CRACOW UNIVERSITY OF ECONOMICS

Quantitative methods in the contemporary issues of economics

Edited by Beata Ciałowicz

- Merger of Populations and Aggregate Relative Deprivation
- Convergence of Conflict Sets and Applications
- The Angle between the 2-dimensional Linear Regression Model Lines
- Analysis of Football Players' Labor Market Migrations Using Panel Gravity Models
- Selected Credit Risk Models
- Optimal Path in Growth Model
- Predictive Power Comparison of Bayesian Homoscedastic vs. Markov-switching Heteroscedastic VEC Models
- MCMC Method for the IG-MSF-SBEKK Model
- Indebted Households' Self-assessment of their Financial Situation: Evidence from Poland

edu-Libri

print pdf

Quantitative methods in the contemporary issues of economics

CRACOW UNIVERSITY OF ECONOMICS

Quantitative methods in the contemporary issues of economics

Edited by Beata Ciałowicz

edu-Libri

Kraków–Legionowo 2020

Edition and correction: Dominika Drygas
Cover design by: GRAFOS

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

List of reviewers:
Jan Acedański, *University of Economics in Katowice*, **Poland**
Łukasz Balbus, *University of Zielona Góra*, **Poland**
Wojciech Gamrot, *University of Economics in Katowice*, **Poland**
Alicja Ganczarek-Gamrot, *University of Economics in Katowice*, **Poland**
Stanisław Heilpern, *Wroclaw University of Economics*, **Poland**
Beata Kocel-Cynk, *Cracow University of Technology*, **Poland**
Grzegorz Kończak, *University of Economics in Katowice*, **Poland**
Robert Kruszewski, *SGH Warsaw School of Economics*, **Poland**
Łukasz Kwiatkowski, *Cracow University of Economics*, **Poland**
Agnieszka Lipieta, *Cracow University of Economics*, **Poland**
Anna Pajor, *Cracow University of Economics*, **Poland**
Stanisław Wanat, *Cracow University of Economics*, **Poland**

Suggested citation:
AUTHOR A. (2020). Title of the paper. In: B. Cialowicz (ed.), *Quantitative methods in the contemporary issues of economics*. **Krakow: Wydawnictwo edu-Libri s.c., pp.xx-xx. ISBN: 978-83-66395-01-5**
The editor is not responsible for the language used in papers.

Scientific monograph was financed from means of Cracow University of Economics

Publishing house: edu-Libri s.c.
ul. Zalesie 15, 30-384 Kraków
e-mail: edu-libri@edu-libri.pl

DTP: GRAFOS
Printing and binding: OSDW Azymut Sp. z o.o.
Łódź ul. Senatorska 31

ISBN (print) 978-83-66395-01-5
ISBN e-book (PDF) 978-83-66395-02-2

Table of contents

Introduction

Beata Ciałowicz[1]

Quantitative methods in economics include mathematical modelling, game theory, optimization techniques, statistical methods and econometrics. The methodological status of these methods in analyzing the issues of economics has been discussed for years. At the most fundamental level quantitative methods are universally and primarily aimed at answering contemporary economic questions at testing economic theories, ideas or hypotheses. These methods help to extend and formalize a broad range of empirical and theoretical problems in economics and influence the development and refinement of formal models in economics. Moreover, quantitative methods increase research efficiency by making it possible to confront theories with empirical data, to apply a formal theory to many different subject matters and to indicate similarities and differences in a comparative analysis of the theories on the same problem.

This monograph presents some interesting applications of quantitative methods in studying the phenomena of economic processes using mathematical knowledge and tools. The area of scientific research is diversified and covers topics relating to macroeconomics, consumer theory, socio-economic development, households quality of life, heteroskedastic Vector Autoregressive models and credit risk models.

In the first chapter by Jakub Bielawski, *Merger of populations and aggregate relative deprivation*, the problem of merging populations and its consequences is analyzed. In particular, this chapter shows that in some situations it is sufficient to change the weight the individuals attach to the comparison with the richest individuals in the population to obtain that the social stress decreases after the merger.

The following chapter by Anna Denkowska entitled *Convergence of conflict sets and applications* presents some new results concerning the semicontinuity of the

[1] Beata Ciałowicz, Cracow University of Economics, e-mail: beata.cialowicz@uek.krakow.pl

conflict set and discusses them for applications. Specifically, it discusses the conflict sets of a finite family of pairwise disjoint closed subsets of the Euclidean space evolving in time.

The chapter by Michał Górnik, *Analysis of football players labor market migrations using panel gravity models*, gives an example of using panel gravity models for estimating the size of players movement between pairs of countries and verifies hypotheses of the impact of the sport level difference between leagues on the number of transfers as well as the correlation between the overall country economy and its top-tier league capability to attract football players.

Albert Gardoń, in his contribution *The Angle between the 2-dimensional Linear Regression Model Lines*, analyzes an impact of the goodness-of-fit and the ratio of sample variances on the angle between the linear regression lines. This chapter shows that the angle depends not only on the correlation between features but also on the ratio of their sample dispersions.

The next chapter by Stanisław Heilpern, *Selected credit risk models*, is devoted to credit risk in two kinds of models based on the generalized binomial distributions. Firstly, the dependent credit risk, using copulas, mainly Archimedean is investigated. Secondly, a case with the uncertain probability of the insolvent obligors is studied.

Marta Kornafel in her contribution *Optimal path in growth model*, considers the general Ramsey-Koopmans-Cass growth model where some of the parameters depend on time. This work is focused on the dependence of the model and its solution to the perturbance of parameters.

The following chapter by Łukasz Kwiatkowski, *Predictive power comparison of Bayesian homoscedastic vs Markov-switching heteroskedastic VEC models,* develops a framework for modelling the forecasting performance of Bayesian vector error correction models featuring two- and three-state Markovian breaks in the conditional covariance matrix to capture time-varying volatility, typically recognized in macroeconomic data.

Anna Pajor, in her contribution *MCMC method for the IG-MSF-SBEKK model* proposes specific numerical method applied to estimate the hybrid IG-MSF-SBEKK model for daily exchange rate returns. In this method a Markov chain Monte Carlo simulation tool is adapted to obtain a sample from the posterior distribution of parameters and latent variables.

The last chapter by Agnieszka Wałęga, *Indebted households' self-assessment of their financial situation: evidence from Poland,* stresses the importance of the level of debt and over-indebtedness risk for self-assessment financial situation of the household. To examine the relationship between the respondents' self-assessment of their financial situation and commonly used objective measures of over-indebtedness, the ordered probit model was used.

All these research findings were planned to be presented in April 2020, postponed due to the pandemic of COVID-19, at the 56$^{\text{th}}$ Conference of Statisticians, Econometricians and Mathematicians of South Poland (SEMPP 2020 conference). The SEMPP conference is one of the oldest scientific conferences in the field of economic sciences in Poland. It was first held in Katowice in 1965, on the initiative of Professor Zdzisław Hellwig, Professor Zbigniew Pawłowski and Professor Kazimierz Zając. It takes place continuously every year and its organizers are alternately economic universities in Katowice, Krakow and Wroclaw. The main aim of this conference is to present the scientific achievements of the employees of economic universities of South Poland in the field of statistics, econometrics and mathematics and their applications in various fields of science including economics, finance and management. Moreover, it gives the opportunity to start and strengthen cooperation between the research centers from Katowice, Krakow and Wroclaw.

Merger of Populations and Aggregate Relative Deprivation

Jakub Bielawski[1]

Abstract

The problem of merging populations and its consequences was widely studied by economists. Stark [Stark, 2013] showed that a merger of populations will always impose an increase in the discomfort of individuals resulting from comparing their incomes with incomes of other members of the group. More precisely, the aggregate relative deprivation of the merged population mustn't be less than the sum of aggregate relative deprivation of populations prior to the merger. In other words, the aggregate relative deprivation is superadditive.

Our major question is whether the superadditivity result is independent of the choice of the relative deprivation measure. We show that in some situations it is sufficient to change the weight the individuals attach to the comparison with the richest individuals in the population to obtain that the social stress decreases after the merger.

Keywords: merger of populations, aggregate relative deprivation, income distribution, social distress

JEL classification: D31, D63, I31

[1] Jakub Bielawski, Cracow University of Economics, e-mail: bielawsj@uek.krakow.pl

1. Introduction

In this paper we study the problem of a merger of two populations, namely we ask how this operation affects the social stress of both groups. We measure social stress by aggregating the levels of relative deprivation experienced by members of these populations.

The relative deprivation of a member of a population can be measured in a variety of ways. One of the most common and natural is the Yitzhaki index defined as the aggregate of the excesses of the incomes of the other individuals in the population divided by the number of individuals in the population (essentially an operationalization of Runciman's 1966 relative deprivation concept by Yitzhaki, 1979, Hey, Lambert, 1980, Chakravarty, 1999, Ebert, Moyes, 2000, Bossert, D'Ambrosio, 2006, Stark, Hyll, 2011). Another method of measuring relative deprivation is the maximal index, where the only comparison with the richest individual matters.

In 2013 Stark showed that merger of populations will always yield an increase in the discomfort of individuals resulting from comparing their incomes with incomes of other members of the group [Stark, 2013]. More precisely, the aggregate relative deprivation of the merged population has to be at least as the sum of aggregate relative deprivation of populations prior to the merger. In other words, the measure of aggregate relative deprivation is superadditive. This fact gives another insight into the tension which individuals experience when their group is combined with another.

The superadditivity result was proved by Stark [2013] for measures of aggregate relative deprivation constructed on the basis of the Yitzhaki index of relative deprivation and for the maximal relative deprivation index. Note that in the Yitzhaki measure an individual compares his income with incomes of all members of his population. In this paper we follow a strand of literature [Ebert, Moyes, 2000, Bossert, D'Ambrosio, 2014, Stark et al. 2017] that excludes comparison of an individual's income with his own income. Therefore in this approach, we assume that the reference group of an individual (the group of individuals the individual compares his income with) does not contain the entire population. Moreover, we assume that the impact on the relative deprivation of each of the excesses of the incomes is not identical. In Stark et al. [2017] we introduced a class of measures of relative deprivation of order $p>0$, which, among other advantages, unifies the Yitzhaki measure and the maximal measure. The parameter p can be seen as the weight the individuals attached to the comparison of their income with the incomes of other individuals in the reference group.

In this article we study the properties of the measure of aggregate relative deprivation of order $p>0$ (ARD_p). Most notably, we show that the aggregate relative deprivation of order $p\geq 1$ satisfies the Pigou-Dalton Transfer Principle[2]. Then we

[2] An index satisfies Pigou-Dalton Transfer Principle if any transfer of income from poorer individual to richer individual imposes increase of the value of the index.

characterize the measure which, among others, regards the superadditivity property of the measure ARD_p. In particular, when individuals put high enough weights on comparisons with the richest individuals in their reference group, if for two populations the highest incomes are different, then the merger of these populations results in increased aggregate relative deprivation beyond the sum of levels of this index for populations prior to the merger.

The paper is organized as follows: In Section 2 we introduce measures of aggregate relative deprivation for the cases of the Yitzhaki measure of relative deprivation, the measure of maximal relative deprivation and the relative deprivation of order $p>0$. Later on, we study the properties of the measure of aggregate relative deprivation of order $p>0$. In Section 3 we show the results on superadditivity property of the measure ARD_p when two populations are merging. Section 4 concludes. The proofs of the results can be found in the Appendix.

2. Aggregate relative deprivation

Let Γ be a population of $n\geq 2$ individuals indexed by $i=1,2,\ldots,n$. Let $X=(x_1,x_2,\ldots,x_n)$ denote the vector of incomes of Γ, where $x_i\geq 0$ is an income of individual indexed by i. We summarize the population and incomes by (Γ, X).

The Yitzhaki measure of *relative deprivation* of individual indexed by i in population (Γ, X) is defined as

$$RD(i,X)\equiv \frac{1}{n}\sum_{j=1}^{n}\max\{x_j-x_i,0\}. \tag{1}$$

The *maximal relative deprivation* index arises when the individual compares his income only with the richest (wealthiest) individual in his reference group, that is

$$RD_{\max}(i,X)\equiv \max_{j=1,\ldots,n} x_j - x_i\,. \tag{2}$$

Summing up the levels of relative deprivation of every individual in the population Γ we obtain the index of *aggregate relative deprivation* for the Yitzhaki measure and for the maximal measure respectively:

$$ARD(X)\equiv\sum_{i=1}^{n}RD(i,X) \qquad \text{and} \qquad ARD_{\max}(X)\equiv\sum_{i=1}^{n}RD_{\max}(i,X)\,.$$

The *relative deprivation of order* $p>0$ of individual indexed by i is defined by

$$RD_p(i,X)\equiv\left(\frac{1}{n-1}\sum_{j\neq i}(\max\{x_j-x_i,0\})^p\right)^{\frac{1}{p}}=\left(\frac{1}{n-1}\sum_{x_j>x_i}(x_j-x_i)^p\right)^{\frac{1}{p}}.$$

This measure unifies two measures of relative deprivation – the measure arising from the comparison of an individual's income with incomes of all other individuals in his population (a measure that is proportional to (1)) and the measure resulting from the comparison of his income with the income of the richest individual (2). Properties of the class $\{RD_p\}_{p>0}$ are discussed in Stark et al. [2017].

We define the *aggregate relative deprivation of order* $p>0$ of population (Γ, X) as the sum of the levels of relative deprivation of order p of all individuals in Γ

$$ARD_p(X) \equiv \sum_{i=1}^{n} RD_p(i, X) = \sum_{i=1}^{n} \left(\frac{1}{n-1} \sum_{j \neq i} (\max\{x_j - x_i, 0\})^p \right)^{\frac{1}{p}}.$$

There is extensive literature on the subject of axiomatic properties of the aggregate relative deprivation measures [Chakravarty, Chakraborty, 1984, Ebert, 1988, Wang, Tsui, 2000]. It is commonly assumed that such measures should fulfill the properties of normalization, symmetry, and Pigou-Dalton Transfer principle. In the following propositions we show, among others, that the measure ARD_p has first two of these properties. We show that ARD_p satisfies Pigou-Dalton Transfer Principle in a separate theorem.

Firstly, we show that the measure ARD_p is symmetric, that is the measure ARD_p is independent of the order of individuals.

Proposition 1. If the set $(\Gamma, \bar{X})$ is obtained from (Γ, X) by permutation of the incomes of members of Γ, then $ARD_p(\bar{X}) = ARD_p(X)$.

By Proposition 1 we can assume that the incomes of members of Γ are ordered. Thus from now on we assume that $x_1 \leq x_2 \leq \ldots \leq x_n$.

Proposition 2. (Properties of ARD_p)

1. (*Monotonicity*) If $p<q$ then $ARD_p(X) \leq ARD_q(X)$ and $ARD_p(X) = ARD_q(X)$ if and only if $x_1 = x_2 = \ldots = x_n$.
2. $\lim_{p \to 0^+} ARD_p(X) = \left(\prod_{j=2}^{n} (x_j - x_1) \right)^{\frac{1}{n-1}}$ and $\lim_{p \to \infty} ARD_p(X) \equiv \sum_{i=1}^{n-1} (x_n - x_i) = ARD_{\max}(X)$.
3. (*Normalization*) If $x_1 = \ldots = x_n$ then $ARD_p(X) = 0$ for every $p>0$.

The first property informs us that the measure ARD_p is increasing with respect to parameter p. Property 2) regards the limit cases, that is, if p converges to zero then the ARD_p converges to the geometric mean of the set of excesses of the incomes over the income of the poorest member of the population, and if p converges to infinity then ARD_p converges to $ARD_{\max}$. By property 3) we have that the measure ARD_p attains its minimal value (zero) when all members of the population Γ have equal incomes.

Before proceeding to the main result of this Section we introduce the *Pigou-Dalton Transfer Principle*.

Definition 1. (Pigou-Dalton Transfer Principle)

Let X be the vector of incomes of members of a population Γ of size n sorted in increasing order, that is $X = (x_1, x_2, \ldots, x_n)$ and $0 \leq x_1 \leq x_2 \leq \ldots \leq x_n$. Let $\bar{X} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$

be the vector of incomes of the population Γ obtained from X by one transfer of some amount of income from a poorer individual to a richer such that this transfer does not affect the order. That is, there exist indexes k,l, $k<l$, and an amount of income $\delta>0$ such that $\bar{x}_k,=x_k-\delta\geq x_{k-1}$, $\bar{x}_l,=x_l+\delta\leq x_{l+1}$ and $\bar{x}_i,=x_i$ for $i=1,\dots,n$, $i\neq k$ and $i\neq l$. We say that an index I satisfies Pigou-Dalton Transfer Principle if $I(\bar{X})>I(X)$.

Below we discuss whether ARD_p satisfies the Pigou-Dalton Transfer Principle.

Theorem 1. If the set $(\Gamma,\bar{X})$ is obtained from (Γ,X) by a transfer of income from a poorer individual to a richer individual, then $ARD_p(\bar{X})> ARD_p(X)$ for every $p\geq 1$.

The result of Theorem 1 is a generalization of the Pigou-Dalton Transfer Principle for the case $p=1$. This result is not a surprise because for $p>1$ changes in incomes of richer individuals are more important for the members of Γ than for $p=1$. Note that for $p\in(0,1)$ we have the opposite effect, therefore, we should not expect the Pigou-Dalton Transfer Principle to take place in this case. In the next two examples we show that if $p\in(0,1)$ then ARD_p may not satisfy the Pigou-Dalton Transfer Principle. In fact, when $p\in(0,1)$ neither the inequality from Pigou-Dalton Transfer Principle nor the opposite inequality will hold.

Example 1. Let $X=(1,5,145)$ be the vector of incomes of population Γ of three individuals. Suppose there was a transfer of $x\in(0,4)$ amount of income from individual indexed by 2 to individual indexed by 3. Let $\bar{X}=(1,5-x,145+x)$ be the vector of incomes of population Γ after the transfer. Then for every $x\in(0,4)$ we have that

$$ARD_{\frac{1}{2}}(X)=\left(\frac{1}{2}\right)^2\left(\left((145-1)^{\frac{1}{2}}+(5-1)^{\frac{1}{2}}\right)^2+\left((145-5)^{\frac{1}{2}}\right)^2\right)=\frac{336}{4}$$

and

$$ARD_{\frac{1}{2}}(\bar{X})=\left(\frac{1}{2}\right)^2\left(\left((145+x-1)^{\frac{1}{2}}+(5-x-1)^{\frac{1}{2}}\right)^2+\left((145+x-(5-x))^{\frac{1}{2}}\right)^2\right).$$

We define

$$f(x)\equiv ARD_{\frac{1}{2}}(\bar{X})-ARD_{\frac{1}{2}}(X)=\frac{1}{4}\left[\left((144+x)^{\frac{1}{2}}+(4-x)^{\frac{1}{2}}\right)^2+2x-196\right].$$

In order to prove that $ARD_{\frac{1}{2}}(\bar{X})<ARD_{\frac{1}{2}}(X)$ it is sufficient to show that $f(x)<0$ for every $x\in(0,4)$. Firstly, note that $f(0)=0$. Secondly, we compute the derivative of the function $f(\cdot)$

$$f'(x)=\frac{1}{4}\left(\frac{\sqrt{4-x}}{\sqrt{144+x}}-\frac{\sqrt{144+x}}{\sqrt{4-x}}+2\right).$$

One can verify that $f'(x)<0$ for every $x\in(0,4)$. This shows that the Pigou-Dalton Transfer Principle is not satisfied for $p=\frac{1}{2}$.

Example 2. Let $X=(1,2)$ be the vector of incomes of population Γ of two individuals. Suppose there was a transfer of $x\in(0,1)$ amount of income from individual indexed by 1 to individual indexed by 2. Let $\bar{X}=(1-x, 2+x)$ be the vector of incomes of population Γ after the transfer. We define

$$f(x)\equiv ARD_p(\bar{X})-ARD_p(X)=\Big((2+x-(1-x))^p\Big)^{1/p}-\Big((2-1)^p\Big)^{1/p}=2x$$

Observe that $f(x)>0$ for every $x\in(0,1)$. Therefore, the inequality opposite to the one shown in Theorem 1 does not hold for $p\in(0,1)$.

By the last three results we have that the Pigou-Dalton Transfer Principle is satisfied by the measure ARD_p if and only if $p\geq 1$. Note that a measure of aggregate relative deprivation can be regarded as a measure of income inequality in a population. Because the Pigou-Dalton Transfer Principle is an important property of any such measure, we have that the measure ARD_p for $p\geq 1$ has the right properties.

3. Merger of populations

In this section we ask whether the merger of two populations will result in increased aggregate relative deprivation beyond the sum of levels of this index for populations prior to the merger. The first results on this subject were obtained in 2013 by O. Stark.

Consider populations Γ_1, Γ_2. Let X denote the vector of incomes of members of population Γ_1, and let Y be the vector of incomes of members of population Γ_2. We denote the merged population by $\Gamma_1\cup\Gamma_2$ and the resulting vector of incomes by $X\circ Y$.

Theorem 2. [Stark, 2013]

$$ARD(X\circ Y)\geq ARD(X)+ARD(Y) \tag{3}$$

and

$$ARD_{max}(X\circ Y)\geq ARD_{max}(X)+ARD_{max}(Y). \tag{4}$$

The property of an index that the value of the sum of arguments is at least the sum of values of the index of the arguments is called *superadditivity*. Therefore, upon (3) and (4) we call the measures ARD and ARD_{max} superadditive.

We show that when the aggregate relative deprivation of the population is measured by using the index ARD_p, then the superadditivity property does not always hold. Firstly, we show that if in the two populations the highest incomes differ, then the measure ARD_p has the superadditivity property provided that p is sufficiently large.

Theorem 3. Let $X=(x_1,\ldots,x_n)$ be the vector of incomes of population Γ_1 of n individuals, and let $Y=(y_1,\ldots,y_m)$ be the vector of incomes of population Γ_2 of m individuals. Suppose that $\max\limits_{i=1,\ldots,n} x_i\neq\max\limits_{j=1,\ldots,m} y_j$. If p is sufficiently large, then

$$ARD_p(X\circ Y)>ARD_p(X)+ARD_p(Y).$$

The result of Theorem 3 can be explained on the basis of the properties of the measure *RD*. Namely, for $p>1$ an individual puts high weights on comparisons with the richest individuals in his reference group. If even richer individuals appear after merging, then his relative deprivation will be affected significantly.

Secondly, we show that a merger of two identical populations results in a decrease of aggregate relative deprivation below the sum of levels of this index for the populations prior to the merger.

Theorem 4. Let populations (Γ_1, X), (Γ_2, Y) be such that $X=Y$. Then for every $p>0$ we have that

$$ARD_p(X \circ Y) \leq ARD_p(X) + ARD_p(Y).$$

Moreover, $ARD_p(X \circ X) = 2 \cdot ARD_p(X)$ if and only if $x_1 = x_2 = \ldots = x_n$.

Finally, because the measures *ARD* and ARD_1 are proportional we can deduce the following result about the measure *ARD*.

Remark 1. By Theorem 2 we have that $ARD(X \circ Y) \geq ARD(X) + ARD(Y)$, while by Theorem 4 we obtain that $ARD_1(X \circ Y) \leq ARD_1(X) + ARD_1(Y)$ whenever $X=Y$. Moreover, because $n \cdot ARD(X) = (n-1) \cdot ARD_1(X)$ we have that *ARD* is proportional to ARD_1. By these facts we deduce that if $X=Y$, then

$$ARD(X \circ Y) = ARD(X) + ARD(Y).$$

4. Conclusions

In this paper we studied how the merger of two populations affects the aggregate relative deprivation of order $p>0$ of these groups. The results we obtained differ from the results of Stark [2013]. In Stark [2013] the measure *ARD* is always superadditive, in other words, the merger of populations will always increase social stress experienced by the members of these populations. By using the measure of aggregate relative deprivation of order $p>0$, we showed that the superadditivity property does not have to occur. In particular, for two identical populations the measure ARD_p is subadditive, and not superadditive. These differences can be explained by the choice of the reference group for individuals. Stark [2013] uses the aggregated Yitzhaki measure of relative deprivation, and in this index the reference group of an individual consists of the entire population. We exclude an individual from his reference group, thus the individual does not compare his income with his own income. This approach is in line with the modern studies in income inequalities [Bowles, Carlin, 2020] and with the axiomatic basis of the relative deprivation measures [Ebert, Moyes, 2000, Bossert, d'Ambrosio, 2014, Stark et al., 2017].

Acknowledgments

We acknowledge support from a subsidy granted to Cracow University of Economics. We acknowledge the invaluable help of dr Fryderyk Falniowski.

Appendix

Before presenting the proofs of the main results we show a lemma that will be useful in proving these results.

Lemma 1. Let I be an interval and let $f{:}I\to\mathbf{R}$ be convex. If $x_1,x_2\in I$, $x_1<x_2$ and $\delta<x_1$, then

$$f(x_1)-f(x_1-\delta)\leq f(x_2+\delta)-f(x_2).$$

Proof

It is well-known that a function $f(\cdot)$ is convex if and only if the function $R(x,y)\equiv\frac{f(y)-f(x)}{y-x}$ is increasing in x, for y fixed (or vice versa). Observe that $x_1-\delta<x_1<x_2<x_2+\delta$. By using the monotonicity of functions $R(x,\cdot)$, $R(\cdot,y)$ we have that

$$\frac{f(x_1)-f(x_1-\delta)}{\delta}\leq\frac{f(x_2)-f(x_1-\delta)}{x_2-x_1+\delta}\leq\frac{f(x_2)-f(x_1)}{x_2-x_1}\leq\frac{f(x_2+\delta)-f(x_1)}{x_2-x_1+\delta}\leq\frac{f(x_2+\delta)-f(x_2)}{\delta}.$$

By multiplying the leftmost and the rightmost sides of this sequence of inequalities by δ we obtain the thesis.

Proof of Proposition 1

The proof is a direct consequence of the definition of the aggregate relative deprivation of order $p>0$.

Proof of Proposition 2

Properties 1) – 3) follow immediately from the properties of the measure RD_p.

Proof of Theorem 1

The proof of case $p=1$ is straightforward. Thus, we assume that $p>1$. Let Γ be a population of n individuals with incomes represented by the vector $X=(x_1,\ldots,x_k,\ldots,x_l,\ldots,x_n)$, where $x_i\leq x_{i+1}$ for $i=1,2,\ldots,n-1$, $k<l$ and $x_k<x_l$. Let $\bar{X}=(x_1,\ldots,x_k-\delta,\ldots,x_l+\delta,\ldots,x_n)$ denote the vector of incomes of population Γ after a transfer of $\delta<\min\{x_k-x_{k-1},x_{l+1}-x_l\}$ amount of income from individual indexed by k to individual indexed by l. The aggregate relative deprivation of order $p>1$ of the populations (Γ,X) and $(\Gamma,\bar{X})$ have the following forms:

$$ARD_p(X)\equiv\sum_{i=1}^{n}RD_p(i,X)\quad\text{and}\quad ARD_p(\bar{X})\equiv\sum_{i=1}^{n}RD_p(i,\bar{X}).$$

We analyze the results of the transfer of income on the levels of relative deprivation of order $p>1$ for members of Γ. By the properties of the measure RD_p [see Stark et al., 2017] we have for $p>1$ that:

$RD_p(i,\bar{X})=RD_p(i,X)$ for $i\in\{l+1,\ldots,n\}$,

$RD_p(l,\bar{X})<RD_p(l,X)$,

$RD_p(i,\bar{X})>RD_p(i,X)$ for $i\in\{1,\ldots,l-1\}$.

To complete the proof it is sufficient to show that

$$RD_p(k,\bar{X})+RD_p(l,\bar{X})+\sum_{\substack{i=1\\ i\neq k}}^{l-1}RD_p(i,\bar{X})>RD_p(k,X)+RD_p(l,X)+\sum_{\substack{i=1\\ i\neq k}}^{l-1}RD_p(i,X). \quad (5)$$

Because $\sum_{\substack{i=1\\ i\neq k}}^{l-1}RD_p(i,\bar{X})>\sum_{\substack{i=1\\ i\neq k}}^{l-1}RD_p(i,X)$, in order to prove (5) it is sufficient to show that

$$RD_p(k,\bar{X})+RD_p(l,\bar{X})\geq RD_p(k,X)+RD_p(l,X).$$

We define $f(x)\equiv\left(\frac{1}{n-1}\sum_{i=1}^{n}(\max\{x_j-x,0\})^p\right)^{\frac{1}{p}}$ for $x\geq x_1$ and for $p>0$. From properties of the measure RD_p we know that the function $f(\cdot)$ is decreasing, strictly convex on (x_1,x_{n-1}), and affine on (x_{n-1},x_n). Moreover, we have that $f'_-(x_i)=f'_+(x_i)$ for $i=2,3,\ldots,n-1$, therefore the function $f'(\cdot)$ exists and is continuous for $x>x_1$. Consequently, we have that $f(\cdot)$ is convex on (x_1,x_n). Thus, by using Lemma 1 we obtain that

$$f(x_k-\delta)+f(x_l+\delta)\geq f(x_k)+f(x_l). \quad (6)$$

Moreover, we observe that

$$RD_p(l,\bar{X})=\left[\frac{1}{n-1}\left(\sum_{j\neq k,\, j\neq l}(\max\{x_j-x_l-\delta,0\})^p+(\max\{x_k-x_l-2\delta,0\})^p\right)\right]^{\frac{1}{p}}$$

$$=\left(\frac{1}{n-1}\sum_{j\neq l}(\max\{x_j-x_l-\delta,0\})^p\right)^{\frac{1}{p}}=f(x_l+\delta) \quad (7)$$

and

$$RD_p(k,\bar{X})=\left[\frac{1}{n-1}\left(\sum_{j\neq k,\, j\neq l}(\max\{x_j-x_k+\delta,0\})^p+(\max\{x_j-x_k+2\delta,0\})^p\right)\right]^{\frac{1}{p}}$$

$$>\left(\frac{1}{n-1}\sum_{j\neq k}(\max\{x_j-x_k+\delta,0\})^p\right)^{\frac{1}{p}}=f(x_k-\delta). \quad (8)$$

Note that equality in (7) follows from the fact that $\max\{x_k - x_l - 2\delta, 0\} = 0$. By using (6), (7) and (8) we obtain that

$$RD_p(k, \bar{X}) + RD_p(l, \bar{X}) > f(x_k - \delta) + f(x_l + \delta) \geq f(x_k) + f(x_l) = RD_p(k, X) + RD_p(l, X),$$

which completes the proof of the Theorem.

