the algorithm of Dijkstra resulting in a new high-dimensional distance $D^*(l,j)$. Here, the compact approach is used, where the two clusters with the minimal variance S are merged to together until given the number of clusters defined by the topographic map is reached.

Let $c_r \subset I$ and $c_q \subset I$ be two clusters such that $r,q \in \{1,\dots,k\}$ and $c_r \cap c_q = \{\}$ for $r \neq q$, and let the data points in the clusters be denoted by $j_i \in c_q$ and $l_i \in c_r$, with the cardinality of the sets being $k = |c_q|$ and $p = |c_r|$ and

$$\Delta Q(j,l) = \frac{k * p}{k + p} D^*(l,j)$$

then, the variance between two clusters is defined as

$$S(c_r, c_k) = \sum_{i=1,j=1,j\neq i}^{k,p} \Delta Q(j,l) \quad .$$

A dendrogram can be shown additionally. The clustering is valid if mountains do not partition clusters indicated by colored points of the same color and colored regions of points. The algorithm was run using the CRAN package in R "DatabionicSwarm".

3 Application: PPP-converted gross domestic product (GDP) per capita

The World GDP data set was logarithmized, and countries with missing values were not considered. As a result, 160 countries[3] remain for which the optimal alignment between two time series (Giorgino, 2009) is calculated and the DBS algorithm applied.

[3] For overview see (Leister, 2016, pp. 105-107)

In contrast to most conventional clustering algorithms, the topographic map allows identifying that clustering of the data is meaningless if the data contains no (natural) clusters (Thrun, 2018). Thus, Fig. 1 demonstrates a clear (natural) cluster structure. The homogeneity of the cluster structures of DBS is visualized in a silhouette plot in Fig. 3, and it is confirmed by the heat map (Fig. 2). In Fig. 4 the Classification and Regression Tree (CART) analysis is shown. The clusters are defined mainly by an event that occurred in 2001. The rules generated from the CART are presented in Table 1, and applied as colored labels to a world map in Fig. 5 with the same colored points as Fig. 1.

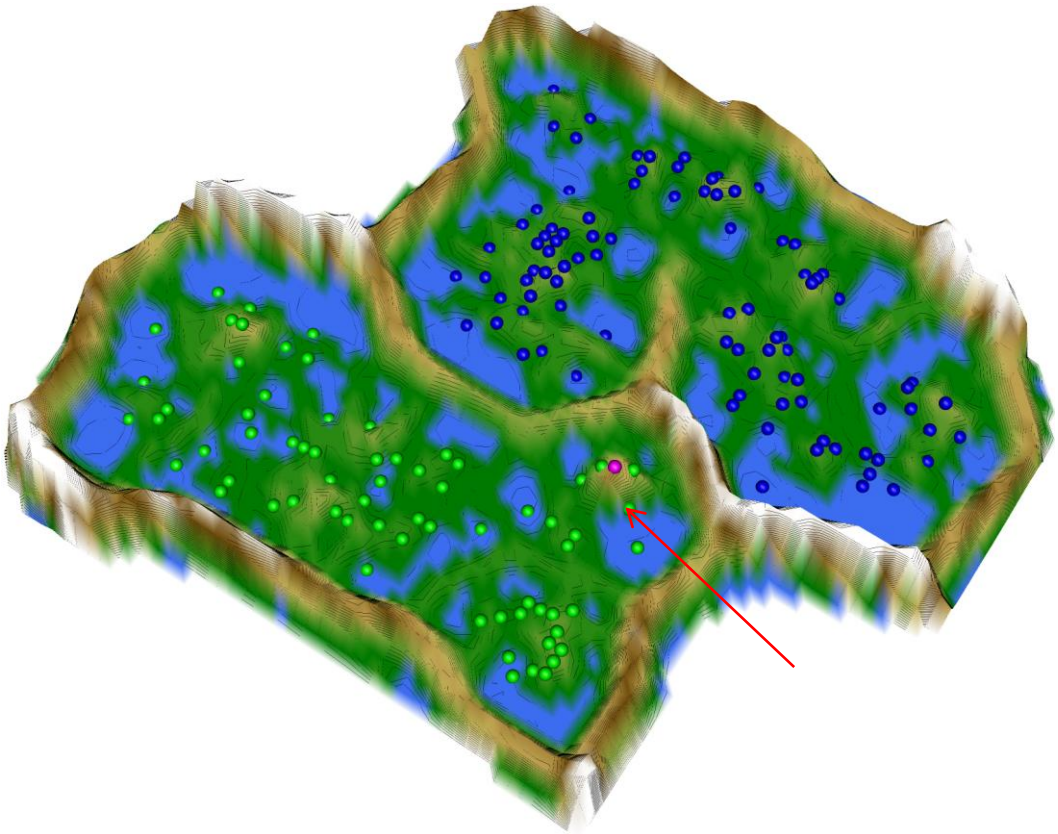


Fig. 1. Topographic map of the DBS clustering of the World GDP data set shows two distinctive clusters. There is one outlier, colored in magenta and marked with a red arrow. The visualization was generated by the CRAN package in R "DatabionicSwarm".

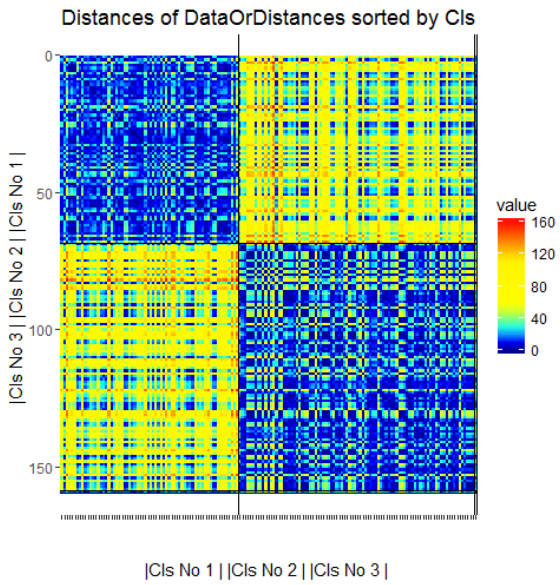


Fig. 2. Heatmap of the dynamic time warping (DTW) distances for the World GDP data set shows a small variance of intracluster distance. The visualization was generated by the CRAN package in R "DataVisualizations".

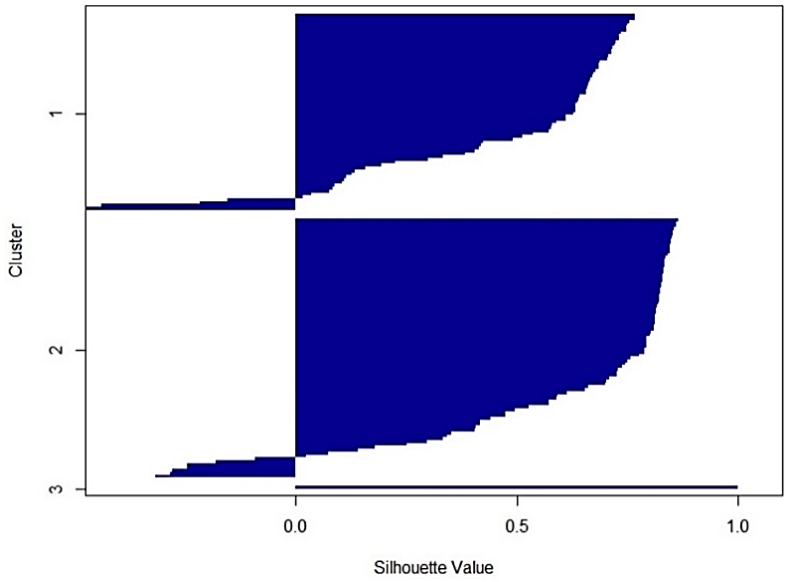


Fig. 3. Silhouette plot of the DBS clustering results for the World GDP data set indicates that data points (y-axis) above a value of 0.5 (x-axis) have been assigned to an appropriate cluster. The visualization was generated by the CRAN package in R "DataVisualizations".

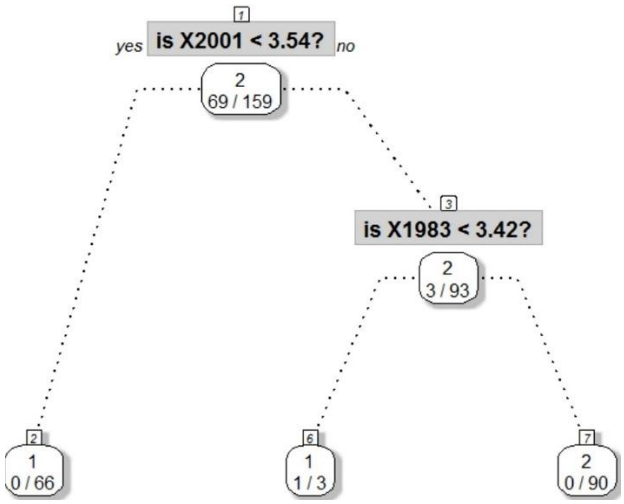


Fig. 4. Classification and Regression Tree (CART) analysis rules for the clusters. The two main clusters are defined only by an event in 2001.

Table 1. The CART rules based on Fig. 4 in which cluster of Fig.1 is used. Egypt, Micronesia and the outlier Equatorial Guinea classified incorrectly by these two rules.

Rule No.	DBS Cluster No.	No. of Nations	Rules
R1	1	66	BIP lower than 3469 U in the year 2001
R2	2	93	BIP higher than 3469 U in the year 2001

Notes: Abbreviations - U: PPP-converted GDP per capita.

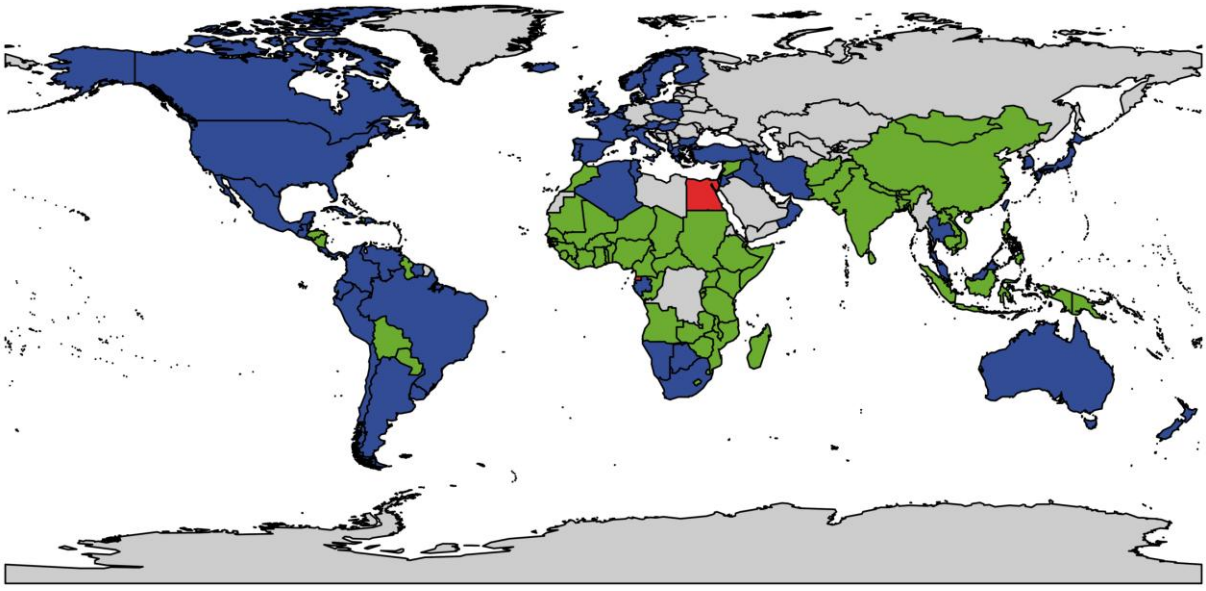


Fig. 5. Two rules of Table 1 classify countries in blue and green. For grey countries, no data was available in (Leister, 2016). The rules result from the clustering of Fig. 1. In red are the Outlier Equatorial Guinea as well as the incorrectly classified countries Egypt and Micronesia by these two rules.

4 Discussion

Regional cluster analysis on GDP datasets was performed for Latin American countries in (Redelico et al., 2009) and European countries in (Gallo and Ertur, 2003). To the knowledge of the author, no cluster analysis of the whole world was performed with the goal to explain the clusters by rules and through a spatial world map (Fig. 5). Here, both clusters found by the Databionic swarm are spatially separated. The first cluster consists mostly of African and Asian countries and the second clusters of industrialized countries of predominantly Europe and America. Due to the correlations between the human development index (HDI) and PPP-converted log(GDP) per capita shown in Fig. 3.2 on page 69 in (UNDP, 2003), the second cluster in Fig. 5 is highly similar to the HDI map of Fig. 1 in (Birdsall and Birdsall, 2005) with HDI higher than 0.7.

The two rules of the (CART) analysis which are presented in Table 1, demonstrate that the clusters are defined by an event that occurred in 2001, which could be the crashing of airplanes into the World Trade Center. In its aftermath, "the world economy was experiencing its first synchronized global recession in a quarter-century" (Makinen, 2002, p. 17). Therefore, the results indicate that the first cluster of African and Asian countries was unaffected by this event, and the second cluster of American and European countries was affected. As published in Vollmer et al. (2013), the GDP is sensitive by economic shocks, e.g., oil-price of excludingly oil-exporting countries or countries with a low number of inhabitants (Vollmer et al., 2013). The data regarding the PPP-converted GDP per capita of Egypt may be misrepresented, because "during the twentieth century the population of Egypt has increased by more than 5 times" (El Araby, 2002). The outlier in Fig. 1 describes the data of Equatorial Guinea. This small country with an area of 28,000 square kilometers is mostly based on oil and is, one of sub-Saharan Africa's largest oil producers. The Federated States of Micronesia is a subregion of Oceania has only a low number of inhabitants (105.000). Thus, it could also be an outlier.

Conclusions

The Databionic swarm (DBS) resulted in a coherent spatiotemporal clustering of the multivariate time series of the PPP-converted gross domestic product (GDP) per capita of 160 countries in the years 1970 to 2010. It seems that 157 countries can be classified by using two rules extracted from the CART with only one threshold for the GDP in the year 2001. This indicates that the economic achievement of these countries were profoundly affected in the year 2001. DBS can be by downloaded as the CRAN package in R "DatabionicSwarm".

References

Birdsall, S. & Birdsall, W. (2005). *Geography matters: Mapping human development and digital access. 2005*. doi:10.5210/fm.v10i10.1281.

Duda, R. O., Hart, P. E. & Stork, D. G. (2001). *Pattern Classification (2nd ed.)*. Ney York, USA: John Wiley & Sons.

El Araby, M. (2002). Urban growth and environmental degradation: The case of Cairo, Egypt. *Cities, 19*(6), 389-400.

Gallo, J. & Ertur, C. (2003). Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995. *Papers in regional science, 82*(2), 175-201.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: the dtw package. *Journal of statistical Software, 31*(7), 1-24.

Handl, J., Knowles, J. & Kell, D. B. (2005). Computational cluster validation in post-genomic data analysis. *Bioinformatics, 21*(15), 3201-3212.

Hennig, C., Meila, M., Murtagh, F. & Rocci, R. (2015). *Handbook of cluster analysis*. New York, USA: CRC Press.

Heston, A., Summers, R. & Aten, B. (2012). Penn World Table Version 7.1 Center for International Comparisons of Production. *Income and Prices at the University of Pennsylvania*.

Jain, A. K. & Dubes, R. C. (1988). *Algorithms for Clustering Data* (Vol. 3). Englewood Cliffs, New Jersey, USA: Prentice Hall College Div.

Kleinberg, J. (2003). *An impossibility theorem for clustering.* Paper presented at the Advances in neural information processing systems.

Leister, A. M. (2016). *Hidden Markov models: Estimation theory and economic applications.* Dissertation, Marburg, Germany.

Lötsch, J. & Ultsch, A. (2014). *Exploiting the Structures of the U-Matrix.* Paper presented at the Advances in Self-Organizing Maps and Learning Vector Quantization, Mittweida, Germany.

Makinen, G. (2002). *The economic effects of 9/11: A retrospective assessment*. Library of congress, Washington D.C. congressional research service.

Mazumdar, K. (2000). Causal flow between human well-being and per capita real gross domestic product. *Social Indicators Research, 50*(3), 297-313.

Mörchen, F. (2006). *Time series knowledge mining*. Dissertation, Marburg, Germany: Citeseer/Görich & Weiershäuser.

Redelico, F. O., Proto, A. N. & Ausloos, M. (2009). Hierarchical structures in the Gross Domestic Product per capita fluctuation in Latin American countries. *Physica A: Statistical Mechanics and its Applications, 388*(17), 3527-3535.

Rogoff, K. (1996). The purchasing power parity puzzle. *Journal of Economic literature, 34*(2), 647-668.

Thrun, M. C. (2018). *Projection Based Clustering through Self-Organization and Swarm Intelligence.* A. Ultsch & E. Hüllermeier Adv.. Dissertation, Heidelberg: Springer.

Thrun, M. C., Lerch, F., Lötsch, J., & Ultsch, A. (2016). *Visualization and 3D Printing of Multivariate Data of Biomarkers*. Paper presented at the International Conference on Computer Graphics, Visualization and Computer Vision (WSCG), Plzen.

Ultsch, A. (2000). *Clustering with DataBots*. Paper presented at the Int. Conf. Advances in Intelligent Systems Theory and Applications (AISTA), Canberra, Australia.

Ultsch, A. & Lötsch, J. (2017). Machine-learned cluster identification in high-dimensional data. *Journal of biomedical informatics, 66*, 95–104. DOI:10.1016/j.jbi.2016.12.011.

UNDP. (2003). *Human development Report: Millennium Development Goals: a compact among nations to end poverty*. United Nations Development Programme (UNDP), New York: Oxford University Press.

Vollmer, S., Holzmann, H. & Schwaiger, F. (2013). Peaks vs components. *Review of Development Economics, 17*(2), 352-364.

Effects of the pay-out system of income taxes to municipalities in Germany

Michael C. Thrun[1], Alfred Ultsch[2]

Abstract

The methods and possibilities of data mining for knowledge discovery in economic data are demonstrated on data of the German system of allocating tax revenues to municipalities. This system is complex and not easily understandable due to the involvement of several layers of administration and legislation. The general aim of the system is that a share of income tax revenue for a municipality (system output) should be a fixed part of the total income tax yield of each municipality (system input). Tools for the scientific exploration of empirical distributions are applied to the input and output variables of the system. The main finding is that, although the critical input variables are unimodally distributed, the output variable showed a bimodal distribution. The conclusion from this finding is that the system works in two distinct states: municipalities receive either a large share or a small share of the input. Relating these states of the system to the location of the municipality a distinct east-west gradient is found. A significantly larger percentage of East-German municipalities receive in average 5% less of the taxes share. This paper focuses on the methods and tools for such type of knowledge discovery.

Keywords: *machine learning, Gaussian mixture models, density estimation, statistical analysis, income tax*
JEL Classification: C310, C380, P170
DOI: 10.14659/SEMF.2018.01.54

1 Introduction

Federal legislation determines the basic rules on how to calculate the income tax share a municipality receives (Ultsch and Behnisch, 2017). A fixed percentage of the total income tax a municipality pays to the state is restituted back to this municipality (details in Ultsch and Behnisch, 2017). The actual share of income tax is determined by the state in which the municipality is located by allocation keys which are used to specify this percentage (Ultsch and Behnisch, 2017). The overall effects of this system is investigated for all municipalities in Germany in the form of an input-output analysis by investigating distributions of extracted features and Bayesian classification of two-dimensional density.

Descriptions of distributions using a single distribution, like Lognormal or Gamma are often quite weak in describing the tails of the distribution (Dagum, 1977). They lead to separate models for the upper vs. lower parts of income distributions (c.f. Thrun and Ultsch,

[1] Corresponding author: University of Marburg, Data Bionics Research Group, Hans-Meerwein-Straße 6, D-35032 Marburg, mthrun@mathematik.uni-marburg.de.
[2] University of Marburg, Data Bionics Research Group, Hans-Meerwein-Straße 6, D-35032 Marburg, ultsch@mathematik.uni-marburg.de.

2015). These approaches can be improved and simplified using Gaussian mixture models (Thrun and Ultsch, 2015).

This approach is applied to a one- and two-dimensional analysis of the German pay-out system. This paper focuses on methods and tools. Spatial planning results were published primarily in (Ultsch and Behnisch, 2017).

2 An Approach for Distribution Analysis

A scientifically sound procedure for the identification and analysis of empirical distributions is a comparison to a known theoretic distribution. The quantile/quantile plot (QQ-plot) allows comparing an empirical distribution to a known distribution (Michael, 1983). Here, in 100 quantiles the model of a Gaussian distribution is compared to the data, and a straight line confirms a good data fit of the model. The Gaussian distribution is the canonical starting point for such a comparison. If the distribution is not Gaussian, the box-cox transformation can be applied. It aims to construct a nonlinear transformation of a variable x such that the transformed variable comes as close as possible to a Gaussian distribution (Asar et al., 2014; Box and Cox, 1964). However, the estimated powers of the box-cox transformation are seldom integers. The box-cox values may be orderly arranged by Tuckey's ladder of power relative to the nearest whole number to get understandable nonlinear transformations of a variable (Tukey, 1977).

In the case of this work, the precise form, i.e., the type, nature and parameters of the formal model of the probability density function (pdf) is the ultimate goal of this analysis. Usually, this is performed using kernel density estimators. The simplest of such a density estimation is the histogram. However, histograms are often misleading and require critical parameters such as the width of the bin (Keating and Scott, 1999). A specially designed density estimation, which has been successfully proved in many practical applications is the "Pareto Density Estimation" (PDE). PDE consists of a kernel density estimator representing the relative likelihood of a given continuous random data (Ultsch, 2005). PDE has been shown to be particularly suitable for the discovery of structures in continuous data hinting at the presence of distinct groups of data and particularly suitable for the discovery of mixtures of Gaussians (Ultsch, 2005). The parameters of the kernels are auto-adopted to the date using an information theoretic optimum on skewed distributions (Ultsch et al., 2015). PDE can be used as an effective method to estimate density in two dimensions (Ultsch, 2005) (scatter density plots) showing the relationships and dependencies between two features. In general, bivariate relations are visualized using scatterplots where two features are plotted against each

other as points. In a scatter density plot the densities of these points can be estimated and visualized (c.f. Berthold et al., 2010, p. 45).

If the distribution of data is more complicated than a single distribution, the underlying process which produces the data may operate in different states, which have their factors to influence the data. Then mixture models can be applied as the standard statistical tool (Fraley and Raftery, 2002; Ultsch et al., 2015). Gaussian mixture models (GMM) are a superposition of a weighted sum of single modes. Each mode (component) consists of a single Gaussian parameterized with the mean and standard deviation (Bishop, 2006). The GMM is a weighted sum of *M* component Gaussian densities as given by the equation

$$p(\vec{r}) = \sum_{i=0}^{M} w_i \, N(\vec{r}|\vec{m}_i, \vec{s}_i) = \sum_{i=1}^{M} w_i \cdot \frac{1}{\sqrt{2\pi s_i}} \cdot e^{-\frac{(\vec{r}-\vec{m}_i)^2}{2*\vec{s}_i^2}} \quad (1)$$

where $N(\vec{r}|\vec{m}_i, \vec{s}_i)$ denotes the two-dimensional Gaussian probability densities (component, mode) with means $\vec{m}_i$ and standard deviations, $\vec{s}_i$. The w_i are the mixture weights indicating the relative contribution of each component Gaussian to the overall distribution, which add up to a value of 1. *M* denotes the number of components in the mixture. Usually, the parameters of the GMM, including the number of components M, are optimized using the expectation maximization (EM) algorithm (Dempster et al., 1977). A GMM represents the presence of subclasses within a complete data set. Precise limits for these classes can be calculated using the theorem of Bayes (Duda et al., 2001). A GMM can be visually verified by a QQ-plot and statistically by Xi-Quadrat test (cf. Thrun and Ultsch, 2015).

Datasets

The dataset of (Ultsch and Behnisch, 2017) has been used. In 2010 there were 11,669 municipalities, and 228 so-called "unincorporated areas" generally forested areas, lakes and larger rivers, of 16 states resulting in thirteen territorial states: Schleswig-Holstein and Hamburg (1,117), Lower Saxony and Bremen (1,026), North Rhine-Westphalia (396), Hessen (426), Rhineland-Palatinate (2,306), Baden-Wuerttemberg (1,101), Bavaria (2,056), Saarland (52), Brandenburg and Berlin (420), Mecklenburg-Vorpommern (814), Saxony (485), Saxony-Anhalt (300) and Thuringia (942). The number of taxpayers per municipality has been given in this data set.

