METODY ILOŚCIOWE
W BADANIACH EKONOMICZNYCH

QUANTITATIVE METHODS
IN ECONOMICS

Vol. XV, No. 2

Warsaw University of Life Sciences – SGGW
Faculty of Applied Informatics and Mathematics
Department of Econometrics and Statistics

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH

QUANTITATIVE METHODS IN ECONOMICS

Volume XV, No. 2

Warsaw 2014

ISSN 2082 – 792X

Warsaw 2014, Volume XV, No. 2

The original version is the paper version

Journal homepage: qme.sggw.pl

Published by Warsaw University of Life Sciences Press

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014

CONTENTS

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 7 – 16

MANAGEMENT OF AN AGRICULTURAL ENTERPRISE ON THE BASIS OF ITS ECONOMIC STATE FORECASTING

Igor Atamanyuk
Department of Higher and Applied Mathematics
Mykolaiv National Agrarian University
e-mail: atamanyuk_igor@mail.ru

Yuriy Kondratenko
Department of Intelligent Information Systems
Petro Mohyla Black Sea State University
e-mail: y_kondrat2002@yahoo.com

Natalia Sirenko
Department of Finances and Credit
Mykolaiv National Agrarian University
e-mail: sirenko@mnau.edu.ua

Abstract: On the basis of the mechanism of accidental sequences canonical expansions the algorithm of the economic state of agricultural enterprise forecasting is obtained which allows to estimate the results of its work in future under the realization of a certain reorganization (change of land resources, labour resources, fixed assets).

Keywords: accidental sequence, canonical expansion, algorithm of extrapolation.

INTRODUCTION

Economic state is the most important criterion of business activity and reliability of an enterprise that determines its competitive ability and potential in effective realization of the economic interests of all participants of economic activity. For ensuring of successful work the management of an enterprise needs to be able to estimate and forecast realistically its economic state as well as partners and competitors. The models of forecasting are the one instrument of the

determination of current enterprise state or possibilities of its development. But such practice in the management of Ukrainian enterprises is absent. Mainly the experts whose methods don't have clear scientific ground and have the name "nosology" which means intuitive approaches basing on personal working experience are occupied with the forecasting at enterprises and banks. Western specialists have the priority in the investigation of the possibilities of the management on the basis of the forecasting of enterprise economic state. Bever started theoretical development and building of prognostic models, then it was continued in the works of Altman (the USA) [Altman 1994], [Altman and other 1997], Alberichi (Italy), Misha (France) and others [Granger and other 1986], [Hall 1994]. More contemporary trend in the building of the algorithms of economic indices forecasting is the usage of stochastic methods of extrapolation. The relevance of such approach is explained with the influence of great number of accidental factors on the results of enterprise functioning (weather conditions, accidental variations of demand and supply, inflation etc.), under the influence of which the change of economic state indices obtains accidental character. But the existing models of prognosis impose considerable limitations on the accidental sequence describing the change of economic indices [Trifonov and other 1998], [Ryabushkin 1987], [Teyl 1971], [Szmuksta-Zawadzka 2013], [Prędki 2013], [Połoński 2012] (Markovian property, stationarity, monotony, scalarity etc.). Thereupon the problem of the building of the forecast model under the most general assumptions about the stochastic properties of the accidental process of the change of the indices of enterprise economic state arises.

AIM AND THE RAISING OF PROBLEM

The aim of this work is the development of the technology of agricultural enterprise management on the basis of the algorithm of the forecasting of the indices of its work. The main requirement to the forecasting algorithm is the absence of any essential limitations on the stochastic properties of the accidental process of economic indices change.

THE SOLVING OF PROBLEM

The most universal from the point of view of the requirements to the investigated accidental sequence is the method that bases on the mechanism of canonical expansions [Pugachev 1962], [Kudritskiy 2001]. The main primary indices of the economic state of agricultural enterprises are the gross profit, gross output, land resources, labour resources, fixed assets that is why the object of the investigation is the vector accidental sequence with five dependant constituents (if necessary the number of figures and their qualitative composition may be changed). Preliminary investigations (the check of dependence of accidental values) showed that the accidental sequences describing the change of the

economic state of the enterprises which relate to the intensive [Sirenko 2010] type of the development during the interval of eleven years that corresponds to the processing of twelve annual indices for the great number of the enterprises of the mentioned type have the most stable and significant stochastic relations. For such vector accidental sequence the canonical expansion has the following look:

$$X_h(i) = M\left[X_h(i)\right] + \sum_{\nu=1}^{i}\sum_{\lambda=1}^{5} V_\nu^{(\lambda)} \varphi_{h\nu}^{(\lambda)}(i),\ i=\overline{1,12},\ h=\overline{1,5}, \tag{1}$$

where $X_1(i)$, $i=\overline{1,12}$ - gross profit;

$X_2(i)$, $i=\overline{1,12}$ - gross output;

$X_3(i)$, $i=\overline{1,12}$ - land resources;

$X_4(i)$, $i=\overline{1,12}$ - labour resources;

$X_5(i)$, $i=\overline{1,12}$ - fixed assets.