Proof of Theorem 3

Without loss of generality, we can assume that $\max\limits_{i=1,\ldots,n} x_i < \max\limits_{j=1,\ldots,m} y_j$. Consider

$$ARD_{\max}(X \circ Y) - ARD_{\max}(X) - ARD_{\max}(Y)$$

$$= \sum_{i=1}^{n} \left(\max_{j=1,\ldots,m} y_j - x_i \right) + \sum_{i=1}^{m} \left(\max_{j=1,\ldots,m} y_j - y_i \right) - \sum_{i=1}^{n} \left(\max_{j=1,\ldots,n} x_j - x_i \right) - \sum_{i=1}^{m} \left(\max_{j=1,\ldots,m} y_j - y_i \right)$$

$$= \sum_{i=1}^{n} \left(\max_{j=1,\ldots,m} y_j - x_i \right) - \sum_{i=1}^{n} \left(\max_{j=1,\ldots,n} x_j - x_i \right) = n \cdot \left(\max_{j=1,\ldots,m} y_j - \max_{j=1,\ldots,n} x_j \right) > 0.$$

Note that $\lim\limits_{p\to\infty} ARD_p(X) = ARD_{\max}(X)$ for any (Γ, X). Therefore there exists $p_0 > 0$ such that for every $p > p_0$ we have that $ARD_p(X \circ Y) > ARD_p(X) + ARD_p(Y)$.

Proof of Theorem 4

Let $X = (x_1, \ldots, x_n)$ denotes the vector of incomes of population Γ_1 and $Y = (y_1, \ldots, y_n)$ denotes the vector of incomes of population Γ_2. Then the aggregate relative deprivation of order $p > 0$ of the combined group $\Gamma_1 \cup \Gamma_2$ has the following general form

$$ARD_p(X \circ Y) = \sum_{i=1}^{n} \left[\frac{1}{n+m-1} \left(\sum_{j \neq i} (\max\{x_j - x_i, 0\})^p + \sum_{k=1}^{m} (\max\{y_k - x_i, 0\})^p \right) \right]^{\frac{1}{p}}$$

$$+ \sum_{i=1}^{m} \left[\frac{1}{n+m-1} \left(\sum_{j \neq i} (\max\{y_j - y_i, 0\})^p + \sum_{k=1}^{n} (\max\{x_k - y_i, 0\})^p \right) \right]^{\frac{1}{p}}.$$

If $X = Y$ (and consequently $m = n$) and there exists at least one pair of indexes $i, j = 1, \ldots, n$, $i \neq j$ such that $x_i \neq x_j$, then

$$ARD_p(X \circ X) = 2 \sum_{i=1}^{n} \left(\frac{1}{2n-1} 2 \sum_{j \neq i} (\max\{x_j - x_i, 0\})^p \right)^{\frac{1}{p}} < 2 \sum_{i=1}^{n} \left(\frac{1}{2n-2} 2 \sum_{j \neq i} (\max\{x_j - x_i, 0\})^p \right)^{\frac{1}{p}}$$

$$= 2 \sum_{i=1}^{n} \left(\frac{1}{n-1} \sum_{j \neq i} (\max\{x_j - x_i, 0\})^p \right)^{\frac{1}{p}} = 2 \cdot ARD_p(X).$$

References

Bossert W., D'Ambrosio C. [2006], *Reference groups and individual deprivation*, "Economics Letters", 90(3), pp. 421–426.

Bossert W., D'Ambrosio C. [2014], *Proximity-sensitive individual deprivation measures*, "Economics Letters", 122(2), pp. 125–128.

Bowles S., Carlin W. [2020], *Inequality as experienced difference: A reformulation of the Gini Coefficient*, "Economics Letters", 108789.

Chakravarty S.R., Chakraborty A.B. [1984], *On indices of relative deprivation*, "Economics Letters", 14(2), pp. 283–287.

Chakravarty S.R. [1999], *Measuring inequality: An axiomatic approach*, [in:] Silber J. (ed.), *Handbook of Income Inequality Measurement*, Kluwer: Dordrecht, pp. 163–186.

Ebert U. [1988], *Measurement of inequality: An attempt at unification and generalization*, "Social Choice and Welfare", 5(2/3), pp. 147–169.

Ebert U., Moyes P. [2000], *An axiomatic characterization of Yitzhaki's index of individual deprivation*, "Economics Letters", 68(3), pp. 263–270.

Hey J.D., Lambert P.J. [1980], *Relative deprivation and the Gini coefficient: Comment*, "Quarterly Journal of Economics", 95(3), pp. 567–573.

Runciman W.G. [1966], *Relative Deprivation and Social Justice*, University of California Press: Berkeley.

Stark O., Hyll W. [2011], *On the economic architecture of the workplace: Repercussions of social comparisons among heterogeneous workers,* "Journal of Labor Economics", 29(2), pp. 349–375.

Stark O. [2013], *Stressful Integration*, "European Economic Review", (63), pp. 1–9.

Stark O., Bielawski J., Falniowski F. [2017], *A class of proximity-sensitive measures of relative deprivation*, "Economics Letters",160, pp. 105–110.

Wang Y., Tsui K. [2000], *A new class of deprivation-based generalized Gini indices*, "Economic Theory", 16(2), pp. 363–377.

Yitzhaki S. [1979], *Relative deprivation and the Gini coefficient*, "Quarterly Journal of Economics", 93(2), pp. 321–324.

Convergence of Conflict Sets and Applications

Anna Denkowska[1]

Abstract

The conflict set of a finite family of pairwise disjoint closed subsets of the Euclidean space is the set of those points whose distance to the union of these sets is realized in more than one of them. The conflict set is always closed. We are interested in the situation when the sets in question evolve in time. The best language which allows us to express such a deformation in time is that of the Kuratowski convergence of closed sets. The notion of conflict set has a wide range of applications, from pattern recognition to geographical or economic issues such as testing distribution and availability patterns (e.g. for hospitals, schools or shopping centers distribution; it belongs to civics and planning) like in central place theory or location models, or in correlating sources of infections in epidemics, and so on. In this note we present and announce some new results concerning the semicontinuity of the conflict set and discuss them from applications' point of view.

Keywords: conflict sets, Voronoi diagrams, spatial economic analysis

JEL classification: C02, C30, O10

[1] Anna Denkowska, Cracow University of Economics, e-mail: anna.denkowska@uek.krakow.pl

1. Introduction

The notion of conflict set and the similar notion of medial axis (skeleton) – the latter being one of the cornerstones of pattern recognition – have a large range of applications in robotics, physics (e.g. wavefront propagation) or mathematics (e.g. in singularity theory). The conflict set of two disjoint subsets of $\mathbb{R}^n$ is the collection of points that are equidistant to both sets. This definition may be extended to more than two sets. Even though we consider here the Euclidean metric, we could replace it successfully by another one. The conflict set may be interpreted as the place where the 'influences' of two 'centers' meet or clash. This yields a new interpretation of some equilibrium models. Skeletons and also conflict sets are considered to be highly unstable. We are interested in the behavior of the conflict set when the sets that define it are subject to deformation. This deformation is conveniently expressed using the Kuratowski convergence. Such an approach shows that there is a kind of stability, namely a lower semi-continuity (in some specific cases even continuity) both of the skeletons and conflict sets.

A particular example of conflict sets are the *Voronoi diagrams* defined for singletons. They are used in the study of geoeconomic, logistical and macroeconomic issues. The present note put forward the theoretical background for further studies in applications of conflict sets from the dynamic point of view. It should be pointed out that the range of already existing applications of conflict sets in geoeconomic problems is rather restricted. They are met within modelling the relative size of each country's economy in terms of nominal GDP (where the weighted Voronoi diagram is used) or in logistics (e.g. the distribution pattern problem for discount chains in urban planning). However, according to the remarks in [Kisiała et al., 2017] in Polish scientific literature these issues are hardly covered, as opposed to western publications. In any case, a dynamic approach to Voronoi diagrams has almost not been exploited yet even in the worldwide literature. What we mean here by 'dynamic approach' is the study of Voronoi diagrams of sets of points evolving in time (points that can split or merge). The only publications in this field up to now seem to be [Reem, 2011] and [Roos, 1993].

A natural tool and a one that has not been used before for the study of time-evolving conflict sets are provided by the *Painlev ́e-Kuratowski convergence* of closed sets. It allows to describe properly the changes occurring in time and keeping track of the distribution as well as detecting the possible merging of points in the case of Voronoi diagrams. Moreover, it constitutes a different approach to the problem than the one presented in [Reem, 2011] or [Roos, 1993]. Among the first theoretical questions to answer is the question concerning the stability of the diagram. This has a particular impact on the applications, especially as the convergence of Voronoi diagrams is of importance in the statistical clustering method known as the k-means clustering, or in the peculiar example of a Nash equilibrium – the Hotelling location model (principle of minimum differentiation on the market). From these remarks we can propose two directions for further studies. One is the stability question and its applications in stability analysis and modelling of market sharing, while a second one would be the possibility of describing equilibria of some market or logistical models with continuously varying data.

2. Origins of the basic notions

The convergence of closed sets was introduced in 1902 by the French mathematician *Paul Painlev ´e* as a development of an earlier work of Felix Hausdorff. *Kazimierz Kuratowski* gathered all the results, set them in order and presented in his book *Topology*. The Kuratowski convergence, as it is often called nowadays, has found many applications in optimal control [Frol´ık, De Giorgi, Franzoni, Dal Maso] in the '80s of the 20th century and since the '90s it has made its way to singularity theory. The skeleton (medial axis) of a domain in $\mathbb{R}^n$, i.e. the set of points whose distance to the boundary is realized by more than one point, was introduced in 1967 by Henry Blum as a basic tool for pattern recognition. The related notion of conflict set can be traced back to such classical geometric constructions as that of the parabola: the conflict set of a point and a line.

3. Medial axes, conflict sets and Voronoi diagrams

Consider a nonempty, closed set X in $\mathbb{R}^n$, where $\mathbb{R}^n$ is endowed with the Euclidean metric. Given $a \in \mathbb{R}^n$, we put $m(a) := \{x \in X \mid ||a - x|| = d(a, X)\}$ for the set realizing the Euclidean distance $d(a, X)$ of a to X. Note that in our setting, $m(a) \neq \emptyset$.

Definition 3.1. The set $M_X := \{a \in \mathbb{R}^n \mid \#m(a) > 1\}$ is called the *medial axis* of X (originally: of the open set $\mathbb{R}^n \setminus X$).

Let now X_1, X_2 in $\mathbb{R}^n$ be two nonempty, disjoint sets.

Definition 3.2. The set $\mathrm{Conf}(X_1, X_2) := \{x \in \mathbb{R}^n \mid d(x, X_1) = d(x, X_2)\}$ is called the conflict set of X_1, X_2.

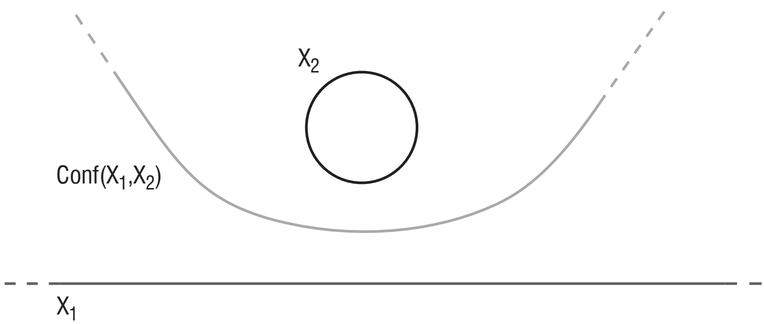


Figure 1. Conflict set

Source: own work.

Additionally, we introduce open sets called territories.

Definition 3.3. $\mathrm{Ter}(X_1, X_2) :=\{x \in \mathbb{R}^n \mid \mathrm{d}(x, X_1) < d(x, X_2)\}$called the *territory* of X_1 (w.r.t. X_2) and the analogous territory $\mathrm{Ter}(X_2, X_1)$ of X_1 w.r.t. X_2.

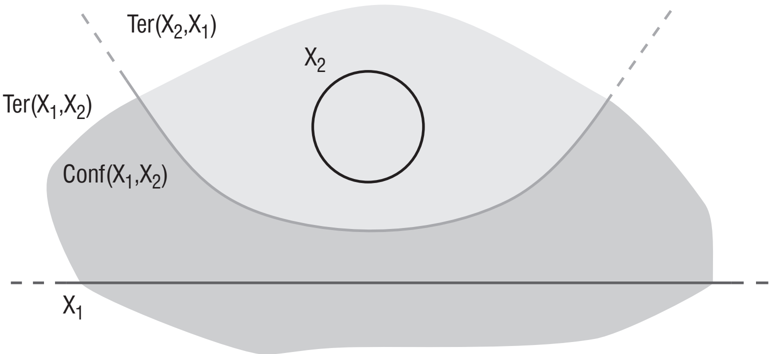


Figure 2. Territories of two sets

Source: own work.

Then $\mathrm{Conf}(X_1,X_2)$ is the common boundary of the two territories. For more than two sets: $X_1,\ldots, X_p$ in $\mathbb{R}^n$ (again: closed, nonempty, pairwise disjoint), the conflict set is defined using a slightly modified notion of the territories for the family of sets $X:=\{X_1, \ldots, X_p\}$.

Definition 3.4. The set $\mathrm{Ter}^-(X_i, X) := \{x \in \mathbb{R}^n \mid d(x, X_i) \leq d(x, X_j), j = 1,\ldots , p\}$ is called the *closed territory* of X_i w.r.t. the family X and then the *conflict set* of this family is defined to be $\mathrm{Conf}(X) := \cup_{i\neq j} (\mathrm{Ter}^-(X_i, X) \cap \mathrm{Ter}^-(X_j, X))$

We also introduce the *open territory* $\mathrm{Ter}(X_i, X) := \{x \in \mathbb{R}^n \mid d(x, X_i) < \min_{i\neq j} d(x,X_j)\}$.

Definition 3.5. If all the sets $X_i =\{x_i\}$ are singletons the medial axis of their union coincides with their conflict set and is called the *Voronoi diagram* of the system of points $\{x_1,\ldots, x_p\}$.

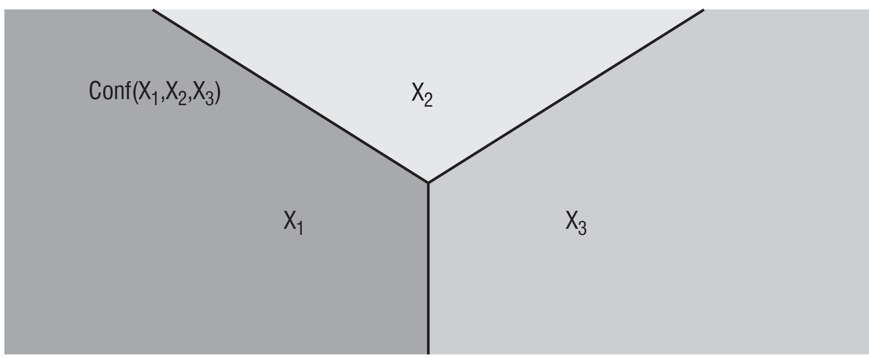


Figure 3. Voronoi diagram of three points

Source: own work.

Observation 3.6. There is a natural relation between the conflict set of the family $X:=\{X_1,\ldots,X_p\}$ and the medial axis of the set $\cup X = X_1 \cup \ldots \cup X_p$, namely:

$Con\,f(X)=M_{\cup X} \setminus \cup_{i=1}^{p} Ter(X_i,X)$

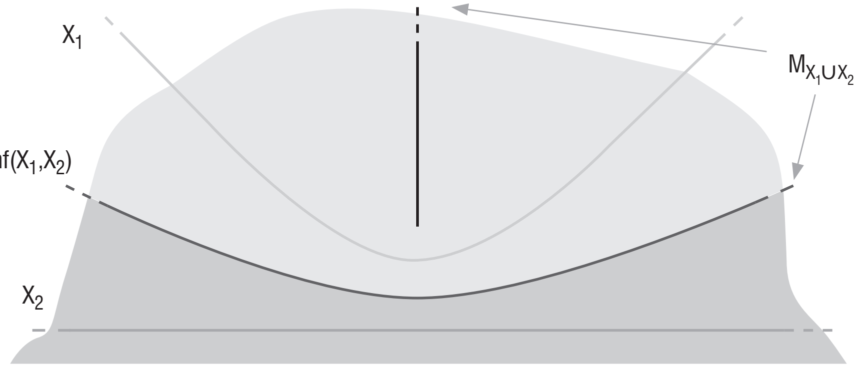


Figure 4. Conflict set and medial axis

Source: own work.

Remark 3.7

The definition of the conflict set makes sense also if the sets – instead of being disjoint – are pairwise distinct and none of them is contained in the union of the remaining ones. However, we may lose the *thinness* of the conflict set in this case.

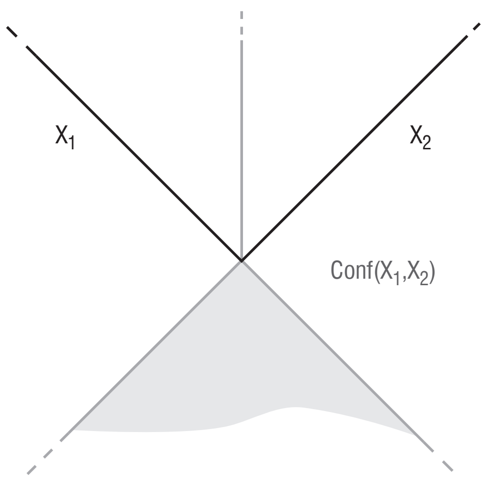


Figure 5. A ‘fat’ conflict set of non-disjoint sets

Source: own work.

4. The lower and upper Kuratowski limits

Consider a set $E \subset \mathbb{R}^k \times \mathbb{R}^n$ in the variables (t, x), where t plays the role of the *parameter*.

Let $\pi(t, x) = t$ and for $t \in \pi(E)$, $E_t := \{x \in \mathbb{R}^n \mid (t, x) \in E\}$ denotes the *section* or *fibre* of E at t.

Fix an accumulation point t_0 of $\pi(E)$.

Definition 4.1. (Lower and upper Kuratowski limits).

$x_0 \in \liminf_{t \to t_0} E_t \Leftrightarrow \forall$ neighborhood U of x_0, $\exists$ a neighborhood V of t_0, such that

$\forall t \in V \cap \pi(E) \setminus \{t_0\}$, $E_t \cap U \neq \emptyset$;

$x_0 \in \limsup_{t \to t_0} E_t \Leftrightarrow \forall$ neighborhood U of x_0, $\forall$ neighborhood V of t_0,

$\exists t \in V \cap \pi(E) \setminus \{t_0\}$ such that $E_t \cap U \neq \emptyset$.

From the definition it follows that $\liminf_{t \to t_0} E_t \subset \limsup_{t \to t_0} E_t$.

Definition 4.2. The set $E \subset \mathbb{R}^n$ is said to be the *Kuratowski limit of E_t* when $t \to t_0$, if

$\limsup_{t \to t_0} E_t \subset E \subset \liminf_{t \to t_0} E_t$

Proposition 4.3.

1. The lower and upper limits are closed sets;
2. The lower and upper limits remain unchanged, if we replace E_t by their closures E_t;
3. If the sets E_t are compact, then $\lim_{t \to t_0} E_t = E \neq \emptyset \Leftrightarrow \lim_{t \to t_0} \mathrm{dist}_H(E_t, E) = 0$

Therefore, from now on we consider only closed sets.

5. O-minimal structures

In order to prove the stability results we are aiming at, we have to *tame* the topology of the sets considered. We will restrict ourselves to sets that are definable in some o-minimal structure cf. [Coste, 2000]. For simplicity, further on we shall focus our attention on *semi-algebraic sets*.

Remark 5.1. The classes of sets introduced in the following two definitions may seem abstract at first sight but actually they are the type of sets most often met within the usual economic models. Indeed, modelling involving polynomials leads directly to the semi-algebraic setting; on the other hand, if the exponential or the logarithm is used, more general o-minimal structures get involved.

Definition 5.2. (O-minimal structure). An *o-minimal structure* on $\mathbb{R}$ expanding the real field (R,+,.) is a sequence $\{S_n\}$ of families $S = S_n \subset P(\mathbb{R}^n)$ of subsets of $\mathbb{R}^n$ satisfying:

$\forall$ n $\in$ N (A, B $\in S_n \Rightarrow$ A$\cap$B, A$\cup$B, $\mathbb{R}^n\setminus A \in S_n$);

$\forall$ n, m $\in$ N ($A \in S_n$, $B \in S_m \Rightarrow A\times B \in S_{n+m}$);

$\forall$ n $\in$ N $\forall$ A $\in S_{n+1}$, $\prod(A)\in S_n$, gdzie $\prod$: $\mathbb{R}^n\times\mathbb{R} \ni (x,t)\longmapsto x\in\mathbb{R}^n$;

$\forall$ n $\in$ N, for all $P \in \mathbb{R}[x_1,\dots,x_n]$, $P^{-1}(0) \in S_n$; S_1 consists exactly of all the possible finite unions of intervals of any type (in particular, points).

The sets belonging to the classes S_n are said to be *definable* in the structure S.

Definition 5.3. (Semi-algebraic set). A set $X \subset \mathbb{R}^n$ is called *semi-algebraic*, if it can be described by finitely many polynomial equations and inequalities, which amounts to say that it is of the form $X=\cup_{i=1}^{p} \cap_{j=1}^{r} \{x \in \mathbb{R}^n | p_i(x)=0,\ q_{ij}(x)>0\}$, where p_i, q_{ij} are real polynomials of n variables.

Remark 5.4. Any o-minimal structure is an expansion of the class of semi-algebraic sets, i.e. S_n contains all the semi-algebraic sets.

Semi-algebraic sets (just as all definable ones) have a large bunch of nice geometric properties, e.g. the interior, the closure, the boundary, the section, the connected components of a semi-algebraic set are semi-algebraic as well, and similarly the image under projection stays semi-algebraic (Tarski- Seidenberg Theorem).

6. Stability theorems – pre-published results and further research

Here we discuss the stability result from [Denkowski, 2016] concerning the medial axis, and announce an extension to the case of conflict sets [Denkowska, Denkowski, 2020] with a few illustrating examples. Interpretations are given in the next section. There are three points to note:

High instability of the medial axis: a small deformation of the set may induce a large deformation of the medial axis's structure [Chazal, Soufflet, 2004]; however, in this point of view the deformation is not seen as a continuous process.

Express the deformation using the Kuratowski convergence: it turns out that we have a kind of stability of the medial axis – we control the latter thanks to the lower limit cf. [Denkowski, 2016] see Theorem 6.1 below.

The only additional assumption that we have to make is that the family of sets considered is definable in an o-minimal structure. Let us stress again that this is not very restrictive, since in most practical situation we are actually dealing with such sets (apart from semi-algebraic sets, also a large subclass of subanalytic sets forms an o-minimal structure). Indeed, models are most often described by a polynomial, or at worst analytic equations and inequalities.

Theorem 6.1. [see Denkowski, 2016]. If $X \subset \mathbb{R}^k \times \mathbb{R}^n$ in the variables (t, x) is a closed semi-algebraic set and $X_{t_0}= \lim\limits_{t\to t_0} X_t$, then $M_{X_{t_0}} \subset \liminf_{t\to t_0} M_{X_t}$.

Remark 6.2. This result is optimal, as shown by the example [Denkowski, 2016] of $X = \{(t, x_1, x_2) \in \mathbb{R} \times \mathbb{R}^2 \mid x_2 = t|x_1|\}$, for which

$M_{X_t} = \emptyset$, if t=0 or $M_{X_t} = \{0\} \times (0, +\infty)$, for t>0, or $M_{X_t} = \{0\} \times (-\infty, 0)$, for t<0,

where M_{X_t} do not converge and $M_{X_0} \subset \liminf_{t \to 0} M_{X_t} = \{(0, 0)\}$

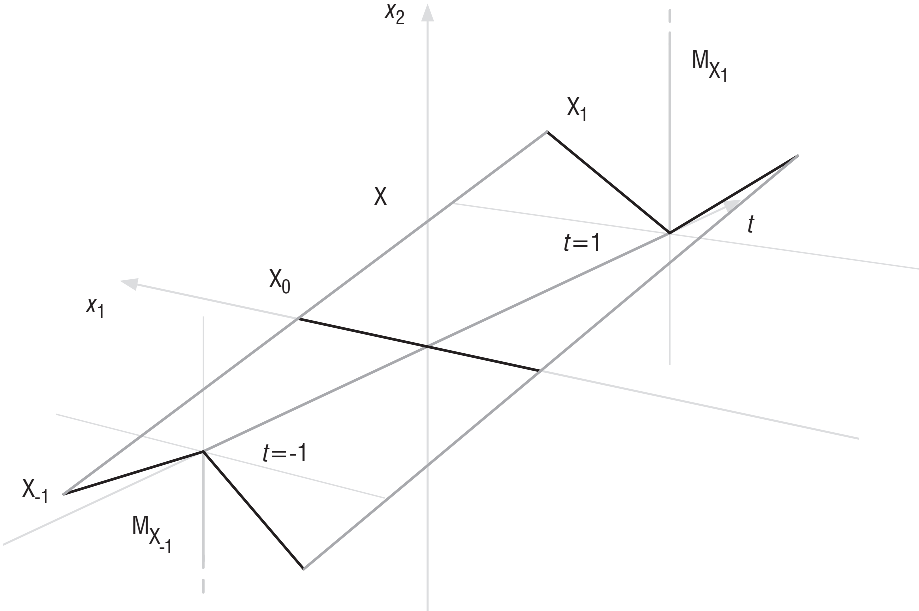


Figure 6. There may be neither convergence nor equality for the lower limit in Theorem 6.1

Source: own work.

Remark 6.2. In accordance with Definition 3.5 of the Voronoi diagram we see that Theorem 6.1 is valid in particular also for Voronoi diagrams of sets of points that can be parametrized in a definable way. Therefore, it is natural to ask the following two questions: (1) is the definability assumption necessary for Voronoi diagrams?

(2) since the Voronoi diagram of a finite set is also the conflict set of the family of the singletons forming this set, can there be a counterpart of Theorem 6.1 for conflict sets?

The proof of the Theorem 6.1 is based essentially on the particular use of the Curve Selection Lemma (a basic tool in the theory of o-minimal structures) and some topological arguments. However, they cannot be directly transposed to the case of conflict sets. Our further theoretical research (to be published in a forthcoming paper) is aimed at answering the two questions stated above. It requires the use of some results from [Denkowska and Denkowski, 2012]. The result we are able to obtain can be interpreted in the following way: at each time *t* we compute the conflict set of a constant number of 'influence centers' (whose shape plays a role) varying continuously with t and that can be described polynomially, then in the limit time t_0 we can identify no less influence centers whose conflict set is contained in the lower limit of the varying conflict sets.

Question (1) is partly inspired by [Roos, 1993]. Actually, the definability seems superfluous in case the number of points stays constant: the Voronoi diagrams of finite sets of points varying continuously and not merging vary continuously, too. This is particularly interesting from applications' point of view.