Data on the income tax per taxpayer collected by each municipality from its population and transferred to the state (Municipality Income Tax Yield, MTY) has not directly been available. To estimate MTY, Tax2007 and Tax2010 of the Regional Database Germany (see Ultsch and Behnisch, 2017) has been used. The sum of the income tax revenues paid by the

state to a municipality has been obtained from Regional Database Germany (see Ultsch and Behnisch, 2017). The income tax share (**ITS**) of a municipality per taxpayer in 2010 has been calculated as TaxShare divided by the number of taxpayers in the year 2010.

3 Results

MTY can be estimated through a scatter density plot between Tax2007 and Tax2010. It is presented in Fig. 1 that the data on the wage and income statistics after-tax return for the year 2007 (Tax2007) as well as for the year 2010 (Tax2010) allows a reasonable estimation of MTY for 2010. The true MTY value is located somewhere between Tax2007 and Tax2010. The Pearson Correlation between the two variables exceeded 92% in a range of up to 10,000 EUR. Therefore, Tax2010 divided by the number of taxpayers in a municipality is taken as a proper estimation for MTY in 2010. However, MTY could not be obtained for n=247 municipalities due to restrictions on data protection and problems of data availability. These municipalities are disregarded in the following distribution analysis. In Fig. 2 and 3 the distributions of ITS and MTY are analyzed. MTY is unimodally distributed, but ITS is multimodal distributed. The first maximum of ITS lies at 300 EUR and the second at 640 EUR.

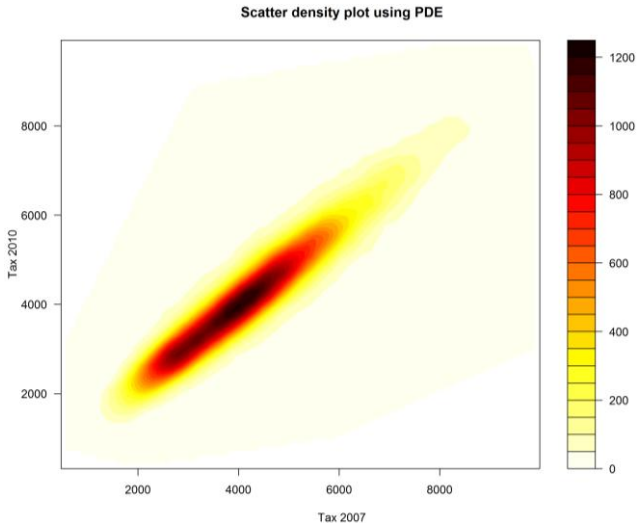


Fig. 1. Scatter density plot of Tax 2010 versus Tax 2007 shows only one mode with a high correlation verified by the Pearson correlation value of 92%.

In the next step, the scatter density plot between the input MTY and the output ITS is presented in Fig. 4. In Contrast to Fig. 1 two modes are visualized by using the PDE. Therefore, a two-dimensional GMM is calculated, and from this GMM a Bayesian

classification is derived. The result of the classification is presented in a simple scatter plot in Fig. 5. It is apparently visible that in the scatter plot the two modes are not recognizable if the PDE is disregarded.

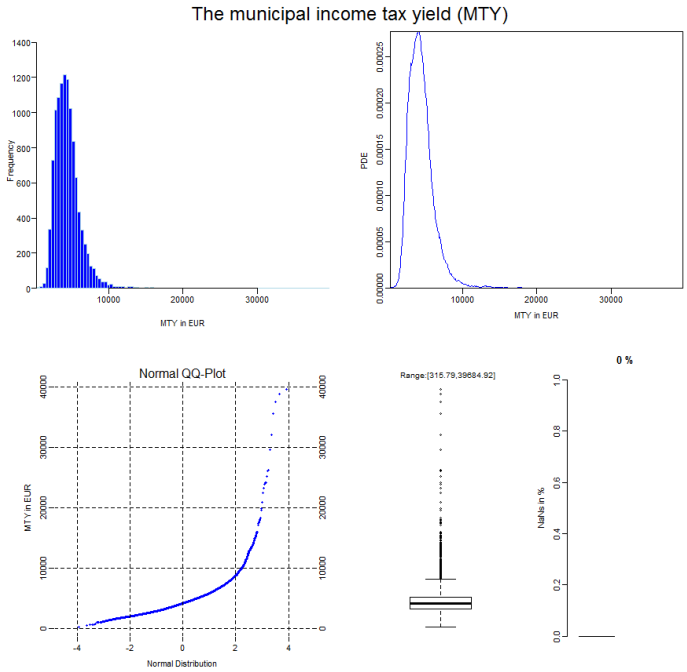


Fig. 2. The distribution of MTY is unimodal shown in the histogram and PDE plot. However, it is not Gaussian because the QQ-plot does not have a straight line. The box plot shows some outliers, and the feature has no missing values.

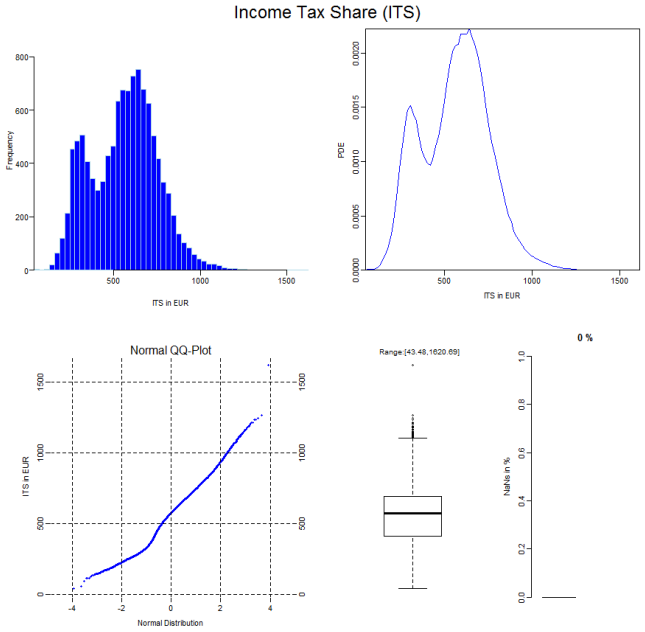


Fig. 3. The Distribution of ITS is multimodal indicated by an s-curve in the QQ plot additionally to the two modes in the histogram and PDE plot. The boxplot shows one large outlier.

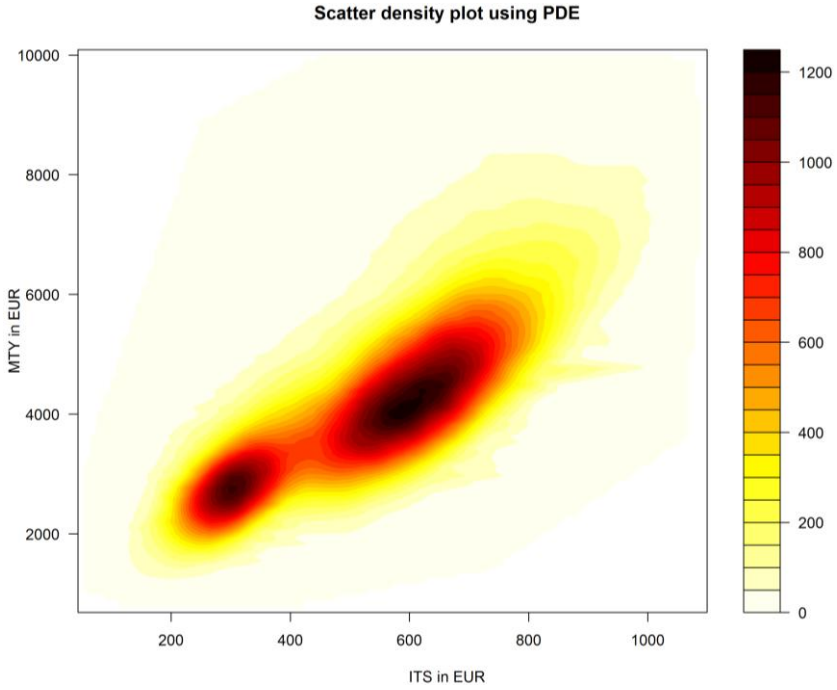


Fig. 4. Scatter density plot using Pareto density estimation of MTY versus ITS shows two distinctive modes.

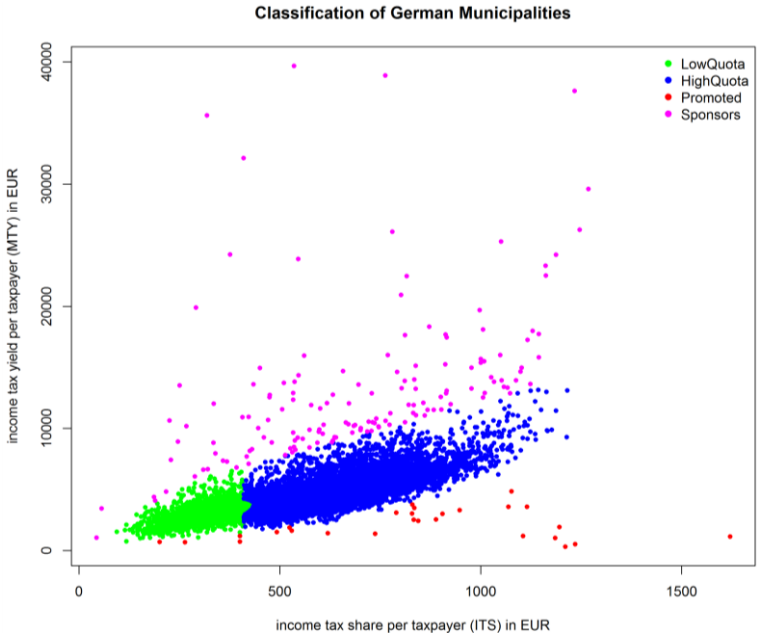


Fig. 5. A scatterplot of MTS versus ITS colored by the labels of a Bayesian classification showing two main groups of low quota and high quota municipalities. Additionally, outliers are manually classified into two separated groups called sponsors and promoted.

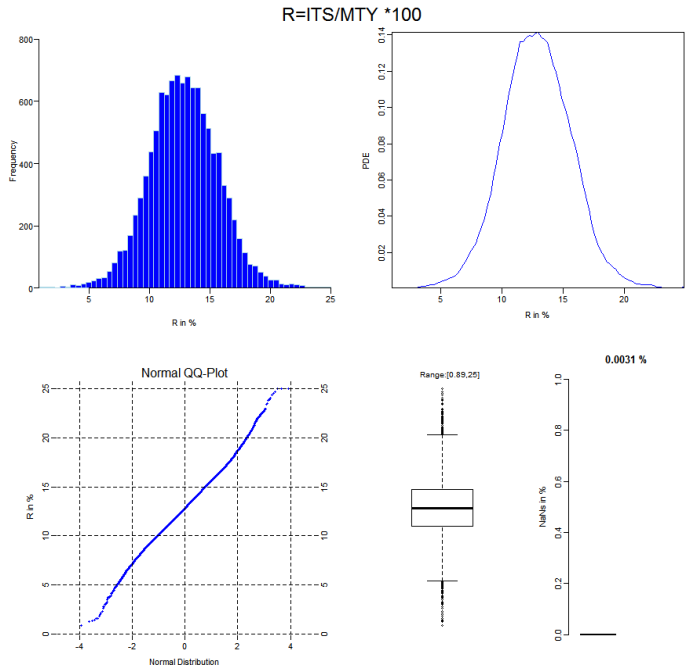


Fig. 6. The ratio R= ITS/MTY*100 is Gaussian distributed. To better visualize the relevant range, values above 25 % as manually set as missing values (0.0031% of data).

Fig. 6 presents the ratio R = ITS/MTY*100 of the payments the municipalities received from the respective state governments (ITS) to the income taxes yield the municipality transferred

from its taxpayers to the state (MTY). The units are percentages [%]. The amount of the collected tax yields described by the ratio has a mean of 13% ± 3%. Accordingly, a German municipality expects to receive 13% funding from the state on average (Fig. 6). Hence, the ratio R allows to describe the four classes as deviations from the expectation (see Fig. 5): high-quota, low-quota, sponsors and promoted.

4 Discussion

Ideally, the system input and output should be proportional, i.e., the income tax share per taxpayer (ITS) should be proportional to the municipal income tax yield (MTY). The fact that this ratio is Gaussian distributed implies that there should be only unintentional deviations from this general percentage. Therefore, the payment of income tax revenues should result in a direct proportionality between tax payments to and funding from the state government. A municipality should expect a certain fixed percentage of the taxes it delivers (Fig. 6). If the deviations from the base percentage ITS are as unintentional, the distribution of ITS should be unimodal. Instead, distribution analysis reveales that the pdf of ITS consists of two modes. Thus, the payout system of income taxes discriminates between low quota and high quota classes of municipalities.

Next, the input of the system (MTY) was connected to the output of the system (ITS) in order to understand this effect of inequity. A clear separation into two distinct distributions can be observed in Fig. 4. For a vast majority of German municipalities that pay income tax per taxpayer of approx. 2,500 to 4,500 EUR to the state, the refund can be either low (~ 10%) or high (~ 20%). Two 2-dimensional Gaussian mixtures could efficiently model the two modus operandi of the pay-out system of income taxes to municipalities. Using the deviation from the expected ratio, the groups of the classification are distinguished as low quota vs. high quota classes. This classification is expanded to include a class of “promoted”-municipalities receiving a substantially larger share of income taxes (above 30%) and a class of “sponsors”-municipalities receiving a considerably smaller share of income taxes (less than 8%).

Conclusions

This work shows that distribution analysis of features itself allows generating insights besides being the prerequisite for cluster analysis (Thrun, 2018, pp. 16-17). Further, the correct estimation of density is crucial for estimating the pdf and improves a scatterplot significantly. The methods are exemplarily applied to Germany’s complex system of allocating tax

revenues. The geographical distribution of the low-quota vs. high-quota municipalities reveals an evident east-west disparity (cf. Ultsch and Behnisch, 2017, p. 29, Fig. 9). The percentage of income tax revenues which a municipality received per taxpayer depends on the location of the municipality. If the municipality is located in western Germany, the municipality could expect about 15-30%. On the contrary, if the municipality is located towards the east, its share is more likely to be only 10% or less. All visualizations are available in the CRAN package "DataVisualizations" in R. The German Research Foundation partly funded the research under grant agreement (BE4234/3-1, UL159/10-1). The DFG had no role in the study design, data selection, and analysis, the decision to publish or preparation of the manuscript.

References

Asar, Ö., Ilk, O. & Dag, O. (2014). Estimating Box-Cox power transformation parameter via goodness of fit tests. *Communications in Statistics-Simulation and Computation, 46*(1), 91-105.

Berthold, M. R., Borgelt, C., Höppner, F. & Klawonn, F. (2010). *Guide to intelligent data analysis: how to intelligently make sense of real data*: Springer Science & Business Media.

Bishop, C. (2006). *Pattern recognition and machine learning*. New York: Springer.

Box, G. E. & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 211-252.

Dagum, C. (1977). New model of personal income-distribution-specification and estimation. *Economie appliquée, 30*(3), 413-437.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B, 39*(1), 1-38.

Duda, R. O., Hart, P. E. & Stork, D. G. (2001). *Pattern Classification. 2nd. Edition.* New York.

Elsner, H. (1979). *Das Gemeindefinanzsystem: Geschichte, Ideen, Grundlagen*. Köln: Kohlhammer Verlag.

Fraley, C. & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of American Statistical Association, 97*(611-631).

Keating, J. P. & Scott, D. W. (1999). A primer on density estimation for the great homerun race of 1998. *STATS, 25*, 16-22.

Michael, J. R. (1983). The stabilized probability plot. *Biometrika, 70*(1), 11-17.

Thrun, M. C. (2018). *Projection Based Clustering through Self-Organization and Swarm Intelligence* (A. Ultsch & E. Hüllermeier Eds.). DOI: 10.1007/978-3-658-20540-9, Heidelberg: Springer.

Thrun, M. C. & Ultsch, A. (2015). *Models of Income Distributions for Knowledge Discovery.* Paper presented at the European Conference on Data Analysis, DOI: 10.13140/RG.2.1.4463.0244, Colchester.

Tukey, J. W. (1977). *Exploratory data analysis* (1st ed.). ISBN: 978-0201076165, Pearson Education (US), United States.

Ultsch, A. (2005). Pareto density estimation: A density estimation for knowledge discovery. In: Baier, D. & Werrnecke, K. D. (Eds.), *Innovations in classification, data science, and information systems* (Vol. 27, pp. 91-100). Berlin, Germany: Springer.

Ultsch, A. & Behnisch, M. (2017). Effects of the payout system of income taxes to municipalities in Germany. *Applied Geography, 81*, 21-31.

Ultsch, A., Thrun, M. C., Hansen-Goos, O. & Lötsch, J. (2015). Identification of Molecular Fingerprints in Human Heat Pain Thresholds by Use of an Interactive Mixture Model R Toolbox (AdaptGauss). *International journal of molecular sciences, 16*(10), 25897-25911.

The quality of life in Poland and Germany

Paweł Ulman[1], Tomasz Kwarciński[2]

Abstract

Due to the fact that Poland aspires to be an economically developed country, it appears to be important to compare the quality of life of Polish citizens with the quality of life of citizens in a fully developed country like Germany. The aim of the paper is to investigate whether the quality of life evaluation scheme in Poland is similar to the scheme in Germany. Statistical data from the European Quality of Life Survey is used to achieve this aim. The subjective self-evaluation of happiness is compare to the objective quality of lifebased on the capability approach. The aggregate measure of quality of life is constructed on the basis of the Total, Fuzzy and Relative approach. The factors of quality of life which are the most divergent from subjective perception of happiness are indicated. Finally, in order to measure the diversity of these assessments the mobility index, applicable to the study of structural changes, is used.

Keywords: *quality of life, happiness, well-being, mobility index*
JEL Classification:I31, I32,
DOI: 10.14659/SEMF.2018.01.55

1 Introduction

The aim of the paper is to investigate whether the quality of life (well-being) evaluation scheme in Poland is similar to the scheme in Germany. Due to the fact that Poland aspires to be an economically developed country, it appears to be important to compare the quality of life of Polish citizens with the quality of life of citizens in a fully developed country like Germany.

Defining the quality of life, we will take into account hedonistic as well as objectivistic approaches to personal well-being. The hedonistic approach focuses on personal happiness or life satisfaction, and is a subjective well-being theory (S), while the objectivistic approach belongs to the objective list theories according to which "certain things are good or bad for us, whether or not we want to have the good things, or to avoid the bad things" (Parfit, 1984).

The objective quality of life refers to Amartya Sen's and Martha Nussbaum's capability approach. According to Sen (2005), a person's capability is defined in terms of the set of valuable "doings" or "beings" i.e. what a person is able to do or to be. For instance, it is important not only that somebody possesses a bike (commodity), even though he/she is

[1]Corresponding author: Cracow University of Economics, Department of Social Statistics, Rakowicka 27, 31-510 Cracow, Poland, pawel.ulman@uek.krakow.pl
[2]Cracow University of Economics, Department of Philosophy, Rakowicka 27, 31-510 Cracow,Poland,tomasz.kwarcinski@uek.krakow.pl

actually biking (functioning) and taking pleasure from that (satisfaction or happiness), but that taking into account their personal characteristics (e.g. health) and natural and social environments (e.g. income) he/she is able to exercise this activity (capability).

While Sen is reluctant to point out any specific list of human capability set, Nussbaum claims that creating such a list is not only possible, but also very important, and useful. Referring to Aristotelian tradition, she proposes a list of ten dimensions of central human capabilities, such as: (1) life, (2) bodily health, (3) bodily integrity, (4) senses, imagination and thought, (5) emotions, (6) practical reason, (7) affiliation, (8) other species, (9) play, (10) control over one's political and material environment. For instance, the life dimension consists of such capabilities as: being able to live to the end of a human life or not dying prematurely. Nussbaum claims that the list is objective and universal because the human capabilities are "central requirements for the life with dignity" (Nussbaum, 2003; Alkire, 2002). Because the list justification is based on moral values this is an openly normative approach.

We will attempt to operationalise Nussbaum's list, linking each dimensions of human capability to specific indicators(variables) from the European Quality of Life Survey (EQLS). On the basis of the Total, Fuzzy and Relative approach (TFR), which has been successfully applied to the research on poverty by Polish and foreign researchers, we will be able to calculate the index of objective quality of life. While there are some research projects applying TFR to measure personal well-being based on human capability sets (Tomer, 2002; Martinetti, 2000; Kapuria, 2016), the novel approach presented in this paper consists in making a comparison between the evaluation of subjective and objective quality of life, and indicating the divergence between these two approaches by using the mobility index.

We believe that this kind of research can be relevant for creating social policy. On the one hand, if people's self-evaluation state of happiness significantly exceeds their objective basis of life quality it can be a sign for the policymakers that some citizens adapt to poor living conditions. On the other hand, if people maintain a low level of happiness, despite living in good objective conditions, it could mean that they develop cost-inefficient tastes or preferences. In both cases a precise measure of these differences delivered bythe mobility index can be important.

2 Procedure

The source of statistical data was the EQLS gathered between 2003 and 2012.The data file contained 484 variables collected for 34 countries in three waves. The data for Poland and

Germany collected in the third wave (2012) were used in this paper. It was a microdata consisting of 2,262 observation units (individuals) for Poland and 3,055 for Germany. After checking the data for completeness and eliminating missing data, the number of observation units was reduced by almost a half. Thus, we decided to complement the missing data. In order to do this, we chose the indicators which had the least deficiencies and were relevant for the present research purposes. Finally, the data contained 2,226 observation units for Poland and 2,990 for Germany and accounted for 34 indicators (variables) without missing data. There is one additional variable from the EQLS which refers to subjective evaluation of happiness (S). This base variable concerns the following question: "Taking all things together on a scale of 1 to 10, how happy would you say you are?". The respondents had to determine their level of happiness on the ten-point scale.

The indicators were grouped into six dimensions of the central human capability: (1) life, (2) health, (3) education, (4) feelings/emotions, (5) social relationships, and (6) income (Table 1). The choice of the central human capability dimensions and the selection of an appropriate set of indicators was related to the availability of statistical data. Due to a shortage of data we had to single out only six dimensions instead of 10 as originally indicated by Nussbaum. We believe that our list of central human capability is objective in a sense that - despite personal tastes or preferences- it is something objectively good to live in a clean and comfortable environment (without excessive noise, crime, shortage of space etc.), being in good health, well educated, emotionally stable, having good relationships with other people, and not having to worry about income shortage.

In order to obtain one, aggregated evaluation of the respondents' quality of life (Q) based on the indicated capability list, we referred to TFR proposed by Zadeh (1965).His approach is typically applied to evaluate people's degree of poverty risk. Without going into the mathematical details of fuzzy sets theory, it is worth noting that the theory was successfully applied to form a membership function of poverty in both monetary and non-monetary approaches. Among those who used this strategy to poverty analysis were Cerioli and Zani (1990), Cheli (1995), Betti et al. (2005),and in Poland: Panek (2011), Ulman and Šoltés (2015).

The main assumption of this approach is to assess a person's degree of poverty risk by means of a function which takes values from a range of [0;1]. In comparison, the classic approach to identify the poor takes only two values: 1 (when someone is poor) or 0 (when someone is not poor) without paying attention to the degree of poverty risk.

Table 1. The set of indicators.

Dimensions of central human capability	Set of indicators
Life	Problems with neighbourhood – noise
	Problems with neighbourhood – air quality
	Problems with neighbourhood – quality of drinking water
	Problems with neighbourhood – crime, violence or vandalism
	Problems with neighbourhood – traffic congestion
	Problems with accommodation – shortage of space
	Problems with accommodation – lack of indoor flushing toilet
	Problems with accommodation – lack of bath or shower
	In my daily life, I seldom have time to do the things I really enjoy
	I feel that the value of what I do is not recognised by others
	Own hobbies, interests
	My daily life has been filled with things that interest me
Health	General self-evaluation of health
	Chronic (long-standing) physical or mental health problem, illness or disability
	Distance to doctors office/hospital/medical centre
	Waiting time to see doctor on day of appointment
Education	Satisfaction from education
	The highest level of education
Feelings/ Emotions	Some people look down on me because of my job situation or income
	I feel close to people in the area where I live
	I have felt lonely
	I have felt downhearted and depressed
	I am optimistic about the future
	Life has become so complicated today that I almost can't find my way
Social relationships	Face-to-face contact with friends or neighbours
	Contact with family members
	Other social contact (not family)

	Take part in sports or physical exercise
	Participate in social activities of a club, society, or an association
	Attended a meeting of a trade union, a political party or political action group
	Attended a protest or demonstration
	Signed a petition, including an e-mail or on-line petition
	Contacted a politician or public official
Income	OECD equivalised household income in PPP

The membership function of poverty is based on poverty symptoms, distinguishing a monetary part (based on incomes or expenses) and a non-monetary part (various factors which can point to poverty risk). Due to the fact that the poverty can be treated as low level of quality of life, we can apply this approach to research on levels and diversities of quality of life (referring to persons, families or households). Thus, we make a membership function of quality of life instead of poverty.