The elements of canonical expansion are the accidental coefficients $V_\nu^{(\lambda)}$, $\nu=\overline{1,12}$, $\lambda=\overline{1,5}$ and nonrandom coordinate functions $\varphi_{h\nu}^{(\lambda)}(i)$, $\nu=\overline{1,12}$, $\lambda=\overline{1,5}$, $h=\overline{1,12}$, $i=\overline{1,5}$:

$$V_\nu^{(\lambda)} = X_\lambda(\nu) - M\left[X_\lambda(\nu)\right] - \sum_{\mu=1}^{\nu-1}\sum_{j=1}^{5} V_\mu^{(j)} \varphi_{\lambda\mu}^{(j)}(\nu) - \tag{2}$$

$$-\sum_{j=1}^{\lambda-1} V_\nu^{(j)} \varphi_{\lambda\nu}^{(j)}(\nu), \nu=\overline{1,12};$$

$$D_\lambda(\nu) = M\left[\left\{V_\nu^{(\lambda)}\right\}^2\right] = M\left[\left\{X_\lambda(\nu)\right\}^2\right] - M^2\left[X_\lambda(\nu)\right] - \tag{3}$$

$$-\sum_{\mu=1}^{\nu-1}\sum_{j=1}^{H} D_j(\mu)\left\{\varphi_{\lambda\mu}^{(j)}(\nu)\right\}^2 - \sum_{j=1}^{\lambda-1} D_j(\nu)\left\{\varphi_{\lambda\nu}^{(j)}(\nu)\right\}^2,\ \nu=\overline{1,12};$$

$$\varphi_{h\nu}^{(\lambda)}(i) = \frac{M\left[V_\nu^{(\lambda)}\left(X_h(i) - M[X_h(i)]\right)\right]}{M\left[\left\{V_\nu^{(\lambda)}\right\}^2\right]} = \frac{1}{D_\lambda(\nu)}\left(M\left[X_\lambda(\nu) X_h(i)\right] - \right. \tag{4}$$

$$-M\left[X_\lambda(\nu)\right]M\left[X_h(i)\right] - \sum_{\mu=1}^{\nu-1}\sum_{j=1}^{5} D_j(\mu)\varphi_{\lambda\mu}^{(j)}(\nu)\varphi_{h\mu}^{(j)}(i) -$$

$$\left. -\sum_{j=1}^{\lambda-1} D_j(\nu)\varphi_{\lambda\nu}^{(j)}(\nu)\varphi_{h\nu}^{(j)}(i),\ \lambda=\overline{1,5}, \nu=\overline{1,i}.\right.$$

Coordinate functions $\varphi_{h\nu}^{(\lambda)}(i), h, \lambda = \overline{1,5},\ \nu, i = \overline{1,12}$ have the following properties:

$$\varphi_{h\nu}^{(\lambda)}(i) = \begin{cases} 1, & h = \lambda \quad \& \quad \nu = i; \\ 0, & i < \nu \ or \ \ h < \lambda \ \& \ \nu = i. \end{cases} \tag{5}$$

The algorithm of extrapolation on the basis of canonical expansion has the look [Atamanyuk 2005], [Atamanyuk and other 2012]:

$$m_h^{(\mu,l)}(i) = \begin{cases} M[X_h(i)], \mu = 0, \\ m_h^{(\mu,l-1)}(i) + \left[x_l(\mu) - m_l^{(\mu,l-1)}(\mu) \right] \varphi_{h\mu}^{(l)}(i), l \neq 1, \\ m_h^{(\mu,5)}(i) + \left[x_1(\mu) - m_1^{(\mu-1,5)}(\mu) \right] \varphi_{h\mu}^{(1)}(i), l = 1, \end{cases} \tag{6}$$