7. Applications

A wide analysis of the spatial economic issues is proposed in Gibas and Heffner [2007]. The authors stress that many economic analyses do not take into account the spatial aspects. In microeconomics, in consumer theory, the choices concern non-localized goods; in producer's theory his position/location is not specified; in price theory it is assumed that supply and demand come together on a market with unspecified range or location in space; the theory of general equilibrium is concerned with interlinkages between partial equilibria, their circumstances and dependencies but with non-spatial features. Similarly, in macroeconomics, when regional and local state aggregate structures are not extracted, then global equilibria can hide lack of partial equilibria. The economic analysis is done in such a way as if the interaction of space did not play any role in the economic processes. Although it is acknowledged that the spatial factor in economic analysis contains a body of knowledge with some explanatory aspects, it is usually assumed that it just corroborates the conclusions in more details and nothing more. The introduction of a spatial dimension or 'spatial factor' into the research on economic, or, more generally, socio-economic processes makes us aware that traditional economic analysis formulates laws based on the assumption that all objects and goods are in the same (one) place. This approach is of fundamental theoretical and practical importance. From an economic point of view, however, the impact of space is not neutral. All forms of human activity are directly or indirectly related to space. This also applies to socio-economic activities. These activities take place in space, depend on the properties of the latter, but also shape these properties to some extent. As a result, almost all arrangements for socio-economic development (except the most general and preliminary) take into account spatial factors. For these reasons, the space with its properties, its variants and elements are one of the aspects of socio-economic planning, both in terms of scientific and research activity as well as practical planning. In spatial economic analysis, we deal with different types of space – geodesic, geographic, economic and social. Economic and social space is important in spatial economic analysis. When determining the research object, it is also important to determine its location in space, as it plays a role in the procedure of combining similar basic spatial units into larger zone-type systems in view of a classification process. Determining zonal arrangements may include both spatially continuous phenomena such as ecological conditioning, land use, demographic characteristics or population distribution, as well as discontinuous ones such as economic activity, or urbanization. Spatial economic analysis should, therefore, be a specialized field of economic analysis, analogous to the analysis of consumption, production or international trade. The introduction of a spatial factor for economic analysis is another approach to any economic analysis.

The importance of a dynamic approach in geoeconomic applications can be illustrated based on Christaller's Central Place Theory which has been long criticized for being static and not taking into account the temporal aspect in the development of central places nor the diversified nature of the services, nor the varied distribution of the resources. Note that Central Place Theory can be (and was) used both in market modelling or shopping centers planification. This is a typical example of a theory where our result could prove useful.

One of the tools that illustrate spatial relationships is Voronoi diagrams, which are a special example of a conflict set. The Voronoi diagram is constructed for a (finite) set of points S on a surface. The surface or domain is divided into cells induced by S, each consisting of points that lie nearer to a given point of S than to any other from this set [Brassel, Reif, 1979]. The first traceable use of such a division is due to Descartes in the 17th century when he used it to present the distribution of the matter in the universe [Okabe et al., 2009]. We encounter Voronoi diagrams in the literature on the subject, e.g. when presenting the relative size of the economy of each country around the world in terms of nominal GDP. "If the global economy were a living, breathing organism, it might look similar to a dynamic Voronoi diagram" wrote Shawn Langlois editor and writer at MarketWatch, Los Angeles, who created a costing site at How-Much.net. Data were collected from the International Monetary Fund and an animation was created showing the evolution of the GDP of the largest countries in the world in the years 1980–2015, which is to illustrate how countries have developed and shrunk relative to each other. In Jastrzębska [2017] Voronoi diagrams are used to construct a spatial distribution map of real estate transaction prices, i.e. market segmentation and to determine price zones and division into expensive and cheap areas. The superiority of Voronoi diagrams over MSI, precinct or hexagon cartograms is indicated. The diagrams also allow to notice the connection between the value of the property and its location, e.g. the distance from the main communication routes or the presence of nuisance in the area. Another example of the use of Voronoi diagrams is the proposal of a location tax on retail commercial activities as a way to rationally shape the service network [Śleszyński, 2019]. The article presents the concept of sales tax for stationery stores, depending on the size of sales area and the number of population in short distances from the place of sale (in short: location tax). It was assumed that it is possible to shape the balance in terms of demand (population) and supply (sales area) in order to avoid market monopolization by the largest entities. Voronoi diagrams can also be applied to any issues connected with logistics planning functions, e.g. path [Long et al., 2011] or location planning; if a concern wants to open a new gas station without too much perturbation in the existing network, it should choose a place sufficiently distant from all the other stations in the area. This can be chosen as a vertex in the Voronoi diagram and it can be determined through a time-linear passing by all the vertices process. In [Kisiała, Rutkiewicz, 2017] the authors studied spatial distribution availability patterns of discounts. In addition, Voronoi diagrams are also used to develop a universal method of exploratory statistical analysis of spatial data. Classic statistics gives various types of algorithms for data analysis that perform both regression and classification tasks. In the work [Fiedukowicz et al. 2014] examples of preliminary concepts of modification and extension of known methods and statistical algorithms are given, including spatial information in these methods. The problem of including spatial data in analyses – through universal to some extent – requires an individual approach due to the specificity of each method. The authors drew attention to the dynamic approach to spatial data analysis. Analysis of time series allows, based on a series of cyclic and regular observations, to identify two components including the observation series, the trend and seasonality. The main result of our work is a mathematical fact that guarantees that along with the spatial changes in the output sets, we have full control over their Voronoi diagram.

Acknowledgements

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

I am grateful to dr hab. Agnieszka Lipieta, prof. UEK, for encouraging me to write this article and to the anonymous reviewers for helpful comments.

References

Brassel K.E., Reif D. [1979], *A procedure to generate Thiessen polygons*, "Geographical Analysis", 11, 3, pp. 289–303.

Coste M. [2000], *An introduction to o-minimal geometry*, Dottorato di ricerca in matematica, Universit`a di Pisa, Dipartimento di Matematica, Istituti editoriali e poligrafici internazionali.

Chazal F., Soufflet R. [2004], *Stability and finiteness properties of medial axis and skeleton*, "Dyn. Control Syst.", 10 (2), pp. 149–170.

Denkowska A., Denkowski M. [2020], *The Kuratowski convergence of medial axes and conflict sets,* preprint arXiv:1602.05422 [math.MG].

Denkowska Z., Denkowski M. [2012], *Kuratowski convergence and connected components*, "Journal of Mathematical Analysis and Applications", 387, pp. 48–65.

Denkowski M. [2016], *The Kuratowski convergence of medial axes*, preprint arXiv:1602.05422 [math.MG].

Fiedukowicz A., Gąsiorowski J., Olszewski R. [2014], *Biznes z szeroką perspektywą, czyli przestrzenne oblicze statystyki. Statystyka-zastosowanie biznesowe i społeczne*, Wydawnictwo Wyższej Szkoły Menedżerskiej w Warszawie im. Prof. L. Krzyżanowskiego.

Gibas P., Heffner K. [2007], *Analiza ekonomiczno-przestrzenna*, Wydawnictwo akademii Ekonomicznej w Katowicach.

Jastrzębska K. [2017], *Mapy średnich cen nieruchomości*, aleBank.pl-Portal Ekonomiczny.

Kisiała W., Rudkiewicz M. [2017], *Zastosowanie diagramu Woronoja w badaniu przestrzennych wzorców rozmieszczenia i dostępności sklepów dyskontowych*, „Przegląd Geograficzny", 89, 2, pp. 187–212.

Long X., Zhang Y., Chen Y. [2011], *Using Voronoi Diagram in Construction the Scope of Logistics Park Hinterland: an Engineering Application*, “Systems Engineering Procedia”, (2), pp. 69–76.

Okabe A., Boots B., Sugihara K., Chiu S.N. [2009], *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, John Wiley and Sons, Chichester-New York-Weinheim-Brisbane-Singapore-Toronto.

Reem M. [2011], *The geometric stability of Voronoi diagrams with respect to small changes of the sites*, preprint arXiv:1103.4125v2.

Roos T. [1993], *Voronoi diagrams over dynamic scenes*, “Discr. Appl. Math.”, 43, pp. 243– 259.

Śleszyński P. [2019], *Proposition of Location Tax on Retailas a Way to Rationally Shaping the Services Network*, “Studies of the Industrial Geography Commission of the Polish Geographical Society”, 33(1).

The Angle between the 2-dimensional Linear Regression Model Lines

Albert Gardoń[1]

Abstract

Applying the Mean Squares method to vertical and horizontal distances between points observed in a 2-dimentional sample (X,Y) and the relevant linear model we obtain two different straight lines, namely $\hat{y}(x)$ and $\hat{x}(y)$, i.e. the regression of Y with respect to X and the regression of X with respect to Y. The lines intersect at the point $(\overline{X},\overline{Y})$ at the angle that is obviously the lower, the greater the determination coefficient between X and Y. But it turns out that the angle depends not only on the correlation between features but also on the ratio of their sample dispersions. Based on the angle the region around $(\overline{X},\overline{Y})$ will be discussed, where the linear model may be reasonably applied, i.e. where the relative difference (with respect to the sample average) between the linear models doesn't exceed an assumed level ε^*, e.g. 5%.

Keywords: linear regression model, model evaluation, quality of prediction

JEL classification: C18, C30, C52, C53

[1] Albert Gardoń, Wroclaw University of Economics, e-mail: Albert.Gardon@ue.wroc.pl

1. Introduction

Let's assume a 2-dimensional sample (X,Y) is collected. The simplest and the most popular way of modelling the common behaviour of the features is the so-called linear regression model [Freedman, 2009; Rencher, Christensen, 2012]. It consists in fitting a straight line to the 2-dimensional data, treated as points with coordinates (x_i , y_i) on the real plane, such that the distances between the points and the line are minimized. Obviously, the shape of the line depends on the minimization criterion. The most popular one is based on the so-called Least Squares method (LS) where the sum of the squared vertical or horizontal distances between the points and the line is minimized, namely:

$$\sum_i (\hat{y}(x_i)-y_i)^2 \to min$$

or

$$\sum_i (\hat{x}(y_i)-x_i)^2 \to min$$

This leads to the solutions $\hat{y}(x)=a_x x+b_x$, called the regression of Y with respect to X, and $\hat{x}(y)=a_y y+b_y$, called the regression of X with respect to Y, with the coefficients:

$$a_x=R\frac{S_Y}{S_X}, \qquad a_y=R\frac{S_X}{S_Y},$$

$$b_x=\overline{Y}-a_x\overline{X},\ \ b_y=\overline{X}-a_y\overline{Y},$$

where R denotes the Pearson's correlation coefficient and S is the sample dispersion. It's easy to show that both lines always intersect at the point ($\overline{X}$,$\overline{Y}$). But since there are 2 different models in fact they will deliver different approximations except for this point. The difference will increase with the distance from ($\overline{X}$,$\overline{Y}$) and will be the greater, the greater angle between $\hat{y}(x)$ and $\hat{x}(y)$. In this paper we calculate the angle and discuss the consequences, including its impact on the relative difference between the models with respect to the sample average.

2. Case report

Since the slope of a straight line is the tangent of the angle between the horizontal X-axis and the line, then the tangent of the angle θ between the linear regressions $\hat{y}(x)$ and $\hat{x}(y)$ is calculated as the tangent of the difference between the angles θ_x and θ_y which the lines cut the X-axis at, respectively. However, the models should be presented firstly based on the same argument, say x, which requires the reformulation of the regression of X with respect to Y to:

$$\tilde{y}(x)=\frac{x-b_y}{a_y}. \tag{1}$$

That means, the second slope is now $\tilde{a}_y=\frac{1}{a_y}=\tan\theta_y$ and

$$\tan\theta=\tan(\theta_y-\theta_x)=\frac{\tan\theta_y-\tan\theta_x}{1+\tan\theta_y\tan\theta_x}=\frac{\frac{1}{a_y}-a_x}{1+\frac{a_x}{a_y}}=\frac{1-R^2}{R}\,\frac{S_X S_Y}{S_X^2+S_Y^2}. \qquad (2)$$

The angle θ may be also calculated as the angle between the direction vectors of the lines, i.e. $u_x=(1,R\frac{S_Y}{S_X})$ and $u_y=(R\frac{S_X}{S_Y},1)$. Without losing the generality let's omit the direction of the angle, thus, it may be calculated using the formula for the absolute scalar product $|u_x\circ u_y|$, namely:

$$\cos\theta=\frac{|u_x\circ u_y|}{\|u_x\|\|u_y\|}=\frac{r(\frac{1}{\sqrt{z}}+\sqrt{z})}{\sqrt{(1+r^2z)(\frac{r^2}{z}+1)}}=\frac{r(z+1)}{\sqrt{(r^2+1)(r^2z+1)}},\quad \theta\in\left[0,\frac{\pi}{2}\right],$$

where $z=\frac{S_Y^2}{S_X^2}$ is the ratio of sample variances and $r=|R|$ the absolute value of the Pearson's correlation coefficient. Let's now investigate the function f:

$$f(r,z)=\frac{r(z+1)}{\sqrt{(r^2+1)(r^2z+1)}},\quad r\in[0,1],\quad z\in(0,\infty),$$

which is equal to $\cos\theta$. Its partial derivatives are equal to:

$$\frac{\partial f(r,z)}{\partial r}=\frac{(1-r^2)z(z+1)}{\sqrt{(r^2+z)^3(r^2z+1)^3}},$$

$$\frac{\partial f(r,z)}{\partial z}=\frac{r(1-r^2)^2(z-1)}{2\sqrt{(r^2+z)^3(r^2z+1)^3}},$$

Since both denominators are always positive, then the signs of the derivatives above depend only on the signs of their numerators. Moreover, $\frac{\partial f}{\partial r}$ is always positive except for $r=1$ where it vanishes identically. In fact, at every point $(1,z)$ the function f reaches its global maximum with the value:

$$f(1,z)=\frac{z+1}{\sqrt{(1+z)(z+1)}}=1$$

It leads to the obvious and well known conclusion that the angle θ between the regression lines drops when the determination coefficient R^2 increases and in case of perfectly linearly correlated features X and Y, i.e. when $R^2=1$, the angle θ is equal to 0. Hence, both linear models become identical in such a case, of course, because all the observations are co-linear in this instance.

More interesting and not that trivial conclusions may be drawn from the behaviour of $\frac{\partial f}{\partial z}$. It vanishes identically again for $r=1$ as well as for $r=0$. At the former r there is the maximum of f calculated above and at the latter $\cos\theta$ reaches its global minimum with the value:

$$f(0,z)=\frac{0}{\sqrt{z}}=0,$$

corresponding to the maximal angle $\theta=\frac{\pi}{2}$ and the case when the linear models are perpendicular one to another and completely useless in practice. Eventually, excluding the trivial cases mentioned before, for any fixed $r\in(0,1)$ the partial derivative $\frac{\partial f}{\partial z}=0$ if $z=1$. Since the derivative is negative for $z<1$ and positive for $z>1$, then for any fixed $r\in(0,1)$ there is a local minimum of $f(r,\cdot)$ at $z=1$ with the value:

$$f(r,1)=\frac{r(1+1)}{\sqrt{(r^2+1)(r^2+1)}}=\frac{2r}{r^2+1}.$$

Thus, for any fixed nontrivial determination R^2, the greater difference between the sample dispersions S_X and S_Y, the lower angle θ between the linear regression models. To sum up, the angle θ between the linear regression models $\hat{y}(x)$ and $\hat{x}(y)$ fulfils the following inequality:

$$\theta\leq\arccos\frac{2|R|}{R^2+1}$$

and for $R^2\in(0,1)$ it reaches this upper bound only if $S_X=S_Y$.

To accomplish the investigation of the function f let's evaluate its limits when z tends to 0^+ or to $+\infty$. Straight calculations yield:

$$\forall r\in(0,1] \qquad \lim_{z\to 0^+} f(r,z)=\lim_{z\to 0^+}\frac{r(z+1)}{\sqrt{(r^2+z)(r^2z+1)}}=\frac{r}{\sqrt{r^2}}=1,$$

$$\forall r\in(0,1] \qquad \lim_{z\to +\infty} f(r,z)=\lim_{z\to +\infty}\frac{r\left(1+\frac{1}{z}\right)}{\sqrt{\left(\frac{r^2}{z}+1\right)\left(r^2+\frac{1}{z}\right)}}=\frac{r}{\sqrt{r^2}}=1.$$

It suggests that the angle will drop to about 0 even in case of insignificant linear dependence between the features X and Y if only the sample dispersions will differ extremely one from another. Besides, the last results mean also that the limits at $(r,z)=(0,0)$ or $(r,z)=(0,+\infty)$ don't exist.

3. Discussion

In this section we'd like to consider the impact of the results presented on the quality of approximations delivered by the linear model. The main purpose of such approximations is the prediction of one feature value (response feature, dependent feature) given the value of another one (explanatory feature, independent feature [Yan, 2009, Cohen et al., 2003]). Nevertheless, such an approximation is usually the most interesting far from the region where the center of the sample is observed, namely far from the point $(\overline{X},\overline{Y})$. A natural measure of the linear prediction quality is the determination coefficient R^2, called also the goodness-of-fit [Draper, Smith, 1998]. However, since there are always 2 equivalent linear models with the same goodness-of-fit then the difference between them should be taken into account when discussing the quality. As it was already mentioned in the previous sections the difference increases with the distance from the center of the sample and, of course, the greater the angle between the model lines, the faster the difference rises.

In order to compare both models it's good to present them based on the same argument, say x, similarly as in (1):

$$\tilde{y}(x)=\frac{x-b_y}{a_y}=\frac{1}{R}\frac{S_Y}{S_X}(x-\overline{X})+\overline{Y},$$
$$\hat{y}(x)=a_x x+b_x=R\frac{S_Y}{S_X}(x-\overline{X})+\overline{Y}.$$

Let's now denote by $\varepsilon(x)$ the absolute value of the relative difference between $\hat{y}(x)$ and $\tilde{y}(x)$ regarding the Y-sample average:

$$\varepsilon(x)=\left|\frac{\hat{y}(x)-\tilde{y}(x)}{\overline{Y}}\right|=\frac{S_Y}{S_X}\left|\frac{R^2-1}{R}\right|\left|\frac{x-\overline{X}}{\overline{Y}}\right|=\frac{S_Y}{S_X}\frac{1-R^2}{|R|}\frac{|x-\overline{X}|}{|\overline{Y}|},$$

that may be understood as a kind of a model error measure. Of course, it'll be required to keep this error at a reasonably low level, say ε^*, which may be equal e.g. to 5%. This leads to the formula for the maximal distance from the X-sample average where the error will be kept at the assumed satisfactory level $\varepsilon(x)\leq\varepsilon^*$:

$$|x-\overline{X}|\leq\varepsilon^*|\overline{Y}|\frac{S_X}{S_Y}\frac{|R|}{1-R^2}, \tag{3}$$

Besides, substituting (2) the evaluation above can be rewritten including the angle θ:

$$|x-\overline{X}|\leq\frac{\varepsilon^*|\overline{Y}|}{\left(1+\frac{S_Y^2}{S_X^2}\right)\tan|\theta|}.$$

As shown in the previous section the angle θ between the regression lines is, in general, the greatest when both sample variances are identical. This may be understood as the most pessimistic case. But at the same time, a high difference between the sample variances may mangle the result delivering virtually better prediction than it is indeed. Therefore, this may be another reason for unification of variances in the sample before the regression analysis, e.g. by means of the standardization. But assuming the sample averages and variances are close one to another the standardization isn't that necessary and a region around X-sample average can be estimated based on (3), within the error $\varepsilon(x)$ doesn't exceed a given level ε^* in case of a given determination R^2. Let's denote this region by $\left[\overline{X}-d\overline{X},\ \overline{X}+d\overline{X}\right]$. Table 1 contains its relative radius d, expressed by the percentage of the X-sample average, for chosen combinations of ε^* and R^2. It shows that in case of typical satisfactory determination coefficient $R^2 = 90\%$ the linear model delivers very good estimates ($\varepsilon(x) < 5\%$) even for arguments different from the mean by almost half of the sample average value.

Table 1. Distance from $\overline{X}$ expressed as the percentage of $\overline{X}$ where the error level ε^* isn't exceeded in case of the given goodness-of-fit R^2

	ε^*	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**	**0.10**
R^2											
0.70		0.028	0.056	0.084	0.112	0.139	0.167	0.195	0.223	0.251	0.279
0.75		0.035	0.069	0.104	0.139	0.173	0.208	0.242	0.277	0.312	0.346
0.80		0.045	0.089	0.134	0.179	0.224	0.268	0.313	0.358	0.402	0.447
0.85		0.061	0.123	0.184	0.246	0.307	0.369	0.430	0.492	0.553	0.615
0.90		0.095	0.190	0.285	0.379	**0.474**	0.569	0.664	0.759	0.854	0.949
0.93		0.138	0.276	0.413	0.551	0.689	0.827	0.964	1.102	1.240	1.378
0.95		0.195	0.390	0.585	0.780	0.975	1.170	1.365	1.559	1.754	1.949
0.97		0.328	0.657	0.985	1.313	1.641	1.970	2.298	2.626	2.955	3.283
0.98		0.495	0.990	1.485	1.980	2.475	2.970	3.465	3.960	4.455	4.950
0.99		0.995	1.990	2.985	3.980	4.975	5.970	6.965	7.960	8.955	9.950

Source: the author.

4. Conclusions

The angle between the linear regression lines depends on both, the goodness-of-fit and the ratio of sample variances and it's maximized when they're identical. It may suggest at least one of the samples should be properly scaled. This can be reached by the standardization as well which is a typical approach for data unification in this sense.

It's turned out that the region where the linear model gives satisfactory predictions doesn't depend on the sample standard deviation and may be expressed by a percentage of the sample average.

The research could be possibly generalized in the future for a multidimensional case, though, it will require the much more complicated matrix calculus and it's not clear if the results will have a compact analytical form.

References

Cohen J., Cohen P., West S.G., Aiken L.S. [2003], *Applied multiple regression/correlation analysis for the behavioral sciences,* (2nd ed.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Draper N.R., Smith H. [1998], *Applied Regression Analysis* (3rd ed.), John Wiley.

Freedman D.A. [2009], *Statistical Models: Theory and Practice,* Cambridge University Press.

Rencher A.C., Christensen W.F. [2012], *Methods of Multivariate Analysis,* Wiley Series in Probability and Statistics, 709 (3rd ed.), John Wiley & Sons.

Yan X. [2009], *Linear Regression Analysis: Theory and Computing,* World Scientific.

Analysis of Football Players' Labor Market Migrations Using Panel Gravity Models

Michał Górnik[1]

Abstract

The transfer of players between different sport clubs is essential and evokes strong emotions in the whole football industry. The objective of the paper is to identify, by means of a panel gravity model, global factors affecting football players' flows between clubs located in different countries. The approach, commonly used in modelling and predicting trade flows, considers spatial and temporal perspective – which can come in handy when determining the right career path for football players. The analysis is based on a transfer history from selected European leagues. The verified hypotheses concerned the impact of the sport level difference between leagues on the number of transfers, as well as the correlation between overall country economy and its top-tier league capability to attract football players.

Keywords: gravity model, football, data analysis, workforce migration

JEL classification: J44, C23

[1] Michał Górnik, Wroclaw University of Economics, e-mail: michal.gornik@ue.wroc.pl

1. Introduction

Football is said to be the most popular sport in the world, with its origins at the beginning of the second half of 19th century – the first football club, Sheffield Football Club, was founded in 1855. Since that time, universities and schools across Great Britain have treated this game as a possibility to compete and promote sports spirit and rivalry. From that earliest moment until now, the game has evolved and gained lots of popularity, as new organizations such as FIFA or UEFA started to organize international competitions in order to select the best national team or a city club. The rising popularity of football can be described by numbers regarding 2014 World Cup Final – it was viewed in real-time by more than billion people across 207 countries. The game has become a huge business and has given rise to multiple markets focused on television rights, marketing, bookmakers, souvenirs and many more with a growing economic importance.

This paper deals with the transfer market. Clubs that want to compete for the best season results need to acquire football players that could perform better than the previously hired athletes. As the form of currently involved players can decrease and lots of them can have a negative impact on clubs' budget due to remuneration, transferring them out is also a tool to remove redundant players from the team. Transfer market is a significant part of economy – according to "The UEFA Club Competition Landscape", the total value of global fees in summer 2019 [UEFA, 2019] transfer window amounted to 6.9 billion euro. Eighty clubs playing in European Cups were responsible for 58% of this value. The outcomes of analyses and methods presented in this paper can be applied in the football area by different users. Team coaches seek the best way to identify factors that could help them distinguish good and bad players in order to make best decisions related to players selling and buying. On the one hand, such an analysis can be carried out on physical aspects of players [Nevill et al., 2009] or their performance history. On the other hand, millions of football fans are looking for information concerning potential incoming transfer, using multiple sources that can spread the information very quickly and affect the player market value or tarnish their image [Caled, Silva, 2019].

One of the most common challenges of analyzing football transfers is the unreliability of the data related to transfer fees. In some cases, the fee paid is not disclosed by teams and it is estimated by means of multiple sources, including regression models or even market rumors. The English football league is very popular to conduct research on as the data was available and published in a consistent way prior to the data regarding leagues from other countries. An attempt to estimate the factors that affect the probability of being transferred with a corresponding fee was made by Carmichael et al. on Premier League transfer data for 1993–1994 season [Carmichael et al., 1999]. The problem of transfer fee estimation was also tackled by Ruijg and van Ophem with OLS model, which included variables concerning player information and performance statistics [Ruijg, van Ophem, 2015].

A continuous workforce movement between different teams and lots of available data, that can be analyzed with the current state of information technology and big data, makes footballers a great case study for labor market research. What makes sport players unique for this purpose is a vast number of data sources – we can retrieve all connections between players and clubs along with all team changes, as well as detailed player information such as name, full performance history, position and many more. Analyses of the labor markets of other professions – with the information sources available, including censuses and surveys, may not yield such accurate or meaningful results [Kahn, 2000].

The market of football player transfers has been analyzed by several researchers. They use mainly network methods to identify all patterns and connection between clubs. Li et.al. used transfer data among 23765 worldwide football clubs in order to construct a network reflecting employees mobility. They concluded that there is no club that acts as a transfer center, and bimodal distributions of node strengths indicate different transfer patterns for players [Li et al., 2019]. Another analysis conducted on top twenty European leagues aimed at recognizing the clubs that constitute a springboard with the use of network analysis. It turned out that Standard Liege is a club that sells most players at a good price to the top clubs from Europe [Tribušon, Lenič, 2016].

It is worth mentioning that transfers in the football labor market can also be quite controversial and may become a subject for studies that focus on social responsibility in sports management. A big debate whether fees paid by clubs to hire new players are ethical emerged after Gareth Bale, a Welsh winger, moved to Real Madrid from Tottenham London back in 2013. A Spanish club paid 100 million euro when the country suffered from an economic crisis and when the unemployment rate reached 28% [López Frías, 2018].

Practically all papers and analyses dealing with the football transfer market use the network algorithms to investigate the strengths of connections between different clubs (modelled as the nodes of the network) [see Li et al., 2019]. To the best of author's knowledge, there have been no attempts to apply gravity models to this market by now. The extension of the gravity models use – applied successfully (e.g. to migrations or labor force movements) – also seems a natural step for this specific market.

In the next section, the foundations of the gravity models are briefly summarized. The following section specifies the model with the explanatory quantities used. It also provides the results of the model estimation. The last section summarizes the findings and relates them to the results obtained within the network models approach.

2. Methods

The development of spatial econometrics and new geographical economy resulted in gravity definition and potential models that use laws of physics to describe and explain various socio-economic phenomena. The gravity is derived from the Newton's law of universal gravitation. It states that every point mass pulls every other point mass by a force directed along the line of centers for the two objects. The force is proportional to the product of the two masses, and inversely proportional to the square of the distance between them. Moving this law into the field of econometrics enabled to perform analyses in a new dimension. A spatial factor that is taken into account can boost the accuracy of the description of economic processes.

The force of influence from the Newton's law can be understood as different processes that are characterized by flows between regional entities. Gravity model is a tool to understand and test the marginal influence of multiple variables on immigration between countries [Lewer, van den Berg, 2008], or interregional migration within one country [Pietrzak et al., 2012]. Another application is modelling and checking the relevance of factors influencing trade exchange between countries [Tinbergen, 1962]. A topic modelled with the use of gravity panel data is also flow of quantified direct investments [Szczepkowska, Wojciechowski, 2002]. The issues related to using and estimating gravity models were addressed in various economic studies, including: [Westerlund, Wilhelmsson, 2011, Burger et al., 2009, Kaluza et al., 2010, Krings et al., 2009].

A regression model for gravity in a multiplicative form is described as follows [Suchecki, 2010]:

$$Y_{ij} = \alpha_0 \frac{M_i^{\alpha_2} P_i^{\alpha_3} M_j^{\alpha_4} P_j^{\alpha_5}}{D_{ij}^{\alpha_1}} \tag{1}$$

where Y_{ij} is the force of influence between i-th and j-th entity, elements of M_i and M_j are vectors of entity masses weights, P_i and P_j are variable vectors with entity masses. D_{ij} is the distance matrix.