The following formula allows for transforming variables (dichotomous, ordinal, interval, ratio) into variables that take values from 0 to 1 (Panek, 2011):

$$e_{hj,i} = \frac{1-F(c_{hj,i})}{1-F(1)},\ h = 1,2,\ldots,m;\ j = 1,2,\ldots,k_h;\ i = 1,2,\ldots,n, \quad (1)$$

where:

$c_{hj,i}$ – is a rank of variant of the j-variable (indicator of quality of life) from *h*-dimension of quality of life for *i*-household (individual),

F(1) – is a value of the cumulative distribution function of ranks of the *j*-variable from *h*-dimension of quality of life for rank equal 1 (variant of *j*-variable indicating the highest level of quality of life). Following this, the membership value equal to 1 means a complete achievement with respect to a given indicator of the quality of life, whereas a value equal to 0denotes the total failure. Intermediate values between these two extremes describes a degree of quality of life.

Applying formula (1), the value of $e_{hj,i}$ was calculated for each indicator. Then, all these values were aggregated by using the arithmetic mean. Therefore, we received the assessment of quality of life degree within six dimensions for each indicator according to the formula:

$$\lambda_{h,i} = \frac{1}{k_h}\sum_{j=1}^{k_h} e_{hj,i}. \quad (2)$$

In the next step, the aggregation of evaluations of quality of life(Q) is performed by calculation of the arithmetic or weighted mean of $\lambda_{h,i}$:

$$\lambda_i = \left(\frac{1}{m}\sum_{h=1}^{m} w_h\, \lambda_{h,i}\right)^{\alpha}. \quad (3)$$

Because we wanted to compare our calculation to subjective evaluations of happiness (S), which was our base variable, we decided to calibrate the function (3) in such a way that the mean of the function (3) was equal to the mean of the base variable (S)[3]. To achieve this goal, the α parameter had to be adjusted and its estimated value, which ensured equality of the means was 0.431 for Poland and 0.452 for Germany. The same values of α parameter can be used for calibrating function (2)[4].

Two approaches were applied in respect of the α parameter. Firstly, it was assumed that the calibration parameter α was calculated on the base of overall evaluation of quality of life (the mean of six dimensions of quality of life) and then was used to calibrate the membership function for each quality of life dimension. Secondly, different calibration parameters α were estimated for each quality of life dimension and then applied to calculate the overall quality of life evaluation as a weighted average of the partial assessment. The coefficients of variation of the partial quality of life evaluations were used to determine the aforementioned weights.

To compare the base variable (S) to calculated values of quality of life (Q), the values of the function (2) or (3) were grouped into 10 categories. We assumed that interval of the function variability would be divided into 10 intervals of equal length. Finally, based on the particular interval of the value of the function (2), (3) the numbers from 1 to 10 were assigned to each observation unit (individuals).

In order to indicate the gap between S and Q we used the Bartholomew's mobility index (*B*), which in the present context can be defined as

$$B = \frac{1}{m-1}\sum_{i=1}^{m}\sum_{j=1}^{m} w_i p_{ij}\,|i-j|, \quad (4)$$

where:

m – is a number of categories,

w_i – is a fraction of people belonging to the*i*-th category of base variable (S)

p_{ij} – is a probability of each element mobility that is calculated by the following formula:

$$p_{ij} = \frac{n_{ij}}{\sum_{j=1}^{m} n_{ij}}, \text{ for} i,j = 1, 2, ...,m, \quad (5)$$

where:

[3] The variable (S) represents the level of happiness on a scale from 1 to 10, while function (3) takes values from 0 to 1 so the average of variable (S) was divided by 10 to compare with the average of values of function (3).

[4] It is also possible to calibrate the membership function for every dimension separately and then yield the total measure of quality of life using the weighted mean, where weights are based on coefficients of variation.

n_{ij} – is a number of people belonging to *i*-th category of base variable (S) and the *j*-th category of the objective quality of life evaluation (Q).

The value of mobility index is based on a matrix of transition probabilities between particular categories of a variables S and Q. The higher the value of this index, the greater is the gap between S and Q taken as a measure of life quality.

3 Results

The comparison of subjectively determined levels of happiness with calculated levels of quality of life is shown in Fig. 1 and 2. The structure of assessment both in subjective (S) and objective (Q) quality of life measure is very similar in Poland and Germany. The respondents more often assess their life satisfaction by indicating higher levels (above 5) – this applies both to the assessment of the happiness level and the aggregate measure of quality of life. It is worth noting that the structure of the quality of life level is more concentrated than it is for the assessment of happiness.

In general, comparing the life satisfaction in Poland and Germany, it can be observed that German respondents assess their happiness slightly better than Polish respondents. A similar conclusion can be drawn using objective measure of quality of life (Q). These outcomes are accompanied by a lack of values in the lower levels of the objective quality of life evaluation.

Looking at the aggregated levels of quality of life for Poland and Germany we can see that in each of those countries less people feel happy or satisfied (S) on the levels seven, eight, and nine than appears to be justified by objective evaluation of their life quality (Q). Also, in both countries significantly more people feel very happy (level ten) than they should be, taking into account objective measure of quality of life (Q).

While the respondents generally assess their happiness (S) at a lower level than it results from the calculated quality of life (Q), the evaluation of quality of life in respect to social relationships (fifth dimension) differs from this pattern both in Poland and Germany. The difference seems to be greater for Poland.

The gap between S and Q which is measured by Bartholomew's mobility index (*B*) is similar for Poland and Germany (Tables 2 and 3).

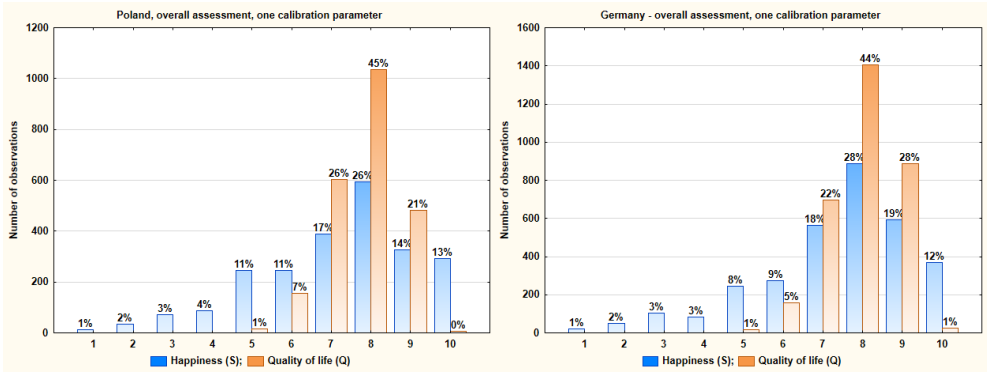


Fig. 1. The structure of assessment of the happiness levels and aggregated quality of life levels for Poland and Germany.

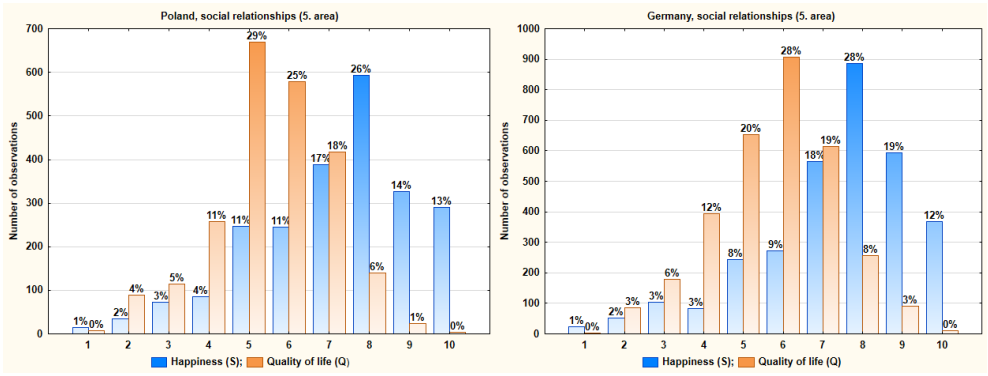


Fig. 2. The structure of assessment of the happiness levels and aggregated quality of life levels for Poland and Germany in respect to social relationships.

Table 2. Mobility index for Poland.

Index	Total	Life	Health	Education	Feelings	Relationships	Income
S>Q	0.0436	0.0199	0.0437	0,0854	0.0586	0.2299	0.1046
S<Q	0.1028	0.1967	0.1727	0.1001	0.0905	0.0339	0.1178
Total	0.1465	0.2167	0.2163	0.1855	0.1491	0.2637	0.2223

Table 3. Mobility index for Germany.

Index	Total	Life	Health	Education	Feelings	Relationships	Income
S>Q	0.0408	0.0130	0.0419	0.1148	0.0264	0.2241	0.1224
S<Q	0.1024	0.2003	0.1616	0.0882	0.1317	0.0377	0.1012
Total	0.1432	0.2133	0.2035	0.2030	0.1581	0.2618	0.2235

The highest level of differentiation is in the case of the fifth dimension (social relationships), while the smallest is for the fourth dimension (feelings/emotions). Looking at the data we can see also the direction of this diversity. There are two possibilities: the first, when the subjective assessment of happiness is greater than the quality of life evaluation (S>Q), and the second, when the happiness assessment is lower than the quality of life evaluation (S<Q).In general, the calculations show that for both countries the objective evaluation of quality of life is higher than subjective declarations of happiness (S<Q) (ca. 0.1), in comparison to the cases when objective calculation of quality of life turned out to be lower than subjective assessments of happiness (S>Q) (ca. 0.04). The mobility index also indicates that the highest move from the objective evaluation of quality of life to subjective assessment of happiness was observed in the relationships dimension (ca. 0.26). In this case, more people should be considered as having lower levels of the quality of life than they subjectively declared, which is illustrated graphically in Fig. 2.Besides the social relationships dimension, we observe a greater differentiation of assessments towards higher evaluation of objective quality of life than the self-evaluation of happiness for Polish respondents. It means that many people have are latively low level of happiness despite quite good objective conditions regarding central human capabilities. We can also observe that the German respondents tend to assess happiness higher than it is shown by the quality of life calculation in respect to education and income. It means that the differentiation between the assessments of happiness and the evaluations of quality of life generated by income is more directed towards a higher assessment of happiness. On the one hand, it seems that even when having low income people can achieve a high level of happiness. On the other hand, the reverse direction is relatively less important; it means that high level of quality of life resulting from high incomes is associated with a lower level of happiness.

Conclusion

Two different measures of quality of life were applied in our research: subjective self-evaluation of happiness (S) and objective assessment (Q) based on the list of central human capability. The comparison of these measures for Poland and Germany revealed very similar schema of the quality of live evaluation in both countries. Due to the calculation of the mobility index, we were also able to notice possibly similar problems regarding the social relationships dimension in both countries. In Poland, as well as in Germany, more people have lower level of quality of life in respect to the social dimension than their declared state of happiness. Despite the fact that Poland is still a less developed country in comparison to

Germany, it appears to be that the schemas of the quality of life in both countries are analogous.

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

Alkire, S. (2002). Dimensions of Human Development. *World Development,30*(2), 181-205.

Betti, G., Cheli, B., Lemmi, A. & Verma, V. (2006). On the Construction of Fuzzy Measures for the Analysis of Poverty and Social Exclusion. *Statistica & Applicazioni, 4*(1), 77-97.

Cerioli, A. & Zani, S. (1990). A Fuzzy Approach To The Measurement Of Poverty. *Income and Wealth Distribution, Inequality and Poverty Studies in Contemporary Economics*,272-284.

Cheli, B. (1995). Totally, Fuzzy and Relative Measures of Poverty in Dynamic Context. *Metron, 53*(3-4), 183-205.

Kapuria, P. (2016). A Human Well-Being Perspective to the Measurement of Quality of Life: Findings From the City of Delhi. *Applied Research in Quality of Life,11*(1), 125-145.

Martinetti, E. (2000). A Multidimensional Assessment of Well-being Based on Sen's Functioning Approach. *Rivista Internazionale di Scienze Sociali,108*(2), 207-239.

Nussbaum, M. (2003). Capabilities As Fundamental Entitlements: Sen And Social Justice.*Feminist Economics,9*(2-3), 33-59.

Panek, T. (2011). *Ubóstwo, wykluczenie społeczne i nierówności. Teoria i praktyka pomiaru*. Warszawa: Oficyna Wydawnicza SGH.

Parfit, D. (1984). Reasons and Persons. Oxford University Press.

Sen, A. (2005). *Commodities and Cappabilities*. Oxford University Press.

Tomer, J. F. (2002). Human Well-Being: A New Approach Based on Overall and Ordinary Functionings. *Review of Social Economy,60*(1), 23-45.

Ulman, P. & Šoltés, E. (2015). The Monetary and Non-Monetary Aspects of Poverty in Poland and Slovakia. *Entrepreneurial Business and Economics Review,3*(1), 61-73.

Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control, 8*, 338-353.

Trade openness and financial development: Granger causality analysis for new EU member states

Marta Wajda-Lichy[1], Paweł Kawa[2], Kamil Fijorek[3], Sabina Denkowska[4]

Abstract

Traditional approach to trade and finance linkage indicates their complementarity. However, causalities between these two variables are ambiguous. The purpose of this paper is to investigate Granger causality between trade openness and financial development in 11 new EU member states. The annual data for the period 1995-2016 comes from WDI database. Openness is measured as the ratio of country's total trade (exports plus imports) to the country's GDP, or of exports to GDP, and of imports to GDP, separately. Financial development is proxied by domestic credit to private sector and domestic credit provided by financial sector (both measures as a percentage of GDP). The results show statistically significant causalities running from trade openness to financial development in half of the analysed countries with the estimated coefficients being negative. In the majority of countries, the linkages from finance to trade were statistically insignificant regardless the proxies considered. The empirical findings are related to theoretical arguments, both traditional and new ones, formulated mainly in the post-crisis era. We find the obtained results consistent with the empirical work of Bordo and Rousseau (2012), who found that finance and trade reinforced each other before 1930, but these effects do not persist after the Second World War. While Bordo and Rousseau analysed the OECD countries, our paper contributes to the finance-trade nexus literature by focusing on emerging markets.

Keywords: *trade openness, financial development, finance-trade nexus, Granger causality*
JEL Classification: F10, O16, C33, C53
DOI: 10.14659/SEMF.2018.01.56

1 Introduction

Since the beginning of the 1990s Eastern European countries started to transform and liberalize their economies. Important part of free-market-oriented reforms covered financial sector and foreign trade. Under socialist system most of the countries performed low level of financial development, as well as international openness which was largely limited to other socialist economies. In the aftermath of transition reforms, both variables increased significantly. For example, in the Baltic economies financialization measured as domestic

[1] Corresponding author: Cracow University of Economics, Department of Macroeconomics, ul. Rakowicka 27, 31-510 Kraków, e-mail: wajdam@uek.krakow.pl.
[2] Cracow University of Economics, Department of Macroeconomics, ul. Rakowicka 27, 31-510 Kraków, kawap@uek.krakow.pl.
[3] Cracow University of Economics, Department of Statistics, ul. Rakowicka 27, 31-510 Kraków, kamil.fijorek@uek.krakow.pl.
[4] Cracow University of Economics, Department of Statistics, ul. Rakowicka 27, 31-510 Kraków, denkowss@uek.krakow.pl.

credit to GDP increased ninefold in Latvia, nearly fivefold in Estonia, and threefold in Lithuania. Levels of financialization also tripled in the Polish and Romanian economies. Twenty years after launching transition reforms, trade openness measured as exports and imports to GDP doubled in the case of Poland and Hungary, whereas in the rest of the examined countries the indexes increased at least by half.

Simple approach to the above-mentioned changes suggests positive correlation of both variables, i.e. trade openness and financial development. This would be consistent with traditional theoretical approach which predicts that trade openness and financial development are complementary. However, as causalities between trade openness and finance are ambiguous, it is worth to examine nexus between these two categories. Another reason for undertaking our research stemmed from the prevailing empirical evidence on the importance of exports, imports and finance for economic development and growth (Frankel and Romer, 1999; King and Levine, 1993; Beck et al., 2000). Understanding the determinants of trade flows or financial development is important, especially for emerging market economies. The trade-finance nexus also raised our concerns in the context of the recent hypothesis that too much finance or excessive openness no longer contribute to growth (Arcand et al., 2015; Rodrik et al., 2004). As theoretical literature suggests that international trade and finance have both direct as well as indirect impact on economic growth, we test for Granger causality between the two determinants to examine channels of these interactions in 11 new EU member states (NMSs).

The paper contains 6 sections. We start with the review of the relevant literature and theoretical approach to the trade-finance nexus. Next we describe research methodology. Sections 4 and 5 include our empirical results and their interpretation. Conclusions are presented in Section 6.

2 Review of literature

The causal linkages between trade openness and finance are relatively rarely discussed in the literature. The prevailing empirical studies in this field focus mainly on testing trade-growth and finance-growth nexus. Although the global financial crisis and permanent external imbalances characteristic for two decades preceding the outset of the crisis triggered off an ongoing debate on finance-trade causal relationship, the results are still inconclusive (Cecchetti and Kharroubi, 2015; Arcand et al., 2015; Rousseau and Wachtel, 2011).

There are two strands of the studies testing the trade-finance nexus. The first one refers to a demand-following hypothesis, which contends that trade creates demand for financial

services. The explanation that trade openness drives financial development is related to investments in tradeable goods and services, as well as direct trade payments and hedging instruments. Rajan and Zingales (2003) delivered interesting argumentation related to political economy which supports for positive correlation between openness and financial development. Baltagi et al. (2009) confirm that trade openness, as well as economic institutions, can explain financial development and a large part of its variations over the period 1980-2003.

The second strand of studies tests supply-leading hypothesis, which suggests that well developed financial markets may constitute a source of comparative advantages for foreign traders. For instance, Beck (2002) explored that mature financial markets not only induce higher volume of trade but also influence its structure. His results from 30-year panel of 65 countries show that sectors with high scale economies profit more from a higher level of financial development than other sectors. There is also a feedback hypothesis which states that international trade and financial development interact with each other. Bordo and Rousseau (2012) used historical data from 1880 for 17 high-income economies to examine the trade-finance nexus. They explored that bidirectional causalities occurred before 1930, but after 1945 these linkages do not persist.

Having reviewed the relevant literature we conclude that the empirical studies provide with ambiguous results on predominance of any of the above-mentioned hypotheses. Some economists confirm that financially developed countries trade more (Beck, 2002), whereas the others emphasize weak or conditioned causality from finance to trade (Chang et al., 2009). There is also evidence for links from international openness to finance, which is conditioned with economic or political institutions (Rajan and Zingales, 2003; Baltagi et al., 2009).

3 Methodology

In order to determine the direction of causality between the variables of interest, we use the panel data framework due to the well-known fact that panel data methods increase the power of statistical tests. In examining causal linkages within the panel framework, two key issues have to be addressed. The first one is to control for cross-sectional dependence across the members of panel, because a shock affecting one country may also affect other countries through the high degree of globalization as well as of international trade and financial integration. Pesaran (2006) shows the importance of testing for cross-sectional dependence in a panel data study. The second issue is to consider whether the data can be pooled across

countries and whether panel estimates account for country-specific heterogeneity (Pesaran and Yamagata, 2008).

Table 1. Results of cross-section dependence and homogeneity tests.

Model[5]	**LM**	$\mathbf{LM_{adj}}$	**CD**	$\widetilde{\Delta}$	$\widetilde{\Delta}_{\mathbf{adj}}$
FDPRIV =	94.95 ***	6.62 ***	4.93 ***	7.90 ***	9.27 ***
L.FDPRIV+L.TO+L.GDPpc	p=0.001	p=0.000	p=0.000	p=0.000	p=0.000
FDPRIV =	72.91 *	1.95 *	3.04 ***	10.59 ***	12.42 ***
L.FDPRIV+L.EX+L.GDPpc	p=0.053	p=0.052	p=0.002	p=0.000	p=0.000
FDPRIV =	113.90 ***	10.67 ***	6.01 ***	4.87 ***	5.71 ***
L.FDPRIV+L.IM+L.GDPpc	p=0.000	p=0.000	p=0.000	p=0.000	p=0.000
FD =	80.02 **	3.51 ***	1.96 *	13.91 ***	16.31 ***
L.FD+L.TO+L.GDPpc	p=0.0154	p=0.000	p=0.051	p=0.000	p=0.000
FD =	72.34 *	1.86 *	0.95	17.91 ***	21.00 ***
L.FD+L.EX+L.GDPpc	p=0.058	p=0.063	p=0.343	p=0.000	p=0.000
FD =	101.40 ***	8.10 ***	3.44 ***	9.05 ***	10.61 ***
L.FD+L.IM+L.GDPpc	p=0.000	p=0.000	p=0.001	p=0.000	p=0.000
TO =	241.00 ***	37.50 ***	13.77 ***	2.40 **	2.82 ***
L.TO+L.FD+L.GDPpc	p=0.000	p=0.000	p=0.000	p=0.016	p=0.005
TO =	271.70 ***	43.93 ***	14.77 ***	3.14 ***	3.68 ***
L.TO+L.FDPRIV+L.GDPpc	p=0.000	p=0.000	p=0.000	p=0.002	p=0.000
IM =	308.30 ***	52.01 ***	15.68 ***	3.09 ***	3.63 ***
L.IM+L.FD+L.GDPpc	p=0.000	p=0.000	p=0.000	p=0.002	p=0.000
IM =	327.60 ***	56.01 ***	16.74 ***	2.64 ***	3.10 ***
L.IM+L.FDPRIV+L.GDPpc	p=0.000	p=0.000	p=0.000	p=0.008	p=0.002
EX =	163.30 ***	20.95 ***	10.08 ***	1.82 *	2.14 **
L.EX+L.FD+L.GDPpc	p=0.000	p=0.000	p=0.000	p=0.068	P=0.033
EX =	173.80 ***	23.16 ***	10.68 ***	3.52 ***	4.13 ***
L.EX+L.FDPRIV+L.GDPpc	p=0.000	p=0.000	p=0.000	p=0.000	p=0.000

Notation: $\boldsymbol{LM}$ – Breusch and Pagan test (1980), $\boldsymbol{LM_{adj}}$ and $\boldsymbol{CD=LM\ CD*}$ are modified versions (Pesaran et al., 2008) of the $\boldsymbol{LM}$ test, $\widetilde{\Delta}$ and $\widetilde{\Delta}_{adj}$ denote Pesaran and Yamagata

[5] Models are given in simplified notation, where L denotes lagged operator.

(2008) tests for slope homogeneity. *** Denotes statistical significance at 1%, **denotes statistical significance at 5%, * denotes statistical significance at 10%.

We start by investigating whether or not there is cross-sectional dependence and heterogeneity across the countries under study. The results reported in Table 1 indicate that the null hypotheses of no cross-sectional dependence for almost all considered models are rejected at the 1% level of significance (only for two of them cross-sectional dependence tests rejected the null hypothesis at 10%). This finding implies that a shock that occurs in one NMS may be transmitted to others. Table 1 also shows that the results from the slope homogeneity tests reject the null hypothesis of slope homogeneity, therefore supporting country-specific heterogeneity.

The variables in models specifications mean: EX – exports as a percentage of GDP; IM – imports as a percentage of GDP; TO – trade openness defined as the ratio of country's total trade (exports plus imports) to the country's GDP; FD – domestic credit provided by the financial sector (it includes all credit to various sectors on a gross basis, with the exception of credit to the central government, which is net. The financial sector includes monetary authorities and deposit money banks, as well as other financial corporations e.g. finance and leasing companies, money lenders, insurance corporations, pension funds, and foreign exchange companies); FDPRIV – domestic credit to private sector (it refers to financial resources provided to the private sector by financial corporations, such as through loans, purchases of nonequity securities, and trade credits and other accounts receivable, that establish a claim for repayment); GDP_{pc} – GDP per capita.