where $m_h^{(\mu,l)}(i) = M\left[X_h(i) / x_\lambda(\nu),\ \lambda = \overline{1,5},\ \nu = \overline{1,\mu-1};\ x_j(\mu), j = 1,l \right],\ h = \overline{1,5},$ $i = \overline{\mu+1,12}$ - is the linear optimal quantity by the criterion of the minimum of the average square of the error of the prognosis is the estimation of the future values of the investigated sequence under the condition that the values for random parameters $X_\lambda(\nu)$ are known $x_\lambda(\nu),\ \lambda = \overline{1,5},\ \nu = \overline{1,\mu-1};\ x_j(\mu),\ j = \overline{1,l}$. In [Atamanyuk 2009] in the form of a theorem is proved that the algorithm has optimal characteristics.

As it follows from (4) the values $\varphi_{h\nu}^{(\lambda)}(i), h, \lambda = \overline{1,5},\ \nu, i = \overline{1,12}$ are determined through auto- and mutually correlated functions of the investigated vector accidental sequence. In the Table 1 the values of autocorrelated function are

Table 1. Autocorrelated function of the accidental constituent $X_1(i),\ i = \overline{1,12}$

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
2002	1	0,99	0,70	0,42	0,79	0,74	0,49	0,72	0,63	0,46	0,55	0,43
2003	0,99	1	0,72	0,42	0,74	0,74	0,52	0,70	0,64	0,48	0,59	0,46
2004	0,70	0,72	1	0,57	0,67	0,58	0,70	0,69	0,70	0,66	0,78	0,60
2005	0,42	0,42	0,57	1	0,38	0,36	0,45	0,21	0,41	0,36	0,19	0,18
2006	0,79	0,74	0,67	0,38	1	0,81	0,55	0,91	0,80	0,72	0,53	0,41
2007	0,74	0,74	0,58	0,36	0,81	1	0,72	0,73	0,92	0,81	0,51	0,44
2008	0,49	0,52	0,70	0,45	0,55	0,72	1	0,51	0,74	0,73	0,49	0,41
2009	0,72	0,70	0,69	0,21	0,91	0,73	0,51	1	0,77	0,80	0,74	0,55
2010	0,64	0,64	0,70	0,41	0,80	0,92	0,74	0,77	1	0,91	0,60	0,59
2011	0,47	0,48	0,66	0,36	0,72	0,81	0,73	0,80	0,91	1	0,71	0,46
2012	0,55	0,59	0,78	0,19	0,53	0,51	0,49	0,74	0,60	0,71	1	0,71
2013	0,43	0,46	0,60	0,18	0,41	0,44	0,41	0,55	0,59	0,46	0,71	1

presented ($M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(i)\right]$, $\nu=\overline{1,12}$, $i=\overline{1,12}$) for the first constituent.

For the period of 2002-2012 the values of the autocorrelated functions $M\left[\overset{o}{X}_h(\nu)\overset{o}{X}_h(i)\right]$, $\nu=\overline{1,11}$, $i=\overline{1,11}$, $h=\overline{1,5}$ determined by means of the processing of statistic data (indices of the activity of Nikolaev region agricultural enterprises during 2002-2012). For 2013 $M\left[\overset{o}{X}_h(\nu)\overset{o}{X}_h(12)\right]$, $\nu=\overline{1,11}$, $h=\overline{1,5}$ are calculated on the basis of the determinate models:

$$M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(12)\right]=0,718M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(11)\right]-0,053M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(10)\right]+$$
$$+0,2128M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(9)\right]-0,105M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(8)\right],\ \nu=\overline{1,11}, \tag{7}$$

$$M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(12)\right]=1,435M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(11)\right]-0,01M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(10)\right]+$$
$$+0,082M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(9)\right]-0,011M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(8)\right]-0,485,\ \nu=\overline{1,11}, \tag{8}$$

$$M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(12)\right]=0,997M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(11)\right]-0,002M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(10)\right]+$$
$$+0,002M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(9)\right]-0,015M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(8)\right],\ \nu=\overline{1,11}, \tag{9}$$

$$M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(12)\right]=0,995M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(11)\right]+0,003M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(10)\right]+$$
$$-0,001M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(9)\right]-0,002M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(8)\right],\ \nu=\overline{1,11}, \tag{10}$$

$$M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(12)\right]=0,786M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(11)\right]-0,056M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(10)\right]+$$
$$-0,017M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(9)\right]+0,059M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(8)\right],\ \nu=\overline{1,11}, \tag{11}$$

The parameters of the equation (7)-(11) satisfy the minimum of the average error of approximation (relative error of the forecast is not more than 1 %).