3. Results

The dataset was collected from transfermarkt.com – a website that is the biggest and most popular data source for any football transfer market analysis. A list of transfers from 28 European countries is considered, taking into account every movement within the first two top-tier leagues in each country. Excluding all retirements and loan finishing a total number of 54222 transfers in the period of 2014–2020 is the basis of this analysis. As the purpose of this paper is to analyze international movement, all transfers that took place in the same country (i.e. a football player changed his team within the same country) were excluded. After performing all transformation and filter steps total number of 16950 observations issued. The variables that were used for a gravity model were chosen to reflect the migration gravity models [Lewer, van den Berg, 2008, Pietrzak et al., 2012] with respect to its football nature. They are as follows:

- GDP_origin, GDP_dest – Gross Domestic Product (EUR) per capita for a country of transfer's origin and destination.
- GEO_Dist – geographical distance between the origin and destination countries capitals.

- UEFA_Rank_diff – a Euclidean distance between UEFA country coefficients. It is a measure that takes into consideration the last 5-year performance of clubs from a country in European club cups. It reflects the distance between countries measured as their football level difference.
- (Lang) – Language – a binary variable indicating the fact of language sharing between a pair of countries.
- League_Size_origin, League_Size_dest – League size – the total number of players registered for a given season from all teams within each country's top-tier league.
- TMV_origin, TMV_dest – the total market value of football players within each country's top-tier league – estimated by transfermarkt.com.

There are multiple model classes that can be used to estimate regression on panel data. Two main kinds include models with fixed effects and models with random effects. In order to check the right model for a football transfer flow, a Hausman test was performed – the null hypothesis was rejected, and thus a model with random effects estimator is biased (the chi-squared statistic value of 222 with p-value lower than 0.0001). In order to decide whether a time, country pair or two ways effect should be included in the model's specification, the Breush-Pagan statistical test was performed. As a result, the model with both country pair and transfer year was selected. To estimate the gravity panel model with variables that are fixed in time, a Hausman-Taylor estimator was employed. Taking into consideration all the variables and logarithmic operation on (1), the following model was specified:

$$\ln(Y_{ij}) = \alpha_1 \ln\left(UEFA_{Rank_{Diff}}\right)_{ij} + \alpha_2 \ln\left(GDP_{Origin}\right)_{it} + \alpha_3 \ln\left(GDP_{Dest}\right)_{jt} + \alpha_4 \ln\left(TMV_{Origin}\right)_{it}$$
$$+ \alpha_5 \ln\left(TMV_{Dest}\right)_{jt} + \alpha_6 \ln\left(LeagueSize_{Origin}\right)_{it} + \alpha_7 \ln\left(LeagueSize_{Dest}\right)_{jt}$$
$$+ \alpha_8 Lang - \alpha_7 \ln\left(GeoDist\right) + B_{ij} + B_t + \varepsilon_{ijt}$$

where Y_{ij} is the football player flow from country i to country j, $\alpha_1 - \alpha_9$ are the model parameters, B_{ij} represents country pair effects, B_t time effects whereas ε is the random noise. Standardized regression coefficients are reported to assess the effect size of each variable on international flow size of football players.

Table 1. Estimated model parameters

	Variable	Estimate	Std. coefficient	p-value
1	UEFA_Rank_Diff	–0.26	–0.06	0.01
2	GDP_Origin	0.95	0.14	<0.01
3	GDP_Dest	1.09	0.16	<0.01
4	TMV_Origin	0.24	0.08	<0.01
5	TMV_Dest	–0.14	–0.05	0.08
6	LeagueSize_Origin	1.36	0.07	0.01

	Variable	Estimate	Std. coefficient	p-value
7	LeagueSize_Dest	0.80	0.04	0.13
8	Language	0.61	–	0.01
9	GeoDist	5.88	0.76	<0.01

Source: own calculations in R.

Determination coefficient is equal to 0.21, which means the goodness of fit is not satisfactory. The parameters associated with the power of top-tier league from the destination country turned out to be statistically non-significant. As far as the distance is concerned – both geographical and UEFA score distance were significant, which means that the spatial factor is important for football player migration modelling, and it also supports the hypothesis that sport level difference between top-tier leagues between countries is significant for the number of transfers. The signs of distance parameters estimates are different. The sign of estimate for difference between UEFA rank score is negative thus it can be concluded that the increase of sport level between leagues decreases the number of transfers between them. The top European leagues are exchanging players between each other, and it is hard for a football player to get transferred from a low-level European league to one of the so-called "Top five" – Spanish, English, German, French, and Italian. Given the negative sign of parameter of ln(GeoDist) in (2) and an estimate of geographic distance between transaction countries – which is positive – the number of transfer increases when the distance decreases. It means that football players tend to change the country to the ones nearby. This conclusion is also supported by a positive sign of the estimate of a language variable – it is easier to make a decision to move to the country with the same language and where adaptation process is faster, which is desired by the acquiring club. The biggest value of standardized regression coefficient for variable with geographic distance (0,76) means that this is the most influential factor.

As far as the issue of country's economy and their inflow of abroad players is concerned, both variables indicating gross domestic product per capita were statistically significant with positive signs. When the economy is in better condition, more players are switching their teams to those in different countries. This phenomenon is associated with poor countries and their inability to attract foreign players. Teams from such countries are forced to contract new players mainly from a domestic league, which is not likely to increase the league level in the long--term future. By analyzing the parameter estimate and their significance, we can conclude that a given country's economy indeed has an impact on attracting players. Analysis of standardized coefficients gives the conclusion that overall economy condition has bigger impact on predicted variable than market value of football players.

By analyzing a league size measured as the number of total registered players and their total estimated market values, we can observe that both variable parameters regarding destination league are statistically non-significant, as opposed to those associated with the origin country. As the total number of players in a given country increases, more of them are being pushed away to foreign countries. The same applies to the total market value of the origin country league – with this value growth we can observe a bigger outflux of football players.

4. Discussion

Gravity model for football transfers can provide some insights into factors affecting the transfer market. It can be concluded that an economic situation of a given country is a significant factor in terms of the influx and outflux of football players, as well as the difference in football level between top-tier leagues from given countries, measured by UEFA country coefficient.

Although this simple model has rather low predictive power $R^2 = 0.21$, it can be further increased, e.g. by including Poisson fixed-effect within the gravity model and variables concerning transfer market rumors and football management agencies that play an important role in the football marketplace. Another way to extend the research is to include the information about transfer fees – due to low public data credibility it was not included in the paper. The model that has been investigated here can also be applied to other kinds of sports, where transfer fees are not such a big issue for acquiring clubs. The flow of money for transfer fees can be estimated by means of methods used in this paper – under the condition of finding a reliable data source – it would be counterpart for the total trade value gravity models that are well adopted when modelling international trade relationships. Another factor that is yet to be verified when more data becomes available, is the impact on big international tournaments on the transfer market.

Acknowledgement

The author is very grateful for reviewers' comments which enabled to significantly improve his work.

References

Burger M., van Oort F., Linders G.J. [2009], *On the specification of the gravity model of trade: Zeros, excess zeros and zero-inflated estimation*, "Spatial Economic Analysis", 4(2), pp. 167–190, doi: 10.1080/17421770902834327.

Caled D., Silva M.J. [2019], *FTR-18: Collecting rumours on football transfer news*, CEUR Workshop Proceedings, 2482.

Carmichael F., Forrest D., Simmons R. [1999], *The Labour Market in Association Football: Who Gets Transferred and for How Much?*, "Bulletin of Economic Research", 51(2), pp. 125–150. doi: 10.1111/1467-8586.00075.

Kahn L.M. [2000], *The sports business as a labor market laboratory,* "Journal of Economic Perspectives", 14(3), pp. 75–94, doi: 10.1257/jep.14.3.75.

Kaluza P. et al. [2010], *The complex network of global cargo ship movements,* "Journal of the Royal Society Interface", 7(48), pp. 1093–1103 doi: 10.1098/rsif.2009.0495.

Krings G. et al. [2009], *Urban gravity: A model for inter-city telecommunication flows*, "Journal of Statistical Mechanics: Theory and Experiment", 7(2009), doi: 10.1088/1742-5468/2009/07/L07003.

Lewer J.J., Van den Berg H. [2008], *A gravity model of immigration*, "Economics Letters", 14(3), pp.75–94, doi: 10.1016/j.econlet.2007.06.019.

Li M.X., Zhou W.X., Stanley H.E. [2019], *Network analysis of the worldwide footballer transfer market*, "Epl", 125(1), pp. 1–7, doi: 10.1209/0295-5075/125/18005.

López Frías F.J. [2018], *Football transfers and moral responsibility, Soccer and Society*, doi: 10.1080/14660970.2016.1221822.

Moorhouse H.F. [1999], *The Economic Effects of the Traditional Transfer System in European Professional Football*, Research Unit in Leisure, Culture and Consumption.

Nevill A., Holder R., Watts A. [2009], *The changing shape of "successful" professional footballers,* "Journal of Sports Sciences", 27(5), pp. 419–426, doi: 10.1080/02640410802668676.

Pietrzak M.B., Drzewoszewska N. and Wilk J. [2012], *The Analysis of Interregional Migrations in Poland in the Period 2004-2010 Using Panel Gravity Model*, "Dynamic Econometric Models", 12, pp. 111–122, doi: 10.12775/dem.2012.008.

Ruijg J., van Ophem H. [2015], *Determinants of football transfers*, "Applied Economics Letters", 22(1), pp. 12–19. doi: 10.1080/13504851.2014.892192.

Suchecki B. [2010], *Ekonometria przestrzenna – Metody i modele analizy danych przestrzennych*, Wyd. C.H. Beck, Warszawa.

Szczepkowska A., Wojciechowski L. [2002], *Możliwości wykorzystania modelu grawitacji do analizy bezpośrednich inwestycji zagranicznych w Polsce*, Przemiany i perspektywy polskiej gospodarki w procesie integracji z gospodarką światową, pp. 267–282, Instytut Wiedzy, Warszawa.

Tinbergen J. [1962], *Shaping the World Economy Suggestions for an International Economic Policy*, The Twentieth Century Fund, 46(1), pp. 271–273, New York.

Tribušon M., Lenič M. [2016], *Identifying top football players and springboard clubs from a football player collaboration and club transfer networks*, Available at: http://arxiv.org/abs/1602.03664.

UEFA [2019], *The UEFA Club Competition Landscape*, Available at: https://www.uefa.com/MultimediaFiles/Download/uefaorg/Clublicensing/02/62/29/75/2622975_DOWNLOAD.pdf

Westerlund J., Wilhelmsson F. [2011], *Estimating the gravity model without gravity using panel data*, "Applied Economics", 43(6), pp. 641–649, doi: 10.1080/00036840802599784.

Selected Credit Risk Models

Stanisław Heilpern[1]

Abstract

The paper is devoted to credit risk models. Two kind of such models based on the generalized binomial distributions are presented. Firstly, we investigate the dependent credit risk, using copulas, mainly Archimedean. The influence of the degree of dependence on the number and value of lost credits is presented. Secondly, we study a case when the main parameter of model, the probability of the insolvent obligors, is uncertain. We treat such a parameter as a fuzzy number and we combine the randomness and fuzziness in this case. We investigate the number of lost credits, too.

Key words: risk credit, copula, generalized binomial distribution, fuzzy numbers, number of lost credits

JEL: G210

[1] Stanisław Heilpern, Wroclaw University of Economics, e-mail: stanislaw.heilpern@ue.wroc.pl

1. Introduction

The paper investigates credit risk models. We present two kinds of models based on the generalized binomial distributions. Firstly, we investigate the dependent credit risk [Frey et al., 2001; Heilpern, 2007; KMV-Corporation, 1997; McNeil et al., 2005]. In classical financial models the random variables are generally independent. This assumption is very convenient from mathematical point of view. However, it is difficult to expect that the bankruptcies of debtors are independent in practice. Common external factors: economic, climatic and political affect them. In particular, it may be the changes in stock prices and exchange rates, economic and political crises or some catastrophic events. We model the dependent structure using copulas, mainly Archimedean. We also study the value of the lost credits. The influence of the degree of dependence on the number and value of lost credits is presented. We use the Clayton and Spearman copulas to this end.

We also study a case when the main parameter of the model, the probability of insolvent obligors, is uncertain. We treat such parameter as a fuzzy number [Dubois, Prade, 1980; Heilpern, 1992] and we combine randomness and fuzziness in this case. In classical approach, we assume that this parameter is estimated on the basis of a random and representative sample. However, these conditions are not always fully met and we have some doubts as for the accuracy of value p of the obtained parameter. For instance, we can treat the value of this parameter as an uncertain value, a fuzzy number "about p" in this case [Heilpern, 2020]. We investigate the number of lost credits, too.

2. Dependence

Let us study a portfolio consisting of n obligors and define n Bernoulli random variables $Y_1, \ldots, Y_n$. They present the statuses of individual obligors: [McNeil et al., 2005; Heilpern, 2007]

$$Y_j = \begin{cases} 1 & \text{he not pay} \\ 0 & \text{he pays} \end{cases}.$$

We will denote a probability that the j-th obligor does not repay the credit as p_j and that he repays as $q_j = 1 - p_j$, i.e.

$$p_j = \Pr(Y_j = 1).$$

The joint distribution of Y_j is described by the probability mass function (p.m.f.)

$$f(i_1, \ldots, i_n) = \Pr(Y_1 = i_1, \ldots, Y_n = i_n),$$

where $i_j \in \{0, 1\}$ and cumulative distribution function (c.d.f.)

$$F(i_1, \ldots, i_n) = \Pr(Y_1 \le i_1, \ldots, Y_n \le i_n).$$

We will investigate the random variable

$$K = \sum_{j=0}^{n} Y_j .$$

It presents the number of insolvent obligors.

We allow that the random variables Y_j may be dependent, so we can describe the dependent structure by copula functions. The copula C is the link between the joint c.d.f. F and the marginal c.d.f. F_j. It satisfies the following relation:

$$F(i_1, \ldots, i_n) = C(F_1(i_1), \ldots, F_n(i_n)).$$

The marginal c.d.f. F_j are equal to

$$F_j(i_j) = \begin{cases} 1 & i_j = 1 \\ q_j & i_j = 0 \end{cases}.$$

Let $A \subset \{1, \ldots, n\}$ and $j \in A$ iff $i_j = 1$. We will use the notation $\mathbf{1}_A = (i_1, \ldots, i_n)$. Now, we assume, that the random variables Y_j have the same distribution, i.e. $q_j = q$ and that copula C is exchangeable, i.e. $C(u_1, \ldots, u_n) = C(u_{\pi(1)}, \ldots, u_{\pi(n)})$, for any permutation π of set $\{1, \ldots, n\}$. If the number of elements $|A| = |B| = k$, then the c.d.f. satisfies the following relation

$$F(\mathbf{1}_A) = F(\mathbf{1}_B) = F_{k,n}$$

and $F_{k,n} = C(\underbrace{1,\ldots,1}_{k}, \underbrace{q,\ldots,q}_{n-k})$. The p.m.f. is equal to [Cossette et al., 2002; Heilpern, 2007]

$$f_{k,n} = \sum_{j=0}^{k} (-1)^j \binom{k}{j} F_{k-j,n}$$

and we obtain

$$\Pr(K=k) = \sum_{j=0}^{k} (-1)^j \frac{n!}{(n-k)!j!(k-j)!} F_{k-j,n}.$$

The expected value and variance of K equal [Heilpern, 2007]:

$$E(K) = np, \qquad V(K) = npq + (n^2 - n)(f_{2,2} - p^2).$$

If the random variables Y_j are independent then the corresponding copula $\prod$ takes the form of $\prod(u_1, \ldots, u_n) = u_1 \cdot \ldots \cdot u_n$. The random variable K has the classical binomial distribution and we obtain

$$F_{k,n} = q^{n-k}, \quad f_{k,n} = p^k q^{n-k}, \quad \Pr(K = k) = \binom{n}{k} p^k q^{n-k}.$$

The strict, positive dependence, the opposite of independence, called the comonotonicity is done by copula $M(u_1, \ldots, u_n) = \min\{u_1, \ldots, u_n\}$. Then we have

$$F_{k,n} = \begin{cases} q & k<n \\ 1 & k=n \end{cases}, \quad f_{k,n} = \Pr(K = k) = \begin{cases} q & k=0 \\ 0 & 0<k<n \\ p & k=n \end{cases}.$$

Joe [1997] introduced the copula, called a Spearman copula, equal to the convex combination of the independency and comonotonicity: $(1 - \rho)\prod + \rho M$. The c.d.f. takes the form

$$F_{k,n} = \begin{cases} (1-\rho)q^{n-k} + \rho q & k<n \\ 1 & k=n \end{cases},$$

where ρ is a Spearman coefficient of the correlation.

The Archimedean copulas are the more popular copulas and often used in practice. They take a simple form induced by the generator φ:

$$C(u_1, \ldots, u_n) = \varphi^{-1}(\varphi(u_1), \ldots, \varphi(u_n)),$$

where generator is the decreasing, convex function, such that $\varphi(0) = \infty$ and $\varphi(1) = 0$. The c.d.f. is equal to $F_{k,n} = \varphi^{-1}((n-k)\varphi(q))$ in this case. The Clayton copula, one of such copulas, is defined by the following formula:

$$C(u_1, \ldots, u_n) = (u_1^{-\alpha} + \ldots + u_n^{-\alpha} - n + 1)^{-1/\alpha},$$

where parameter $\alpha > 0$ reflects the degree of dependence. The Kendall coefficient of correlation is equal to $\tau = \alpha/(\alpha + 2)$ and we obtain

$$F_{k,n} = (1 + (n-k)(q^{-\alpha} - 1))^{-1/\alpha}$$

in this case.

Example 1. We analyze a portfolio consisting of $n = 30$ obligors. Let the probability of solvency of every obligors is equal to $q = 0.8$ and the dependent structure of Y_j is described by Clayton copula. The distribution of K, the number of insolvent obligors, for different values of parameter α reflecting the degree of dependence is presented in Figure 1. We chose α equal to 0 (independency), 1.3, 6 and ∞ (comonotonicity). The corresponding values of Kendal coefficient of correlation and variance are equal to 0, 0.39, 0.75, 1 and 4.8, 33.57, 82.76, 144.

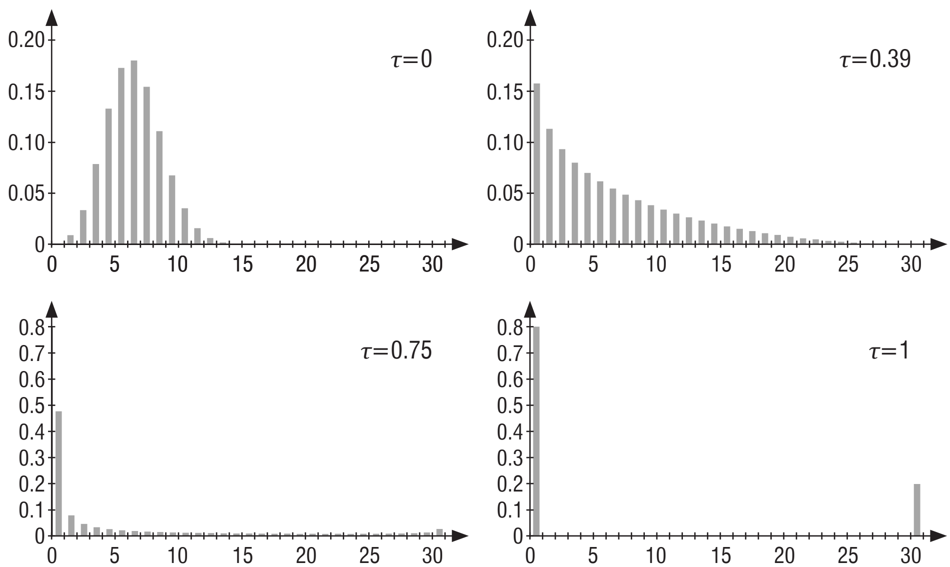


Figure 1. Distribution of the random variable K for different degrees of dependence – Clayton copula

Source: own elaboration.

We can see that the growth of value of parameter α (growth of the degree of dependence) implies a change in shape of the graph of p.d.f. of the distribution of insolvent obligors. It changes from the classical unimodal distribution with a small asymmetry to the distribution focusing on two points 0 and 1 only through the distributions with a right-sided asymmetry. For the independence, the distribution of K is condensed around the expected value equals 6. As the relationship increases, the mass of probability shifts to the left side until the lack of insolvent obligors obtains the most value of probability. For bigger dependencies the extreme values (there are no insolvents obligors or all of them are insolvent) are the most probable. In addition, the variance increases as the dependence of Y_j increases.

Now, we will study the value of insolvent credits. Let the random variables Z_j represent the value of lost credits and S is a global value of the lost credit, i.e.

$$S = \sum_{j=1}^{n} Y_j Z_j .$$

Then, if random variables Z_j and Y_j are independent and Z_j have the same distribution, then the moment generating function (m.g.f.) of the random variable S takes the form of [Cossette et al., 2002; Heilpern, 2007]

$$M_S(t) = \sum_{k=0}^{n} \binom{n}{k} f_{k,n} \left(M_Z(t)\right)^k .$$

Therefore S is the combination of the convolutions of the random variables Z_j and we have

$$F_S(x) = \sum_{k=0}^{n} \binom{n}{k} f_{k,n} F_Z^{*k}(x),$$

where F_S and F_Z are the c.d.f. of S and Z_j. In case of comonotonicity we obtain

$$F_S(x) = q + pF_Z^{*n}(x).$$

It depends on the n-th convolution of Z_j only. The expected value of random variable S is equal to $E(S) = npE(Z)$. We can compute the variance of S using the m.g.f. of this random variable.

Example 2. (continuation of Example 1) Let the value of a lost credit Z_j have the exponential distribution and $E(Z_j) = 2$. The graphs of the p.d.f. of random variable S, the global value of the lost credit, are presented in fig. 2 for α equal to 0, 1.3, 6, 20 and ∞. The corresponding values of Kendal coefficient equal 0, 0.39, 0.75, 0.91 and 1. We also present a graph of the p.d.f. of S for $q = 0.3$ and $\alpha = 4$ ($\tau = 2/3$).

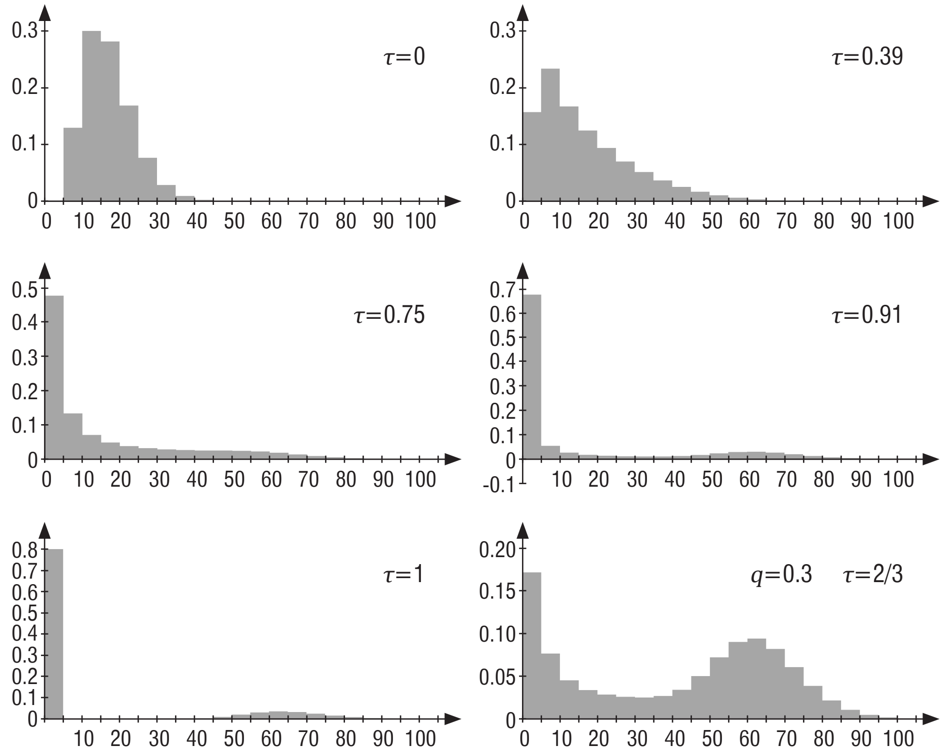


Figure 2. Distribution of the random variable S for different degrees of dependence – Clayton copula

Source: own elaboration.

We obtain $E(S) = 12$ and the variance of such random variable for different values of the parameter α is equal to 43.25, 158.07, 354.95, 503.78 and 600.00 respectively. We see, that the variance increases as the dependence (the value of parameter α) increases.

For smaller and medium dependencies we have a similar situation as for the random variable K. However, for bigger dependencies we obtain a smaller mass probability focused around 60. This effect is more visible for smaller probabilities of solvency obligors q, see the last graph on Figure 2.

When the dependence structure of Y_j is described by another copula we can obtain more irregular distributions of random variables K and S.

Example 3. (continuation of Example 1 and 2) Let the dependence structure of Y_j is characterized by Spearmen copula. We study four cases, when the Spearman coefficient of correlation ρ is equal to 0.1, 0.4, 0.6 and 0.8. The distributions of the random variables K and S are presented on Figures 3 and 4.

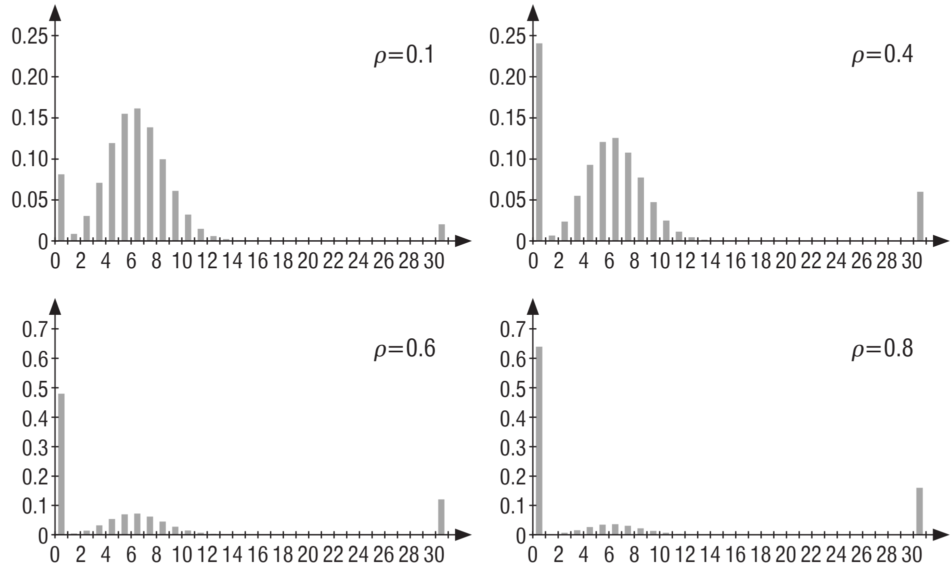


Figure 3. Distribution of the random variable K for different degrees of dependence – Spearman copula

Source: own elaboration.

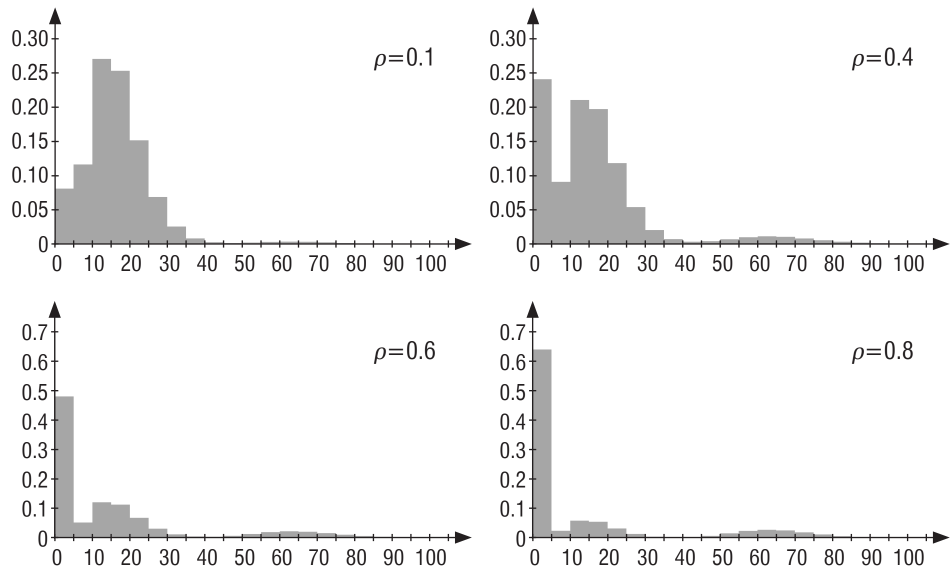


Figure 4. Distribution of the random variable *S* for different degrees of dependence – Spearman copula

Source: own elaboration.

We can see that the graph of the distribution of random variable K has the local maximum in near the expected value and the graph of S has two local maxima.

3. Fuzziness

First we recall some definition and notion connected with the fuzzy sets [Zadeh, 1965; Dubois, Prade, 1980]. The fuzzy set A defined on the space Z is described by its membership function μ_A: $Z \to [0, 1]$. The crisp sets $A_\alpha = \{z \in Z \mid \mu_A(z) \geq \alpha\}$, when $0 < \alpha \leq 1$, are called the α-cuts of the fuzzy set A and they univocally characterize it. Set A_1 is the core of A and A_0, the closure of set $\{z \in Z \mid \mu_A(z) > 0\}$, is the support of fuzzy set A.