In order to test the panel Granger causality we use the bootstrap panel causality method proposed by Kónya (2006), which accounts for both cross-sectional dependence and slope heterogeneity. This method is based on seemingly unrelated regression (SUR). This approach is also robust to the unit root and cointegration properties of the variables. Therefore the testing procedure does not require any pretesting for unit root and cointegration, and the variables are therefore used in their levels. The first step of the bootstrap panel causality approach requires estimating the equation system specified below:

$$\begin{aligned} X_{1t} &= \alpha_{11} + \sum_{l=1}^{p_1} \beta_{11l} X_{1t-l} + \sum_{l=1}^{p_2} \delta_{11l} Y_{1t-l} + \sum_{l=1}^{p_3} \varphi_{11l} Z_{1t-l} + \varepsilon_{11t} \\ &\vdots \\ X_{Nt} &= \alpha_{1N} + \sum_{l=1}^{p_1} \beta_{1Nl} X_{Nt-l} + \sum_{l=1}^{p_2} \delta_{1Nl} Y_{Nt-l} + \sum_{l=1}^{p_3} \varphi_{1Nl} Z_{Nt-l} + \varepsilon_{1Nt} \end{aligned} \quad (1)$$

$$Y_{1t} = \alpha_{21} + \sum_{l=1}^{p_4} \beta_{21l} X_{1t-l} + \sum_{l=1}^{p_5} \delta_{21l} Y_{1t-l} + \sum_{l=1}^{p_6} \varphi_{21l} Z_{1t-l} + \varepsilon_{21t}$$
$$\vdots \qquad (2)$$
$$Y_{Nt} = \alpha_{2N} + \sum_{l=1}^{p_4} \beta_{2Nl} X_{Nt-l} + \sum_{l=1}^{p_5} \delta_{2Nl} Y_{Nt-l} + \sum_{l=1}^{p_6} \varphi_{2Nl} Z_{Nt-l} + \varepsilon_{2Nt}$$

where[6] X is the trade openness (measured as a sum of exports and imports to GDP, or exports to GDP, or imports to GDP), Y is the financial development (which is proxied by domestic credit to private sector and domestic credit provided by financial sector (both measures as a percentage of GDP)), Z is GDP *per capita*, N is the number of countries of panel (i = 1,…, N), t is the time period (t = 1,…, T), and l is the lag length.

In testing for Granger causality, alternative causal relations for a country are likely to be found. For example, there is one-way Granger causality from Y to X if not all $\delta_{1,i}$ are zero, but all $\beta_{2,i}$ are zero; there is one-way Granger causality from X to Y if all $\delta_{1,i}$ are zero, but not all $\beta_{2,i}$ are zero; there is two-way Granger causality between Y and X if neither $\delta_{1,i}$ nor $\beta_{2,i}$ is zero; there is no Granger causality between Y and X if all $\delta_{1,i}$ and $\beta_{2,i}$ are zero. To determine the direction of causality, the Wald statistics for Granger causality are compared with the country-specific critical values that are obtained from the bootstrap sampling procedure.

4 Empirical results

Demand-following hypothesis implies positive signs of the regression coefficients in the causality tests. It means that trade and finance are expected to be complementary. Our results show statistically significant causalities (with the bootstrap probability less than 0.05) from exports to finance development in the case of six countries, when the latter was proxied by domestic credit provided by financial sector (FD), and in four countries, when it was measured as domestic credit to private sector (FDPRIV) (Table 2). In both of these cases the estimated coefficients revealed to be negative, which does not support the demand-following hypothesis. Positive causal nexus between trade openness and finance was confirmed only in Bulgaria when openness was measured as imports to GDP and finance was proxied by FD.

Causality from finance development to trade openness was found to be weak (Table 2). The supply-leading hypothesis was proved as valid only in Lithuania and Croatia, when finance was proxied with FDPRIV, and Latvia when we tested FD. Robust results were also

[6] p_i (for $i = 1, \ldots 6$) in formulas (1) and (2) denotes number of lags.

explored for Poland, regardless of measures used for openness or finance development. In this case, however, negative regression coefficients indicated the reverse of supply-leading hypothesis. In the majority of NMSs causalities from finance to trade openness were statistically insignificant. This means that supply-leading hypothesis was not confirmed for most of the countries under consideration.

Table 2. Panel Granger causality test results FDPRIV←EX and EX←FDPRIV.

COUNTRY	FDPRIV ← EX				EX ← FDPRIV			
	Coef. Wald	Stat	boot 95%	*p* boot	Coef. Wald	Stat	boot 95%	*p* boot
Bulgaria	-0.48	2.79	11.49	0.3250	0.09	1.46	11.30	0.4850
Croatia	-0.15	0.98	13.80	0.5990	0.22	29.33	13.38	0.0064
Czech Republic	0.06	0.15	9.73	0.8130	-0.13	7.49	12.99	0.1440
Estonia	-0.32	4.84	10.92	0.1680	0.20	7.86	11.40	0.0970
Hungary	-0.36	17.80	12.99	0.0264	0.10	1.62	23.16	0.5900
Latvia	-1.05	27.81	16.50	0.0144	0.10	8.64	12.93	0.1003
Lithuania	-0.51	20.16	16.10	0.0305	0.30	12.53	11.62	0.0430
Poland	1.05	4.19	10.78	0.1950	-0.28	61.70	15.59	0.0012
Romania	-0.09	0.59	12.73	0.6490	0.12	0.60	14.13	0.6710
Slovak Republic	0.10	0.28	10.22	0.7350	-0.24	7.89	13.93	0.1220
Slovenia	-0.61	38.13	11.99	0.0018	0.03	0.83	15.60	0.6520

5 Interpretation of empirical results

Several explanations can be given to explain why the results we obtained show in the majority of the examined countries no statistical significance of the finance-trade nexus or a significance with (in most cases) negative coefficients. The explanations can be derived from theoretical point of view, but may also follow from empirical shortcomings.

Financial development reflects a notion that is probably too complex to determine its linkages with trade openness when using aggregated categories (like whole domestic credit without dividing it into consumer and investment credit). The composition of credit may have a crucial importance for explaining the finance-trade links. Credit to firms removes financing constraints that exporters could tackle with, thus leading to greater investment and potentially greater export. On the contrary, credit to households can increase domestic consumption (e.g. housing) and even weaken trade related sectors. As the NMSs tried to aspire to consumption

patterns characteristic for Western European economies, the increased credit availability could have been channelled mainly to reach the consumption convergence, while trade sectors remained relatively unaffected by an easy access to credit.

Financial system can generate positive, neutral or negative trade links with time-varying intensity depending mainly on the type of financial market and the variable specifications that are used as proxies. With an increase in GDP per capita, faster growth of domestic private bond markets and stock markets relative to the banking system can be observed, as well as proliferation of mutual funds and pension funds. Therefore, relying only on bank credit measures may not cover all channels of finance-trade links. While in recent 20 years the size of such defined financial sector remained largely stable or grew gradually, the share of shadow banking system (assets of nonbanks) increased significantly. The diversity of financial markets and instruments across countries implies the need to reassess the adequacy of financial development measures.

A large financial sector or its incommensurable high dynamics may lead to increasing volatility and to a suboptimal allocation of resources. The social returns of the financial sector may be lower than its private returns, and a large financial sector may "steal" talents from trade-oriented sectors of the economy and therefore be inefficient from trade point of view. Cecchetti and Kharroubi (2015) emphasize the "crowding out" of human capital away from the real sector to the financial sector when a rapid financialization occurs. When taking into account the data from sector level, it is apparent that before the 2008 financial crisis, resources were shifted from more productive sectors towards financial sector.

A new strand in literature and empirical analysis suggests also that there could be a threshold (reflecting the "too much finance" concept developed by Arcand et al., 2015) above which financial development brings on negative returns. In fact, the countries we analysed have financial markets depth well below the levels reached in advanced economies and significantly below a threshold above which the financial sector is supposed to harm trade (Arcand et al., 2015) suggest the threshold of about 80-100% of credit to GDP, and Gächter and Gkrintzalis (2017) estimate the vertex of the parabola at 115%-120% of credit to GDP ratio as critical for exports and import, respectively). However, the threshold is not universal for all countries at all times. Assuming that the relationship between finance and trade is actually characterized by an inverse U-shape, it should be remembered that the location of the parabola may depend on specific country's characteristics like income level, institutions, and regulatory and supervisory quality. It means that countries with weak institutions, below average quality of regulations and not satisfactory supervision arrangements may have this

threshold below the estimates given. It might be the case of NMSs.

The results we obtained are basically coherent with recent literature suggesting that after 1990 the conventional positive linkages between finance and real side of economy have weakened. The increased frequency of banking crises[7] has also contributed to "disappearing" empirical links which included finance as one of variables (Rousseau and Wachtel, 2011).

Conclusions

Argumentation for negative causality from trade openness to financial development is not evident. We interpret our findings taking into consideration specific features of NMSs, as well as changes in global economy. First, the examined countries should be considered not only as new members of the EU, but also as relatively new market economies. Both statuses imply important structural changes and adjustments in the field of economic openness and financial sector. Second, as transformation launched dynamic effects, the final results seem not to be completed, especially when euro adoption consequences are considered. Third, NMSs are still catching-up economies as their GDP per capita is lower than the EU average. Moreover, they are classified as emerging markets. These features are important for interpretation of reversed causalities explored between trade openness and financial development.

Acknowledgement

The authors acknowledge support from research funds granted to the Faculty of Management and Faculty of Economics and International Relations at Cracow University of Economics.

References

Arcand, J. A., Berkes, E. & Panizza, U. (2015). Too Much Finance? *Journal of Economic Growth*, *20*(2), 105-148.

Baltagi, B., Demetriades, P. & Law, S. H. (2009). Financial development and openness: Evidence from panel data. *Journal of Development Economics*, *89*(2), 285-296.

Beck, T. (2002). Financial development and international trade: Is there a link?. *Journal of International Economic*s, *57*(1), 107-131.

[7] In countries we analysed the banking crisis started: Bulgaria in 1996, Croatia in 1998, the Czech Republic in 1996, Estonia in 1992, Hungary in 1991 and 2008, Latvia in 1995 and 2008, Lithuania in 1995, Poland in 1992, Romania in 1990, the Slovak Republic in 1998, and Slovenia in 1992 and 2008.

Beck, T., Levine, R. & Loayza, N. (2000). Finance and the sources of growth. *Journal of Financial Economics*, *58*(1-2), 261-300.

Bordo, M. & Rousseau, P. (2012). Historical evidence on the finance-trade growth nexus. *Journal of Banking & Finance*, *36*(4), 1236-1243.

Breusch, T. & Pagan, A. (1980). The Lagrange Multiplier Test and its Application to Model Specifications in Econometrics. *The Review of Economic Studies*, *47*(1), 239-253.

Cecchetti, S. & Kharroubi, E. (2015). Why Does Financial Sector Growth Crowd Out Real Economic Growth?. *BIS Working Paper 490*, Bank for International Settlements, Basel.

Chang, R., Kaltani, L. & Loayza, N. V. (2009). Openness can be good for growth: The role of policy complementarities. *Journal of Development Economics*, *90*, 33-49.

Frankel, J. & Romer, D. (1999). Does Trade cause Growth*?*. *American Economic Review*, *89*(3), 379-399.

Gächter, M. & Gkrintzalis, I. (2017). The finance-trade nexus revisited: Is the global trade slowdown also a financial story? *Economics Letters*, *158*, 21-25.

King, R. G. & Levine, R. (1993). Finance and growth: Schumpeter might be right. *The Quarterly Journal of Economics*, *108*(3), 717-737.

Kónya, L. (2006). Exports and growth: Granger causality analysis on OECD countries with a panel data approach. *Economic Modelling*, *23*(6), 978-992.

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panel with a multifactor error structure. *Econometrica*, *74*(4), 967-1012.

Pesaran, M. H. & Yamagata, T. (2008). Testing slope homogeneity in large panels. *Journal of Econometrics*, *142*(1), 50-93.

Pesaran, M. H., Ullah, A. & Yamagata, T. (2008). A bias-adjusted LM test of error cross-section independence. *The Econometrics Journal*, *11*(1), 105-127.

Rajan, R. & Zingales, L. (2003). The great reversals: the politics of financial development in the twentieth century. *Journal of Financial Economics*, *69*, 5-50.

Rodrik, D., Subramanian, A. & Trebbi, F. (2004). Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development. *Journal of Economic Growth*, *9*(2), 131-165.

Rousseau, P. & Wachtel, P. (2011). What is Happening to the Impact of Financial Deepening on Economic Growth? *Economic Inquiry*, *49*(1), 276-88.

Does debt improve housing conditions? Evidence from Polish households

Agnieszka Wałęga[1], Grzegorz Wałęga[2]

Abstract

The present paper is aimed at validating the role of debt in improving housing conditions in Poland. It is common knowledge that consumer debt, especially mortgage, is a major factor of quantitative and qualitative changes of living conditions of households. Housing conditions belong to most commonly investigated aggregative social data in many respects. Research on the role of indebtedness in changing housing conditions is a rather neglected area in this field. Recent significant increase in household debt is a good opportunity to investigate its relationship with household conditions. We use data from the household budget survey conducted by the Polish Central Statistical Office during the period 2005–2015. Using microlevel data, we assess various symptoms (diagnostic features) related to both flat size and quality. We have adopted the mathematical theory of fuzzy sets to construct a multidimensional index to examine housing conditions. Similar to the analysis of the risk of poverty, the hazard indicator for poor housing has been calculated for each household. On average, we have found improvement of housing conditions in Poland after accession to the European Union. As expected, our research demonstrates that indebted households are characterised by better quality of housing.

Keywords: *housing conditions, household, indebtedness*
JEL Classification: D10, D14, D31, R29
DOI: 10.14659/SEMF.2018.01.57

1 Introduction

Every human being aims first at the satisfaction of the need of protection against external factors and at ensuring their safety. The pursuance of such needs entails the necessity to guarantee shelter. This is the reason why flats have always been, in a more primitive or less primitive form, important resources for humans in the process of need satisfaction (consumption). Although civilisation progress has caused that households look at contemporary properties from a completely different perspective than they used to in the past (e.g. as a form of investment), the function of satisfying housing needs still remains most important.

The difficulty of satisfying housing needs results from considerable capital intensity and long duration of the manufacturing cycle. Poland has faced the problem of incomplete satisfaction of housing needs for a long time. Even though the initiation of economic

[1] Corresponding author: Cracow University of Economics, Department of Statistics, Rakowicka 27, 31–510 Kraków, agnieszka.walega@uek.krakow.pl.
[2] Cracow University of Economics, Department of Microeconomics, Rakowicka 27, 31–510 Kraków, grzegorz.walega@uek.krakow.pl.

transformation after 1989 allowed for establishing the foundations for a modern housing market, it brought no breakthrough or solution to the problem. A real boom in the housing market was triggered by Poland's accession to the European Union. The process was undoubtedly accelerated by increased availability of mortgages from financial institutions (Wałęga, 2015b). The purpose of the article is to assess changes in housing conditions of households in Poland during the period 2005–2015, in particular taking into account the influence of debt on the housing market. Research involved an attempt to verify a hypothesis that a boom in mortgages following the accession to the European Union had a positive impact on the satisfaction of housing needs of Polish people, both in quantitative and qualitative terms. The analyses used secondary data of the National Bank of Poland and unidentified individual data from the examination of household budgets in Poland carried out by the Central Statistical Office of Poland between 2005 and 2015.

2 Impact of the mortgage market on the housing market

The demand in the housing market is complex and depends on factors fundamental for this market (housing needs as a function of income and demographic changes as well as availability of mortgages), policy of the state (direct subsidies and regulations affecting risk level in the sector as well as costs and income of business entities) and speculations related to further expected increase in prices (Eickmeier and Hofmann, 2013).

The accession of Poland to the European Union has given an impulse to an increase in debts of households. A similar situation has occurred in almost all countries of average and low level of development (Greece, Ireland and Portugal), which joined the European Community. These countries witnessed a boom in mortgages, with an annual mortgage growth rate for the household sector amounting to a more than a dozen or even several dozen percent per annum during the first period after accession (Brzoza-Brzezina, 2005). The belief in long-term stability of economic development and structural reforms in the scope of regulations and institutions are incentives for incurring debts and investing in properties after the accession of Poland to the European Union (Backé, 2009). Research suggests that European integration is also associated with more predictable monetary and FX policy.

A possibility to use lower interest rates was another factor that encouraged households to increase debt to purchase properties after 2005. Due to a relatively free convergence of real percentage rates in the mortgage market in recently joined countries, there emerged pressure on the replacement of mortgages in a domestic currency with mortgages denominated in foreign currencies of lower exchange rates, especially during the early period. The pressure

rises with growing disparity in interest rates and greater share of foreign capital in the banking sector (due to easier access to cheaper refinancing sources in international financial markets). The accession to the EU structures has also strengthened consumer expectations as to the decrease in real interest rates.

Another undisputed factor which had a positive influence on both the mortgage market and the housing market in Poland after Poland joined the EU was an increase in real income of households and improved moods of consumers (Wałęga, 2015a). Research by Leszczyński and Olszewski (2017) also confirms the considerable influence of salaries, unemployment rate and interest rates on the Polish real estate market. An increase in income has not only boosted purchasing power of consumers but, most importantly, had a positive influence on credit worthiness and expanded the possibilities of household indebtedness.

Furthermore, social and cultural transformations, and especially changes in social perception of mortgages, also need to be emphasised. The taking of credit or loan facilities is no longer perceived in a negative way. Credit and loan facilities are treated as an instrument of inter-period consumption management. In this regard, Polish households become more and more similar to consumption patterns of highly developed countries where consumption, especially the purchase of flats, is financed with mortgages to a much greater extent than in Poland. The perception of real estate by households also changes (Widłak and Łaszek, 2016).

The real estate market is characterised by low flexibility in supply accompanied by simultaneous relatively long period of adaptation to market shocks (resulting to a large extent from technological, legal and institutional conditions). This entails consequences in case of sudden increased demand. During the first decade of the 21st century, the situation in the housing market in Poland was also affected by additional demand from people born in the periods of baby boom who wanted to satisfy their housing needs by having their own flat. Another group that contributed to the market situation in the early 21st century were foreign investors who pursued their purchase transactions on a mass scale after 2004, encouraged by relatively low prices of real estates in Poland. The first symptoms of a housing boom gave an additional impulse to increase mortgage debt. In fact, it was a self-perpetuating process which was not stopped until global economic slowdown after 2009.

When analysing the causes of boom in the market of loans to households and in the market of residential properties, one should take into account the state policy referring to the real estate market. During the period 2007–2013, it was a factor of cyclical phenomena in the real estate market in the scope of financial support for families (Wałęga, 2013).

Financial phenomena (increased debt of households on account of mortgages) have an influence on the real sphere (housing market) and on housing conditions of the population. The financial flow directed by households into the housing market includes accumulated funds (savings), and credit and loan facilities. Quantitative and qualitative improvement of housing resources should be the outcome of these investments.

3 Data and research method

A fuzzy approach has been used to evaluate housing conditions of indebted households and households having no debt; it has been relatively developed for the purposes of poverty analysis. The approach is based on the concept of fuzzy sets proposed by Zadeh in 1965 and developed for the purpose of analysing impoverishment of society by Ceriolii Zani (1990), Cheli and Lemmi (1995) and Betti et al. (2015) as well as Verma et al. (2017).

With reference to the evaluation of the housing conditions of individual households and entire social groups, methods for the evaluation of poverty in the non-monetary dimension have been adopted (Ulman, 2016). This approach should specify symptoms of the risk of a poor housing conditions. These should be grouped into areas of assessment and finally aggregated into a single function of a risk of a poor housing conditions. Even though the proposed method has been originally used to assess poverty[3] in the scope of housing conditions, it may be successfully used as a synthetic measure to enable the assessment of a housing conditions of households. The possibility of taking the poverty ratio into account in the quantitative and qualitative estimations is a significant advantage of this approach.

Indicators of a risk of poor housing conditions have been determined in a similar manner, as in case of poverty measurement. The starting point is the assessment of the degree of risk of poor housing conditions of *i*-th household as part of *j*-the symptom belonging to *h*-the area (Panek, 2011):

$$e_{hj,i} = \frac{1 - F(c_{hj,i})}{1 - F(1)},\ h = 1, 2, ..., m;\ j = 1, 2, ..., k_h;\ i = 1, 2, ..., n \qquad (1)$$

where:

$F(c_{hj,i})$ – value of the distribution function of ranks of *j*-th variable (symptom of poor housing conditions) from *h*-th area of these conditions of *i*-th household,

[3] It can be also find application of the fuzzy set approach to eg. quality of life measurement (Betti et al., 2016).

$F(1)$ – value of the distribution function of ranks of *j*-th variable from *h*-th area of poor housing conditions for rank equal to 1 (variant of *j*-th variable indicating higher ratio of a risk of poor housing conditions).

In the subsequent step, the assessment of the absence of risk of poor housing conditions is aggregated with the use of the following formula:

$$e_{h,i} = \frac{\sum_{j=1}^{k_h} w_{hj}(1 - e_{hj,i})}{\sum_{j=1}^{k_h} w_{hj}}. \tag{2}$$

Values w_{hj} (weight of *j*-th symptom of poor housing conditions from *h*-th area) are obtained using the following formula:

$$w_{hj} = \ln \frac{1}{\lambda(z_{hj})}, \tag{3}$$

where:

$$\lambda(z_{hj}) = \frac{1}{n} \sum_{i=1}^{n} \lambda_i(z_{hj}), \tag{4}$$

whereas $\lambda_i(z_{hj}) = (1 - F(c_{hj,i}))$.

The analysis of the housing conditions of Polish households was conducted based on individual data from the examination of household budgets carried out by the Central Statistical Office of Poland during the years 2005–2015. Full data sets cover from 34,767 to 37,427 observations of households for each year.

Individual data on household budgets enable cross-sectional analysis of characteristics describing the housing conditions of families considering its selected characteristics. Due to the need to have identical features in the analysed years, the following variables are proposed for the analysis: surface of a flat per person, number of persons per room, period during which the building was constructed, equipment of the flat: bathroom, flushed toilet, water supply system, hot water, intercom, use of garage, use of another house. Housing conditions have been analysed depending on the fact whether the household has a credit or loan facility and depending on the education level of the head of household and place of residence[4].

[4] For the purposes of research, it has been assumed that a household in debt is a household which meets at least one of the following conditions: it has taken a loan or mortgage, a bank loan, has a credit card or another bank loan, a loan or credit facility in a different institution, a money loan from a private individual or has repaid a principal instalment and/or interest on the aforementioned forms of debt during the analysed period.

4 Empirical results

Almost all conditions for increased demand in the residential market and resulting housing boom appeared after Poland joined the European Union. A considerable increase in long-term debt on account of housing loans and large numbers of purchasers (including foreign ones) had a double impact on the market situation. Firstly, thanks to a broader access to external funding, households could pursue their housing needs. This translated into stimulation of the construction sector. During the first three years after Poland joined the European Union, the speed of building new flats increased by even up to 40% (year by year).

Fig. 1 presents the situation in the domestic real estate market in terms of the dynamics of the number of flats under construction and the number of flats commissioned for use. A transmission of impulses from the mortgage market to the residential market is noticeable even though in this case the mechanism has a delay of 2–3 years.

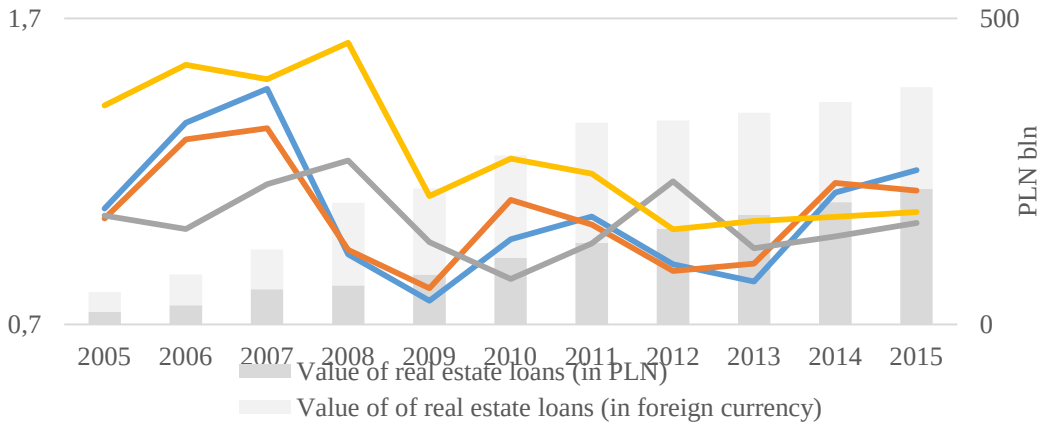


Fig. 1. Dynamics of flats under construction and dynamics of flats commissioned for use (previous year = 1) against the value of household debt on account of real estate loans in total (in PLN billion) during the years 2005–2015.

During the period of 2005–2015, housing conditions have generally improved (Table 1). On average, Poles (both those in debt and those who have no debts) live in larger and larger flats (both in absolute terms and in conversion into the surface of flat per capita). During the analysed years, the median of the size of flat rose from 60 sq m to 64 sq m in case of households which do not have a debt and from 59 sq m to 64 sq m for households which have debt. The above is consistent with macroeconomic data. Considerable diversification in terms of the housing conditions is also worth noting (measured by coefficient of variation CV). During the analysed period, also the number of rooms in leased flats increased and the ratio of the number of persons per room decreased.

Table 1. Selected characteristics of the housing conditions of households in Poland during the period 2005–2015.