We will use the fuzzy numbers in this section. They are the fuzzy subsets of the real line R and every α-cut A_α of them is the compact interval $[A_\alpha^L, A_\alpha^U]$ [Dubois, Prade, 1980]. The membership function of the trapezoidal fuzzy number $A = (a, b, c, d)$ is linear on the intervals $[a, b]$ and $[c, d]$. The interval $[a, d]$ is the support of the fuzzy number while $[b, c]$, when the membership function takes value one, is the core of it. If $b = c$, then we get the triangular fuzzy number $A = (a, b, d)$.

Let f: $Z \to$ R and A be a fuzzy subset of Z, then $f(A)$ has the following membership function

$$\mu_{f(A)}(x) = \sup_{f(z)=x} \mu_A(z).$$

It is so-called extension principle [Zadeh, 1975]. We can define the arithmetic operations * on fuzzy numbers using such a principle. The membership function of the fuzzy number $A * B$ is equal to $\mu_{A*B}(z) = \sup_{x*y=z} \{\min\{\mu_A(x), \mu_B(y)\}\}$. The borders of α-cuts of $A * B$ are defined by the borders of A and B, e.g. $(A+B)_\alpha^L = A_\alpha^L + B_\alpha^L$, $(A+B)_\alpha^U = A_\alpha^U + B_\alpha^U$.

The mean value of the fuzzy number A equals [Campos, Gonzalez, 1989;Heilpern, 1992]

$$MV(A) = \frac{1}{2}\int_0^1 \left(A_\alpha^L + A_\alpha^U\right) d\alpha.$$

The borders of α-cuts and the mean value for the trapezoidal fuzzy number $A=(a, b, c, d)$ take the following form

$$A_\alpha^L = (b-a)\alpha + a,\ A_\alpha^U = (c-d)\alpha + d,\ MV(A) = \frac{a+b+c+d}{4}.$$

Now we assume that the random variables Y_i, which described the status of an obligor, are independent. Then the random variable K, the number of insolvent obligors, has the Binomial distribution B(n, p). Let us also assume that we cannot make valid estimation of the main parameter p, the probability that the obligor does not repay the credit. We know the imprecision value of such a parameter only. In this case we can treat it as the fuzzy number P. This fuzzy number induces the fuzzy subset $\mathbf{K}$ on the family of binomial random variables, with the following membership function [Heilpern, 2020]:

$$\mu_{\mathbf{K}}(K) = \mu_P(p),$$

where $K \sim \mathrm{B}(n, p)$, using the extension principle. The sample size n, the number of obligors in our case, is fixed. If we know only, that the probability p is equal to "about p_0", then we can treat such information as the fuzzy number $P = (p_1, p_0, p_2)$.

We define the expected value of $\mathbf{K}$ using the extension principle. Then we obtain

$$\mu_{E(\mathbf{K})}(m) = \mu_P\left(\frac{m}{n}\right).$$

In a similar way we can determine the variance of $\mathbf{K}$ and the fuzzy probability $\Pr(\mathbf{K} \in B)$, where B is a crisp event. Let $f(p) = \Pr(K \in B)$, where $K \sim B(n, p)$, then

$$\mu_{\Pr(K \in B)}(p_0) = \sup_{f(p)=p_0} \mu_P(p).$$

Example 3. Let $n = 30$. We obtain information that the probability of event that the obligor does not repay the credit is equal to "about 0.2". We treat it as the fuzzy number $P = (0.15, 0.2, 0.3)$. The value of the membership function of $\mathbf{K}$ at $K \sim B(n, p)$ equals

$$\mu_{\mathbf{K}}(K)=\mu_P(p)=\begin{cases} 20p-3 & 0.15\le p\le 0.2 \\ -10p+3 & 0.2<p\le 0.3 \\ 0 & \text{otherwise} \end{cases}$$

The expected value of $\mathbf{K}$ is the triangular fuzzy number (4.5, 6, 9) with the mean value 6.375. We can interpret $E(\mathbf{K})$ as "about 6". But the variance is not the triangular fuzzynumber. Its α-cuts are equal to

$$V(\mathbf{K})_\alpha = [-0.075\alpha^2 + 1.05\alpha + 3.825, -0.3\alpha^2 - 1.2\alpha + 6.3],$$

because the function $g(p) = 30p(1 - p)$ defining the variance is increasing on [0.15, 0.3] and mean value of the variance is equal to 4.9625. The graph of the membership function of this fuzzy set is presented in Figure 3a.

Now, we compute the probability $\Pr(K = 7)$, i.e. the probability that there are seven insolvent obligors. Let $K \sim \mathrm{B}(n, p)$ and $f(p) = \Pr(K = 7)$. We obtain a maximal value of this function for $p_0 = 7/30$. We have $f(p_0) \approx 0.1700$ and $\mu_P(p_0) = 2/3$. Then the α-cuts of the fuzzy probability are equal to

$$\Pr(\mathbf{K}=7)_\alpha=\begin{cases} [f(0.05\alpha+0.15),\ f(7/30)] & 0\le\alpha\le 2/3 \\ [f(0.05\alpha+0.15),\ f(-0.1\alpha+0.3)] & 2/3<\alpha\le 1 \end{cases}.$$

The graph of the membership function of such a fuzzy set is presented on Figure 3b. The mean value of such fuzzy probability is equal to 0.1448. The graph is not continuous in $p = f(p_0)$.

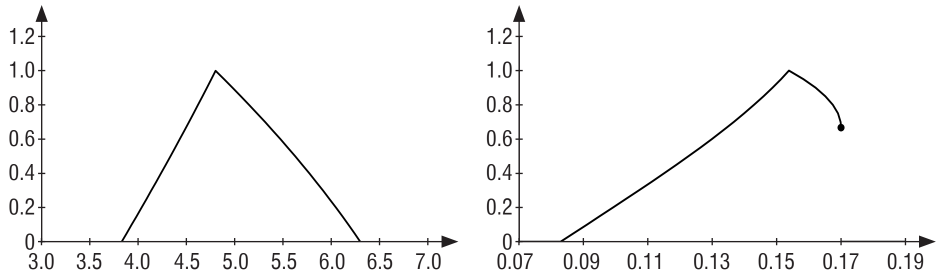


Figure 3. The membership functions of $V(\mathbf{K})$ and $\Pr(\mathbf{K} = 7)$

Source: own elaboration.

4. Conclusion

We presented two credit risk models. They are a generalization of the classical models. Firstly, we assumed that the statuses of individual obligors are dependent Bernoulli random variables. Therefore, we obtained a generalized, dependent binomial distribution. We investigated the number and the value of lost credits. We studied the distribution of such random variables with respect to the different degrees of dependency. These distributions are significantly different from the classical distributions with an independent assumption. In future works, we wish to widely investigate the model with a random number of obligors.

Next, we analyzed a situation when the main parameter of the model, the probability of insolvent obligors, is uncertain. We regarded it as the fuzzy number and investigated the number of lost credits in this case. We derived the fuzzy parameters of such a fuzzy variable, the fuzzy expected value and variable, and the probability of a fuzzy event. In the future, we would like to broadly expand this topic and conduct the empirical research.

References

Campos L.M., Gonzalez A. [1989], *A subjective approach for ranking fuzzy numbers*, “Fuzzy Sets and Systems”, Vol. 29, pp. 145–153.

Cossette H., Gaillardetz P., Marceau E., Rioux J. [2002], *On two dependent individual risk models*. “Insurance: Mathematics and Economics”, Vol. 30, pp. 153–166.

Dubois D., Prade H. [1980], *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, San Diego, CA.

Frey R., McNeil A.J., Nyfeler N.A. [2001], *Copulas and credit models* “Risk”, Vol. 10, pp. 111–114.

Heilpern S. [1992], *The expected value of a fuzzy number*, “Fuzzy Sets and Systems”, Vol. 47, pp. 81–86.

Heilpern S. [2007], *Dependent binomial distribution and its application in reinsurance and credits*, „Badania Operacyjne i Decyzje", 1, pp. 45–61. (in Polish).

Heilpern S. [2020], *Random variables with fuzzy parameters. Application in insurance.* Unpublished paper.

Joe H. [1997], *Multivariate Models and Dependence Concepts*, Chapman and Hall, London.

KMV-Corporation [1997], *Modelling Default Risk*, Technical Document.

McNeil A.J., Frey R., Embrechts P. [2005], *Quantitative Risk Management*, Priceton University Press, Priceton.

Nelsen R.B. [1999], *An Introduction to copulas*, Springer, New York.

Zadeh L.A. [1965], *Fuzzy sets*, "Inform. Contr.", pp. 338–353.

Zadeh L.A. [1975], *The concept of linguistic variable and its application to approximate reasoning,* "Inform. Sciences", Vol. 8, pp. 199–249, 301–357; Vol. 9, pp. 43–80.

Optimal Path in Growth Model

Marta Kornafel[1]

Abstract

In the paper we consider the Ramsey-Koopmans-Cass growth model where some of the parameters depend on time. The main focus of the work is on the dependence of the model and its solution to the perturbance of parameters. The conditions under which the solution of the perturbed model approximates the solution of the original model, are formulated.

Keywords: stability of optimal path, Γ-convergence, growth model

JEL classification: C02, C61, C62

[1] Marta Kornafel, Cracow University of Economics, e-mail: marta.kornafel@uek.krakow.pl

1. Introduction

The economic models often depend on parameters whose values may be given or estimated with some accuracy. Usually, having a set of data and using some statistical methods or numerical methods one produces a value or function representing these data. Therefore the method of how such an approximation is obtained may influence the solution of the considered model.

In this paper we study the influence of approximation of parameters in the Ramsey model on the stability of optimal path. Some of the parameters are constants, the other ones depend on time. They are approximated by sequences of values or sequences of functions that converge in an appropriate sense to their limits.

In order to provide the answer to the problem of stability of the optimal consumption path we recall the concept of Γ -convergence and the basic framework of its application in the context of the model considered. Next, the description of the Ramsey model is cited after Barro and Sala-i-Martin [2004]. The approximation of the model is constructed by considering the approximations of: the wage rate, the rate of returns, the growth rate of population and the degree of relative risk aversion, defining CRRA utility. Under the appropriate assumptions on those approximations we prove stability of the optimal consumption path.

2. Γ-convergence as a tool in perturbed optimization

In the perturbed optimization problems the crucial point is to assure the convergence (in a proper sense) of minima and minimizers of functionals to the respective minimum and minimizer of limiting functional. It was De Giorgi who studied this issue as the first. In his monograph [1984] he summarized his earlier works in this area and the developed concept of Γ-convergence. Parallely, Kazimierz Kuratowski extended the concept of Hausdorff metric, defined for nonempty and compact (therefore, in $\mathbb{R}^n$: closed and bounded) sets, to the case of closed sets. Moreover, he showed the equivalence between the Γ-convergence of functionals and the Kuratowski convergence of their epigraphs [Kuratowski, 1961].

Basing on Dal Maso [1993], we cite below the basic definitions and theorems on the epigraphical convergence to give a brief outline of both approaches. In what follows we denote by X a topological space[2] and $f_n\colon X \to \overline{\mathbb{R}}$ (where $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ is the extended real line) is a sequence of functions defined in it, while $N(x)$ denotes the family of open neighbourhoods of the point x.

[2] In the definitions and theorems we consider X to be a topological space, metric space or topological vector space. They are mathematical structures, which allows usage of different operations on the elements of set X, like measuring the distance or operating on vectors, respectively. For the sake of our further applications the reader may consider $X=\mathbb{R}^n$ with standard metric or space of continuous functions $X=C\big((0,\infty)\big)$ with the topology of uniform convergence, which both satisfy all the necessary requirements of each case.

Definition 1. Let function f: $X \to \mathbb{R}$ be given, where X is a topological space. The *epigraph* of function f is the set:

$$epi(f) \coloneqq \{(x,v) \in X\times\mathbb{R}: v \geq f(x)\}$$

Geometrically, the epigraph of function is the part of the Cartesian product $X\times\mathbb{R}$ above the graph of f. For a lower semicontinuous function f its graph is a closed subset of the Cartesian product $X\times\mathbb{R}$.

Definition 2. [Dal Maso, 1993, p. 38] The *lower* and *upper Γ-limits* of the sequence (f_n) are defined as follows:

$$f_{low-\Gamma} \coloneqq \Gamma-\liminf_{n\to\infty} f_n(x) \coloneqq \sup_{U\in N(x)} \liminf_{n\to\infty} \inf_{y\in U} f_n(y).$$

$$f_{upp-\Gamma} \coloneqq \Gamma-\limsup_{n\to\infty} f_n(x) \coloneqq \sup_{U\in N(x)} \limsup_{n\to\infty} \inf_{y\in U} f_n(y).$$

If $f_{low-\Gamma} = f_{upp-\Gamma} \coloneqq f_\Gamma$, then sequence (f_n) is said to be *Γ-convergent* to f_Γ and function f_Γ is *Γ-limit* of it.

We define now the *lower* and *upper Kuratowski limits* of the family of sets (net) $(E_n) \subset X$, when $n\to\infty$.

Definition 3. [Dal Maso, 1993, p. 41] The *lower* and *upper Kuratowski limits* of a sequence of sets are respectively:

$$x \in K-\liminf_{n\to\infty} E_n \Leftrightarrow \forall\, U \in N(x)\ \exists\, k \in \mathbb{N}\ \forall\, h\geq k:\quad U \cap E_h \neq \emptyset$$

$$x \in K-\limsup_{n\to\infty} E_n \Leftrightarrow \forall\, U \in N(x)\ \forall k \in \mathbb{N}\ \exists\, h\geq k:\quad U \cap E_h \neq \emptyset.$$

Clearly, $K-\liminf_{n\to\infty} E_n \subset K-\limsup_{n\to\infty} E_n$. If the converse inclusion holds as well, we denote the resulting set by $E \coloneqq K-\lim_{n\to\infty} E_n$ and call it the *Kuratowski limit* of (E_n) when $n\to\infty$. Therefore, (E_n) converges to some set E as $n\to\infty$, iff

$$K-\limsup_{n\to\infty} E_n \subset E \subset K-\liminf_{n\to\infty} E_n.$$

The next theorem establishes the relation between the Γ-convergence of a sequence of functions and the Kuratowski convergence of their epigraphs.

Theorem 1. [Dal Maso, 1993, p. 44] Let $f_{low-\Gamma}$ and $f_{upp-\Gamma}$ be respectively the lower and upper limits of a sequence of functions (f_n). Then:

$$epi(f_{low-\Gamma}) = K-\limsup_{n\to\infty} epi(f_n),$$

$$epi(f_{upp-\Gamma}) = K-\liminf_{n\to\infty} epi(f_n).$$

Therefore, (f_n) Γ-converges to f_Γ if and only if $epi(f_\Gamma) = K-\lim_{n\to\infty} epi(f_n)$.

The next theorem describes the fundamental role of Γ-convergence in optimization theory. This is a particular case of more general theorem 7.12 in Dal Maso [1993, p.73].

Let us denote by $M(f)$ the set of minimizers (possibly empty) of a function $f: X \to \overline{\mathbb{R}}$, i.e.:

$$M(f) := \left\{ x \in X:\ f(x) = \inf_{y \in X} f(y) \right\}.$$

Theorem 2. Assume that a sequence (f_n) is Γ-convergent to f_Γ. Then:

a. $K-\liminf_{n\to\infty} M(f_n) \subset M(f_\Gamma)$,

   i.e. any limit of a sequence of minima y_n is a minimizer of f_Γ.

b. if $K-\lim_{n\to\infty} M(f_n) \neq \emptyset$, then $M(f_\Gamma) \neq \emptyset$ and $\min_{x\in X} f_\Gamma(x) = \lim_{n\to\infty} \left(\inf_{x\in X} f_n(x)\right)$,

   i.e. if there exists a limit of a sequence of minima y_n, then the function f_Γ has at least one minimizer (this limit itself, maybe also some other) and minimum of f_Γ is approximated by minima y_n.

c. if f_Γ is a proper function[3], then

$$M(f_\Gamma) \subset K-\limsup_{n\to\infty} M(f_n),$$

   i.e. if f_Γ has at least one finite value, then any minimizer of this function is limit of a sequence of minimizers y_n.

It was shown in Kornafel [2018] that both pointwise and Γ - limits may exist, but may be different. Generally, they are independent concepts and it may happen that one of those limits exists while the other does not. However, under some assumptions both kinds of limits do exist and coincide. Below we gathered the theorems from Dal Maso [1993] describing the most important cases, usual in economic modelling. The subscript "p" is related to the pointwise limit.

Theorem 3.

a. $f_{low-\Gamma} \leq f_{low-p}$ and $f_{upp-\Gamma} \leq f_{upp-p}$.
   In particular, if both the Γ-limit f_Γ and the pointwise limit f_p exist, then $f_\Gamma \leq f_p$.

b. If each function f_n is continuous and sequence (f_n) converges uniformly[4] to a function f, then f is continuous and $f = f_\Gamma$.

c. If (f_n) is an increasing sequence of continuous functions, then $f_\Gamma = \sup_{n\in\mathbb{N}} f_n$.

d. Let X be a normed vector space. If (f_n) is a sequence of equi-bounded[5] in a neighbourhood of a point $x \in X$ and convex functions, then – provided the sequence (f_n) is convergent – $f_\Gamma = f_p$.

[3] In general theory, we may consider functions whose values are in the extended set of reals: $f: X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$, i.e. function which may take the value $+\infty$. A function is proper if it takes at least one finite value at some point $x \in X$, i.e. $\exists\, x \in X: f(x) \in \mathbb{R}$.

[4] Having (X,d) a metric space with distance function d, a sequence of functions $f_n: X \to \overline{\mathbb{R}}$ is *uniformly convergent* to a function $f: X \to \overline{\mathbb{R}} =$ iff $\lim_{n\to\infty} \sup_{x\in X} d(f_n(x), f(x)) = 0$.

[5] Family (or sequence) of functions (f_n), $f_n: X \to \mathbb{R}$, is equi-bounded iff there exists a constant $M > 0$, which bounds any function, i.e. for any $x \in X: |f_n(x)| < M$.

The next theorem determines when the properties of convexity and homogeneity are inherited by Γ-limits.

Theorem 4. Let X be a topological vector space over the real numbers. Then:

a. if (f_n) is a sequence of convex functions, then $f_{upp-\Gamma}$ is a convex function. In particular, for a Γ-convergent sequence the Γ-limit f_Γ is convex.

b. if (f_n) is a sequence of positively homogeneous of degree k functions[6], then both $f_{low-\Gamma}$ and $f_{upp-\Gamma}$ are positively homogeneous of degree k. In particular, if a sequence (f_n) Γ-converges to f_Γ then f_Γ is positively homogeneous of degree k.

Remembering the obvious facts:

1. if a function f attains minimum at a point x_0, then the function $-f$ attains maximum at this point;
2. if a function f is increasing on a set A, then the function $-f$ is decreasing on this set;
3. if a function f is convex on a set X, then the function $-f$ is concave on the same set;
4. $\limsup_{n\to\infty}\left(-f_n(x)\right)=-\liminf_{n\to\infty} f_n(x)$ and $\liminf_{n\to\infty}\left(-f_n(x)\right)=$ $-\limsup_{n\to\infty} f_n(x)$;

all the theorems above may be easily reformulated to characterize maximization problems in the context of Γ-convergence.

3. Modified Ramsey model

We consider now the household's optimization problem of the Ramsey model (in the spirit of [Barro, Sala-i-Martin, 2004]), in which the parameters are given with some approximation. In contrast to Kornafel [2018] we allow some of them to be continuous functions of time. All the introduced functions are assumed to be smooth.

The households are treated as identical – their preferences are the same, they share the wage rate $\omega=\omega(t)$, have the same rate of returns $r=r(t)$ and the same assets per person. The population grows at the rate $n>0$, so $L(t)=L(0)\cdot e^{nt}$. For simplicity we take $L(0)=1$. The total consumption is denoted by $C=C(t)$, while consumption per capita is $C(t):=\frac{C(t)}{L(t)}$. We consider CRRA utility, i.e.:

$$u(c)=\frac{c^{1-\theta}-1}{1-\theta}, \qquad \theta\in(0,1) \tag{1}$$

[6] Having X a vector topological space, function $f\colon X\to\overline{\mathbb{R}}$ is positively homogeneous of degree k if for any $t>0$ and any $x\in X$ it holds that $f(tx)=t^k f(x)$.

which satisfies the usual monotonicity and concavity assumptions and meets Inada conditions. Therefore the households face the problem of choosing such a consumption path $c(t)$ to maximize the intertemporal utility functional (with a discount rate $\rho(t)>n$, $\forall t\geq 0$):

$$U[c]=\int_0^{\infty} u\Big(c(t)\Big)e^{-(\rho(t)-n)t}\,dt, \tag{2}$$

taking into account the budget constraints and $c(t)\geq 0$. The budget constraints are given by the dynamics of household's assets per person $a(t)$:

$$\frac{da}{dt}=[r(t)-n]\cdot a(t)+\omega(t)-c(t). \tag{3}$$

The transversality condition is that the present value of assets is asymptotically nonnegative:

$$\lim_{t\to\infty} a(t)\cdot e^{-\left(\int_0^t r(\tau)d\tau-nt\right)}\geq 0, \tag{4}$$

Thanks to (4) and Pontryagin Maximum Principle we can derive the optimal consumption path:

$$c^*(t)=c(0)\,\exp\left(\frac{1}{\theta}\Big(\int_0^t r(\tau)d\tau-\rho(t)\Big)\right) \tag{5}$$

For the derivation of the constraints and detailed solution of the model see [Barro, Sala-i-Martin, 2004, p. 88–93].

Consider now the Ramsey model with "disturbed" parameters. The rates $r(t)$, n, $\rho(t)$, $\omega(t)$ and parameter θ are given with some approximation, which may depend on measurement rules. Denote those approximate values by $r_\varepsilon(t)$, n_ε, $\rho_\varepsilon(t)$, $\omega_\varepsilon(t)$ and θ_ε, respectively. Assume that for any ε and any $t\geq 0$, $\rho_\varepsilon(t)>n_\varepsilon$ and $\theta_\varepsilon\in(0,1)$. With the increase in the accuracy of measurement, the approximate values b_ε tend to the actual value b for $b\in\{n,\theta\}$ and f_ε tend to f for $f\in\{r,\rho,\omega\}$. With approximate values b_ε and f_ε the household's maximization problem is to maximize the functional $U_\varepsilon[c]=\int_0^{\infty} u_\varepsilon\Big(c(t)\Big)e^{-(\rho_\varepsilon(t)-n_\varepsilon)t}\,dt$, where $u_\varepsilon(c)=\dfrac{c^{1-\theta_\varepsilon}-1}{1-\theta_\varepsilon}$ subject to the constraints analogous to (3) and (4). By analogous reasoning as above, for each $\varepsilon>0$ we obtain the optimal solution: the consumption per capita path

$$c_\varepsilon{}^*(t)=c(0)\,\exp\left(\frac{1}{\theta_\varepsilon}\Big(\int_0^t r_\varepsilon(\tau)d\tau-\rho_\varepsilon(t)\Big)\right).$$

The next theorems give a positive answer for the basic questions that arise: whether or not $c_\varepsilon{}^*$ converges to c^* and $U_\varepsilon[c_\varepsilon{}^*]$ converges to $U[c^*]$, i.e. whether or not the optimal path $c_\varepsilon{}^*$ approximates the actual optimal path c^* and $U_\varepsilon[c_\varepsilon{}^*]\approx U[c^*]$.

First, notice that the sequence of assets allocations $(a_\varepsilon)\subset C^1$, whose terms satisfy for any $\varepsilon>0$ analogon of (3), is uniformly convergent to $a(t)\in C^1$ (also thanks to uniform convergence of (ω_ε). Therefore the constraint holds in the limit. Having this, it is easy to justify that if $r_\varepsilon\rightrightarrows r$, then transversality condition (4) holds in the limit as well.

Theorem 5. If $\lim_{\varepsilon\to 0} n_\varepsilon = n$ and $\lim_{\varepsilon\to 0}\theta_\varepsilon=\theta$, and $r_\varepsilon \rightrightarrows r$, $\rho_\varepsilon \rightrightarrows \rho$, then U_ε converges to U uniformly.

Proof. The sequence of utility functions (u_ε) converges uniformly to the function u. Indeed:

$$|u_\varepsilon(c)-u(c)|=\left|\frac{c^{1-\theta_\varepsilon}-1}{1-\theta_\varepsilon}-\frac{c^{1-\theta}-1}{1-\theta}\right|=$$

$$=\left|\frac{\left(c^{1-\theta_\varepsilon}-c^{1-\theta}\right)+\left(\theta_\varepsilon\cdot c^{1-\theta}-\theta\cdot c^{1-\theta_\varepsilon}\right)+\left(\theta-\theta_\varepsilon\right)}{\left(1-\theta_\varepsilon\right)\left(1-\theta\right)}\right|\leq$$

$$\leq\frac{|c^{1-\theta_\varepsilon}-c^{1-\theta}|+|\theta_\varepsilon\cdot c^{1-\theta}\pm\theta_\varepsilon\cdot c^{1-\theta_\varepsilon}-\theta\cdot c^{1-\theta_\varepsilon}|+|\theta-\theta_\varepsilon|}{\left(1-\theta_\varepsilon\right)\left(1-\theta\right)}\leq$$

$$\leq\frac{\left(1+\theta_\varepsilon\right)|c^{1-\theta_\varepsilon}-c^{1-\theta}|+\left(c^{1-\theta_\varepsilon}+1\right)|\theta-\theta_\varepsilon|}{\left(1-\theta_\varepsilon\right)\left(1-\theta\right)}.$$

If $\theta_\varepsilon\to\theta$, then the numerator of the fraction tends to zero, so $u_\varepsilon \rightrightarrows u$. We are ready to prove the uniform convergence of operators U_ε to U:

$$|U_\varepsilon[c]-U[c]|=\left|\int_0^\infty u_\varepsilon\left(c(t)\right)e^{-(\rho_\varepsilon(t)-n_\varepsilon)t}\,dt-\int_0^\infty u\left(c(t)\right)e^{-(\rho(t)-n)t}\,dt\right|\leq$$

$$\leq\int_0^\infty |u_\varepsilon\left(c(t)\right)e^{-(\rho_\varepsilon(t)-n_\varepsilon)t}\pm u\left(c(t)\right)e^{-(\rho_\varepsilon(t)-n_\varepsilon)t}-u\left(c(t)\right)e^{-(\rho(t)-n)t}|\,dt\leq$$

$$\leq\int_0^\infty\left[\left|u_\varepsilon\left(c(t)\right)-u\left(c(t)\right)\right|\cdot e^{-\left((\rho_\varepsilon(t)-\rho(t))-(n_\varepsilon-n)\right)t}+u\left(c(t)\right)\right]$$

$$\cdot\left|e^{-\left((\rho_\varepsilon(t)-\rho(t))-(n_\varepsilon-n)\right)t}-1\right|\Big]\cdot\left|e^{-(\rho-n)t}dt\right.$$

Due to the uniform convergence of (u_ε) and (ρ_ε), with the convergence of sequences (n_ε) the integrand tends to zero function, so $U_\varepsilon \rightrightarrows U$, what proves the theorem.

Theorem 6. The functional U is the Γ-limit of the sequence U_ε when $\varepsilon\to 0$.

Proof. The theorem 6 is the immediate consequence of the theorems 5 and 3b.

Corollary

If $\lim_{\varepsilon\to 0} n_\varepsilon=n$ and $\lim_{\varepsilon\to 0}\theta_\varepsilon=\theta$, and $r_\varepsilon \rightrightarrows r$, $\rho_\varepsilon \rightrightarrows \rho$, $\omega_\varepsilon \rightrightarrows \omega$, i.e. if the parameters of the model are given with approximations whose absolute errors tend to zero, then the corresponding consumption path indeed approximates the optimal "theoretical" consumption path and the obtained value of utility is close to actual maximum.

4. Conclusions

The aim of this paper is to study the stability of optimal solution in the general Ramsey-Koopmans-Cass model. We showed that if the approximation of parameters in the model is such that the absolute error of approximation tends to zero, then the approximate solution of the model is a good approximation of actual solution and can be trustworthy for economic usage as it leads to "almost"-maximization of CRRA utility functional.

Acknowledgements

The research, whose results are presented in this paper, is supported by subsidy granted to Cracow University of Economics.

References

Attouch H. [1984], *Variational Convergence for Functions and Operators,*. Pitman, London.

Barro R.J., Sala-i-Martin X. [2004], *Economic Growth*, MIT, 2nd ed.

Dal Maso G. [1993], *An introduction to* Γ*-convergence*, Birkhäuser.

De Giorgi E. [1984], *G-operators and* Γ*-convergence*, Proc. of the International Congress of Mathematicians (Warszawa 1983), 1175–1191, North Holland, Amsterdam.

Kornafel M., [2018], *When do Optima Converge to Optimum?* "Argumenta Oeconomica Cracoviensia", 19, pp. 23–34.

Kuratowski K. [1961], *Topologie I et II*, PWN, 2nd edition, Monogr. Mat.

Predictive Power Comparison of Bayesian Homoscedastic vs. Markov-switching Heteroscedastic VEC Models

Łukasz Kwiatkowski[1]

Abstract

In the paper we examine the forecasting performance of Bayesian vector error correction models (allowing for long-term relationships between modelled variables) featuring two- and three-state Markovian breaks in the conditional covariance matrix to capture time-varying volatility, typically recognized in macroeconomic data. Predictive performance of the models is evaluated within the probabilistic paradigm of forecasting, with the accuracy of density forecasts measured through the log predictive score and Bayes factors, while also using Probability Integral Transform (PIT) to assess the calibration of the forecasts. In the empirical study, we conduct a series of *ex-post* prediction experiments within the so-called small models of monetary policy, using macroeconomic data for the US economy. The results indicate considerable gains in the predictive power of VEC models with Markov-switching heteroscedasticity (in comparison to the homoscedastic VEC systems), although the three-state models provide no further improvements of density predictions as compared to the two-state specifications.

Keywords: cointegration, regime switching, probabilistic forecasting, predictive score, predictive Bayes factor

JEL Classification: C11, C32, C53

[1] Łukasz Kwiatkowski, Cracow University of Economics, e-mail: lukasz.kwiatkowski@uek.krakow.pl

1. Introduction

For an effective forecasting of any time series by means of some statistical model the latter needs to capture key characteristics of the data at hand. Macroeconomic time series, which are of this paper's main focus, typically display two features[2]: non-stationarity (due to the presence of stochastic trends) and conditional heteroscedasticity, with the latter one having already been commonly associated not only with financial and commodity markets. Dealing with non-stationary processes jointly for different variables usually requires the use of cointegration analysis, with the underlying vector autoregression (VAR) model in its vector error correction (VEC) form. Then, to account also for the other feature some time-variability needs to be introduced into the conditional covariance matrix of the observations, with typical choices including a variety of multivariate GARCH (MGARCH) or stochastic volatility (MSV) processes, both classes enabling continuously-valued (rather than discrete) changes of conditional variances and/or correlations. Recently, Wróblewska and Pajor [2019] examined the predictive performance of VEC models equipped with hybrid MSV-MGARCH structures (introduced by Osiewalski [2009], Osiewalski, Pajor [2009]), using macroeconomic data for the Polish economy. As evidenced in the cited work, extending homoscedastic VEC models with time-varying volatility dramatically improves their forecasting abilities model (as measured by the log predictive score, energy score, and also mean squared forecast error).

In this paper we shift the attention to a qualitatively different and simpler (as compared with MGARCH, MSV or their hybrids) approach to modelling conditional heteroscedasticity in cointegrated VAR/VEC systems. We conjecture that in the case of macroeconomic (as opposed to financial) time series it may be empirically 'sufficient' (for prediction) to enable discrete rather than continuously-valued shifts in the multivariate volatility. Following this line of reasoning, we allow the conditional covariance matrix to switch between either two or three regimes according to a homogenous and ergodic Markov chain.

Although the concept of Markov-switching (MS) time series models has been well--established and present in the literature for a long time (since the seminal paper by Hamilton 1989), papers devoted to forecast evaluation of MS-VEC models for macroeconomic data are almost exclusively limited to the point (rather than density) prediction [Clarida et al., 2003; Sarno et al., 2005; Psaradakis, Spagnolo, 2005]; in the latter the authors, apart from the point forecasts, also evaluate the calibration of density forecasts *via* the Probability Integral Transform (PIT). Therefore, following the recent shift of the forecasting paradigm from the point to probabilistic prediction, the aim of this study is an empirical evaluation of predictive densities performance of VEC models with Markov-switching heteroscedasticity (VEC-MSH) in comparison with homoscedastic VEC structures.

As for the statistical inference framework, we resort to the Bayesian approach to estimation and prediction, similarly as Wróblewska and Pajor [2019]. Admittedly,

[2] Potentially, along with some seasonal or cyclical patterns, which remain beyond the scope of this research.

the Bayesian setting is the most suitable while dealing with latent processes (like stochastic volatility or hidden Markov chains). Moreover, it handles coherently the parameter uncertainty while producing predictive densities, which in the case of regime-changing models may be essential for the sake of their forecasting performance (as conjectured by Psaradakis and Spagnolo [2005]). The latter one is evaluated here through the log predictive score (LPS), which underlies the so-called predictive Bayes factor. We also examine PIT histograms to assess forecasts' densities calibration [Geweke, Amisano, 2010; Gneiting, Raftery, 2007; Gneiting et al., 2007].

2. VEC models with Markov-switching heteroscedasticity

An n-variate VAR(k) model with Markov-switching conditional covariance matrix can be written in its VEC (henceforth VEC-MSH) form as:

$$\Delta x_t = \widetilde{\Pi} x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi D_t + \varepsilon_t, \quad t = 1, 2, \ldots, T, \tag{1}$$

$$\varepsilon_t | \psi_{t-1}, S_t, \theta \sim N(0, \Sigma_t), \tag{2}$$

where x_t is an n-variate random vector, $\{\varepsilon_t\}$ is a vector white noise with some unconditional covariance matrix Σ, matrix D_t comprises deterministic variables (such as the constant, trend and seasonal dummies), $\widetilde{\Pi}$, Γ_i and Φ are real-valued matrices of parameters, all collected in θ, and ψ_{t-1} denotes the past of the process $\{x_t\}$ up to time $t-1$. The matrix $\widetilde{\Pi}$ decomposes as $\widetilde{\Pi} = \alpha\widetilde{\beta}'$, with α $(n \times r)$ storing the adjustment coefficients and $\widetilde{\beta}$ $(m \times r,\ m \geq n)$ pertaining to the cointegration relationships, once they exist (then $r < n$ is their number). Note that $m > n$ only under the deterministic components restricted to the cointegration relationships (otherwise, $m = n$). The initial conditions $x_{-k+1}, x_{-k+2}, \ldots, x_0$ are assumed to be known and set as pre-sample observations; see [Wróblewska, Pajor, 2019] and the references therein.

The sequence $\{S_t\}$, where $S_t \in \{1, 2, \ldots, K\}$, forms a K-state homogenous and ergodic Markov chain with (time-invariant) transition probabilities $p_{ij} \equiv \Pr(S_t = j | S_{t-1} = i, \theta) \in (0, 1)$, $\sum_{j=1}^{K} p_{ij} = 1$ (for all $i, j \in \{1, 2, \ldots, K\}$) forming the transition matrix $P = [p_{ij}]_{i,j=1,2,\ldots,K}$, the $K(K-1)$ free elements of which are also stored in θ. This latent process governs the switches between K regimes, each featured by 'its own' conditional (given ψ_{t-1} and θ) covariance matrix $\Sigma_t \equiv \Sigma_{S_t}$ of the error term ε_t (see Eq. 2). Note that in our setting we restrict the regime changes only to the volatility, thereby restricting the other parameters of the VEC structure to be time-invariant, assuming that possible long-term relationships and short-term adjustments hold constant over the entire sample.

3. Bayesian model specification, estimation and prediction

The methodology of Bayesian Markov-switching VEC models has been developed by Jochmann and Koop [2015] and we follow their approach to a large extent, with minor modifications to tailor our framework to the one presented in [Wróblewska, Pajor, 2019][3]. As Bayesian modelling requires specification of the prior distributions for model parameters, we adopt their structure for the VEC part from the latter of the two above-mentioned papers, while also assuming that Σ_i's (i=1, 2, …, K) follow the inverse Wishart distribution – the same as the one considered in [Wróblewska, Pajor, 2019] for Σ in the homoscedastic VEC models. For the rows of the transition matrix in all the models with Markov regime changes we impose the uniform (over the unit simplex) priors, which in the two-state models boil down to the uniform distributions for p_{11} and p_{22}, whereas in the models with $K>2$ – to the Dirichlet (1, 1, …, 1) distributions for $p_i = (p_{i1}\ p_{i2}\ \dots\ p_{iK})$, i=1, 2, …, K.

Bayesian estimation of the models at hand necessitates the use of MCMC methods, including the Gibbs sampler (in all the models) and the *Forward-Filtering-Backward--Sampling* scheme (developed by Carter and Kohn [1994]) for sampling latent Markov chain's state variables [Jochmann, Koop, 2015]. Additionally, to handle the label switching, a problem inherent to mixture models, we use the permutation sampler designed by Frühwirth-Schnatter [2006], enforcing an inequality restriction for conditional variances of a selected variable across the regimes, $Var(x_{ti}|S_t=1,\psi_{t-1},\theta)>Var(x_{ti}|S_t=2,\psi_{t-1},\theta)>\ldots>Var(x_{ti}|S_t=K,\psi_{t-1},\theta)$ for a given $i\in\{1, 2, \ldots, n\}$.

Although requiring additional simulations at each MCMC step, prediction within the VEC and VEC-MSH models is quite straightforward, owing to a sequential structure of the likelihood.

4. Empirical analysis

The empirical analysis to follow is based on various VEC and VEC-MSH specifications of the so-called small models of monetary policy for the US economy. The models comprise three variables (n=3): inflation rate of consumer prices, unemployment rate and short-term interest rate, which relate this analysis to an influential paper by Primiceri [2005] (see also [Wróblewska, Pajor, 2019] for an analysis for Polish data). As regards the latter variable, instead of nominal rates we choose the so-called shadow interest rates, calculated according to Wu and Xia [2015], to avoid the restraint of the zero lower bound (with the nominal rates hovering near the bound since 2011). In this study we use quarterly, seasonally adjusted data, covering the period 1960:Q1–2015:Q4 (T=224 observations). Modelled time series are displayed in the top panel of Fig. 1.

[3] Details can be provided by the author upon request.

The predictive performance (in the sense of density forecasts) of the models under consideration is evaluated *via* series of *ex-post* one-quarter-ahead density predictions, based on a sequence of expanding (recursive) samples, with each model being reestimated upon the arrival of each new observation. For the sake of the *ex-post* prediction analysis, we spare the final $N=56$ observations, so that the experiment covers 2002:Q1–2015:Q4. As can be inferred from Fig. 1, this period witnesses some evident regime changes (resulting in particular from sharp movements of the unemployment and inflation rates), with a clear regime distinction in the two-state model, and somehow less unequivocal assignment in the three-state case[4].

In each of the models under study we assume that the order of the underlying VAR process equals $k=2$ [Primiceri, 2005; Jochmann, Koop, 2015; Wróblewska, Pajor, 2019][5]. We consider two alternative specifications of the constant term in Eq. (1): either an unrestricted constant (conventionally denoted as $d=3$), or a constant restricted to the cointegration relationships ($d=4$). As for the number $r<n$ of cointegration relations, we consider all of its possible values, i.e. $r \in \{0, 1, 2\}$, as well as the case of a stationary VAR system (i.e. $r=n=3$). Models with given d and r are labelled as VEC(d, r) and VEC(d, r)-MSH(K), with K indicating the number of regimes in the Markov-switching specifications. In the latter we allow for switches between either two ($K=2$) or three ($K=3$) states, thereby discriminating between either high and low volatility regimes, or also allowing for an additional, intermediate state. To address the problem of label switching in the VEC-MSH(K) models we impose an identification restriction enforce the conditional variances of the interest rates, $Var(Int.rate_t \mid S_t=i,\psi_{t-1},\theta)$ ($i=1, \ldots, K$) in descending order, so that the first regime features a higher ($K=2$) or the highest ($K=3$) volatility[6].

Out of all thirteen possible specifications of the VEC-MSH models (with different values of d, r and K, excluding the methodologically irrelevant case of VEC(4, $3=n$)-MSH), only for six of them a full analysis could be successfully conducted, with the remaining eight being therefore omitted from further considerations. We encountered some numerical problems during the estimation of VEC(3, $r\in\{2, 3\}$)-MSH ($K\in\{2, 3\}$) and VEC(4, $r\in\{1, 2\}$)-MSH(3), which hindered the MCMC sampler. Although no such issues have arisen in the estimation of the VEC(4, $r\in\{1, 2\}$)-MSH(2) models, at some data points across the prediction period the zero value of predictive score has been obtained (probably due to its well-known sensitivity to tail events [Gneiting, Raftery, 2007]), thereby prohibiting calculation of the log predictive score and Bayes factors, utilized here for density forecasts *ex-post* evaluation[7].

[4] The posterior probabilities presented in Fig. 1 have been obtained in the best VEC-MSH(2) and VEC-MSH(3) models (in terms of the forecasting performance to be analyzed in further part of the section).

[5] In setting the lag length (k) we follow numerous empirical studies (some of which are cited in the main text) in which $k=2$ proves the most valid choice, particularly for the US economy.

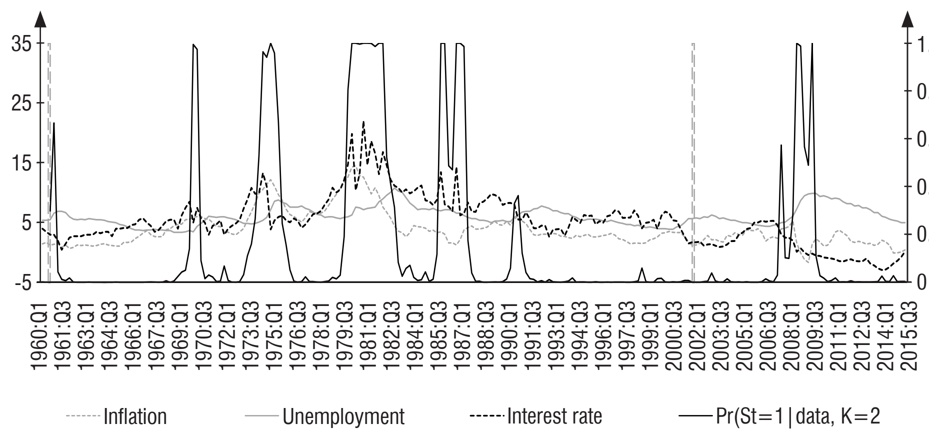
[6] The results presented in this study are 'robust' to the choice of the variable underlying the state-identification restriction.

[7] Other measures, such as the energy score, which are less sensitive to observations materializing in the tails of a predictive distribution, could be applied to assess multivariate density predictions [Gneiting, Raftery, 2007]. However, such scoring rules lack strict Bayesian foundations and therefore are not considered in this study.

For numerical reasons, estimation of the Markov-switching models required tightening of the priors for the adjustment coefficients and the elements of the matrix *B*, related to the cointegrating vectors [Wróblewska, Pajor, 2019], with their standard deviations reduced from 1 (in the homoscedastic VECs) to 0.1. However, the modification does not affect the prior for the cointegration space.

Each of the predictive densities in this study is based upon 200 000 MCMC posterior draws, preceded by either 400 000 burn-in passes – for the first of *N* forecasts (to achieve convergence) – or 20 000 cycles for the subsequent $N-1$ predictions, with the sampler each time initiated at the final draw of the previous run. Density forecasts are evaluated through the log predictive score (LPS, computed with decimal log; the higher the value, the better), with the difference between LPS's for two alternative models defining the log predictive Bayes factor (LPBF). Their values cumulated over the entire *ex-post* forecasting period are denoted as CLPS and CLPBF, respectively, and presented in Table 1.

As can be inferred from Table 1, the Markov-switching models substantially outrank the homoscedastic VECs in terms of LPS, with the cumulated predictive scores for VEC-MSH specifications topping by as much as ca. 11 orders of magnitude. On the other hand, the differences in CLPS among the VEC-MSH and, similarly, VEC models are fairly negligible, although models with one (or also two in the case of VEC) cointegration relations prove marginally better. As indicated by Fig. 2, this superiority of Markov-switching models hinges directly upon evident occurrences of volatility breaks over the prediction period (particularly 2008: Q4 through 2010: Q1), which is quite intuitive. However, the number of regimes in the VEC-MSH specifications does not appear crucial here to enhancing the predictive power of the regime-switching models. This may result from the apparently higher uncertainty as to regime distinction within the three-state structures (see Fig. 2).



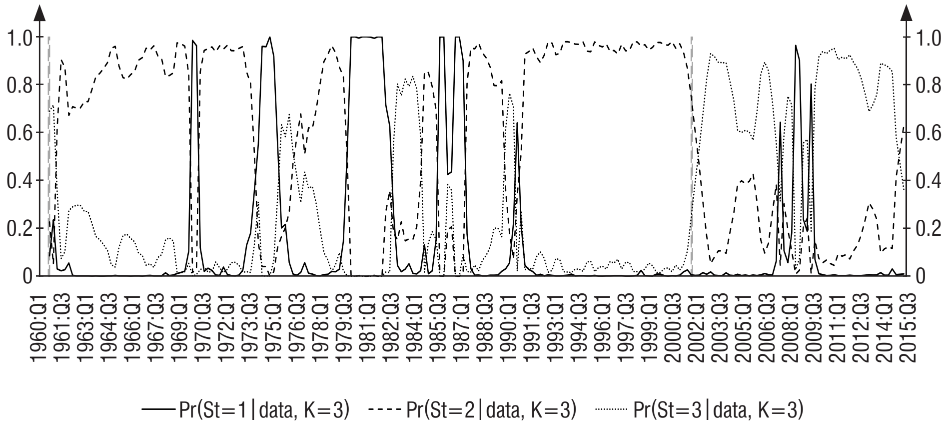


Figure 1. Data and posterior probabilities of regimes

Note: Data presented in the top panel (the LHS axis): quarterly consumer price index (3-month average), unemployment rate (in the last month of quarter), and the Wu-Xia shadow federal funds rate (in the last month of quarter). Posterior probabilities of regimes (the RHS axis) calculated in the VEC(3,1)-MSH(2) (top) and VEC(3,1)-MSH(3) (bottom) models. Vertical lines demark the initial conditions (violet) and the period of predictive performance evaluation (red).

Source: Own elaboration based on data: inflation and unemployment rates from the Bureau of Labor Statistics (www.bls.gov), and shadow interest rates from https://www.frbatlanta.org/cqer/research/wu-xia-shadow-federal-funds-rate.aspx

Table 1. Ranking of the models.

Ranking (i)	d	r	K	Model	CLPS$_i$	CLPBF$_{1i}$
1	3	1	3	VEC-MSH	–71.487	0
2	3	1	2	VEC-MSH	–71.590	0.103
3	3	0	3	VEC-MSH	–71.664	0.177
4	4	0	3	VEC-MSH	–71.750	0.263
5	3	0	2	VEC-MSH	–71.997	0.510
6	4	0	2	VEC-MSH	–72.066	0.579
7	4	2	–	VEC	–82.768	11.281
8	4	1	–	VEC	–82.875	11.388
9	3	2	–	VEC	–82.887	11.400
10	3	1	–	VEC	–82.974	11.487
11	4	0	–	VEC	–83.067	11.580
12	3	0	–	VEC	–83.180	11.693
13	3	3	–	VEC	–83.569	12.082

Note: Cumulated (decimal) log predictive scores (CLPS) and cumulated log predictive Bayes factors (CLPBF) in favour of the best model; alternative specifications of the constant include: an unrestricted constant (d=3) and the constant restricted to the cointegration relations (d=4); the numbers of cointegration relations include $r \in \{0,1,2\}$; r=3 indicates a stationary VAR; $K\in\{2,3\}$ indicates the number of regimes.

Source: Own elaboration.

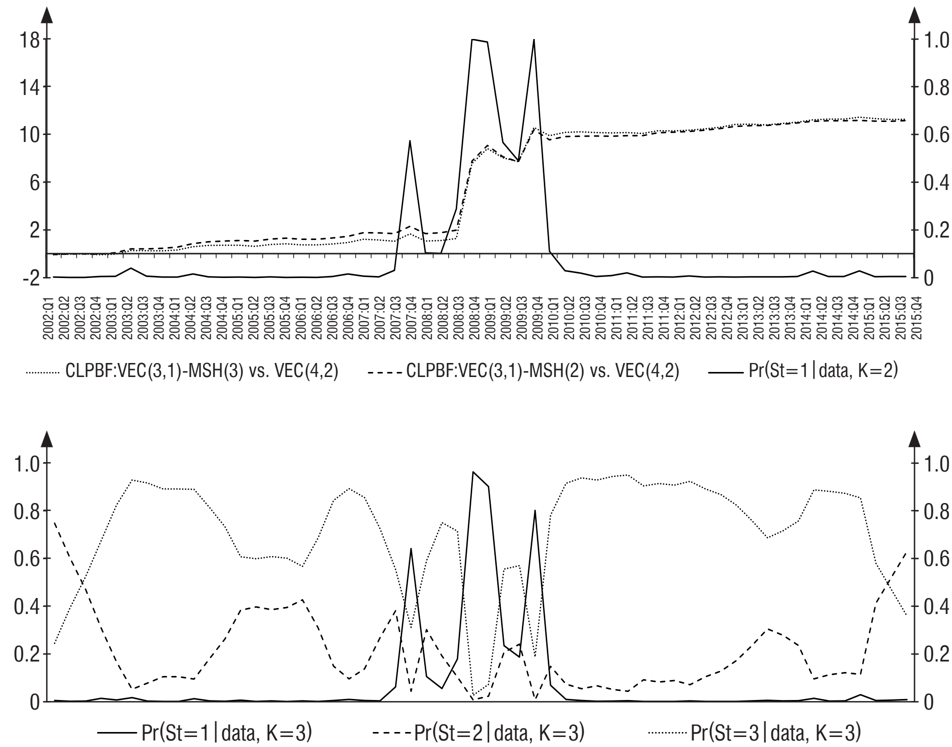
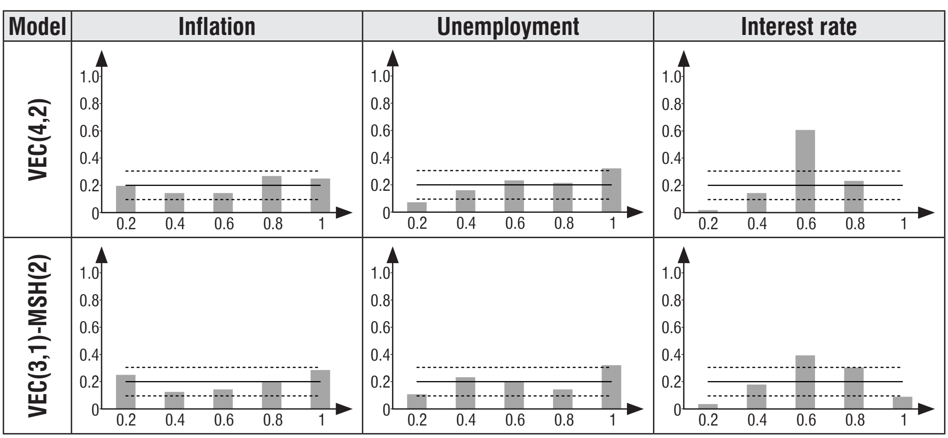


Figure 2. Cumulative log predictive Bayes factors and the posterior probabilities of regimes across the forecasting period.

Note: Cumulative (decimal) log predictive Bayes factors in favour of the best VEC-MSH models against the best VEC model (top panel, blue and red lines; the LHS axis), along with the posterior probabilities of the regimes (black lines; in VEC(3,1)-MSH(2) – the top panel, and VEC(3,1)-MSH(3) – the bottom panel).

Source: Own elaboration.



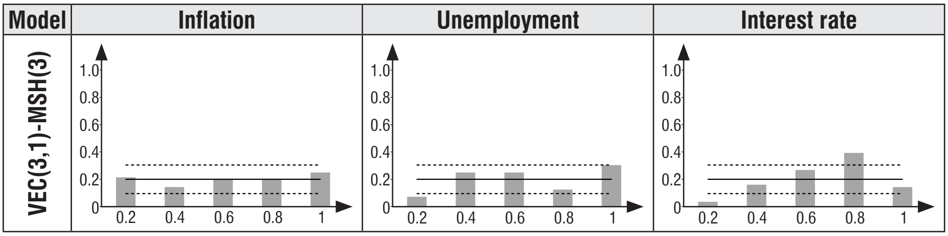


Figure 3. PIT histograms in the best VEC, VEC-MSH(2) and VEC-MSH(3) models.

Note: Horizontal lines represent 95% confidence bands around the value of 0.2 (corresponding, in our setting, to the ideal case of PIT uniformity)

Source: Own elaboration.

Finally, we examine the calibration of density forecasts *via* PIT histograms, displayed in Fig. 3, along with 95% (non-Bayesian) confidence bands constructed as in [Wróblewska, Pajor, 2019] around the value of 0.2 (representing, in our setting, the perfect PIT uniformity). Overall, although none of the models considered in these figures proves ideal for all the variables, it appears that introducing two-state Markovian breaks in the conditional covariance matrix of VEC models refines predictive densities calibration for the interest rates, with allowing for yet another, third regime, however, resulting in no further improvement.

5. Conclusions

In the paper we examined probabilistic predictive performance of Bayesian homoscedastic VEC models with their extensions allowing for either two- or three-state Markov-switching heteroscedasticity. To this end, the log predictive score (LPS) and Bayes factors, as well as Probability Integral Transform were employed, which are typically used for such assessments.

In general, the results obtained within small models of monetary policy for the US economy indicate that enabling Markovian shifts in the conditional covariance matrix of the VEC models leads to a substantial improvement of density forecasts performance (as measured by cumulated LPS), although allowing for two regimes appears just enough, for no additional increase of the prediction power is observed upon adding yet another state. Intuitively enough, the superiority of the VEC-MSH specification is gained only upon volatility breaks occurrence in the forecasts *ex-post* evaluation period.

Since the results of this study indicate that allowing for some sort of time-varying conditional volatility in otherwise homoscedastic VEC models may improve considerably their forecasting performance, it appears due for future research to extend our framework by including also other conditional volatility processes, such as

GARCH, SV or their hybrids [Osiewalski, 2009; Osiewalski, Pajor, 2009], in which volatility evolution is continuously-valued rather than discrete (as in the case of the Markov-switching models).

Acknowledgements

Research financed from a subvention granted to Cracow University of Economics. The author is greatly indebted to Justyna Wróblewska for her support and help with developing relevant codes in GAUSS.

References

Carter C.K., Kohn R. [1994], *On Gibbs sampling for state space models*, "Biometrika", Vol. 81, No. 3, pp. 541–553.

Clarida R.H., Sarno L., Taylor M.P., Valente G. [2003], *The out-of-sample success of term structure models as exchange rate predictors: a step beyond*, "Journal of International Economics", Vol. 60, pp. 61–83.

Frühwirth-Schnatter S. [2006], *Finite mixture and Markov switching models*, "Springer Series in Statistics", Springer.

Geweke J., Amisano G. [2010], *Comparing and evaluating Bayesian predictive distributions of asset returns*, "International Journal of Forecasting", Vol. 26, pp. 216–230.

Gneiting T., Balabdaoui F., Raftery A.E. [2007], *Probabilistic forecasts, calibration and sharpness*, "Journal of the Royal Statistical Society: Series B (Statistical Methodology)", Vol. 69, pp. 243–268.

Gneiting T., Raftery A.E. [2007], *Strictly proper scoring rules, prediction, and estimation,* "Journal of the American Statistical Association", Vol. 102, pp. 359–378.

Hamilton J.D. [1989], *A new approach to the economic analysis of nonstationary time series and the business cycle*, "Econometrica", Vol. 57, pp. 357–384.

Jochmann M., Koop G. [2015], *Regime-switching cointegration,* "Studies in Nonlinear Dynamics & Econometrics", Vol. 19, No. 1, pp. 35–48.

Osiewalski J. [2009], *New hybrid models of multivariate volatility (a Bayesian perspective],* "Statistical Review", Vol. 56, No. 1, pp. 15–22.

Osiewalski J., Pajor A. [2009], *Bayesian analysis for hybrid MSF–SBEKK models of multivariate volatility,* "Central European Journal of Economic Modelling and Econometrics", Vol. 1, No. 2, pp. 179−202.

Primiceri G.E. [2005], *Time varying structural vector autoregressions and monetary policy*, "Review of Economic Studies", Vol. 72, pp. 821–852.

Psaradakis Z., Spagnolo F. [2005], *Forecast performance of nonlinear error-correction models with multiple regimes,* "Journal of Forecasting", Vol. 24, pp. 119–138.

Sarno L., Thornton D.L., Valente G. [2005], *Federal Funds Rate Prediction*, "Journal of Money, Credit, and Banking", Vol. 37, No. 3, pp. 449–471.

Wróblewska J., Pajor A. [2019], *One-period joint forecasts of Polish inflation, unemployment and interest rate using Bayesian VEC-MSF models*, "Central European Journal of Economic Modelling and Econometrics", Vol. 11, pp. 23–45.