Specification		Households with no debt				Households with debt			
		size of flat (m^2)	no. of rooms	area of the flat per person (m^2)	no. of persons per room	size of flat (m^2)	no. of rooms	area of the flat per person (m^2)	no. of persons per room
2005	*AV*	68.45	2.71	28.34	1.27	68.19	2.75	24.94	1.38
	ME	60.00	3.00	23.33	1.00	59.00	3.00	20.00	1.00
	CV	54.81	47.78	68.34	66.81	53.32	45.41	66.59	62.66
2007	*AV*	71.98	2.77	30.63	1.20	72.05	2.80	26.90	1.32
	ME	60.00	3.00	25.00	1.00	60.00	6.00	21.70	1.00
	CV	56.12	47.88	67.99	65.34	55.42	45.57	69.21	63.42
2009	*AV*	73.96	2.80	32.44	1.15	74.83	2.88	28.52	1.23
	ME	60.00	3.00	26.50	1.00	60.00	3.00	23.75	1.00
	CV	56.62	47.77	67.19	66.91	56.97	45.97	65.91	62.41
2012	*AV*	78.33	2.86	34.09	1.09	80.23	2.98	30.17	1.15
	ME	63.00	3.00	28.00	1.00	63.00	3.00	25.00	1.00
	CV	58.09	47.16	66.71	66.38	59.28	45.5	66.78	60.37
2015	*AV*	78.24	2.95	36.16	1.01	81.07	3.07	31.87	1.09
	ME	64.00	3.00	30.00	1.00	64.00	3.00	25.50	1.00
	CV	56.84	46.76	66.27	64.32	58.85	44.17	68.70	57.74

Note: *AV* – average; *ME* – median; *CV* – coefficient of variation.

When considering the housing conditions in terms of selected symptoms presented in Table 2, the results confirm varied threats to housing conditions due to the education level of reference person and place of residence. If the reference person in the household has higher education or if the household is located in a large city, it is characterised by the lowest risk of poor housing conditions. The higher the education level of the household head and the size of the location where the household lives, the more noticeable are improvements in housing conditions.

Table 2. Indicator of the lack of risk of poor housing conditions, considering the education level of reference person and class of locality in 2005 and in 2015.

Specification	2005	2015	2005	2015
	households with no debt		households with debt	
education of a reference person				
lower secondary or lower	50.14	63.02	49.73	63.08
basic vocational	56.26	68.41	56.13	69.07
upper secondary general	63.65	70.62	62.25	70.75
upper secondary vocational	63.77	71.30	62.87	71.48
tertiary	68.81	73.30	68.29	73.67
class of locality				
rural	53.68	67.59	54.66	69.57
towns below 100 thous.	61.31	70.03	61.47	70.99
cities 100-499 thous.	62.05	70.61	61.61	71.44
cities 500 thous. and over	62.97	71.30	63.15	72.18
total	58.66	69.18	59.31	70.75

During the analysed years, an increase may also be noted in the ratio of the absence of risk of poor housing conditions in each household category, both in the case of households in debt and those which are not in debt. When comparing households in debt and those that have no debt, it should be stated that in 2005 the former were in a slightly worse conditions in terms of housing conditions while in 2015 the indicators of the absence of a risk of poor conditions were higher. The findings indicate that differences in the risk of poor housing conditions of indebted and non-indebted households do not generally occur.

Conclusions

The conducted research suggests that during the period from 2005 to 2015 Poland has witnessed improved housing conditions, in both quantitative and qualitative terms. An average household in Poland has a greater average surface, number of rooms and surface in square metres per person. A boom in the mortgage market also had a positive impact on housing conditions, which expressed itself especially in an increase in the indicator of the absence of risk of a poor housing conditions. Based on this, it may be concluded that there has been an improvement in housing conditions not only in quantitative terms but also in the scope of the technical condition of flats and buildings. Well-educated persons in the largest

Polish cities may be treated as the main beneficiaries of the boom in the housing and mortgage markets.

However, considering the increase in the debt of households in Poland during the analysed period (estimated at a level of more than PLN 330 billion), it seems that changes in general housing conditions do not differ considerably from the conditions of households which are not in debt. For this reason, an evaluation of the impact of the mortgage boom on housing conditions is not so clearly positive. Potential benefits in the housing conditions which may only be obtained thanks to mortgages have been partly neutralised by an increase in prices in the real estate market. Prices in the primary and resale markets during the period before 2008 crisis rose by approx. 50% in the largest cities of Poland (in comparison to the period before accession to the European Union) and have not changed in subsequent years. This has made the prices of properties in Poland reach levels of other European countries. This was one of the most important sectors of Polish economy to undergo such speedy and deep real convergence. A permanent burden of loan repayment on household budgets was the consequence of increased debt of households in the years 2005–2015. Even though Polish households improved their housing conditions thanks to debts, the boom resulted in a still unresolved problem of mortgages denominated in foreign currencies and the exposure of households to FX risk which is a threat to their financial stability.

Acknowledgements

This paper was supported by funds from the National Science Centre (NCN, Poland) through grant No. UMO-2015/19/D/HS4/02569.

References

Backé, P. (2009). Credit Growth and Financial Deepening: Implications for Euro Adoption? In: *Raport na temat pełnego uczestnictwa Rzeczypospolitej Polskiej w trzecim etapie Unii Gospodarczej i Walutowej. Eseje zagranicznych i polskich ekonomistów*. Warszawa: Narodowy Bank Polski.

Betti, G., Gagliard, F., Lemmi A. & Verma V. (2015). Comparative measures of multidimensional deprivation in the European Union. *Empirical Economics, 3*(49), 1071-1100.

Betti, G., Soldi, R. & Talev, I. (2016). Fuzzy Multidimensional Indicators of Quality of Life: The Empirical Case of Macedon. *Social Indicators Research, 1*(127), 39-53.

Brzoza-Brzezina, M. (2005). *Lending Booms in the New EU Member States. Will Euro Adoption Matter?* European Central Bank, Working Paper Series no. 543: 11–18.

Cerioli, A. & Zani, S. (1990). A Fuzzy Approach to the Measurement of Poverty. In: C.

Cheli, B. & Lemmi, A. (1995). A Totally Fuzzy and Relative Approach to the Multidimenlsional Analysis of Poverty. *Economic Notes, 24*(1), 115-134.

Eickmeier, S. & Hofmann, B. (2013). Monetary Policy, Housing Booms, and Financial (Im)Balances. *Macroeconomic Dynamics, 17*, 830-860.

Leszczyński, R. & Olszewski, K. (2017). An analysis of the primary and secondary housing market in Poland: evidence from the 17 largest cities. *Baltic Journal of Economics, 17*(2), 136-151.

Panek, T. (2011). *Ubóstwo, wykluczenie społeczne i nierówności. Teoria i praktyka pomiaru*. Warszawa: Wyd. SGH.

Ulman, P. (2016). Ocena zagrożenia złymi warunkami mieszkaniowymi w Polsce. *Studia i Prace WNEiZ US, 44/2*, 475-487.

Verma, V., Betti, G. & Gagliardi, F. (2017). Fuzzy Measures of Longitudinal Poverty in a Comparative Perspective. *Social Indicators Research, 130*, 435-454.

Wałęga, A. (2015a). *Spójność ekonomiczna gospodarstw domowych w Polsce. Analiza statystyczna*, Monografie: Prace Doktorskie nr 23, Kraków: Wydawnictwo Uniwersytetu ekonomicznego w Krakowie.

Wałęga, G. (2013). *Kredytowanie gospodarstw domowych*, Warszawa: Polskie Wydawnictwo Ekonomiczne.

Wałęga, G. (2015b). Zmiany w polityce kredytowej banków a zadłużenie polskich gospodarstw domowych. *Konsumpcja i Rozwój, 1*(10), 44-55.

Widłak, M. & Łaszek, J. (2016). Milestones of Housing Finance in Poland. Milestones in European Housing Finance, 291.

Evaluation of economic efficiency of small manufacturing enterprises in districts of Wielkopolska province using interval-valued symbolic data and the hybrid approach

Marek Walesiak [1], Grażyna Dehnel[2]

Abstract

The article describes a hybrid approach to evaluating economic efficiency of small manufacturing enterprises (employing from 10 to 49 people) in districts of Wielkopolska province. The analysis was based on data prepared in a two-stage process. First, a dataset of 2,162 observations was obtained for three metric variables describing economic efficiency of small manufacturing enterprises. These unit-level data were aggregated at district level and turned into interval-valued symbolic data. Economic efficiency of small manufacturing enterprises was evaluated using a hybrid approach. In the first step, multidimensional scaling (see Borg and Groenen, 2005; Mair et al., 2017) is applied to obtain a visual representation of objects in a two-dimensional space. In the next step, a set of objects is ordered linearly based on the Euclidean distance from the pattern (ideal) object. The proposed approach provides new possibilities for interpreting linearly ordered results of a set of objects.

Keywords: *small enterprises, interval-valued symbolic variables, multidimensional scaling ,composite measures*
JEL Classification: C38, C43, C63
DOI: 10.14659/SEMF.2018.01.58

1 Introduction and motivation

The SME sector plays an important role in the development of the Polish economy, and the group of small businesses (employing between 10 and 49 people) is of particular interest in this respect. What makes small companies noteworthy is their ability to compete even with the largest enterprises thanks to a strict control of costs, their flexibility, which enables them to react quickly to changing market requirements and the ability to implement innovative solutions relatively quickly. At present about 57,000 small businesses are active in Poland, most of which tend to operate locally. In terms of industrial classification, manufacturing is one of the most important and also most numerous category of activity in this group: manufacturing companies account for 26% of all small businesses, generate 20% of total revenue, and provide 30% of jobs in this sector (Główny Urząd Statystyczny, 2017).

[1]Corresponding author: Wrocław University of Economics, Department of Econometrics and Computer Science, ul. Nowowiejska 3, 58-500 Jelenia Góra, Poland, marek.walesiak@ue.wroc.pl.
[2] Poznan University of Economics and Business, Department of Statistics, al. Niepodległości 10, 61-875 Poznań, Poland, g.dehnel@ue.poznan.pl.

Economic efficiency is defined as a relationship between effects and outlays, which, in this case, is measured on an operational level using efficiency ratios to assess the company's performance (Jaki, 2012; Koliński, 2011).

The article describes a study designed to evaluate the economic efficiency of small manufacturing enterprises in districts of Wielkopolska province. The evaluation was conducted using a hybrid approach combining multidimensional scaling (MDS) and linear ordering. Studies of this type are typically based on a classical data matrix. The novelty of the present study is the fact that is based on interval-valued symbolic data obtained in a two-stage process. Interval-valued variables describe objects of interest more precisely than classical metric variables. For classical metric data, an observation on the *j*-th variable for the *i*-th object in a data matrix is expressed as one real number. In contrast, for symbolic interval-valued data, observations on each variable are expressed as intervals$x_{ij} = [x_{ij}^l, x_{ij}^u](x_{ij}^l \leq x_{ij}^u)$, where$x_{ij}^l$denotes the lower bound and x_{ij}^u 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. The empirical study described in this article was based on official statistics from a survey of small businesses. The survey is carried out to collect information about basic measures of economic activity in companies (Dehnel, 2015). The reference period for survey data was 2012. Another data source was the register maintained by the Ministry of Finance.

2 Research methodology

A two-step hybrid approach, presented by Walesiak (2016), enables the visualisation of linear ordering results. In this study it was adopted to order analysed objects. A research procedure that takes into account the specificity of interval-valued symbolic variables includes steps:

1. Select the research problem.

2. Select objects and interval-valued variables substantively related to the research problem. A pattern object (upper pole) and an anti-pattern object (lower pole) are added to the set of objects. Preference variables (stimulants, destimulants and nominants) are included among the interval-valued variables. Definitions of these variables are available in the study e.g. (Walesiak, 2016). Nominants are transformed into stimulants.

3. Collect data and construct data table$\mathbf{X} = [x_{ij}]_{nxm}$for $x_{ij} = [x_{ij}^l, x_{ij}^u]$, where $x_{ij}^l \leq x_{ij}^u$, $i, k = 1, \ldots, n$ and $j = 1, \ldots, m$. The pattern object includes the most favourable variable values, whereas the anti-pattern – the least favourable values of the preference variables

(separately for lower and upper bounds of the interval).

4. Select the variable normalization method and the construction of normalized data table$\mathbf{Z} = [z_{ij}]_{nxm}$ for $z_{ij} = [z_{ij}^{l}, z_{ij}^{u}]$, where $z_{ij}^{l} \leq z_{ij}^{u}$($z_{ij}$ –normalized observation). Interval-valued symbolic data require a special normalization approach. The lower and upper bound of the interval of the *j*-th variable for *n* objects are combined into one vector containing *2n* observations. This approach makes it possible to apply normalization methods used for classical metric data. The data were normalized using the interval_normalization function from the cluster Simpackage (Walesiak and Dudek, 2017a).

5. Select the distance measure for interval-valued data (4 distance measures were taken into account: Ichino-Yaguchi, Euclidean Ichino-Yaguchi, Hausdorff, Euclidean Hausdorff – see Billard and Diday, 2006; Ichino and Yaguchi, 1994)and construct a distance matrix in *m*-dimensional space$\boldsymbol{\delta} = [\delta_{ik}(\mathbf{Z})]_{nxn}$for $i, k = 1, \ldots, n$.

6. Perform multidimensional scaling (MDS):$f: \delta_{ik}(\mathbf{Z}) \rightarrow d_{ik}(\mathbf{V})$for all pairs (i, k)– mapping distances in m-dimensional space $\delta_{ik}(\mathbf{Z})$into corresponding distances$d_{ik}(\mathbf{V})$in q-dimensional space ($q < m$) by a representation function *f*. The distances $d_{ik}(\mathbf{V})$are always unknown. That is, MDS must find a configuration **V**of predetermined dimensions *q* on which the distances are computed. To enable graphic presentation of linear ordering results $q = 2$. Iterative procedure in the smac of algorithm is presented in the study by (Borg and Groenen, 2005).The solution allowing the choice of an optimal MDS procedure was used to account for the methods used to normalize the variables, the distance measure for interval-valued variables and scaling models, according to the procedure available in mdsOpt package (Walesiak and Dudek, 2017b, 2017c), which applies the smac of Sym function from the smacof package (Mair et al. 2017; De Leeuw and Mair, 2009).

7. Finally, after applying the optimal multidimensional scaling procedure, a data matrix in two-dimensional space $\mathbf{V} = [v_{ij}]_{nxq}$($q$ equals 2) is generated.

8. Depending on the position of the pattern and anti-pattern in two-dimensional space$\mathbf{V} = [v_{ij}]_{nx2}$ the coordinate system needs to be rotated by an angle of φ according to the formula$[v'_{ij}]_{nx2} = [v_{ij}]_{nx2} \times D$ ($[v'_{ij}]_{nx2}$ – data matrix in two-dimensional scaling space after rotating the coordinate system by an angle of φ,$D = \begin{bmatrix} cos\varphi & -sin\varphi \\ sin\varphi & cos\varphi \end{bmatrix}$– rotation matrix).

9. Graphic presentation and interpretation of the results in a two-dimensional space. Two points, representing the anti-pattern and pattern, are joined by a straight line to form the so-called set axis in the diagram. Isoquants of development (curves of equal development) are

drawn from the pattern point. Objects located between the isoquants represent a similar level of development. The same level can be achieved by objects located at different points along the same isoquant of development (due to a different configuration of variable values).

10. Objects are ordered linearly using an aggregated measure (composite indicator) d_i based on the Euclidean distance from the pattern object (Hellwig, 1981):

$$d_i = 1 - \sqrt{\sum_{j=1}^{2}(v_{ij} - v_{+j})^2} \Big/ \sqrt{\sum_{j=1}^{2}(v_{+j} - v_{-j})^2}, \quad (1)$$

where: v_{ij} – j-th coordinate for the i-th object in a two-dimensional MDS space, $v_{+j}(v_{-j})$ – j-th coordinate for the pattern object (anti-pattern) in the 2-dimensional MDS space.

Values of aggregated measure d_i are included in the interval $[0; 1]$. The higher the value of d_i, the higher the economic efficiency of small manufacturing enterprises in districts. Target objects are ranked according to the descending values of the aggregated measure (1).

3 Empirical results

The empirical study uses statistical data about the economic efficiency of small manufacturing enterprises in districts of Wielkopolska province in 2012. The target dataset was prepared in two stages. First, a dataset of 2,162 observations was compiled with three metric variables describing economic efficiency of small manufacturing enterprises (employing 10-49 people): x1 – return on sales in % (net profit as a percentage of sales revenue); x2 – sales revenue in thousand PLN per one employee; x3 – costs in thousand PLN per one employee.

The study did not cover more variables due to unavailability of data. Variables x1 and x2 are stimulants and x3 is a destimulant. In the second step, the observations were aggregated at the level of districts of Wielkopolska province, producing a set of symbolic interval-valued data. The lower bound of the interval for each symbolic interval-valued variable in each district was given by the first quartile of the entire dataset. The upper bound of the interval was obtained by calculating the third quartile. Table 1 presents interval-valued symbolic data describing the economic efficiency districts of Wielkopolska province.

The selection of an optimal scaling procedure was made after testing combinations of ten normalization methods (n1, n2, n3, n5, n5a, n8, n9, n9a, n11, n12a – see Walesiak and Dudek, 2017a; Jajuga and Walesiak, 2000), four distance measures for interval-valued data (Ichino-Yaguchi, Euclidean Ichino-Yaguchi, Hausdorff, Euclidean Hausdorff) and four MDS models (ratio, interval, mspline of second and third degree – see Borg and Groenen, 2005) – altogether 160 MDS procedures. MDS was performed for each procedure separately. Next,

the procedures were arranged in ascending order taking into account values of the *Stress*-1, which measures goodness-of-fit (see e.g. Borg et al. 2013). The percentage shares of objects in the value of *Stress*-1 (*spp* – stress per point) measure, was used to calculate the *HHI* index (Herfindahl, 1950; Hirschman, 1964):$HHI_p = \sum_{i=1}^{n} spp_{pi}^2$. The HHI_p index takes values in the interval$[10{,}000/n\,;10{,}000]$. From the perspective of MDS the lowest value of the HHI_p index is desirable. Of the acceptable MDS procedures, for which$Stress\text{-}1_p \leq$ mid-range$(Stress\text{-}1)$, we selected one which meets the condition $\min_p\{HHI_p\}$. It wasprocedure 95: normalization n5 (normalization in range [–1; 1]); mspline MDS model of second degree; Euclidean Ichino-Yaguchi distance.

Table 1. Interval-valued data for three variables describing the economic efficiency of small manufacturing enterprises in districts of Wielkopolska province in 2012.

No.	District	x1	x2	x3
1	chodzieski	[1.86, 10.36]	[85, 265.21]	[80.06, 247.58]
2	czarnkowsko-trzcianecki	[1.17, 15.49]	[92.04, 215.08]	[79.24, 208.82]
3	gnieźnieński	[1.44, 12.49]	[67.61, 198.79]	[65.35, 175.54]
4	gostyński	[2.28, 12.01]	[65.48, 205.99]	[59.82, 168.22]
5	grodziski	[2.3, 9.72]	[129.02, 341.85]	[112.96, 323.22]
6	jarociński	[2.08, 17.09]	[64.92, 153.01]	[53.69, 135.89]
7	kaliski	[1.07, 5.77]	[104.95, 394.23]	[112.33, 358.85]
8	kępiński	[1.87, 8.9]	[76.8, 161.89]	[73.4, 150.87]
9	kolski	[1.48, 7.91]	[75.86, 437.76]	[73.17, 433.45]
10	koniński	[1.72, 7.97]	[99.07, 267.03]	[89.85, 246.49]
11	kościański	[2.41, 14.53]	[98.37, 217.48]	[87.7, 195.97]
12	krotoszyński	[1.83, 10.67]	[81.67, 181.89]	[73.61, 153.8]
13	leszczyński	[1.09, 9.45]	[100.63, 197.59]	[95.72, 191.07]
14	międzychodzki	[2.29, 9.96]	[71.29, 178.49]	[67.49, 172.63]
15	nowotomyski	[3.56, 12.94]	[73.95, 250.52]	[71.61, 219.97]
16	obornicki	[0.63, 8.03]	[91.14, 197.87]	[88.30, 196.24]
17	ostrowski	[2.05,11.46]	[72.59, 217]	[67.2, 186.12]
18	ostrzeszowski	[1.83, 9.12]	[79.34, 270.57]	[70.31, 261.16]
19	pilski	[1.03, 12.08]	[82.57, 227]	[75.69, 194.45]
20	pleszewski	[1.39, 14.4]	[62.39, 178.2]	[63.67, 173.23]

21	poznański	[1.53, 12.6]	[104.53, 262.46]	[96.56, 246.48]
22	rawicki	[2.78, 8.27]	[69.74, 183.93]	[61.31, 171.96]
23	słupecki	[1.34, 10.09]	[74.83, 333.98]	[64.62, 327.98]
24	szamotulski	[0.01, 11.5]	[129.28, 267.24]	[120.39, 253.81]
25	średzki	[1.5, 10.31]	[136.71, 275.28]	[125.13, 269.17]
26	śremski	[0.69, 9.82]	[55.12, 281.88]	[55.46, 273.49]
27	turecki	[1.91, 10.64]	[71.71, 136.18]	[69.26, 127.26]
28	wągrowiecki	[2.23, 7.04]	[77.38, 231.82]	[74.72, 227.62]
29	wolsztyński	[2.17, 10.03]	[124.61, 493.23]	[107.75, 457.19]
30	wrzesiński	[1.21, 9.76]	[71.62, 187.06]	[68.09, 180.97]
31	złotowski	[2.39, 7.78]	[88.54, 263.99]	[82.16, 249.96]
32	m. Kalisz	[2.17, 10.3]	[47.08, 187.11]	[56.67, 196.88]
33	m. Konin	[3.73, 13.63]	[67.23, 210.7]	[60.84, 170.29]
34	m. Leszno	[3.29, 11.25]	[74.92, 266.16]	[70.53, 227.62]
35	m. Poznań	[0.77, 12.8]	[94.52, 268.76]	[77.55, 251.88]
P	Pattern	[3.73, 17.09]	[136.71, 493.23]	[53.69, 127.26]
AP	Anti-pattern	[0.01, 5.77]	[47.08, 136.18]	[125.13, 457.19]

Fig. 1 (left panel) shows the Shepard diagram which confirms the correctness of the selected MDS model. The right panel (Stress plot) shows that the MDS configuration represents all proximities almost equally well. Finally, after applying the optimal MDS procedure a data matrix in a two-dimensional space was obtained. Fig. 2 presents results of MDS of 37 objects (35 districts, the pattern and anti-pattern), in terms of the economic efficiency of small manufacturing enterprises.The coordinate system was rotated by an angle $\varphi = 0.9\pi$. The anti-pattern (AP) and pattern (P) were connected by a straight line to form the so-called set axis. Six isoquants of development were defined by dividing the set axis into 6 equal parts. Next, the values of the composite measure (1) were calculated. Table 2 presents the ordering of 35 districts in terms of the economic efficiency of small manufacturing enterprises, in descending order of values of (1). The calculations were performed using R (R Core Team, 2017).

By presenting results in this way it is possible to:

– order districts by the economic efficiency of small manufacturing enterprises measured by three variables based on values of measure (1) and present them graphically in Fig. 2,

– distinguish classes of districts (districts between isoquants) sharing a similar level of

economic efficiency (see Fig. 2),

– identify districts characterized by a similar level of economic efficiency, but having a different location on the isoquant of development (see Fig. 2). For example, Kępiński District (8) and Kolski District (9) have a similar level of economic efficiency, but are located at different points on the isoquant of development and in different parts of the province (see Fig. 3). A similar situation occurs for Wolsztyński District (29) and Turecki District (27): while these districts achieved a similar level of development, they were characterized by quite different configurations of variable values.

Table 2. Ordering of districts of Wielkopolska province according to the economic efficiency of small manufacturing enterprises in 2012.