Wu C.J., Xia F.D. [2015], *Measuring the macroeconomic impact of monetary policy at the zero lower bound,* "Chicago Booth Research Paper", No. 13–77, available at: http://dx.doi.org/10.2139/ssrn.2321323

MCMC Method for the IG-MSF-SBEKK Model

Anna Pajor[1]

Abstract

In this paper a Markov chain Monte Carlo (MCMC) simulation tool used in the hybrid IG-MSF-SBEKK is described. The MCMC method is adapted to obtain a sample from the posterior distribution of parameters and latent variables. The Gibbs sampler with Metropolis-Hastings steps is used. The proposed numerical method is applied to estimate the hybrid IG-MSF-SBEKK model for daily exchange rate returns.

Keywords: Markov chain Monte Carlo, stochastic volatility, Bayesian inference

JEL classification: C11, C15

[1] Anna Pajor, Cracow University of Economics, Jagiellonian University in Cracow, e-mail: eopajor@cyf-kr.edu.pl

1. Introduction

Hybrid MSV-MGARCH models have been introduced as useful tools for modelling highly dimensional time series. These models formally belong to the Multivariate Stochastic Volatility (MSV) class, but in fact they constitute some hybrids of the Multivariate GARCH (MGARCH) and MSV specifications. Hybrid models can be parsimoniously parameterized, and simultaneously they can allow for capturing the time-varying conditional variances, covariances and correlations. The hybrid MSV-MGARCH models have been introduced by Osiewalski and Pajor [2007, 2009], Osiewalski [2009] and Osiewalski and Osiewalski [2016]. To make it possible to analyze large portfolios Osiewalski [2009] has proposed a relatively simple multivariate volatility model that uses one latent process and has a non-trivial covariance structure. In the MGARCH class a practical tool for analyzing large portfolios is the scalar BEKK process [Baba, Engle, Kraft, Kroner, 1989], whereby in the MSV class the Multiplicative Stochastic Factor (MSF) process can be considered. Finally, Osiewalski [2009] has proposed some LN-MSF-SBEKK hybrids. The LN-MSF-SBEKK structure is obtained by multiplying the SBEKK conditional covariance matrix H_t by a scalar random variable g_t such that $\{\ln g_t\}$ is a Gaussian AR(1) latent process with auto--regression parameter φ.

In [Osiewalski, Pajor, 2018; 2019] the IG-MSF-SBEKK specification has been proposed as a natural hybrid extension of the SBEKK process with the Student t conditional distribution. In the new specification the latent process $\{g_t\}$ is no longer marginally log-normal (LN), but for $\varphi=0$ the variable g_t has an inverted gamma (IG) distribution that leads to the t-SBEKK process. If $\varphi\neq 0$, the unconditional distribution of the latent variables g_t remains unknown.

In this paper a Markov chain Monte Carlo simulation tool used in the hybrid IG-MSF--SBEKK is described. The MCMC method is developed to obtain a sample from the posterior distribution of parameters and latent variables. An empirical example is also presented to illustrate that our sampler performs well.

2. The IG-MSF-SBEKK model

Osiewalski and Pajor [2018; 2019] have introduced the hybrid MSV-MGARCH model, which is based on the assumption that the natural logarithm of the latent variable follows an autoregressive process with the inverted gamma innovations. They introduced the following specification for log-returns:

$$r_t=\delta_0+r_{t-1}\Phi_1+\varepsilon_t\,,\; t=1,\dots,T, \tag{1}$$

where $r_t=(r_1, r_2, \dots, r_n)$ is an $1\times n$ vector of observations, δ_0 is an $1\times n$ vector of parameters, Φ_1 is an $n\times n$ matrix of real coefficients, and T is the length of the observed time series. The hybrid IG-MSF-SBEKK structure for the $1\times n$ 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

$$H_t = (1-\beta_1-\beta_2)A + \beta_1(\varepsilon'_{t-1}\,\varepsilon_{t-1}) + \beta_2 H_{t-1}, \quad (3)$$

$$\ln g_t = \varphi \ln g_{t-1} + \ln \gamma_t,\ \{\gamma_t\} \sim iiIG\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \quad (4)$$

$$\{\zeta_t\} \sim iiN\,(0, I_n),\ \zeta_t \perp \gamma_s \text{ for all } t, s \in \{1, \ldots, T\},\ 0 < |\varphi| < 1, \quad (5)$$

$\{\zeta_t\}$ is a Gaussian white noise sequence with the mean vector zero and the unit covariance matrix, H_t is a square matrix of order n, symmetric and positive definite for each t and having a scalar BEKK (SBEKK) form, $\{g_t\}$ is a scalar stochastic latent process, $\{\gamma_t\}$ is a sequence of independent positive random variables, γ_t is inverted gamma distributed with mean $\frac{\nu}{\nu-2}$ for $\nu>2$, the notation $\zeta_t \perp \gamma_s$ denotes that random variables ζ_t and γ_s are independent.

Under (1) - (5), 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; r_t it has the normal distribution with the mean vector $\mu_t = \delta_0 + r_{t-1}\Phi_1$ and the covariance matrix $\Sigma_t = g_t H_t$, which depends on both g_t and the past of r_t, so the distribution of r_t given only its past is the scale mixture of $N(\mu_t,\ g_t H_t)$ distributions with an unknown marginal distribution of g_t. However, for $\varphi=0$ $g_t = \gamma_t$, so the distribution of g_t is known by assumption. Since $\varphi=0$ corresponds to the t-SBEKK we may view the IG-MSF-SBEKK structure as a natural hybrid extension of the popular SBEKK specification with the conditional Student t distribution.

For $u_t = \ln \gamma_t$ we have $E(u_t) = \ln\frac{\nu}{2} - \psi_0\left(\frac{\nu}{2}\right)$ and $Var(u_t) = \psi_1\left(\frac{\nu}{2}\right)$, where $\psi_0(\cdot)$ and $\psi_1(\cdot)$ denote the digamma and trigamma function, respectively. Thus, $\{\ln \gamma_t\}$ does not have zero mean. Since $\ln g_t$ can be expressed as $\ln g_t = \varphi^t \ln g_0 + \sum_{j=0}^{t-1} \varphi^j u_{t-j}$, then (for $\ln g_0$ constant and $|\varphi|<1$):

$$E(\ln g_t) = \varphi^t \ln g_0 + \frac{1-\varphi^t}{1-\varphi^t}\left[\ln\left(\frac{\nu}{2}\right) - \psi_0\left(\frac{\nu}{2}\right)\right],\quad Var(\ln g_t) = \frac{1-\varphi^{2t}}{1-\varphi^2}\,\psi_1\left(\frac{\nu}{2}\right).$$

Thus, $\lim\limits_{t\to+\infty} E(\ln g_t) = \dfrac{\left[\ln\left(\frac{\nu}{2}\right) - \psi_0\left(\frac{\nu}{2}\right)\right]}{1-\varphi}$ and $\lim\limits_{t\to+\infty} Var(\ln g_t) = \dfrac{\psi_1\left(\frac{\nu}{2}\right)}{1-\varphi^2}$.

The Bayesian statistical model amounts to specifying the joint distribution of all observations, latent variables and parameters. The assumptions presented so far determine the conditional distribution of the observations and latent variables given the parameters. Thus, what remains to be done is to formulate the marginal distribution of the parameters (the prior or *a priori* distribution). We assume independence among groups of parameters and use the same prior distributions as Osiewalski and Pajor [2019] for the same parameters. The $n(n+1)$ elements of $\delta = (\delta_0, vec(\Phi_1)')$ are assumed to be *a priori* independent of other parameters, with the $N(\mu_\delta, \Sigma_\delta)$ prior. Matrix A is a free symmetric positive definite matrix of order n, with an inverted Wishart prior distribution; the elements of $\beta = (\beta_1, \beta_2)'$ are 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, exponentially distributed *a priori* with mean $\frac{1}{\lambda}$; φ has the uniform distribution over (–1, 1), and for ν we assume the gamma distribution with mean λ_a/λ_ν and variance λ_a/λ_ν^2. The most popular prior distributions for ν are: exponential [Geweke, 1993; Fernández, Steel, 1998; 1999; Osiewalski, Pajor, 2018], and uniform [Jacquier, Polson, Rossi, 2004]. It is worth noting that the exponential distribution is a special case of the gamma distribution (when $\lambda_a=1$).

Now we can write the full Bayesian model as:

$$p(r_1, \dots, r_T, g_1, \dots, g_T, \delta, A, \beta, \varphi, \nu, h_0)=\prod_{t=1}^{T} f_N^n\,(r_t|\mu_t, g_t H_t) \times$$
$$\prod_{t=1}^{T} \frac{\left(\frac{\nu}{2} g_{t-1}^{\varphi}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{1}{g_t}\right)^{\frac{\nu}{2}+1} e^{-\frac{\frac{\nu}{2} g_{t-1}^{\varphi}}{g_t}} \times p(\delta)p(A)p(\beta)p(\varphi)p(\nu)p(h_0), \tag{6}$$

where $f_N^n(\cdot|a, B)$ denotes the density function of the n-variate normal distribution with the mean vector a and the covariance matrix B;

$p(\delta) \propto e^{-\frac{1}{2}(\delta-\mu_\delta)\Sigma_\delta^{-1}(\delta-\mu_\delta)'}\, I_{(0,1)}(|\lambda_R|_{max})$, λ_R is the vector of eigenvalues of the companion matrix, connected with the VAR form, in the paper $\mu_\delta=0$, $\Sigma_\delta= I_{n(n+1)}$; the symbol $I_s(\cdot)$ denotes the indicator function of the set S;

$p(A)=f_{IW}\,(A|\Omega, \mu_A, n) \propto \det(A)^{-\frac{\mu_A+n+1}{2}} e^{-\frac{1}{2}tr(\Omega A^{-1})}$, $E(A)=\frac{\Omega}{\mu_A-n-1}$ for $\mu_A>n+1$; here $\Omega=nI_n$, $\mu_A=n$, thus $E(A)$ does not exist and A^{-1} has Wishart prior distribution with mean I_n; $p(\beta) \propto I_{(0,1)}\,(\beta_1+\beta_2)I_{(0,1)}(\beta_1)I_{(0,1)}(\beta_2)$; $p(\varphi) \propto I_{(-1,1)}(\varphi)$;

$p(\nu)=f_G(\nu|\lambda_a, \lambda_\nu)=\frac{\lambda_\nu^{\lambda_a}}{\Gamma(\lambda_a)}\, \nu^{\lambda_a-1}e^{-\lambda_\nu \nu}I_{(0,+\infty)}(\nu)$, here $\lambda_a=3$, $\lambda_\nu=0.1$;

$p(h_0)=f_{Exp}(h_0|\lambda)=\lambda e^{-\lambda h_0}I_{(0,+\infty)}(h_0)$, with $\lambda=1$.

3. Numerical implementation

The posterior density function, proportional to (6), is highly dimensional and non-standard. To make inference about parameters and latent variables numerical methods are needed. We propose a Markov chain Monte Carlo method, namely the Gibbs algorithm, i.e. the sequential sampling from the full conditional distributions obtained from (6).

3.1. The full conditional distributions of δ, A, β and h_0

The conditional posterior density functions of the parameters of VAR and SBEKK structures are the following:

$$p(\delta|r_1, \dots, r_T, g_1, \dots, g_T, A,\beta,\varphi,\nu,h_0) \propto p(\delta) \prod_{t=1}^{T} f_N^n(r_t|\mu_t,g_t H_t), \tag{7}$$

$$p(A|r_1, \dots, r_T, g_1, \dots, g_T, \delta,\beta,\varphi,\nu,h_0) \propto p(A) \prod_{t=1}^{T} f_N^n(r_t|\mu_t,g_t H_t), \tag{8}$$

$$p(\beta,h_0|r_1, \ldots, r_T, g_1, \ldots, g_T, \delta, A,\varphi,\nu) \propto p(\beta)\, p(h_0) \prod_{t=1}^{T} f_N^n(r_t|\mu_t,g_tH_t). \quad (9)$$

Since densities of the full conditional distributions in (7) – (9) have none of any known closed forms, we must simulate $\delta, A,\beta,$ and h_0 using the Metropolis-Hastings algorithm. These groups of parameters can be generated separately in three steps of the Gibbs sampler or in one step by blocking components. The sequential Metropolis-Hastings algorithm with truncated Student *t* distribution (with 3 degrees of freedom) centred at the previous values of the chain (similarly as in [Osiewalski, Pajor, 2009]) can be applied. The covariance matrix of the Student *t* distribution can be determined by initial draws of the algorithm.

3.2. The conditional distribution of φ

It is immediate from (6) that the full conditional distribution of φ is given as follows:

$$p(\varphi|r_1, \ldots, r_T, g_1, \ldots, g_T,\delta, A,\beta,\nu,h_0) \propto e^{\frac{\varphi\nu}{2}\sum_{t=1}^{T}\ln g_{t-1}-\frac{\nu}{2}\sum_{t=1}^{T}\frac{g_{t-1}^{\varphi}}{g_t}}\, I_{(-1,1)}(\varphi). \quad (10)$$

The distribution in (10) has a non-standard form. Therefore, we use the sequential Metropolis-Hastings algorithm, drawing from the truncated normal distribution with variance 10^{-2}, centred at the previous state of the chain.

3.3. The conditional distribution of ν

The conditional distribution of ν given $(r_1, \ldots, r_T, g_1, \ldots, g_T,\delta, A,\beta,\varphi,h_0)$ has the following kernel density function:

$$p(\nu|r_1, \ldots, r_T, g_1, \ldots, g_T,\delta, A,\beta,\varphi,h_0) \propto p(\nu|\kappa,T,\lambda_a)=\left(\frac{\nu}{2}\right)^{\frac{T\nu}{2}+\lambda_a-1}\Gamma\left(\frac{\nu}{2}\right)^{-T}e^{-\kappa\nu}. \quad (11)$$

where $\kappa= \lambda_\nu-\frac{1}{2}\sum_{t=1}^{T}\ln\frac{g_{t-1}^{\varphi}}{g_t}+\frac{1}{2}\sum_{t=1}^{T}\frac{g_{t-1}^{\varphi}}{g_t}$.

Although the full conditional posterior of ν is non-standard, the acceptance--rejection method described by Geweke [1994] in the context of the trend stationary model can be applied here. The sampling distribution used for candidate draws is exponential with probability density function: $g(\nu;\alpha)=\alpha e^{-\alpha\nu}$. The draw is accepted with probability $\frac{p(\nu|\kappa,T,\lambda_a)}{c(\alpha)g(\nu;\alpha)}$, where $c(\alpha)=\sup_\nu \frac{p(\nu|\kappa,T,\lambda_a)}{g(\nu;\alpha)}$, with $p(\nu|\kappa,T,\lambda_a)$ defined in (11). As has been pointed out by Geweke [1994], the maximization of the acceptance probability is equivalent to the minimization of $c(\alpha)$. Let $Q(\nu,\alpha)=\ln\frac{p(\nu|\kappa,T,\lambda_a)}{g(\nu;\alpha)}$, thus

$$Q(\nu,\alpha)= \left(\frac{T\nu}{2}+\lambda_a-1\right)\ln\frac{\nu}{2}-\mathrm{T}\ln\Gamma\left(\frac{\nu}{2}\right)+(\alpha-\kappa)\nu-\ln\alpha. \quad (12)$$

The choice of α is determined by the solution to the following problem: $\inf_\alpha\{\sup_\nu Q(\nu,\alpha)\}$.

The partial derivative of $Q(\nu,\alpha)$ with respect to ν is:

$$\frac{\partial Q(\nu,\alpha)}{\partial \nu} = \frac{\mathrm{T}}{2}\left[\ln\frac{\nu}{2} + 1 - \Psi\left(\frac{\nu}{2}\right)\right] + \frac{\lambda_a - 1}{\nu} + (\alpha-\kappa) \tag{13}$$

where $\Psi(x)=\dfrac{\partial \ln \Gamma(x)}{\partial x}$ is the digamma function.

Note that functions $f_1(\nu)=\ln\frac{\nu}{2}+1-\Psi\left(\frac{\nu}{2}\right)$ and $f_2(\nu)=\dfrac{\lambda_a - 1}{\nu}$ for $\lambda_a>1$ are monotone decreasing on the interval $(0,+\infty)$ from $+\infty$ to 1, and from $+\infty$ to 0, respectively. Additionally, since the function $f(x)=x-\ln x$ is minimized at $x=1$, $\kappa\geq\lambda_\nu$ +T/2. Since $\kappa>\frac{T}{2}$, for small $\alpha>0$ there exists a unique maximum of $Q(\nu,\alpha)$. On the other hand, the partial derivative of $Q(\nu,\alpha)$ with respect to α is $\dfrac{\partial Q(\nu,\alpha)}{\partial \alpha}=\nu-\dfrac{1}{\alpha}$, thus $\dfrac{\partial Q(\nu,\alpha)}{\partial \alpha}=0$ for $\alpha=\frac{1}{\nu}$. Substituting $\frac{1}{\nu}$ for α (i.e. $\alpha=\frac{1}{\nu}$) in (13) yields the equation: $\frac{\mathrm{T}}{2}\left[\ln\frac{\nu}{2} + 1 - \Psi\left(\frac{\nu}{2}\right)\right] + \frac{\lambda_a}{\nu} + \kappa=0$, which can be solved numerically or by using the following approximation: $\Psi(z) \approx \ln z - \dfrac{1}{2z} - \dfrac{1}{12z^2}$ (see Abramowitz and Stegun, 1968). The resulting expression is: $\nu^*\approx\dfrac{-(3T+6\lambda_a)+\sqrt{(3T+6\lambda_a)^2-4T(3T+6\kappa)}}{6T-12\kappa}$. Note in particular that $\nu^*>0$, because of $\kappa>\frac{T}{2}$. Finally, the sampling distribution of candidate draws of ν is exponential: $g\left(\nu;\frac{1}{\nu^*}\right)=\frac{1}{\nu^*}e^{-\frac{1}{\nu^*}\nu}$.

The draw is accepted with probability:

$$\left(\frac{\nu}{2}\right)^{\frac{T\nu}{2}+\lambda_a-1}\left(\frac{\nu^*}{2}\right)^{-\left(\frac{T\nu^*}{2}+\lambda_a-1\right)}\Gamma\left(\frac{\nu}{2}\right)^{-T}\Gamma\left(\frac{\nu^*}{2}\right)^{T}e^{\left(\kappa-\frac{1}{\nu^*}\right)(\nu^*-\nu)}.$$

3.4 The conditional distribution of $g=(g_1, \ldots, g_T)$

We analyze each element of the vector g in a separate Gibbs step. The conditional posterior density of g_t, $t \in \{1, \ldots, T\}$, is defined by

$$p(g_t|r_1, \ldots, r_T, g_1, \ldots, g_{t-1}, g_{t+1}, \ldots, g_T,\delta, A,\beta,\varphi,\nu,h_0) \propto$$
$$e^{-\frac{\nu}{2}\frac{g_t^{\varphi}}{g_{t+1}}} f_{IG}\left(g_t\middle|\frac{n}{2}+\frac{\nu}{2}(1-\varphi), \frac{1}{2}(r_t-\mu_t)H_t^{-1}(r_t-\mu_t)'+\frac{\nu}{2}g_{t-1}^{\varphi}\right), \text{ for } t=1, \ldots, T-1 \tag{14}$$

$$p(g_T|r_1, \ldots, r_T, g_1, \ldots, g_{T-1},\delta, A,\beta,\varphi,\nu,h_0)=$$
$$f_{IG}\left(g_T\middle|\frac{n}{2}+\frac{\nu}{2}(1-\varphi), \frac{1}{2}(r_T-\mu_T)H_T^{-1}(r_T-\mu_T)'+\frac{\nu}{2}g_{T-1}^{\varphi}\right), \tag{15}$$

where $f_{IG}(\cdot|a, b)$ denotes the probability density function of the inverted gamma distribution (since called $IG(a, b)$) with mean $\frac{b}{a-1}$ (for $a>1$) and variance $\frac{b^2}{(a-1)^2(a-2)}$ (for $a>2$).

In order to simulate latent variables, $g_1, \ldots, g_{T-1}$, the independence Metropolis – Hastings algorithm is applied, similarly to Jacquier, Polson and Rossi [1994]. A proposal distribution used to simulate g_T is the inverted gamma:

$$IG\left(\frac{n}{2}+\frac{\nu}{2}, \frac{1}{2}(r_t-\mu_t)H_T^{-1}(r_T-\mu_T)'+\frac{\nu}{2}g_{T-1}^{\varphi}\right).$$

An alternative proposal distribution such as the inverted gamma:

$$IG\left(\frac{n}{2}+\frac{\nu}{2}(1-\varphi), \frac{1}{2}(r_T-\mu_T)H_T^{-1}(r_T-\mu_T)'+\frac{\nu}{2}g_{T-1}^{\varphi}\right)$$

was found to be less efficient in practice. Drawing from the conditional distribution of g_T is straightforward, because it is an inverted gamma distribution. As regards initial conditions for $\{\ln g_t\}$, it is assumed that $\ln g_0 = 0$.

The full conditional distributions presented in (7)–(11), (14) and (15) are used in the Gibbs algorithm with $T+3$ steps. Convergence of the induced Markov chain in distribution to the posterior distribution is ensured by ergodic theorems, in particular those pertaining to the Gibbs sampler [Gelfand, Smith, 1990; Gilks, Richardson, Spiegelhalter, 1996].

4. Empirical illustration

In order to illustrate the proposed numerical method we use the growth rates of two exchange rates that are most important for the Polish economy, namely the zloty (PLN) values of the US dollar and Euro. The data covers the period from January 2, 2007 till December 20, 2019. The first three observations (January 2, 3 and 4, 2007) are initial conditions. Thus, T, the length of the modelled vector time series of daily growth rates (logarithmic return rates) is equal to 3328. The data set was obtained from the website http://stooq.pl.

Table 1. Sample characteristics

Series (return rates)	Mean	Standard deviation	Skewness	Kurtosis	Correlation coefficient
USD/PLN	0.0087	0.9053	0.1033	9.8161	0.7832
EUR/PLN	0.0032	0.5546	0.1311	12.9013	

Source: calculated by the author.

Table 1 summarizes descriptive statistics for the logarithmic returns. Both series are centred about zero, with several outliers and changing volatility. The sample kurtosis is much higher than 3, indicating non-Normal empirical distribution for return rates. Both series are quite strongly positively correlated, their sample correlation coefficient is equal to 0.7832.

Table 2. Estimates of posterior means and standard deviations of the parameters of the IG-MSF-SBEKK model

Parameter	Posterior mean	Standard deviation
δ_{01}	–0.011	0.011
δ_{02}	–0.011	0.006
δ_{11}	0.005	0.024
δ_{12}	0.008	0.013
δ_{21}	–0.017	0.040
δ_{22}	–0.029	0.024
a_{11}	0.183	0.027
a_{12}	0.058	0.012
a_{22}	0.057	0.007
φ	0.296	0.107
ν	9.997	1.267
β_1	0.035	0.004
β_2	0.948	0.005
h_0	0.053	0.026

Note: Here $\delta_0=(\delta_{01}, \delta_{02})$, $\Phi_1=[\delta_{ij}]$, $A=[a_{ij}]$, i=1,2, j=1,2.

Source: calculated by the author.

In Table 2 estimates of the posterior means and standard deviations of the parameters of the IG-MSF-SBEKK model are presented. These estimates are based on 1 million dependent draws from the posterior distribution, which is the stationary distribution of the Markov chain used here. The average acceptance rate in the Metropolis-Hastings algorithm for the latent variables is about 88%, about 77% for parameters of the SBEKK and VAR structures, and 18% for φ.

To informally assess the convergence of any sampler Yu and Mykland [1998] proposed visual inspection of the CUSUM statistics for some scalar function of parameters and latent variables (univariate quantity of interest, denoted by $f(\boldsymbol{\theta}, g)$, where $\boldsymbol{\theta}$ stands for the vector of parameters, g is the vector of latent variables):

$$S_n = \sum_{q=n_0+1}^{n}[f(\boldsymbol{\theta}^{(q)}, g^{(q)}) - \mu_f],\ n = n_0+1, \dots, N, \tag{16}$$

where $\{\boldsymbol{\theta}^{(q)}, g^{(q)}\}_{q=1}^{N}$ are drawn from the posterior distribution, μ_f is the arithmetic mean of $f(\boldsymbol{\theta}, h)$, calculated on the basis of the last $N-n_0$ draws; n_0 is the number of initial discarded iterations. Commonly used functions are $f(\boldsymbol{\theta}, g) = \theta_i$ and $f(\boldsymbol{\theta}, g) = g_t$. Of course, the CUSUM path plot ends at zero. But if the Markov chain converges, then the S_n path (against n) converges smoothly to zero. On the other hand, a smooth plot of the CUSUM indicates slow mixing of the chain, while a "hairy" plot indicates that the chain is mixing well. Brooks [1998] suggested the following measure of "hairiness":

$$D_n = \frac{1}{n-n_0-1}\sum_{i=n_0+1}^{n-1} d_i \qquad n_0+2 \le n \le N \tag{17}$$

where

$$d_i = \begin{cases} 1 & if\ (S_{i-1} > S_i\ \ and\ S_i < S_{i+1})\ or\ (S_{i-1} < S_i\ \ and\ S_i > S_{i+1}) \\ 0 & else \end{cases}$$

for $i = n_0+1, \dots, N-1$.

D_n can be interpreted as an average number of times $f(\boldsymbol{\theta}, g)$ crosses μ_f, and takes values between 0 and 1. $D_n = 0$ indicates a totally smooth plot, whereas $D_n = 1$ indicates maximum "hairiness". Under assumptions that the sequence of $\{d_i\}$ is i.i.d. and distributed symmetrically about the mean, Brooks [1998] shows that D_n has an asymptotic normal distribution with mean $\frac{1}{2}$ and variance $\frac{1}{4(n-n_0-1)}$.

Thus, the convergence is diagnosed once the sequence $\{D_n\}$ lies within the bounds: $\frac{1}{2} \pm q_{\alpha/2}\sqrt{\frac{1}{4(n-n_0-1)}}$ in approximately $100\left(1-\frac{\alpha}{2}\right)$ % of $N-n_0$ cases. The symbol $q_{\alpha/2}$ stands for the quantile of order $\alpha/2$ of the standardized normal distribution. Of course, the assumption that the $\{d_i\}$ sequence is i.i.d. is not satisfied for MCMC sampler, but can be made approximately true by thinning of the chain [see Brooks 1998].

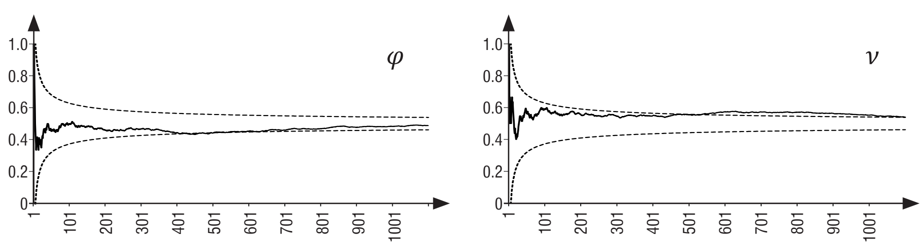


Figure 1. CUSUM diagnostic plots for selected parameters

Note: The solid line represents the sequence of D_n values, the dotted line represents the 99% confidence bounds.
Source: calculated by the author.

Our chain was run for 1.5 million steps, and the first 400,000 steps were discarded. The chain has been thinned by a factor of 1000, because of the independence assumption underlying the method by Brooks [1998]. Considering only every 1000th step reduces autocorrelation among the samples. It seems that our Markov chain is mixing well, because we see a lot of “hairiness” in the plots of S_n (not presented to save space). In Figure 1 we present the plots of D_n (a quantitative representation of the mixing rate of the chain) for selected two parameters. The mixing rates are satisfactory, since the sequence of D_n alues remains inside the 99% confidence bounds. For each parameter, the values of D_N are between 0.477 and 0.536. These are reasonably close to $\frac{1}{2}$.

In turn, Bauwens, Lubrano and Richard [1999] consider a standardized version of the CUSUM defined as follows: $CS_n = \frac{\frac{1}{n-n_0}\sum_{q=n_0+1}^{n} f(\theta^{(q)}, g^{(q)}) - \mu_f}{\sigma_f}$, $n = n_0+1,\ldots,N$, μ_f and σ_f are the arithmetic mean and standard deviation of $f(\theta, g)$, respectively, calculated on the basis of the last $N-n_0$ draws. They point out that if the Markov chain converges, then the graph of CS_n against n should converge smoothly to zero. In turn, “long and regular excursions away from zero are an indication of the absence of convergence” [Bauwens, Lubrano, Richard, 1999].

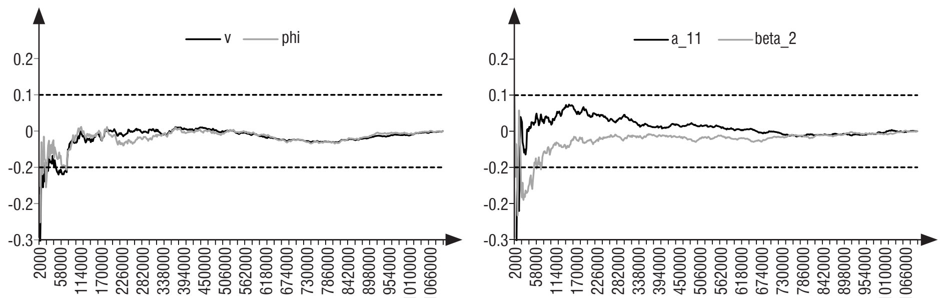


Figure 2. Standardized CUSUM plots for selected parameters

Source: calculated by the author.