Rank	**District**	**No.**	d_i	**Rank**	**District**	**No.**	d_i
1	czarnkowsko-trzcianecki	2	0.5956	19	śremski	26	0.4305
2	kościański	11	0.5589	20	ostrzeszowski	18	0.4158
3	nowotomyski	15	0.5569	21	krotoszyński	12	0.4013
4	m. Poznań	35	0.5490	22	m. Kalisz	32	0.3911
5	szamotulski	24	0.5426	23	międzychodzki	14	0.3810
6	poznański	21	0.5376	24	wrzesiński	30	0.3807
7	m. Konin	33	0.5247	25	leszczyński	13	0.3772
8	m. Leszno	34	0.4963	26	koniński	10	0.3741
9	średzki	25	0.4865	27	złotowski	31	0.3634
10	pleszewski	20	0.4848	28	wolsztyński	29	0.3616
11	pilski	19	0.4808	29	turecki	27	0.3551
12	jarociński	6	0.4767	30	kępiński	8	0.3340
13	gnieźnieński	3	0.4648	31	kolski	9	0.3258
14	grodziski	5	0.4596	32	wągrowiecki	28	0.3252
15	chodzieski	1	0.4536	33	obornicki	16	0.3189
16	gostyński	4	0.4521	34	rawicki	22	0.3157
17	ostrowski	17	0.4517	35	kaliski	7	0.2775
18	słupecki	23	0.4315				
Mean			0.4324	Median			0.4315
Standard deviation			0.0818	Median absolute deviation			0.0960

The results of the MDS of districts combined with information about their geographical

location seem to confirm the assumptions of the theory of growth poles. This is true not only for its original application, which was limited to economic entities, but also in the current interpretation, which accounts for the spatial dimension (Isard, 1960). One can clearly see the impact of Poznań, which acts as a growth pole, on the neighbouring districts (see Fig. 3). Districts located further away from Poznań are found lower in the ranking based on measure (1), except for two districts – obornicki and wągrowiecki – for which the value of measure d_i was very low (0.32 and 0.33 respectively). This discrepancy can be explained in a number of ways. For one things, these two districts are characterised by high unemployment rate, a relatively small number of working persons; moreover, their inhabitants mainly work in Poznań. According to a study of commuting flows, the largest number of commuters working in Poznań come from districts located north of the city (Główny Urząd Statystyczny, 2014).

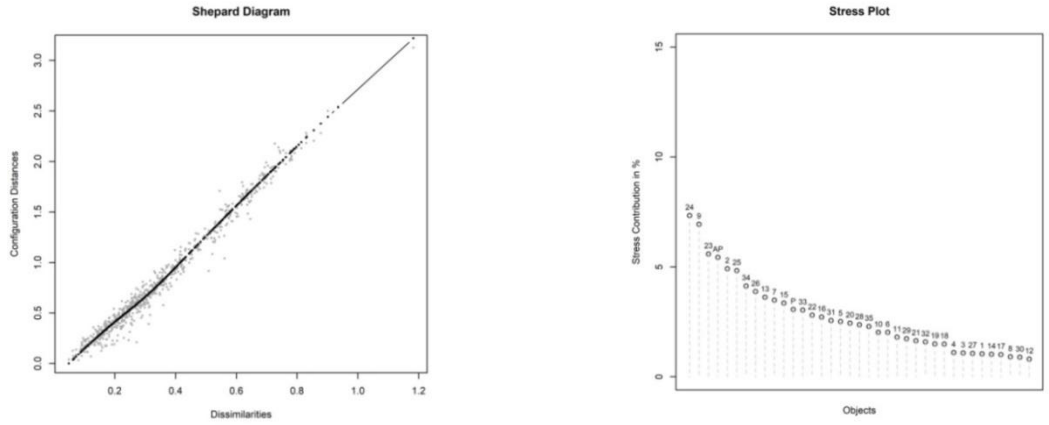


Fig. 1. Shepard diagram and Stress plot.

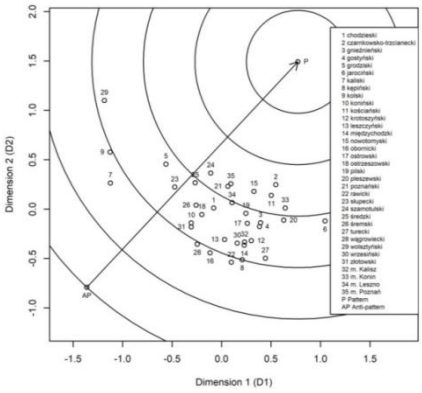


Fig. 2. Results of MDS of 35districtsof Wielkopolska byeconomicefficiency of small manufacturing enterprises.

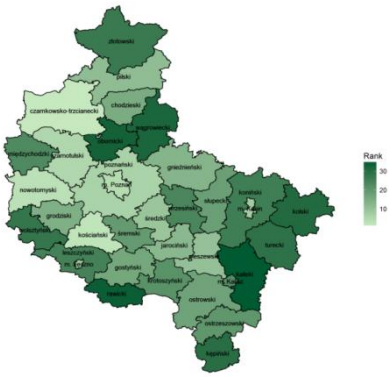


Fig.3. Spatial distribution of districts of Wielkopolska based on the ranking Table 2.

Conclusions

The study described above was an attempt to compare districts of Wielkopolska province in terms of the economic efficiency of small manufacturing companies.The authors used a hybrid approach combining multidimensional scaling and linear ordering.The empirical study was based on interval-valued symbolic data. Districts were evaluated according to the economic efficiency of small manufacturing companies measured by three variables. Thanks to the methodological approach used in the study, it was possible to present the results of linear ordering graphically in a two-dimensional space. In this way districts could be arranged in terms of the economic efficiency of small manufacturing companies and divided into groups sharing a similar level of economic efficiency.The graphical presentation also facilitated the identification of groups of similar districts characterised by similar values of the target variables and those with a different configuration of variable values.The empirical results presented on a map confirm the influence of Poznan as a growth pole on the neighbouring districts. 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 implement a particular methodological approach.

Acknowledgements

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

References

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. (2013). *Applied Multidimensional Scaling*. 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.

Dehnel, G. (2015). Robust regression in monthly business survey. *Statistics in Transition new series*, *16*(1), 1-16.

De Leeuw, J. & Mair, P. (2009). Multidimensional scaling using majorization: SMACOF in R. *Journal of Statistical Software*, *31*(3), 1-30.

Gioia, F. & Lauro, C. N. (2006). Principal component analysis on interval data.

Computational Statistics, *21*(2), 343-363.

Główny Urząd Statystyczny.(2014). *Dojazdy do pracy*. Warszawa: GUS.

Główny Urząd Statystyczny. (2017). *Działalność przedsiębiorstw niefinansowych w 2015 roku*. Warszawa: GUS.

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*(pp. 46-68). Warszawa: PWE.

Herfindahl, O. C. (1950). *Concentration in the US steel industry*. Unpublished PhD. Dissertation, Columbia University.

Hirschman, A. O. (1964). The paternity of an index. *Am Econ Rev.*, *54*(5), 761-762.

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.

Isard, W. (1960). Methods of Regional Analysis: An Introduction to Regional Science. Cambridge, Massachusetts: The M.I.T. Press.

Jajuga, K. & Walesiak, M. (2000). Standardisation of data set under different measurement scales. In: Decker, R. & Gaul, W., (Eds.), *Classification and Information Processing at the Turn of the Millennium*(pp. 105-112). Berlin, Heidelberg: Springer-Verlag.

Jaki, A. (2012). *Mechanizmy procesu zarządzania wartością przedsiębiorstwa*. Zeszyty Naukowe UE w Krakowie. Seria Specjalna, Monografie (215).

Koliński, A. (2011). Przegląd metod i technik oceny efektywności procesu produkcyjnego. *Logistyka*, 5, 1083-1091.

Mair, P., De Leeuw, J., Borg, I. & Groenen, P. J. F. (2017). smacof: Multidimensional Scaling. *R package*, version 1.9-6. http://CRAN.R-project.org/package=smacof

R Core Team. (2017). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. http://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. & Dudek, A. (2017a). clusterSim: Searching for Optimal Clustering Procedure for a Data Set. *R package*, version 0.47-1. http://CRAN.R-project.org/package=clusterSim

Walesiak, M. & Dudek, A. (2017b). mdsOpt: Searching for Optimal MDS Procedure for Metric Data. *R package*, version 0.2-2. http://CRAN.R-project.org/package=mdsOpt

Walesiak, M. & Dudek, A. (2017c). Selecting the optimal multidimensional scaling procedure for metric data with R environment. *Statistics in Transition new series*, *18*(3), 521-540.

Estimation of VaR bounds under dependence uncertainty and their use for the SCR calculation in Solvency II

Stanisław Wanat[1], Krzysztof Guzik[2]

Abstract

The subject of the article is part of the discussion regarding the VaR estimation for aggregated risk under conditions of uncertain structure of dependence, i.e. VaR for the sum of individual risk with known marginal distributions and unspecified dependence structure. It demonstrates that the use of a standard formula in accordance with Article 115 of the Delegated Regulation (EU) 2015/35, without identifying the actual dependency structure, may lead to an erroneous estimation of solvency capital requirements for non-life premium and reserve risk. It also indicates how large the errors may be. The analysis was carried out using the dependence uncertainty spread estimation methods known in the literature.

Keywords: *dependence structure, Solvency Capital Requirements, risk aggregation, VaR bounds,*
JEL Classification: C150, C580, G220
DOI: 10.14659/SEMF.2018.01.59

1 Introduction

The Value-at-Risk (VaR) has been mainly used for measuring risk at banks and insurance companies for the last three decades. This measure constitutes basis for making decisions on limiting risk exposure to an acceptable level, depending on available economic capital, and it is also used for determining regulatory capital which are to guarantee the solvency. For example, VaR is used in Solvency II in the standard method of determining the solvency capital requirements (SCR). So it is not surprising that in the subject literature there have been and are still wide-ranging debates being conducted on issues related to diversification, aggregation, dependence uncertainty, economical interpretation, optimization, extreme behavior, robustness, and back-testing of VaR; see, among others, Embrechts et al. (2015), Embrechts et al. (2014) and Emmer et al. (2014) for more details.

The article subject fits into the discussion on assessing VaR for the aggregated risk in the conditions of the unspecified dependence structure, that is VaR for the sum of individual risks with known marginal distributions but the unknown dependence among them. It broadens the existing subject literature in a scope of quantitative, multivariate modeling of the risk in the

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

process of determining the SCRs. The work shows that the use of a standard formula in accordance with Article 115 of the Delegated Regulation (EU) 2015/35, without identifying the actual dependency structure, may lead to an erroneous estimation of solvency capital requirements for non-life premium and reserve risk. It also indicates how large the errors may be. The analysis was carried out using the dependence uncertainty spread estimation methods known in the literature. To the authors' best knowledge, the results of that type of studies have not been presented in the literature yet.

The other part of the article is organized in the following way. In the next chapter we present the methods of assessing the limitations for VaR (VaR bounds). In chapter 3 we present briefly the standard method assessing the solvency capital requirements for the non-life premium and reserve risk used in Solvency II. Then in chapter 4 we discuss the assumptions and the results of the simulation study performed, and in the last fifth chapter we present conclusions.

2 Methods of assessing the bounds for VaR – a literature review

We consider a portfolio consisting of n random variables $L_1, L_2 \dots, L_n$ (individual risks associated with a given business line or a risk type) with the finite mean and the variance. We assume marginal distributions F_j $(j = 1, \dots, n)$ corresponding to them are provided. For random variable $L = \sum_{j=1}^{n} L_j$, modelling the aggregate portfolio loss, at the established confidence level 1-α , we define Value-at-Risk (VaR):

$$VaR_{1-\alpha}(L) = \inf\{x \epsilon R : F_L(x) \geq 1 - \alpha\},$$

where F_L is the distribution of L. The Value-at-Risk depends only on the dependence structure of the multivariate random variable $(L_1, \dots, L_n)$, and the copula C describes this structure on basis of the Sklar's theorem. When we know the copula C we are able to determine VaR for the random variable L in a simple way, otherwise we can give only its bounds.

The first results for assessing VaR in the case of the unspecified dependence structure for two random variables are presented in the works (Makarov, 1981) and (Frank et al., 1987) and independently in (Rüschendorf, 1982). This issue has been discussed recently in many works (Puccetti and Rüschendorf, 2012a, 2012b, 2014; Embrechts et al., 2013; Puccetti et al., 2013; Bernard et al., 2016, 2017). It should be pointed out that the results obtained for three or more random variables require certain assumptions such as e.g. identically distributed risks having monotone densities (Puccetti, 2013; Puccetti and Rüschendorf, 2013). For the inhomogeneous case there are no analytical solutions enabling to determine „reasonable" VaR bounds. The

studies on assessing VaR resulted in proposing numerical algorithms which do not require the assumptions mentioned above. We present briefly three of them.

Rearrangement Algorithm *(RA)* was proposed by Puccetti and Rüschendorf (2012b) and Embrechts et al. (2013). On its basis the VaR assessments can be determined at known, not necessarily homogeneous, marginal distributions. In the RA algorithm the Value-at-Risk measure such as Expected Shortfall (*ES*) was used, *ES* unlike VaR meets the property of sub-additivity and is thus coherent:

$$ES_{1-\alpha}(L) = \frac{1}{\alpha}\int_{1-\alpha}^{1} VaR_q(L)dq \ .$$

Value $ES_{1-\alpha}(L)$ is the average loss of all upper $VaRs$ from level $1-\alpha$, when those losses at the established confidence level 1-α exceed $VaR_{1-\alpha}(L)$. Therefore, the upper estimate of VaR by this measure is following:

$$ES_{1-\alpha}(L) \leq \overline{ES_{1-\alpha}}(L) = \sum_{j=1}^{n} ES_{1-\alpha}(L_j).$$

Since it results from the ES definition that $\overline{VaR_{1-\alpha}}(L) \leq \overline{ES_{1-\alpha}}(L)$ is the bounds for VaR we obtain by estimating ES. In the *RA* algorithms at the established number of discretization points N from the tail of distribution $[1-\alpha, 1]$ for each random variable L_j we determine N of realization of this variable by making matrix X of dimension $N \times n$. Then, the value of each random variable we permute countermononic in relation to other variables so that realizations of the sum of those random variables in the N cases were similar to each other. In this iterative way we obtain the sought estimate at the fixed level absolute tolerance ε

In turn, Adaptive Rearrangement Algorithm was an answer to weak efficiency of the RA algorithm at a greater number of the random variables. It was proposed in work (Hofert et al., 2017). The authors have considered in it at what level values N and ε should be established for the algorithm to be effective (the RA algorithm authors do not take it into consideration). A course of the ARA algorithm is analogous to RA one, with the difference that a number of N discretization points is selected depending on the algorithm convergence degree. Additionally, in contrast to the absolute tolerance ε in the RA algorithm convergence at the assumed level of the relative convergence tolerance ε is examined in the ARA algorithm. Owing to these modifications the searched bounds are obtained much faster. The performed numerical simulations indicate high accuracy of algorithms RA and ARA, although the VaR estimate interval obtained in such a way is mostly very wide. It results from the fact of the unspecified dependence structure and random vector$(L_1, L_2 \ldots, L_n)$.

At the end we mention the Extended Rearrangement Algorithm (ERA) which was proposed by Bernard et al. (2017), at an additional constraint that a bound on the variance of L

is known, that is $D^2(L) \leq s^2$. As it can be expected this assumption improves the VaR estimate of bounds comparing to algorithms RA and ARA. In the ERA algorithm the authors use the risk measures of TVaR (Tail Value at Risk) and LTVaR (Left Tail Value at Risk), which are defined analogically to ES and they are sometimes called so. In this algorithm for fixed *k*, we divide the distribution random variable *L* into two parts corresponding to the fixed confidence level 1-α, determining the matrix $X_{N\times n}$ in such a way. Then, we use the RA algorithm to the upper and lower part of the distribution respectively, obtaining the matrix X^* of dimension $N \times n$. When inequality $D^2(L) \leq s^2$ is satisfied we obtain the lower and upper $VaR_{1-\alpha}(L)$ bounds. On basis of the *ERA* algorithm the authors prove that the models used so far can underestimate VaR. Moreover, they state that establishing the capital requirements at the high confidence level at e.g. 99,5% is justified.

3 The standard method of assessment of the SCR for the non-life premium and reserve risk

In the standard approach of Solvency II, the overall solvency capital requirement for insurer is calculated with use of the following formula (Solvency II Directive, 2009, Art. 103):

$$SCR = BSCR + Adj + SCR_{Op}, \tag{1}$$

where: $BSCR$ - basic solvency capital requirement, Adj - adjustment for the risk absorbing effect of technical provisions and deferred taxes, SCR_{Op} - the capital requirement for operational risk.

The *BSCR* value is determined in the aggregation process of SCRs established for the main risk modules (namely market risk, counterparty default risk, life underwriting risk, non-life underwriting risk, health underwriting risk, intangible assets risk), SCRs for the modules are determined by aggregation of SCRs for the sub-modules, and the last ones as a result of the SCR aggregation for risk carriers. Therefore, this process includes 3 levels of aggregation, they are presented in details in (Commission Delegated Regulation, 2015, CHAPTER V). The solvency capital requirements for each risk (that is the module, the sub-module and the carrier) at the specific aggregation level should correspond to the economic capital (*EC*) established for one year at the confidence level of 0.995 (Solvency II Directive, 2009, Note 64, s. 7). Thus, according to the definition *ECs* (Lelyveld, 2006) should be equal:

$$SCR_i^{(l)} = VaR_{0.995}\left(L_i^{(l)}\right), \;\; L_i^{(l)} = X_i^{(l)} - E\left(X_i^{(l)}\right), \tag{2}$$

where $X_i^{(l)}$ means the random variable which models losses, over an annual time horizon, associated with the *i*-risk at *l*-level of the aggregation (thus, $L_i^{(l)}$ means unexpected losses).

In the further part of our paper we concentrate on the non-life premium and reserve risk sub-module[3]. The standard formula for SCR for this sub-module is following (Commission Delegated..., Art. 115):

$$SCR_{nl\ prem\ res} = 3 \cdot \sigma_{nl} \cdot V_{nl} \tag{3}$$

where: σ_{nl} - the standard deviation for non-life premium and reserve risk determined in accordance with (Commission Delegated Regulation, 2015, Art. 117); V_{nl} - the volume measure for non-life premium and reserve risk determined in accordance with (Commission Delegated Regulation, 2015, Art. 116):

$$V_{nl} = \sum_s V_s, \tag{4}$$

V_s - the volume measure for the premium and reserve risk of segment *s* (a list of the segments is provided in the (Commission Delegated Regulation, 2015, Annex II).

The key role in estimating SCR on basis of formula (2) is played by the standard deviation σ_{nl}, which is given as (Commission Delegated Regulation, 2015, Art. 117, Para.1):

$$\sigma_{nl} = \frac{1}{V_{nl}} \cdot \sqrt{\sum_{s,t} CorrS_{s,t} \cdot \sigma_s \cdot V_s \cdot \sigma_t \cdot V_t} \tag{5}$$

where: $CorrS_{s,t}$- the correlation parameter for the non-life premium and reserve risk for segment *s* and segment *t* set out in (Commission Delegated Regulation, 2015, Annex IV); σ_s, σ_t- standard deviations for the non-life premium and reserve risk of segments s and t respectively; V_s, V_t - volume measures for the premium and reserve risk of segments *s* and *t*, referred to in (Commission Delegated Regulation, 2015, Art. 116), respectively.

For all segments set out in Annex II, the standard deviation for the non-life premium and reserve risk shall be equal to the following (Commission Delegated Regulation, 2015, Art. 117, Para.2):

$$\sigma_s = \frac{\sqrt{\sigma^2_{(prem,s)} \cdot V^2_{(prem,s)} + \sigma_{(prem,s)} \cdot V_{(prem,s)} \cdot \sigma_{(res,s)} \cdot V_{(res,s)} + \sigma^2_{(res,s)} \cdot V^2_{(res,s)}}}{V_{(prem,s)} + V_{(res,s)}} \tag{6}$$

where: $\sigma_{(prem,s)}$ - the standard deviation for the non-life premium risk of segment *s* determined in accordance with (Commission Delegated Regulation, 2015, Art. 117, Para. 3); $\sigma_{(res,s)}$ - the standard deviation for the non-life reserve risk of segment *s* as set out in (Commission Delegated Regulation, 2015, Annex II); $V_{(prem,s)}$ - the volume measure for the premium risk of segment *s* referred to in (Commission Delegated Regulation, 2015, Art. 116);

[3] As a result of SCR aggregation for this sub-module with SCRs for the non-life catastrophe risk sub-module and the non-life lapse risk sub-module SCR for non-life underwriting risk module is obtained. This is the second aggregation level.

$V_{(res,s)}$ - denotes the volume measure for the reserve risk of segment s referred to in (Commission Delegated Regulation, 2015, Art. 116).

It results from the above that in the standard formula on $SCR_{nl\ prem\ res}$ there are not any explicit hypotheses for the distribution of the random variable that describe unexpected losses. This variable will be still marked by L_{nl} (to keep transparency of the entry we omit the superscript denoting the aggregation level). However, considering the general principle saying that SCR for the specific risk should secure unexpected losses at 0,995 confidence level (cf. formula (2)) and a form of formula (3), such an assumption is needed. In particular, for a normal distribution it holds that: $VaR_{0.995}(X) = c \cdot \sigma_X$, where $c = \Phi^{-1}(0.995) = 2.58$. In the standard formula constant $c = 3$. It means that SCR is estimated at 0.9987 confidence level, thus at a higher level than generally accepted. According to legislators it is to prevent SCR underestimate when L_{nl} does not have a normal distribution.

Another doubt connected with applying the standard formula on a $SCR_{nl\ prem\ res}$ concerns the dependence structure between the non-life premium and reserve risks in the particular segments. From a way of determining the parameter for the aggregated non-life premium and reserve risk (formula (5)) it results that it is described only by Pearson correlation coefficients. Since the same correlation coefficients may correspond to different dependence structures obvious questions arise. To what extent is SCR estimated in such a way reliable? How much can it differ from SCR estimated taking the proper dependence structure into account? In order to answer these questions for a specific case a study was performed and its results are presented in the next chapter.

4 SCR for the non-life premium and reserve risk - empirical results

This chapter presents the results of the study in which SCR sensitivity for the premium and reserve risk for two segments of non-life insurance and reinsurance on the dependence structure between these segments was analyzed. The segments in question are:

- motor vehicle liability insurance and proportional reinsurance (segment *s*).
- other motor insurance and proportional reinsurance (segment *t*).

The same values of the volume measures were assumed for the premium and reserve risk of segments *s* and *t*, whereas the values of the necessary standard deviations and correlation parameter were taken as in the (Commission Delegated Regulation, 2015, Annex II and Annex IV) (see Table 1).

Table 1. Parameters used in the study.

Segment *s*	**Segment *t***
$V_{(prem,s)} = 1.0$	$V_{(prem,t)} = 1.0$
$V_{(res,s)} = 1.2$	$V_{(res,t)} = 1.2$
$\sigma_{(prem,s)} = 0.10$	$\sigma_{(prem,t)} = 0.08$
$\sigma_{(res,s)} = 0.09$	$\sigma_{(res,t)} = 0.08$
$CorrS_{s,t} = 0.5$	

Then, according to the standard formula (cf. formula (3)) SCR was calculated, obtaining values $SCR_{nl\ prem\ res} = 0.8656$ (cf. Table 2). Next it has been assumed that random variables L_s, L_t, which model the unexpected losses for the non-life premium and reserve risk of segments *s* and *t* respectively, have normal distributions: $L_s \sim N(0, V_s \cdot \sigma_s)$, $L_t \sim N(0, V_t \cdot \sigma_t)$. For the assumed parameters (Table 1) the following results were obtained (after performing necessary calculations using formulas (3) and (6)): $L_s \sim N(0, 0.1802)$, $L_t \sim N(0, 0.1526)$. At those assumptions, using formula (2) for variable $L_{nl} = L_s + L_t$ SCR was determined for the following cases:

- Case A: Variables L_s and L_t are independent.
- Case B: Variables L_s and L_t are comonotonic.
- Case C: Unknown dependence structure between L_s and L_t. The unconstrained SCR upper bound was determined using TVAR, in accordance with theorem 1 in (Bernard et al., 2017).
- Case D: Unknown dependence structure. The unconstrained SCR upper bound was determined using the ARA algorithm.
- Case E: The partially known dependence structure, namely it is only known that the correlation coefficient between L_s and L_t is 0.5 (cf. Table 1). The constrained SCR upper bound was determined in accordance with theorem 5 in (Bernard et al., 2017)

The obtained results are presented in Table 2. The following conclusions can be drawn on basis of the research made:

- In the process of determining the solvency capital requirements for the non-life premium and reserve risk, knowledge about only the distribution of random variables L_s and L_t without knowledge about the dependence structure between them is not

sufficient. The interval of possible values for SCR from[4] 0.6060 to 0.9342 (case D) obtained in this case is useless from the practical point of view. The same applies to case C, where the length of the interval is even greater due to the method used.

- Using only the correlation coefficient to describe the dependence structure (as it is in the Solvency II standard formula) is insufficient. The same correlation coefficient between L_s and L_t may correspond to different dependence structures, and thus different SCR values. In the performed study (case E) the upper bound of SCR was obtained at 0.9251 level. It means that SCR can be higher than one determined by the standard formula about 6.9%.
- By the standard formula higher SCR was obtained than at the assumption of the co-monotonic dependence structure between L_s and L_t (cf. case B), namely higher than the sum of SCRs for segments s and t. It means that using the standard formula on SCR for the non-life premium and reserve risk a diversification effect is not considered.

Table 2. Research results.

Cases	Capital requirement
Solvency II standard formula	0.8656
Case A	0.6060
Case B	0.8573
Case C	(0.6060*, 0.9625)
Case D	(0.6060*, 0.9342)
Case E	(0.6060*, 0.9251)

(*)The lowest SCR for independent random variables L_s and L_t has been assumed.

Conclusions

The paper demonstrates, using a specific example, that the use of the standard formula in accordance with the Committee Delegated Regulation (EU) 2015/35 included in Article 115 may lead to the improper level of solvency capital requirements for non-life premium and reserve risk. The conducted analysis shows that the correct estimation of SCRs depends on the correct identification of the structure of dependencies between random variables modeling unexpected losses. The application of only linear correlation coefficients for this purpose may

[4] The lowest value was SCR for independent variables L_s and L_p.

lead to erroneous results, since they describe only linear dependencies, in a straightforward manner. In the general case, different dependency structures may have the same value of this coefficient. The work assumes normal distribution for unexpected losses. In practice, however, these are very often asymmetrical distributions. The next stage of the research will be the analysis of the effect of skewness on the SCR evaluation.

Acknowledgements

The study is supported with subsidies for maintaining research capacity granted to the Faculty of Finance and Law at the Cracow University of Economics by the Polish Ministry of Science and Higher Education.

References

Bernard, C., Denuit, M. & Vanduffel, S. (2016). Measuring Portfolio Risk under Partial Dependence Information. *Journal of Risk and Insurance*. In press.

Bernard, C., Rüschendorf, L. & Vanduffel, S. (2017). Value-at-Risk Bounds With Variance Constraints. *Journal of Risk and Insurance, 84(3)*, 923-959.

Commission Delegated Regulation (EU) 2015/35 of 10 October 2014 supplementing Directive 2009/138/EC of the European Parliament and of the Council on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II), Official Journal of the European Union, 17.1.2015.

Directive of the European Parliament and Council 2009/138/EC of 25th November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II)**,** Official Journal of the European Union, 17.12.2009.

Embrechts, P., Puccetti, G. & Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. *Journal of Banking & Finance, 37(8)*, 2750-2764.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R. & Beleraj, A. (2014). An Academic Response to Basel 3.5. *Risks, 2(1)*, 25-48.

Embrechts, P., Wang, B. & Wang, R. (2015). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance and Stochastics, 19(4*), 763-790.

Emmer, S., Kratz, M. & Tasche, D. (2015). What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk, 18(2)*, 31-60.

Frank, M. J., Nelsen, R. B. & Schweizer, B. (1987). Best-possible bounds for the distribution of a sum – a problem of Kolmogorov. *Probability Theory and Related Fields, 74(2)*, 199-211.

Hofert, M., Memartoluie, A., Saunders, D. & Wirjanto, T. (2017). Improved algorithms for computing worst Value-at-Risk. *Statistics and Risk Modeling, 34(1-2)*.

Lelyveld, I. V. (2006). Economic capital modelling: concepts, measurement and implementation. London, England: Risk Books.

Makarov, G.D. (1981). Estimates for the distribution function of the sum of two random variables when the marginal distributions are fixed. *Theory of Probability and its Applications, 26(4),* 803–806.

Puccetti, G. (2013). Sharp bounds on the expected shortfall for a sum of dependent random variables. *Statistics & Probability Letters, 83(4),* 1227-1232. doi:10.1016/j.spl.2013.01.022

Puccetti, G., & Rüschendorf, L. (2012a). Bounds for joint portfolios of dependent risks. *Statistics and Risk Modeling, 29(2),* 107-132.

Puccetti, G. & Rüschendorf, L. (2012b). Computation of sharp bounds on the distribution of a function of dependent risks. *Journal of Computational and Applied Mathematics, 236(7),* 1833-1840.

Puccetti, G. & Rüschendorf, L. (2013). Sharp bounds for sums of dependent risks. *Journal of Applied Probability, 50(1),* 42-53.

Puccetti, G. & Rüschendorf, L. (2014). Asymptotic equivalence of conservative VaR- and ES-based capital charges. *The Journal of Risk, 16(3),* 3-22.

Puccetti, G., Wang, B. & Wang, R. (2013). Complete mixability and asymptotic equivalence of worst-possible VaR and ES estimates. *Insurance: Mathematics and Economics, 53(3),* 821-828.

Rüschendorf, L. (1982). Random variables with maximum sums. *Advances in Applied Probability, 14(03),* 623-632.

Estimation of indirect demographic losses in Ukraine due to armed conflict

Antonina M. Yerina[1], Zinaida O. Palian[2]

Abstract

Any military actions lead to population losses. Direct losses – killed and dead as a result of armed confrontation, forced emigration, as well as the loss of residents through the alienation of the territory in which they remained. The estimate of the extent of direct demographic losses could partly be done by current estimation and more completely – by census. But any statistical survey is not able to determine the size of indirect losses – the number of unborn children due to the hostilities. In this case, it's possible to apply statistical modeling.

In this article we presented the results of statistical estimation of the impact of armed conflict on reduction of the number of births using the Interrupted ARIMA model $(0,1,1)(0,0,1)_4$, which also takes into account the seasonal character of childbearing. The model allows us to conclude that the number of births in Ukraine only in the second quarter of 2014 decreased by 9749 persons, through the annexation of the Crimea and the armed conflict in the Donbas. Continued hostilities, worsening of medical-demographic and socio-economic situation as in combat zone, as well as generally around Ukraine, gives reason to anticipate a further increase in indirect demographic losses.

***Keywords**:* indirect demographic losses, armed conflict, interrupted ARIMA

***JEL Classification**:* C22, C32, J11

DOI: 10.14659/SEMF.2018.01.60

1 Introduction

Today's demographic losses in Ukraine are a consequence of not only recent large-scale tragic events: internal social upheavals and the armed conflict Russia and Ukraine. To this day, the population of Ukraine bears a historical trace of previous numerous losses. If Ukraine had not gone through a series of demographic catastrophes in wake of socio-political disasters and economic crisis, the hypothetical population at the beginning of 1991 would have been 87 million people instead of the actual 52 million (Romaniuk and Gladun, 2015). Thus, according to retrospective calculations of scientists of the Ptoukha Institute for Demography and Social Studies National Academy of Sciences of Ukraine, from 1904 until the beginning of the Russian aggression (2014), the total population losses in Ukraine, according to the current administrative-

[1] Taras Shevchenko National University of Kyiv, Statistics and Demography Department, 90-a, Vasylkivska St. 03022 Kyiv, Ukraine, am_yerina@ukr.net.

[2] Corresponding author: Taras Shevchenko National University of Kyiv, Statistics and Demography Department, 90-a, Vasylkivska St. 03022 Kyiv, Ukraine, zipalyan@ukr.net.

territorial boundaries, amounted to 35.6 million people (Romaniuk and Gladun, 2015). Ukrainian population has suffered such a devastating blow as a result of three wars, the Bolshevik revolution, three stages of mass famine of 1921-1923, 1932-1933, 1946-1947, the cholera epidemic (1910), Stalin's repressions and deportations, compulsory migration, as well as due to the "mortality crisis" (1969-1989) in peaceful times. In fact, the depopulation in Ukraine has been evident over the last 25 years, with its acceleration having started already 11 years before the Russian military aggression. After a prolonged fertility crisis and the survival mode (1992-2002), as well as a powerful wave of emigration (1994-2004), a period of "demographic thaw" came around in wake of a notable increase in fertility, reduced infant mortality, prolonged life expectancy, and a positive surplus in external migration. However, with an aggravation of the socio-political situation and the deployment of military actions (2014-2017), the indicators of natural and migration increase and the process of depopulation accelerated (during 2013-2017 the population decreased annually by an average of 0.58% against -0.51% in 2005-2012). In total, over the years of independence, the population of Ukraine has decreased by 7.3 million persons, also taking into account the loss of residents of Crimea - by 9.2 million persons. As a result of the annexation of Crimea and escalation of the armed conflict in the Donbas region, Ukraine has suffered significant human losses. According to UN data, from April 2014 to August 2017, the total number of deaths (military men on both sides and civilians) reached 10 225 and 24 541 wounded (*Report on the Human Rights*..., 2017).

Formulation of the problem**.** The estimation of the extent of direct demographic losses can be partially done based on current estimation and completely enough during census. Unfortunately, Ukraine has missed the preliminary world round of censuses in 2010-2011; and since the last all-Ukrainian census of 2001, it has been using the current estimates of natural and migration increase. It is clear that during of war it is not possible to ensure reliability of the current estimation, even in the controlled territory because of the unaccounted actual external migration (emigration) and high mobility of internally displaced persons, as well as incompleteness of the registration of deaths and births in the zone of military actions. Partially this gap should be filled by the data of statistical registers: internally displaced persons and taxpayers. But no statistical survey can determine the amount of indirect losses – the number of unborn children as a result of hostilities. Military mobilization, large-scale emigration, worsening living conditions have exacerbated the current regime of population replacement and its gender

balance, especially in the reproductive age, which has adversely affected the intensity of marriage and childrearing.

Therefore, the problem of estimating indirect demographic losses is relevant; In particular, the question arises about the choice and application of statistical estimation methods: interpolation of the parameters of demographic models (Ediev, 2010; Mesle and Valin et al., 2008), reconstruction of interrupted time series using archival data of censuses and accounting (Duthe et al., 2010; Pechholdova et al., 2017; Rudnytskyi et al., 2015), simulation modeling based on certain assumptions, etc. (Palian, 2016). The results of a comprehensive scientific research of direct and indirect hypothetical losses of the population of Ukraine due to three demographic disasters were first published in monograph (Mesle and Vallin et al., 2008). Well-known Russian researchers have been fruitfully working on statistical assessment of the effects of demographic disasters on the territory of the former USSR (Andreev et al., 2006). Results of a multifaceted scientific work by scientists of the Ptoukha Institute for Demography and Social Studies National Academy of Sciences of Ukraine together with Romaniuk and Gladun (2015), Rudnytskyi et al. (2015) deserve special attention. Due to a thorough archival search, diverse scientific developments performed by O. Rudnyts'kyi, time series of population size and number of births and deaths in Ukraine (in its current borders) between 1850 and 2013 were reconstructed. Thus, today we have an idea of the size of losses of the Ukrainian population as a result of social cataclysms of the last century (Rudnytskyi, et al., 2015). It is extremely important for Ukraine to know about demographic consequences of the current armed conflict not only from the point of view of population forecast, but also for the formation of an adequate state demographic policy. Indirect demographic losses as a result of military conflicts affect both natural population increase and its sex-age structure, in particular, the deepening of aging. The topic of impact of military events on demographic losses, infant mortality and deficiency of births was considered in works (Duthe et al., 2010; Krimer, 2015; Palian, 2016). In particulary, an analysis of changes in childbearing behavior of women during and after armed conflicts in the countries of the Balkan Peninsula with reflection to Ukraine (Krimer, 2015) was made, and an attempt to estimate the size of direct and indirect losses of Ukrainian population as a result of the annexation of Crimea was done by means of simulation, cohort component method and side calculations (Palian, 2016).

From the beginning of the socio-political crisis and Russian military aggression, the birth rate between 2013 and 2016 dropped by an average of 2.5% annually and the number of births – by 35.5 thousand children, including the results of complete alienation of the Crimean territory and parts of Donetsk and Luhansk regions.

The purpose of the article is to assess indirect demographic losses of Ukraine due to the annexation of Crimea and escalation of the armed conflict in the Donbas region.

Methodology. To substantiate the model describing the trend of live births in Ukraine, we will consider quarterly data for 2003-2016 (Fig.1).

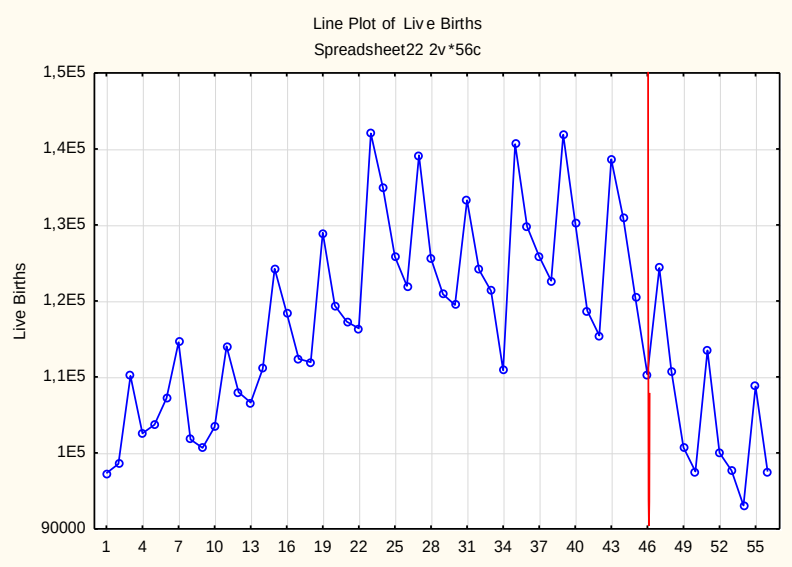


Fig. 1. Dynamics of the quarterly of live births in Ukraine by 2003-2016.

Source: calculated by the data of the State Statistics Service of Ukraine.

It is obvious that the dynamics reveals two peculiarities of the process of childbearing: 1) significant seasonal fluctuations; 2) breaking the positive trend through external influence – the intervention that took place since the beginning of the annexation of Crimea and the armed conflict in the Donbas (in the second quarter of 2014). For analyzing and forecasting processes related to seasonal fluctuations and the broken trend, the Autoregressive Integrated Moving-Average (ARIMA) model is most effective. ARIMA is formed through a combination of autoregressive models of order p and a moving average of order q. The non-stationary series reduces to a stationary type with the help of the operator of finite differences d. The ARIMA model (p, d, q) is quite flexible and describes a wide range of processes. In the presence of

seasonal fluctuations, the model takes into account their periodicity with the lag *s*, which is equal to the seasonal cycle and the analogous content of seasonal parameters (*P, D, Q*) *s*. Therefore, the full seasonal ARIMA model has the form of *ARIMA (p, d, q) (P, D, Q) s.* Depending on the type of intervention, to assess its consequences corresponding additional parameters are introduced to the ARIMA model. In the integrated STATISTICA system, three types of intervention are distinguished: (1) Abrupt, Permanent; (2) Gradual, Permanent; (3) Abrupt, temporary. The type of intervention and values of the parameters of the autoregression *p*, the moving average *q* and the operator *d* are determined at the identification stage of the model.

Principles and concepts of the Interrupted time-series analysis method are elaborately described by McDowall et al. (1980). Numerous scientific studies confirm the analytical capabilities of Interrupted ARIMA for measuring effects the effects of different types of interventions (see more: Linden, 2015; Norpoth, 1981). In the context of analysis of consequences of military conflicts, the study of Norpoth (1987) deserves attention; in it, the Interrupted ARIMA is used to assess the impact of the 1982 Falkland War and the change in macroeconomic policies on the popularity of Thatcher's government and the Conservative Party. The Interrupted time-series analysis techniques tend to improve over time. Thus, Linden (2017) suggests additional Interrupted ARIMA procedures allowing assessment of the effects of interventions on several groups and compares the effects of different interventions.

2 Estimation and forecasting of live births based on the Interrupted ARIMA

Adequacy of the ARIMA model to the real processes and its predictive properties depend on the parameters *p, d, q.* Analysis of correlograms gives grounds to identify the model of the number of births in Ukraine as ARIMA$(0,1,1)(0,0,1)_4$. The filtration of the trend is performed by an operator of differences of the first-order ($d = 1$), the type of the model is an exponential average ($q = 1$). The model takes into account the seasonal nature of fertility. Analysis of the graphically presented dynamics of live births in Ukraine (Fig.1) shows a sharp shift in the positive trend in point 46: the time series is rapidly breaking to a certain level, without changing the seasonal mode of fluctuations. This is a result of the escalation of the armed conflict in the 2nd quarter of 2014. In this regard, the Interrupted ARIMA model was used to predict the live births: the intervention number is 46, the type of intervention is Abrupt, Permanent. The parameters of the Interrupted ARIMA model $(0,1,1)(0,0,1)_4$ provided in Table 1 give grounds to conclude that the

model is adequate to the actual process. Thus, the parameters of smoothing of the series (q(1) = 0.66 and $Q_4(1)$ = -0.57) turned out significant with a probability of 0.999; the parameter of influence of intervention ω with the probability of 0.913 indicates that due to the annexation of Crimea and the armed conflict in the Donbas, the live births in Ukraine per quarter decreased by 9749 persons. Adequacy of the model is also confirmed by results of the analysis of residues.

Table 1. Parameters of the Interrupted ARIMA model $(0,1,1)(0,0,1)_4$ for the time series of live births in Ukraine in 2003-2016.

Input: Live Births								
Transformation: D(1) (Interrupted ARIMA)								
Model (0, 1, 1)(0,0,1) Seasonal lag: 4 MS Residual= 5466E4								
Paramet.	**Param.**	**Asympt. Std. Err.**	**Asympt. t(52)**	**p**	**Lower 95% Conf**	**Upper 95% Conf.**	**Interv. Case No.**	**Interv. Type**
q(1)	0.66	0.132	4.998	0.000	0.40	0.92		
$Q_s(1)$	-0.57	0.099	-5.804	0.000	-0.80	-0.38		
Omega	-9749	5587	-1745	0.087	-20961	-1462	46	Abr/Perm

Table 2. Estimated values of the live births in Ukraine in 2016 and 2017.

Forecasts: Model (0,1,1)(0,0,1) 1 interventions Input Live Births						
Start of origin: 1 End of origin: 52						
Case #	**Forecast**	**Lower 95%**	**Upper 95%**	**Std. Err.**	**Observed**	**Residual**
53	97939	85168	110711	7626	97678	-261
54	95270	81780	108760	8055	93076	-2194
55	103606	89434	117778	8462	108743	5137
56	99539	84716	114361	8851	97542	-1997
57	101736	88940	114411	7604		
58	98633	84780	112486	8272		
59	102762	87874	117650	8890		
60	98556	82721	114430	9467		

Defined on the basis of Interrupted ARIMA $(0,1,1)(0,0,1)_4$, the forecasted values of the number of births in Ukraine for 2016 and 2017 are given in Table 2. The predicted properties of the model are confirmed by a retrospective assessment of the expected number of births in 2016. Relative errors of quarterly forecasts by modulo: for the first quarter – 0.27%; for the second – 2.35%; for the third – 4.72%; for the fourth – 2.04%.

Thus, in a short-term perspective, a slight decrease of the live births is expected. A logical question arises as to how corresponding is the dynamics of the two interrelated processes – childbearing and marriage?

3 Specificity of marriage trends

Despite the current global trends in the formation of marriage and family relations, the complex and unstable socio-economic situation, in Ukraine the preferences of partners for marriage and family are preserved. The main form of marital relations remains the officially registered marriage. As shown on Fig. 2, the dynamics of the number of registered marriages with internally annual minimum and maximum values for 2003-2016 does not show a clear tendency to decrease. At the same time, there is an effect of people's commitment to ancient folk traditions: in leap years (2008, 2012, 2016) that are considered unfavorable for marriage, their number decreases, while in years before the leap year - on the contrary, increases. It is noteworthy that recently (from 2012) there is a contraction of internal annual differences between the maximum and minimum number of registered marriages. That is, the monthly seasonality of marriages is less expressive, and the dependence of motivation of marital behavior on old traditions diminishes.

The armed conflict in the Donbas region and economic difficulties became a major challenge for Ukrainian families: the numbers of widows and orphans increases, family ties weaken due to forced migration (internally displacing from the combat zone, emigration). At the same time, in conditions of instability and unpredictability of the future, it is family solidarity and mutual assistance that allows maintaining life activity, and sometimes even survival. Many pairs want to register their marital relationships, give birth to children. If there were 31200 marriages per year by three pre-war years, but over a period of the escalation of the armed conflict (2014-2016), the average annual number of registered marriages was 27500. Therefore,

even if part of the population remained in the annexed and occupied territory of Ukraine, the number of marriages decreased insignificantly.

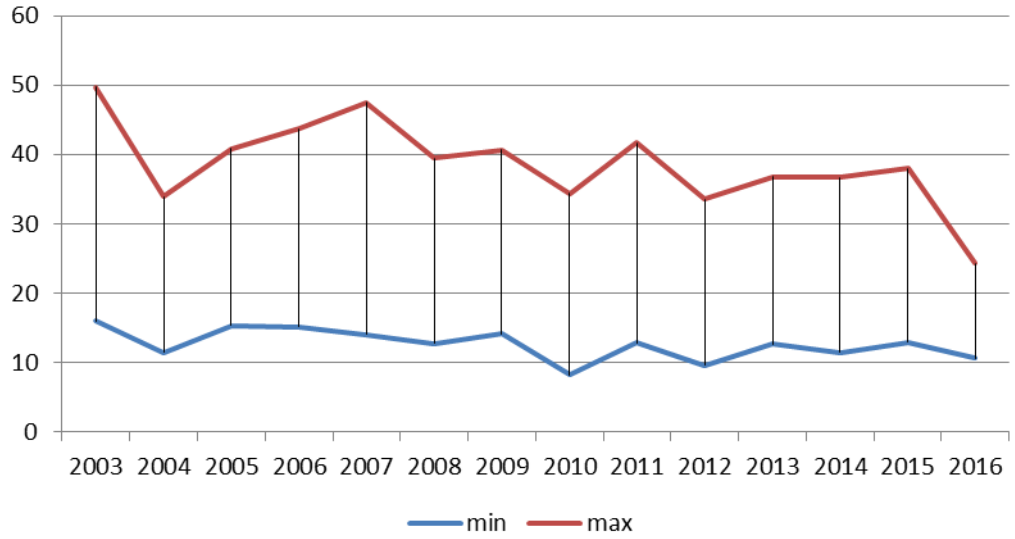


Fig. 2. Number of registered marriages in Ukraine by 2003-2016, 000'.
Source: calculated by the of the SSSU.

The specificity of the modern marriage process in Ukraine is also evidenced by the correlogram: 1) absence of the trend confirms the non-essential nature of the first-order auto-correlation coefficient r_1; 2) statistically significant coefficients of autocorrelation r_k, which at k = 2, 4, 6, 8, 12, alternately change the sign, indicate an intra-annual cyclicity when peaks and recessions of marriage registration are alternating (Fig. 3). Thus, marriages registration during the period of escalation of the armed conflict didn't suffer such significant changes that happened to the number of births. On the other hand, the relative stability of the time series of registered marriages is an important prerequisite for an expected equalization of the dynamics of childbearing. It is obvious, however, that responsible parenting requires a stable internal political and socio-economic situation. As long as the undeclared war on the part of Russia continues in Ukraine, moderate state demographic, environmental, health and sanitary, as well as economic policies should play an important role in the country. This also implies social support of young families, including those with children, as well as employment for young people and middle-

aged people at the regional level, regardless of the status of the territory: whether it is peaceful, de-occupied or a buffer zone.

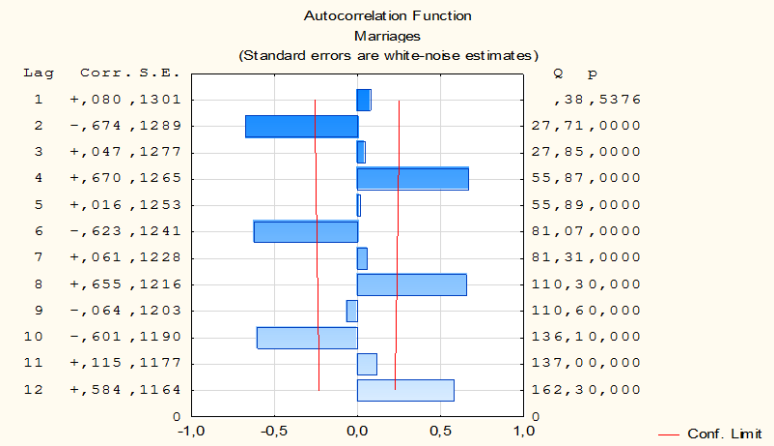


Fig. 3. Correlogram of the number of registered marriages.

Source: calculated by the data of the SSSU.

Conclusions

An armed conflict inevitably leads to human losses, and in the context of modern wars on the post-Soviet territories (Moldova, Armenia, Georgia and Ukraine), it also means losses of the population due to violation of integrity of the countries. Statistical estimation the size of direct and indirect demographic losses is complicated by incomplete registration, territorial incomparability of data and fragmentation of time series. Application of adequate modeling methods can find the effect of occupation the part Ukrainian's territory on indirect population losses. Further research in this area may focus on the estimation of expected indirect population losses as a result of partial deformation the sex-age structure of the reproductive contingent due to the armed conflict in country.

References

Andreev, E., Bohoiavlenskyj, D., Vyshnevskyj, A., Zakharov, S. & Khar'kova, T. (2006). Demographics catastrophes of XX century. In: Vyshnevskyi, A. H. (ed.), *Demographic modernization of Russia, 1900-2000*. Moscow: Novoe yzdatel'stvo, 339 - 447.