Bauwens and Lubrano [1998] observe that the sampler has converged after $n(\epsilon)$ draws for the estimation of $f(\theta, h)$ with a relative error of $100 \times \epsilon$ per cent if CS_n remains within a band of $\pm\epsilon$ for all n larger than $n(\epsilon)$. A value of 0.1 for a CUSUM means that the estimate of the posterior mean diverges from the final estimate by 10 per cent in units of the final estimate of the posterior standard deviation. As can be seen in Figure 2, our chain has converged after about 80000 steps, since the standardized CUSUMs (for $f(\theta, g) = \theta_i$) lie within the band of ± 0.1 for all n larger than 80000.

5. Conclusions

The aim of this paper was to show how to estimate within the Bayesian approach the hybrid IG-MSF-SBEKK generalisation of the *t*-SBEKK model. The Bayesian analysis of the model relies on Gibbs sampling with Metropolis-Hastings steps. We developed the algorithm to estimate all the parameters and latent variables. We illustrated our methods through an empirical application of the growth rates of two exchange rates, namely the zloty (PLN) values of the US dollar and Euro. Our numerical method, proposed in order to obtain a sample from the posterior distribution of parameters and latent variables, has proven to be reliable.

Acknowledgment

The paper is rooted in the author's long-term cooperation with Professor Jacek Osiewalski in the area of the hybrid SV-GARCH models.

Research supported by a grant from the National Science Center under decision no. UMO-2018_31/B/HS4/00730

References

Abramowitz M., Stegun N. [1968], *Handbook of Mathematical Functions*, Dover Publication, New York.

Baba Y., Engle R., Kraft D., Kroner K. [1989], *Multivariate Simultaneous Generalised ARCH*, Manuscript, University of California at San Diego, Department of Economics.

Bauwens L., Lubrano M. [1998], *Bayesian inference on GARCH models using the Gibbs sampler,* "Econometrics Journal", No. 1, pp. C23–C46.

Bauwens L., Lubrano M., Richard J.-F. [1999], *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, Oxford.

Brooks S.P. [1998], *Quantitative convergence assessment for Markov chain Monte Carlo via cusums*, "Statistics and Computing", No. 8, pp. 267–274.

Fernández C., Steel M.F.J. [1998], *On Bayesian Modelling of Fat Tails and Skewness*, "Journal of the American Statistical Association", No. 93 (441), pp. 359–371.

Fernández C., Steel M.F.J. [1999], *Multivariate Student-t regression models: Pitfalls and inference*, "Biometrika", No. 86, pp. 153–167.

Gelfand A.E., Smith A.F.M. [1990], *Sampling based approaches to calculating marginal densities*, "Journal of the American Statistical Association", No. 85, pp. 398–409.

Geweke J. [1993], *Treatment of the Independent Student-t Linear Model*, "Journal of Applied Econmetrics", No. 8 (S1), pp. S19–S40.

Geweke J. [1994], *Priors for Macroeconomic Time Series and Their Applications*, "Econometric Theory", No. 10, pp. 609–632.

Gilks W.R., Richardson S., Spiegelhalter D.J. [1996], *Markov Chain Monte Carlo in practice*, Chapman & Hall, London.

Jacquier E., Polson N., Rossi P. [1994], *Bayesian analysis of stochastic volatility models* [with discussion], "Journal of Business and Economic Statistics", No. 12, pp. 371–417.

Jacquier E., Polson N., Rossi P. [2004], *Bayesian analysis of stochastic volatility models with fat-tails and correlated errors*. "Journal of Econometrics", No. 122, pp. 185–212.

Osiewalski J. [2009], *New hybrid models of multivariate volatility a Bayesian perspective*, "Statistical Review", No. 56, pp. 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”, No. 8(4), pp. 241–271.

Osiewalski J., Pajor A. [2007], *Flexibility and parsimony in multivariate financial modelling: a hybrid bivariate DCC–SV model*, [In:] *Financial Markets. Principles of Modeling, Forecasting and Decision--Making* (FindEcon Monograph Series No. 3), [ed.] W. Milo, P. Wdowiński, Łódź University Press, Łódź.

Osiewalski J., Pajor A. [2009], *Bayesian analysis for hybrid MSF–SBEKK models of multivariate volatility*, “Central European Journal of Economic Modelling and Econometrics”, No. 1(2), pp. 179–202.

Osiewalski J., Pajor A. [2018], *A hybrid MSV-MGARCH generalisation of the t-MGARCH model*, [In:] The 12th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena: Conference Proceedings, No. 1 [ed.] M. Papież, S. Śmiech, pp. 345–354.

Osiewalski J., Pajor A. [2019], *On Sensitivity of Inference in Bayesian MSF-MGARCH Models*, “Central European Journal of Economic Modelling and Econometrics”, No. 11(3], pp. 173–197.

Yu B., Mykland P. [1998], *Looking at Markov samplers through cusum paths plots: a simple diagnostic idea*, “Statistics and Computing”, No. 8(3), pp. 275–286.

Indebted Households' Self-assessment of their Financial Situation: Evidence from Poland

Agnieszka Wałęga[1]

Abstract

The standard of living of households does not only depend on their income. The subjective financial assessments are primarily associated with day-to-day concerns. Therefore, the self-assessed financial situation can provide more detailed information about the household's living conditions. It is particularly crucial for indebted households, which may report difficulties with making ends meet even though their income is high enough to avoid such problems. The paper is based on the results of the questionnaire survey addressed to indebted Polish households in 2018. The ordered probit model was used to examine the relationship between the respondents' self-assessment of their financial situation and commonly used objective measures of over-indebtedness. The subjective financial situation of one's household depends on – among others – the level of debt and over-indebtedness risk. Regardless of the debt burden, the self-assessment of the financial situation among the youngest household cohorts is generally more accurate than among people aged 45+.

Keywords: indebtedness, standard of living, household, financial situation

JEL classification: D12, D14, I31

[1] Agnieszka Wałęga, Cracow University of Economics, e-mail: agnieszka.walega@uek.krakow.pl

1. Introduction

Households' financial situation is an essential determinant of consumption decisions and the standard of living. It can be measured in different ways. In general, objective indicators are based on monetary assessment, whereas subjective ones are based on survey questions about the perceived financial situation of the respondents[2]. Nevertheless, it is not possible to fully assess the financial situation of the household using only objective indicators. Households' subjective opinions matter for financial decision-making processes. It is commonly known that the objective financial situation of a given household might be completely different than it is suggested by the results of subjective measures. Thus, it might lead to different financial decisions within the household [Bialowolski, Wieziak-Bialowolska, 2014]. It means that the subjective and objective indicators capture different information concerning the development of the household's financial situation in the future.

There are many works investigating the subjective assessment of the financial situation, some of which focus on satisfaction with household income [Joo, Grable, 2004; Vera-Toscano et al., 2006; Stanovnik, Verbič, 2006; D'Ambrosio, Frick, 2007; Archuleta et al., 2011; Ranta, Salmela-Aro, 2018]. The research on the subjective assessment of the households' income in Poland was conducted by Ulman [2006], Dudek [2009], Liberda et al. [2012], Dudek [2013] and Kasprzyk [2016].

The review of the literature suggests that the subjective financial situation is determined by socioeconomic factors [Vera-Toscano et al., 2006]. The researchers also find the dependence between the debt and well-being [Baek, DeVaney, 2004; Tay et al., 2017]. The household's debts may exert a negative effect both on the individual's economic well-being [Baek, DeVaney, 2004] or consumption [Kukk, 2016]. Indebtedness may affect the individual's well-being in different aspects e.g., economic, psychological [Brown et al., 2005], physical health [Clayton et al., 2015], and family relationships [Reading, Reynolds, 2001].

The debt burden affects subjective financial well-being and is likely to exert spill-over effects that influence other life domains, such as leisure, relationships, family life, etc. [Tay et al., 2017]. This is because financial well-being is one of the vital life domains.

The paper aims to explore the determinants of Polish indebted households' self-assessment of their financial situation, focusing on the role of over-indebtedness. The analysis is based on a sample of households drawn from the primary survey of indebted households conducted in 2018.

[2] Household finances are economic phenomena occurring within a household that are directly related to the accumulation and spending of funds [Bywalec, 2009]. The financial situation of households is the result of these phenomena.

2. Data and methods

Regression models can be used to identify factors influencing the assessment of the respondents' financial situation. Due to the nature of the dependent variable, the study is based on the ordered probit model built around the latent regression model in the same manner as the binominal probit model. The probability that the variable (discrete variable) takes a particular value is specified with the use of the standardised normal distribution function in the following manner [Greene, 2008]:

$$P(Y=1)=\Phi(\beta' x+\beta_1)$$

$$P(Y=2)=\Phi(\beta' x+\beta_2)-\Phi(\beta' x+\beta_1)$$

$$P(Y=j)=1-\Phi(\beta' x+\beta_{j-1})$$

where: Φ – the distribution function of the standard normal distribution, β – the vector of estimated parameters, x – the vector of independent variables, $\beta_1 \ldots \beta_j$ – estimated constants.

The parameters of the model are estimated using the maximum likelihood. The positive value of the particular β parameter for the adequate independent variable x (with increasing values) increases the probability of the occurrence of the first value of dependent variable Y and reduces the probability of the occurrence of the last value j. In the case of the middle values of the dependent variable, the changes in the probability of the occurrence of these values are not unequivocal. It is impossible to determine (having the estimated parameter) in which direction the change of the probability goes in these classes. We can unequivocally interpret the parameters concerning the first and last class. At the same time, in the case of the whole range of the dependent variable, we can only perform simulations showing the development of the structure of changes in the assessment of the financial situation resulting from the changes of the explanatory variables. Therefore, caution is needed in interpreting the coefficients in this model.

The analysis of the financial situation of indebted households was based on the data obtained from the primary survey conducted in the first half of 2018 using the CATI technique. The respondents were adults aged 18 or above who had at least one loan commitment (either secured or unsecured). Ultimately the database contains 1,107 individuals from all over Poland. The respondents were asked questions connected with their demographic characteristics, their household's debt and income, and their attitudes to debt.

The paper focuses on the influence of over-indebtedness on the assessment of the households' financial situation. To identify over-indebted households, the measures commonly used were adopted [D'Alessio, Iezzi, 2016; Betti et al., 2007]. It was assumed that over-indebted households are those whose spending on total borrowing repayments take them below the poverty line (BPL) – (equal to 60% of the median income using the modified OECD scale of equivalence) and those whose debt-service to income indicator is 30% and more (DSTI30). Besides, the number of credit agreements (4 or more – NL4) and being in arrears (more than three instalments – A3) were taken into account. The additional complementary indicator was the subjective assessment (SB) of over-indebtedness (the answer to the question of whether the respondent feels over-indebted).

3. Empirical results

The analysis of the financial situation was conducted using a variable of a subjective nature. The self-assessed financial situation was determined on the following statements from the questionnaire: we lack money even for necessary daily expenses (1); we have enough money to satisfy only our basic needs and nothing more (2); we economise to satisfy at least our basic needs, and we sometimes manage to save some money on something more expensive (3); we have enough money to meet all our needs (4); we are comfortably well-to-do, we can meet all our needs, even the sophisticated ones (5).

The following characteristics of the households and the respondents have been used as explanatory variables:

- household monthly income (up to 2000 PLN; (2000–4000] PLN; (4000–6000] PLN; (6000–8000] PLN; over 8000 PLN);
- level of repayments (up to 200 PLN; (200–500] PLN; (500–1000] PLN; (1000–2000] PLN; over 2000 PLN);
- household size (number of persons);
- gender (1 – male, 0 – female);
- education level (vocational or lower, secondary, tertiary);
- age (18–24; 25–34; 35–44; 45–54; 55–64; 65+);
- place of residence (rural, town, city);
- over-indebtedness (five dummy variables identifying over-indebted households).

Table 1 presents the structure of the sample. Most households were made up of the respondents with secondary education (41.8%) or higher education (39.4%). The groups of respondents aged between 25, and 34 (23.5%) and 35–44 (23.7%) were the most numerous. Only about 6% of respondents were younger than 25 and 9.5% were older than 64. Approximately 54% of the respondents lived in rural areas (25.7%) or small towns (28.7%), and around 13.6% in big cities. Regarding the household size, 63.7% of the respondents' households were rather small (three persons or fewer) – only 4.3% of households consisted of more than five persons. Taking into consideration the general economic and financial condition, in 51.6% of the households' net monthly income was higher than 4000 PLN, and the income below 2000 PLN declared 15.4% of respondents. As far as their debts were concerned, 73.1% of households spent below 1000 PLN on their monthly debt repayment, while 6,8% – over 2000 PLN.

Table 1. The structure of the sample (%)

Specification	Percent	Specification	Percent
income		level of repayments	
Up to 2000 PLN	15.36	Up to 200 PLN	20.86
(2000–4000] PLN	33.00	(200–500] PLN	28.05
(4000–6000] PLN	28.05	(500–1000] PLN	24.23
(6000–8000] PLN	10.60	(1000–2000] PLN	20.04
Over 8000 PLN	12.98	Over 2000 PLN	6.83
no. of persons		age	
1	10.84	18–24	6.14
2	25.29	25–34	23.49
3	27.55	35–44	23.67
4	24.66	45–54	21.23
5	7.41	55–64	15.99
6+	4.25	65+	9.49
place of residence		level of education	
Rural	25.66	Vocational or lower	18.79
Town	44.61	Secondary	41.82
City	29.74	Tertiary	39.39
over-indebtedness			
4 or more credit commitments (NL4)	4.61	In arrears: more than 3 instalments (A3)	1.90
Debt service to income ratio: 30% and more (DSTI30)	16.71	Self-assessment of over-indebtedness (SB)	17.07

Source: own calculations based on the primary survey data.

The distribution of the assessments of the households' financial situation is negatively skewed, which means that households more often chose the highest values on the scale (4 and 5). This situation indicates households' positive assessment of their financial condition, despite their debt (Table 2).

Table 2. Self-assessment of the respondents' financial situation by selected measures of over-indebtedness

Specification	Total	Self-assessment of over-indebtedness (SB)		4 or more credit commitments (NL4)		In arrears: more than 3 instalments (A3)		Debt service to income ratio: 30% and more (DSTI30)		DSTI30 & SI	
		IN	OIN	IN	OIN	IN	OIN	IN	OIN	IN	OIN
1	0.45	0.22	1.60	0.47	0.00	0.18	14.29	0.33	1.29	0.39	1.37
2	4.62	2.07	17.02	4.27	11.76	4.16	28.57	3.59	9.68	3.49	20.55
3	25.00	22.05	39.36	24.88	27.45	24.84	33.33	23.07	35.48	23.67	43.84
4	39.86	41.92	29.79	39.98	37.25	40.44	9.52	41.57	32.58	40.74	27.40
5	30.07	33.73	12.23	30.39	23.53	30.38	14.29	31.45	20.97	31.72	6.85

IN – indebted, OIN – over-indebted

Source: own calculations based on the primary survey data.

When comparing the assessment of the financial situation of indebted and over-indebted households, it can be concluded that, in general, over-indebted households assess their financial situation as worse than indebted households (Table 2). It happens regardless of the criterion adopted for identifying households as over-indebted. More often, over-indebted households chose lower values on the scale, and even the distribution is positively skewed (in arrears: more than 3 instalments – A3). It stays in line with expectations and previous [Tay et al., 2017] research.

The analysis raises a question about the determinants of the subjective assessment of the household's financial situation. The interesting aspect is the role of over-indebtedness in shaping personal well-being. To achieve this aim the econometric modelling was adopted. The financial situation was as a dependent variable and was expressed in five categories. In the first model (Model 1), the assessment of the financial situation was explained by variables identifying the household as over-indebted obtained by the measures listed in the previous section. All the proposed measures are dummy variables (according to the adopted criterion zero means a household that is indebted and one – an over-indebted household). The estimation proves that some parameters were statistically insignificant in determining the households' financial situation (Table 3). Among other things, four constants turned out to be important. Their number is directly related to the number of categories of the dependent variable and always equals this number minus one. The statistical significance of the three explanatory variables (Model 1) indicates that the assessment of the financial situation depends on the subjective assessment of belonging to the group of over-indebted households (SB). Over-indebted households assess their financial situation as worse than households that do not assess themselves as over-indebted. If the BLP indicator or A3 classifies a household as over-indebted, the probability of a worse assessment of its financial situations is increased.

The estimated parameters allow for analysing the changes in the probability in two extreme classes (the worst financial situation (1) – we lack money even for necessary daily expenses, and the best financial situation (5) – we are comfortably well-to-do, we can meet all our needs, even the sophisticated ones) regarding specific characteristics. The model indicates that about 38% of indebted households – which are not classified as over-indebted ones – would assess their financial situation as very good. In comparison only 0.5% of the over-indebted households (classified as such by all the measures adopted in the study) would assess their situation as very good.

The following variables describing socio-economic characteristics of the respondents and their households were added to the previously used set of explanatory variables in the model (Model 2): gender, age, level of education, place of residence, number of people in the household, income, and debt repayments. Only age and income considerably affected the assessment of the financial situation. A negative sign of the parameter of the income variable indicates a better assessment of the financial situation when the income level increases. Moreover, the older the respondents, the probability that they would assess their financial situation as bad increases.

Table 3. Ordered probit regression: households' self-assessed financial situation (the whole sample)

Specification	Model 1		Model 2	
	Parameter	Standard error	Parameter	Standard error
Constant 1	–3.302***	0.197	–3.301***	0.331
Constant 2	–2.098***	0.080	–2.047***	0.269
Constant 3	–0.819***	0.049	–0.635***	0.260
Constant 4	0.300***	0.046	0.537*	0.260
SB	0.766***	0.094	0.740*	0.101
NL4	–0.108	0.161	0.112	0.173
A3	1.022***	0.245	0.924***	0.264
DSTI30	0.031	0.096	0.010	0.120
BPL	0.452***	0.078	0.168	0.103
Gender	–	–	0.324	0.569
Income	–	–	–0.402***	0.087
No. of persons	–	–	–0.019	0.032
Age	–	–	0.399***	0.050
Education	–	–	0.051	0.054
Place of residence	–	–	–0.015	0.049
Level of repayments	–	–	–0.026	0.041
	Stat.	Stat/Df	Stat.	Stat/Df
Pearson Chi-square	3779.78	0.858	3482.93	0.868
AIC	2598.68	–	2278.89	–
Log(likelihood ratio)	–1290.34	–	–1123.45	–
Counting R^2	41.93%	–	44.39%	–

Note: *** – $p < 0.001$; ** – $p < 0.05$; * – $p < 0.1$

Source: own calculations based on the primary survey data.

A significant effect of age on the assessment of the respondents' financial situation gives the grounds for the analysis of the issue in the cross-section. The increase in the respondents' age is accompanied by the decrease in the percentage of responses identifying their financial situation as very good, and the increase in the percentage of responses with average (3) and bad (2) ratings. The correlation analysis (Cramér's V = 0.21, $p < 0.001$) indicates a weak but significant correlation between the examined variables.

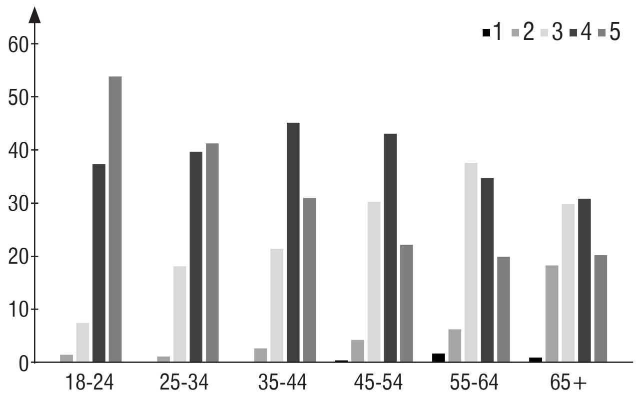


Figure 1. The structure of the assessment of the financial situation of households by age of the respondents

Source: own calculations based on the primary survey data.

Note: 1 – we lack money even for necessary daily expenses; 2 – we have enough money to satisfy only our basic needs and nothing more; 3 – we economise to satisfy at least our basic needs, and we sometimes manage to put some money aside on something more expensive; 4 – we have enough money to meet all our needs; 5 – we are comfortably well-to-do, we can meet all our needs, even the sophisticated ones.

Table 4. Ordered probit regression: households' self-assessed financial situation by age groups

Specification	Up to 34	35–54	55+
	Parameter	Parameter	Parameter
Constant 1	–1.5180**	–2.8333***	–2.1561***
Constant 2	0.0055	–1.4443***	–0.7995*
Constant 3	1.1804**	0.1078	0.5471
Constant 4	–	1.4238***	1.5546***
SB	0.8444***	0.6725***	0.6898***
NL4	–0.2329	0.2698	–0.2354
A3	–0.1192	1.2647**	1.8570**
DSTI30	–0.2238	0.1139	0.0068
BPL	–0.2696	0.3384**	0.4222**
Gender	0.0668	0.0595	–0.2547*
Income	–0.4683**	–0.4885***	–0.2756
No. of persons	0.0006	–0.0384	–0.0902
Education	0.0081	0.1599**	–0.0188

Specification	Up to 34	35–54	55+
	Parameter	Parameter	Parameter
Place of residence	–0.1686*	0.0590	0.0462
Level of repayments	0.0972	–0.0963	–0.0439
Pearson Chi-square	901.88	1498.44	929.43
AIC	646.27	1002.30	636.95
Log(likelihood ratio)	–309.14	–486.15	–303.47
Counting R^2	47.47%	47.90%	38.55%

Note: *** – p < 0.001; ** – p < 0.05; * – p < 0.1

Source: own calculations based on the primary survey data.

The results of the ordered probit model according to age groups are presented in Table 4. The parameters reflect the influence of socio-economic characteristics on the assessment of the financial situation in particular age groups. The subjective assessment of over-indebtedness (SB) has a significant effect on the dependent variable. The greatest influence can be observed in the group of respondents below 34. The below poverty line index (BPL) and being in arrears (A3) considerably affect the group of respondents above 34. The influence of both these variables is greater in the households of respondents aged 55+. The income in the households of respondents below 54 markedly influences the assessment of their financial situation. An increase in income reduces the probability of a "bad" subjective assessment of their financial situation. The place of the residence turned out to be significant ($p < 0.1$) for the respondents below 34 – the larger the place of residence is, the lower the probability of assessment of the bad financial situation is. For the respondents aged 35–54, the level of education was also significant ($p<0.05$) – the higher level of the respondents' education was correlated with the higher probability of a negative assessment of their financial situation. Gender also played a substantial role ($p < 0.1$) in the assessment of the financial situation among the respondents aged 55+ – men less frequently assessed their financial situation as worse than women.

4. Conclusions

The main finding suggests that, in general, indebted households have a positive assessment of their financial situation. Most households have sufficient financial resources to meet all their needs or to be able to satisfy even sophisticated ones. Regardless of the measure, the subjective assessment becomes worse when a household is classified as over-indebted.

Apart from over-indebtedness, the assessment of the financial situation is influenced by the household's income and the respondents' age. Growing income implies an improved assessment of the financial situation of indebted households. Similar observations apply to all households [Dudek, 2013]. The age also proved to be significant – the probability of a negative self-perception of households' financial situation increases with their age. Similar conclusions were drawn by Parlińska and Petrych [2014], who modelled the subjective assessment of farmers' financial situation. Depending on the respondents' age (apart from being classified as over-indebted and the level of income), the place of residence, education level, and gender also exerted a noticeable effect on self-perceived financial situation.

Limitations of this study result from the measurement of the subjective assessment of the financial situation. Using a single item (one question) to assess financial satisfaction may not be the best technique [Baek, DeVaney, 2004]. Using a multi-item scale in further studies could provide a more accurate measure of the respondents' subjective assessment of their financial situation (well-being). A comparative analysis of the financial situation assessed subjectively and measured by objective measures seems to be a promising approach, especially that Giarda [2013] points out that using only subjective indicators may be not sufficiently reliable.

Acknowledgements

This paper was supported by funds from the National Science Centre (NCN, Poland) through grant No. 2015/19/D/HS4/02569.

References

Archuleta K.L., Britt S.L., Tonn T.J., Grable J.E. [2011], *Financial satisfaction and financial stressors in marital satisfaction*, "Psychological reports", 108(2), pp. 563–576.

Baek E., DeVaney S.A. [2004], *Assessing the baby boomer financial wellness using financial ratios and a subjective measure*, "Family and Consumer Sciences Research Journal", 32(4), pp. 321–348.

Betti G., Dourmashkin N., Rossi M., Ping Yin Y. [2007], *Consumer over-indebtedness in the EU: measurement and characteristics*, "Journal of Economic Studies", 34(2), pp. 136–156.

Bialowolski P., Weziak-Bialowolska D. [2014], *The index of household financial condition, combining subjective and objective indicators: An appraisal of Italian households*, "Social Indicators Research", 118(1), pp. 365–385.

Brown S., Taylor K., Price S.W. [2005], *Debt and distress: Evaluating the psychological cost of credit*, "Journal of Economic Psychology", 26(5), pp. 642–663.

Bywalec C. [2009], *Ekonomika i finanse gospodarstw domowych*. PWN, Warszawa.

Clayton M., Liñares-Zegarra J., Wilson J.O. [2015], *Does debt affect health? Cross country evidence on the debt-health nexus.* "Social science & medicine", 130, pp. 51–58.

D'Ambrosio C., Frick J.R. [2007], *Income satisfaction and relative deprivation: An empirical link*, "Social Indicators Research", 81(3), pp. 497–519.

D'Alessio G., Iezzi S. [2016], *Over-indebtedness in Italy: how widespread and persistent is it?* "Bank of Italy Occasional Paper", 319.

Dudek H. [2009], *Subjective aspects of economic poverty-ordered response model approach*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu", 73 (Quality of Life Improvement through Social Cohesion), pp. 9–24.

Dudek H. [2013], *Subiektywne postrzeganie sytuacji dochodowej – mikroekonometryczna analiza danych panelowych*, „Roczniki Kolegium Analiz Ekonomicznych", 30, pp. 219–233.

Giarda E. [2013], *Persistency of financial distress amongst Italian households: Evidence from dynamic models for binary panel data*, "Journal of Banking & Finance", 37(9), pp. 3425–3434.

Greene W.H. [2008], *Econometric analysis* 6th edition, Pearson Education Inc, New Jersey.

Joo S.H., Grable J.E. [2004], *An exploratory framework of the determinants of financial satisfaction*, "Journal of family and economic Issues", 25(1), pp. 25–50.

Kasprzyk B. [2016], *Subiektywne oceny dobrobytu ekonomicznego w gospodarstwach domowych w świetle modelowania dyskryminacyjnego*, „Ekonomista", 2, pp. 233–250.

Kukk M. [2016], *How did household indebtedness hamper consumption during the recession? Evidence from micro data*, "Journal of Comparative Economics", 44(3), pp. 764–786.

Liberda B., Pęczkowski M., Gucwa-Leśny E. [2012], *How do we value our income from which we save?* "Annals of the Alexandru Ioan Cuza University-Economics", 59(2), pp. 93–104.

Parlińska M., Pietryc, Ł. [2014], *Modelowanie subiektywnej oceny sytuacji finansowej oraz warunków pracy rolników*, „Roczniki" (Annals), 16(3), pp. 218–223.

Ranta M., Salmela-Aro K. [2018], *Subjective financial situation and financial capability of young adults in Finland*, "International Journal of Behavioral Development", 42(6), pp. 525–534.

Reading R., Reynolds S. [2001], *Debt, social disadvantage and maternal depression*, "Social science & medicine", 53(4), pp. 441–453.

Stanovnik T., Verbič M. [2006], *Analysis of subjective economic well-being in Slovenia*, "Eastern European Economics", 44(3), pp. 60–70.

Tay L., Batz C., Parrigon S., Kuykendall L. [2017], *Debt and subjective well-being: The other side of the income-happiness coin*, "Journal of Happiness Studies", 18(3), pp. 903–937.

Ulman P. [2006], *Subjective assessment of economic poverty in Poland*, The 25th SCORUS Conference on Regional and Urban Statistics and Research "Globalization Impact on Regional and Urban Statistics", Wrocław, Poland. 30.08–1.09.2006.

Vera-Toscano E., Ateca-Amestoy V., Serrano-Del-Rosal R. [2006], *Building financial satisfaction*, "Social Indicators Research", 77(2), pp. 211–243.

This monograph presents some interesting applications of quantitative methods in studying the phenomena of economic processes using mathematical knowledge and tools. The area of scientific research is diversified and covers topics relating to macroeconomics, consumer theory, socio-economic development, households quality of life, heteroskedastic Vector Autoregressive models and credit risk models.