Ediev, D. (2010). On the reproductive value and the spectrum of a population projection matrix with implications for dynamic population models. *Theoretical population biology*, *78* (2), 67-70.

Krimer, B. (2015). Effect of recent decades wars on fertility in Europe. *Demography and social economy*, *1*(23), 126-136.

Linden, A. (2015). Conducting interrupted time-series analysis for single- and multiple-group comparisons. *Stata Journal*, *15*(2), 480-500.

Linden, A. (2017). A comprehensive set of post estimation measures to enrich interrupted time-series analysis. *Stata Journal*, *17*(1), 73-88.

McDowall, D., McCleary, R., Meidinger, E. E. & Hay, R. A. (1980). *Interrupted Time Series Analysis*. SAGE University Paper, California.

Mesle, F., Vallin, J., Shkolnikov, V., Pyroshkov, S. & Adamets, S. (2008). *Mortality and deaths by causes in Ukraine on XXth century*. Kyiv: VD Stylos.

Norpoth, H. (1981). Interrupted time-series analysis - Mcdowall, D., Mccleary, R., Meidinger, Ee., Hay, Ra. *Quality & Quantity*, *15*(5), 473-474.

Norpoth, H. (1987). Guns and Butter and Government Popularity in Britain. *The American Political Science Review*, *81*(3), 949-959.

Palian, Z. (2016). Statistical estimation of modern demographic losses in Ukraine. *Statistics of Ukraine*, *1*(72), 34–41.

Pechholdova, M., Camarda, C. G., Mesle, F. & Vallin, J. (2017). Reconstructing Long-Term Coherent Cause-of-Death Series, a Necessary Step for Analyzing Trends. *European Journal of Population*, *33*(5), 629-650.

Romaniuk, A. & Gladun, O. (2015). Demographic trends in Ukraine: Past, Present and Future. *Population and Development Review*, *41*(2), 315-337.

Rudnytskyi, O., Levchuk, N., Wolowyna, O., Shevchuk, P. & Kovbasiuk, A. (2015). Demography of a man-made human catastrophe: the case of massive famine in Ukraine 1932–1933. *Canadian Studies in Population*, *42* (1–2), 53–80.

Report on the Human Rights Situation in Ukraine. Retrieved May 16 to August 15, 2017, from https://www.humanitarianresponse.info/system/files/documents/files/20170912_ukraine_-_19th_hrmmu_report_eng_0.pdf.

Estimation of quantile ratios of the Dagum distribution

Wojciech Zieliński[1], Alina Jędrzejczak[2], Dorota Pekasiewicz[3]

Abstract

Inequality measures based on ratios of quantiles are frequently applied in economic research, especially to the analysis of income distributions in different divisions. Simple quantile ratios or quantile dispersion ratios, which can be considered supplementary to the popular Gini and Zenga indices, are often applied in comparison of incomes for various subpopulations or to assess inequality changes over time. They have the advantage of being focused on extremal income groups that are especially interesting from the point of view of economic inequality and polarization. In the paper a confidence interval for such measures, assuming the Dagum distribution, is constructed. The ends of the confidence interval depend on an unknown shape parameter of the underlying income distribution model. In applications this parameter must be estimated from the data. The constructed confidence interval, applied to decile and quintile ratios, was implemented to the analysis of income inequaity in Poland. The quantile-based inequality measures have been estimated for the Polish macro-regions (NUTS1) and for the whole country, on the basis of micro-data coming from the Household Budget Survey 2015.

Keywords: *Dagum distribution, income inequality, quantile ratios*

JEL Classification: C13, C18

DOI: 10.14659/SEMF.2018.01.61

1 Introduction

Distribution quantiles of a random variable X or estimators of these quantiles have frequently been applied to the construction of numerous inequality indices and indicators. Among them the most popular are quintile and decile share ratios (see e.g.: Panek, 2011) mainly for their simplicity and straightforward economic interpretation The income quintile share ratio is calculated as the ratio of income received by the 20% of the population with the highest income to that received by the 20% of the population with the lowest income or as the ratio of the top quintile to the bottom quintile. Similarly we can define decile share ratios. More sophisticated measures of income inequality have been constructed using ratios (or differences) between population and income quantiles. Probably the first of such measures was the Holme's coefficient standardized by Bortkiewicz, which is based on the quantiles of

[1]Corresponding author: Warsaw University of Life Sciences, Department of Econometrics and Statistics, Nowoursynowska 159 St., 02-776 Warsaw, Poland, wojciech_zielinski@sggw.pl

[2] University of Łódź, Institute of Statistics and Demography, Department of Statistical Methods, 41/43 Rewolucji 1905 St., 90-214 Łódź, Poland, jedrzej@uni.lodz.pl.

[3] University of Łódź, Institute of Statistics and Demography, Department of Statistical Methods, 41/43 Rewolucji 1905 St., 90-214 Łódź, Poland, pekasiewicz@uni.lodz.pl.

order 0.5. The concentration curve and corresponding synthetic concentration coefficient proposed by Zenga are also defined in terms of quantiles of a size distribution and the corresponding quantiles of the first-moment distribution. The quantile based inequality measures were found more sensitive to changes of income inequality in particular parts of income distribution, especially in the tails.

Quantile-based inequality measures are traditionally estimated using the classical quantile estimator based on a relevant order statistic. Their estimators can also be obtained in the parametric approach based on a theoretical income distribution model. The parametric estimators of quantile ratios which are based on income distribution models with sensible stochastic or empirical foundations, have the advantage of being robust to irregularities coming from imperfect data collection methods. Moreover, based on these models, the confidence intervals can be derived to provide information about how close the point estimate is to the true parameter with the margin of error. The main objective of this study was to provide a confidence interval of quantile ratio for the three-parameter Dagum distribution.

The second section of this paper is devoted to the non-parametric point estimators of quantile ratios. The third part introduces a parametric estimator of quantile ratios based on the Dagum distribution and finally the confidence interval for this parameter is derived. In the last part of the paper we present the application of the proposed estimation methods to income inequality analysis based on the Polish Household Budget Survey (HBS) data.

2 Point estimators of quantile ratios

Let X be a continuous random variable with a distribution function F and let $\gamma_p = F^{-1}(p)$ be the p-quantile of the random variable X, where $p \in (0,1)$. If F is continuous and strictly increasing distribution function, the p^{th} quantile is uniquely determined.

Among estimators of quantiles γ_p we can distinguish the standard estimator, Huang-Brill estimator, Harrel-Davis estimator and Bernstein estimator, to name only a few (Huang and Brill, 1999; Harrell and Davis, 1982; Zieliński, 2006).

An application of these estimators to the evaluation of quantiles and quantile ratios has recently been presented in Jędrzejczak and Pekasiewicz (2017).

In what follows we apply the well-known estimator of the quantile γ_p:

$$\hat{\gamma}_p = X_{\lfloor pn \rfloor+1:n}, \tag{1}$$

where $X_{1:n} \le X_{2:n} \le ... \le X_{n:n}$ is the ordered sample of a random sample $X_1, X_2, ..., X_n$ and $\lfloor x \rfloor$ denotes the greatest integer not greater than *x*.

In this case the estimator of the quantile ratio $r_{\alpha,\beta} = \frac{\gamma_\beta}{\gamma_\alpha}$, where $0 < \alpha < \beta < 1$, has the following form:

$$\hat{r}_{\alpha,\beta} = \frac{X_{\lfloor \beta n \rfloor+1:n}}{X_{\lfloor \alpha n \rfloor+1:n}}. \tag{2}$$

3 Point and interval estimators of quantile ratios for the Dagum distribution

In our considerations we confine ourselves to the Dagum distribution, i.e. throughout the paper it will be assumed that the distribution of the population income is the Dagum one. As it was mentioned above, the Dagum distribution fits population income quite well for many countries all over the world. It is based on both empirical and stochastic foundations, similarly to the Pareto model (Dagum, 1977).

The probability density function of the Dagum distribution is given by (Kleiber and Kotz, 2003):

$$f_{a,\nu,\lambda}(x) = \frac{a\nu}{\lambda}\left(\frac{x}{\lambda}\right)^{a\nu-1}\left(1+\left(\frac{x}{\lambda}\right)^{\nu}\right)^{-a-1} \quad \text{for } x > 0, \tag{3}$$

where $a, \nu, \lambda > 0$. Its cumulative distribution function equals:

$$F_{a,\nu,\lambda}(x) = \left(1+\left(\frac{x}{\lambda}\right)^{-\nu}\right)^{-a} \quad \text{for } x > 0 \tag{4}$$

and the quantile function is

$$Q_{a,\nu,\lambda}(q) = \lambda\left(q^{-\frac{1}{a}} - 1\right)^{-\frac{1}{\nu}} \quad \text{for } 0 < q < 1, \tag{5}$$

so the ratio of quantiles of the Dagum distribution has the following form:

$$r_{\alpha,\beta} = \frac{\gamma_\beta}{\gamma_\alpha} = \left(\frac{\beta^{-\frac{1}{a}} - 1}{\alpha^{-\frac{1}{a}} - 1}\right)^{-\frac{1}{\nu}} \quad \text{for } 0 < \alpha < \beta < 1. \tag{6}$$

The problem lies in constructing a confidence interval at the confidence level δ for a ratio of quantiles $r_{\alpha,\beta}$ based on $\hat{r}_{\alpha,\beta}$.

For "large" sample sizes, i.e. when $n \to \infty$, it is known that $\hat{r}_{\alpha,\beta}$, defined by the formula (2), is a consistent estimator of $r_{\alpha,\beta}$ (David and Nagaraja, 2003; Serfling, 1980).

Let $Y_i = \ln X_i$ and γ_α^Y, γ_β^Y denote the quantiles of *Y*. We have (David and Nagaraja, 2003; Serfling, 1980):

$$\sqrt{n}\begin{bmatrix} Y_{\lfloor n\alpha \rfloor+1:n} - \gamma_\alpha^Y \\ Y_{\lfloor n\beta \rfloor+1:n} - \gamma_\beta^Y \end{bmatrix} \to N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \dfrac{\alpha(1-\alpha)}{\left(f_Y\left(\gamma_\alpha^Y\right)\right)^2} & \dfrac{\alpha(1-\beta)}{f_Y\left(\gamma_\alpha^Y\right) f_Y\left(\gamma_\beta^Y\right)} \\ \dfrac{\alpha(1-\beta)}{f_Y\left(\gamma_\alpha^Y\right) f_Y\left(\gamma_\beta^Y\right)} & \dfrac{\beta(1-\beta)}{\left(f_Y\left(\gamma_\beta^Y\right)\right)^2} \end{bmatrix} \right), \tag{7}$$

where $f_Y(\cdot)$ is the probability density function of *Y*. Hence

$$\sqrt{n}\left[\left(Y_{\lfloor n\beta \rfloor+1:n} - Y_{\lfloor n\alpha \rfloor+1:n}\right) - \left(\gamma_\beta^Y - \gamma_\alpha^Y\right)\right] \to N\left(0, \sigma^2\right), \tag{8}$$

where

$$\sigma^2 = \frac{\beta(1-\beta)}{\left(f_Y\left(\gamma_\beta^Y\right)\right)^2} + \frac{\alpha(1-\alpha)}{\left(f_Y\left(\gamma_\alpha^Y\right)\right)^2} - 2\frac{\alpha(1-\beta)}{f_Y\left(\gamma_\alpha^Y\right) f_Y\left(\gamma_\beta^Y\right)}. \tag{9}$$

Since $Y_{\lfloor np \rfloor+1:n} = \ln X_{\lfloor np \rfloor+1:n}$ and in Dagum distribution $\gamma_p^Y = \ln \gamma_p$ we have:

$$\sqrt{n}\left(\ln \frac{X_{\lfloor n\beta \rfloor+1:n}}{X_{\lfloor n\alpha \rfloor+1:n}} - \ln \frac{\gamma_\beta}{\gamma_\alpha} \right) \to N\left(0, \sigma^2\right) \tag{10}$$

and applying Delta method (Greene, 2003) with $g(t) = e^t$ we obtain:

$$\sqrt{n}\left(\frac{X_{\lfloor n\beta \rfloor+1:n}}{X_{\lfloor n\alpha \rfloor+1:n}} - \frac{\gamma_\beta}{\gamma_\alpha} \right) \to \left(\frac{\gamma_\beta}{\gamma_\alpha} \right) N\left(0, \sigma^2\right), \tag{11}$$

where

$$\sigma^2 = \frac{1}{(av)^2}\left(\frac{1-\beta}{\beta} \frac{1}{\left(1-\beta^{\frac{1}{a}}\right)^2} + \frac{1-\alpha}{\alpha} \frac{1}{\left(1-\alpha^{\frac{1}{a}}\right)^2} - 2\frac{1-\beta}{\beta} \frac{1}{\left(1-\alpha^{\frac{1}{a}}\right)\left(1-\beta^{\frac{1}{a}}\right)} \right). \tag{12}$$

The parameter *v* can be determined using the parameter *a* and the quantile ratio $r_{\alpha,\beta}$ of the Dagum distribution in the following way: $v = \ln\left(\dfrac{\alpha^{-\frac{1}{a}} - 1}{\beta^{-\frac{1}{a}} - 1} \right) \left(\ln r_{\alpha,\beta}\right)^{-1}$.

Hence

$$\sigma^2 = \left(\ln r_{\alpha,\beta}\right)^2 \omega^2(a), \tag{13}$$

where

$$\omega^2(a)=\left(a\ln\left(\frac{\alpha^{-\frac{1}{a}}-1}{\beta^{-\frac{1}{a}}-1}\right)\right)^{-2}\left(\frac{1-\beta}{\beta}\frac{1}{\left(1-\beta^{\frac{1}{a}}\right)^2}+\frac{1-\alpha}{\alpha}\frac{1}{\left(1-\alpha^{\frac{1}{a}}\right)^2}-2\frac{1-\beta}{\beta}\frac{1}{\left(1-\alpha^{\frac{1}{a}}\right)\left(1-\beta^{\frac{1}{a}}\right)}\right). \tag{14}$$

The confidence interval for the quantile ratio of the Dagum distribution has the following form:

$$P_{r,a}\left\{\sqrt{n}\left|\frac{\hat{r}_{\alpha,\beta}-r_{\alpha,\beta}}{\omega(a)r_{\alpha,\beta}\ln r_{\alpha,\beta}}\right|\le u_{(1+\delta)/2}\right\}=\delta, \tag{15}$$

where δ is a given confidence level and $u_{(1+\delta)/2}$ is the quantile of N(0,1) distribution. Solving the above inequality with respect to $r_{\alpha,\beta}$ we obtain:

$$\left(\frac{\hat{r}_{\alpha,\beta}z_+(a)}{W\left(\hat{r}_{\alpha,\beta}z_+(a)e^{z_+(a)}\right)},\frac{\hat{r}_{\alpha,\beta}z_-(a)}{W\left(\hat{r}_{\alpha,\beta}z_-(a)e^{z_-(a)}\right)}\right), \tag{16}$$

where $z_+(a)=\frac{\sqrt{n}}{u_{(1+\delta)/2}\omega(a)}$, $z_-(a)=\frac{\sqrt{n}}{u_{(1-\delta)/2}\omega(a)}$ and W is the W Lambert function.

Note that the ends of the confidence interval depend on an unknown shape parameter *a*. This parameter should be estimated from the data. As the estimation techniques we can choose for example: Maximum Likelihood Method, Method of Moments or L-moments, Methods of Ordinary Least-Squares or Weighted Least-Squares and the methods based on percentiles (Dey et al., 2017).

4 Application of quantile ratios of the Dagum distribution to the Polish income data

The inequality measures based on deciles and quintiles have been applied to income inequality analysis in Poland by macro-region (NUTS1). The calculations were based on the micro data coming from the Household Budget Survey 2015 conducted by the Central Statistical Office of Poland. The variable of interest was household available income, the basic income category of HBS sample.

To adjust the available income for differences in family size, we adopted the recent OECD equivalence scale, where the household income was divided by the square root of relevant household size. Basic characteristics of the HBS sample by macro-region are presented in Table 1.

Table 2 contains the results of fitting the Dagum distribution to the empirical data. In Fig. 1-2 there are histograms accompanied by fitted Dagum density curves describing income distributions in selected macro-regions.

Analysing the results presented in Fig. 1 and 2 one can observe very high consistency of the empirical distributions with the theoretical ones for the selected *Central* and *Southern* macroregions. Similar results were obtained for the other macro-regions. It can also be confirmed by the values of a goodness-of-fit measure (the overlap coefficient) calculated for each region and the whole country and presented in the last column of Table 2. The overlap coefficient also known as the coefficient of distribution similarity was proposed by Vielrose in 1960, and represents the "common part" of the empirical and theoretical distributions.

Table 1. Numerical characteristics of equivalent income in macroregions.

Macroregion	**Number of households**	**Minimum**	**Maximum**	**Average**	**Standard Deviation**
Central	8058	3.78	69047.65	2763.75	2248.36
Southern	7465	88.00	26400.00	2358.24	1224.19
Eastern	6207	1.77	52702.31	2120.09	1502.48
North-western	5608	36.69	105846.64	2378.83	1909.83
South-western	3914	7.85	28394.00	2428.96	1383.25
Northern	5608	10.00	67370.00	2299.07	1826.83
Poland	36860	1.77	105846.64	2408.43	1760.20

Table 2. Approximation of equivalent income distributions for macroregions.

Macroregion	**Dagum distribution parameters**			**Overlap measure**
	a	v	λ	
Central	0.817	3.175	2512.444	0.989
Southern	0.888	3.953	2223.571	0.994
Eastern	0.775	3.677	2041.918	0.993
North-western	0.829	3.871	2260.756	0.996
South-western	0.697	4.051	2494.583	0.995
Northern	0.825	3.527	2138.206	0.993
Poland	0.821	3.588	2261.110	0.996

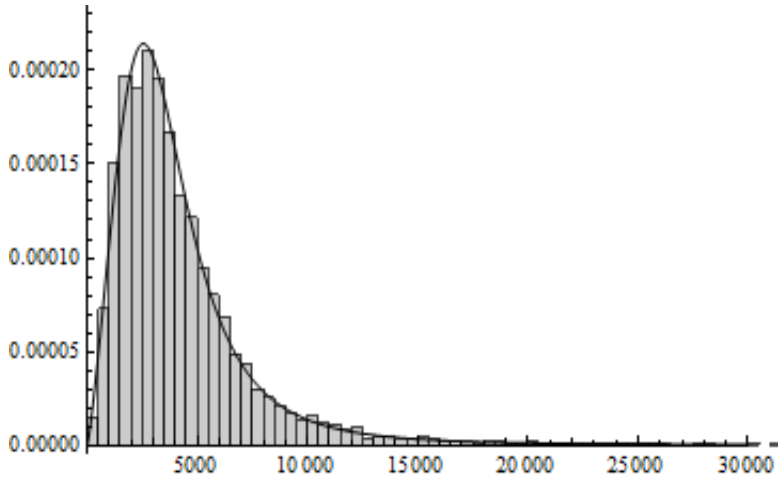


Fig. 1. Equivalent income distribution for Central Macroregion and fitted Dagum density.

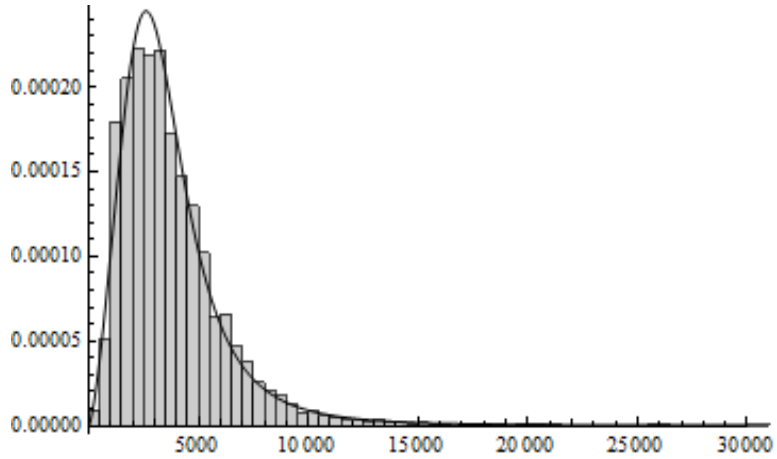


Fig. 2. Income distribution for Southern Macroregion and fitted Dagum density.

Based on the confidence interval for quantile ratios which has been proposed in the section 3 (see: eq. (16)), we constructed the confidence intervals for the income quintile ratio:

$$W_{20:20} = \frac{\gamma_{0.8}}{\gamma_{0.2}}, \tag{17}$$

and the income decile ratio:

$$W_{10:10} = \frac{\gamma_{0.9}}{\gamma_{0.1}}, \tag{18}$$

where $\gamma_{0.8}$, $\gamma_{0.2}$ are quintiles and $\gamma_{0.9}$, $\gamma_{0.1}$ are deciles.

The shape parameter *a* has been estimated by the maximum likelihood method which is the most popular one due to its good asymptotic properties.

The basic results of the inequality analysis have been outlined in Table 3. The estimated values of quintile and decile ratios indicate the Central macroregion as the one with the

highest income inequality level. It is especially visible for the decile share ratio which shows the "social distance" between the rich and the poor of about 4.3 times $\gamma_{0.1}$ (more precisely, it is a number from the interval (4.1673, 4.5685)).

Table 3. Confidence intervals of quintile and decile share ratios.

Macro-region	$\hat{W}_{20:20}$	Confidence Interval of $W_{20:20}$	$\hat{W}_{10:10}$	Confidence Interval of $W_{10:10}$
Central	2.4787	(2.4252, 2.5376)	4.3039	(4.1673, 4.5685)
Southern	2.1319	(2.0917, 2.1760)	3.2186	(3.1324, 3.3143)
Eastern	2.2615	(2.2115, 2.3171)	3.5354	(3.4243, 3.6606)
North-western	2.1751	(2.1282, 2.2270)	3.2982	(3.1985, 3.4104)
South-western	2.2010	(2.1433, 2.2665)	3.4080	(3.2803, 3.5557)
Northern	2.3209	(2.2690, 2.3811)	3.7146	(3.5914, 3.8541)
Poland	2.2863	(2.2646, 2.3089)	3.6293	(3.5800, 3.6811)

5 Conclusions

In the paper an asymptotic confidence interval for a ratio of quantiles of the Dagum distribution was constructed. The ends of this confidence interval depend on the shape parameter *a* of the Dagum distribution. This parameter can be estimated by the maximum likelihood method, which was shown to be efficient for large sample sizes.

The proposed method of the estimation of quantile ratios can be applied to income inequality analysis when a household or personal income follows the Dagum distribution. The confidence interval constructed above is symmetrical in the following sense: the risks of underestimation as well as overestimation of the true population parameter are the same. It can be a useful tool for social-policy makers interested in the evaluation of current income inequality level with the acceptable margin of error.

Income quintile and decile ratios considered in the paper have the advantage of being focused on extremal income groups that are especially interesting from the point of view of economic inequality and polarization. They can assess the "social distance" between rich and poor groups of income receivers. The empirical analysis revealed substantial discrepancies between extremal income groups in the macro-regions of Poland. Income distribution in the most affluent *central* region turned out to be the most unequal what is especially visible in the

extremal decile groups. All the distributions under consideration presented very high consistency with the Dagum model.

References

Dagum, C. (1977). A New Model of Personal Income Distribution: Specification and Estimation. *Economie Appliquee, 30*, 413-437.

David, H. A. & Nagaraja, H. N. (2003). *Order Statistics, Third Edition*. John Wiley & Sons.

Dey, S., Al-Zahrani, B. & Basloom, S. (2017). Dagum Distribution: Properties and Different Methods of Estimation. *International Journal of Statistics and Probability, 6*, 74-92.

Greene, W. H. (2003). *Econometric Analysis (5th ed.)*. Prentice Hall.

Harrell, F. E. & Davis C. E. (1982). A New Distribution-Free Quantile Estimator. *Biometrika, 69*, 635–640.

Huang M. L. & Brill P. H. (1999). A Level Crossing Quantile Estimation Method. *Statistics & Probability Letters, 45*, 111–119.

Jędrzejczak, A. & Pekasiewicz, D. (2017). Analysis of the properties of selected inequality measures based on quantiles with the application to the Polish income data. In: Papież, M. & and Śmiech, S. *11th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena. Conference Proceedings. Cracow: Foundation of the Cracow*, 113-122.

Kleiber, C. & Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*. John Wiley & Sons.

Panek, T. (2011). *Ubóstwo, wykluczenie społeczne i nierówności. Teoria i praktyka pomiaru*. Szkoła Główna Handlowa, Warszawa.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons.

Vielrose E. (1960). *Rozkład dochodów według wielkości,* Polskie Wydawnictwo Gospodarcze, Warszawa.

Zieliński, R. (2006). Small-Sample Quantile Estimators in a Large Nonparametric Model. *Communications in Statistics Theory and Methods*, 35, 1223–1241.