In the Table 2 coordinate function $\varphi_{1\nu}^{(1)}(i)$, $\nu, i = \overline{1,12}$ corresponding to autocorrelated function $M\left[\overset{o}{X}_1(\nu)\overset{o}{X}_1(i)\right]$, $\nu = \overline{1,12}$, $i = \overline{1,12}$ and determining the degree of the influence of former values of gross profit for future values is presented.

Table 2. Value of the coordinate function $\varphi_{1\nu}^{(1)}(i)$, $\nu, i = \overline{1,12}$

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
2002	0,89	0,54	0,55	0,62	0,43	0,45	0,89	0,858	0,90	2,36	2,65
2003	1	2,25	-1,46	-2,40	0,27	5,47	-2,71	3,55	1,83	2,85	4,70
2004	0	1	5,09	1,17	-1,53	-0,03	-2,77	-0,23	-5,52	2,34	5,07
2005	0	0	1	0,17	0,26	0,94	0,18	0,77	1,05	-0,57	-1,17
2006	0	0	0	1	0,48	1,27	1,06	1,05	2,01	-2,37	0,69
2007	0	0	0	0	1	-1,81	0,74	3,53	0,37	9,31	2,86
2008	0	0	0	0	0	1	-0,68	1,44	3,18	-6,74	-3,39
2009	0	0	0	0	0	0	1	1,29	2,21	-3,30	0,93
2010	0	0	0	0	0	0	0	1	3,88	0,19	-8,44
2011	0	0	0	0	0	0	0	0	1	1,99	-4,96
2012	0	0	0	0	0	0	0	0	0	1	0,50
2013	0	0	0	0	0	0	0	0	0	0	1

Additionally to the Table 2 in the model (6) the values $\varphi_{h\nu}^{(\lambda)}(i), h, \lambda = \overline{1,5}$, $h \neq \lambda$, $\nu, i = \overline{1,12}$ which allow to consider mutual stochastic relations between the constituents $X_h(i)$, $h=\overline{1,5}$ (for example, the influence of land resources on gross profit, labour resources on gross output etc.) are used.

The future values of the mathematical expected value of the investigated vector accidental sequence {X} are estimated with the usage of the determinate model

$$M\left[X_h(12)\right] = 2,392M\left[X_h(11)\right] - 1,923M\left[X_h(10)\right] + \\ +1,087M\left[X_h(9)\right] - 0,105M\left[X_h(8)\right],\ h = \overline{1,5}. \qquad (12)$$

The parameters of the equation (12) as well as in (7)-(11) are determined from the condition of the minimum of the average error of approximation. For agricultural enterprises of Nikolaev region related to intensive type of development the mathematical expectations are $M\left[X_1(12)\right] = 4276,9$, $M\left[X_2(12)\right] = 12844,5$.

In all in the algorithm of the prognosis (6) 55 entrance values $x_h(i), h=\overline{1,5},\ i=\overline{1,11}$ and 1775 that are not equal to zero balance coefficients $\varphi_{h\nu}^{(\lambda)}(i), h,\lambda=\overline{1,5},\ \nu,i=\overline{1,12}$ are used.

For the increase of the effectiveness of the calculating processes during the prognosis by extrapolator (6) it is advisably to use the calculating procedure the substance of which is the fulfillment of the following stages:

Step 1. For the fixed point t_ν (initially $\nu=1$) the dispersions $D_\lambda(\nu)$(initially $\lambda=1$) of the accidental coefficients $V_\nu^{(\lambda)}$ with the help of the expression (3) are determined;

Step 2. Using the obtained at the previous step value $D_\lambda(\nu)$coordinate functions $\varphi_{h\nu}^{(\lambda)}(i)$ for $h=\overline{\lambda,5};\ i=\overline{\nu,12}$ by the formula (4) are calculated;

Step 3. The condition $\lambda<5$ is checked. If the outcome is positive, λ is increased by one $\lambda=\lambda+1$and the transition to Step 1 is fulfilled. Otherwise the calculating process is continued by the transition to the next Step 4.

Step 4. The check of $\nu<12$ is fulfilled. If the condition is performed, the value ν is increased by one $\nu=\nu+1$, the parameter λ is given the value one $\lambda=1$ and the transition to Step 1 is fulfilled. If the condition is not carried out, it means that the parameters of the extrapolator are determined for all points of discretization in which accidental process is viewed and the transition to Step 5 is fulfilled;

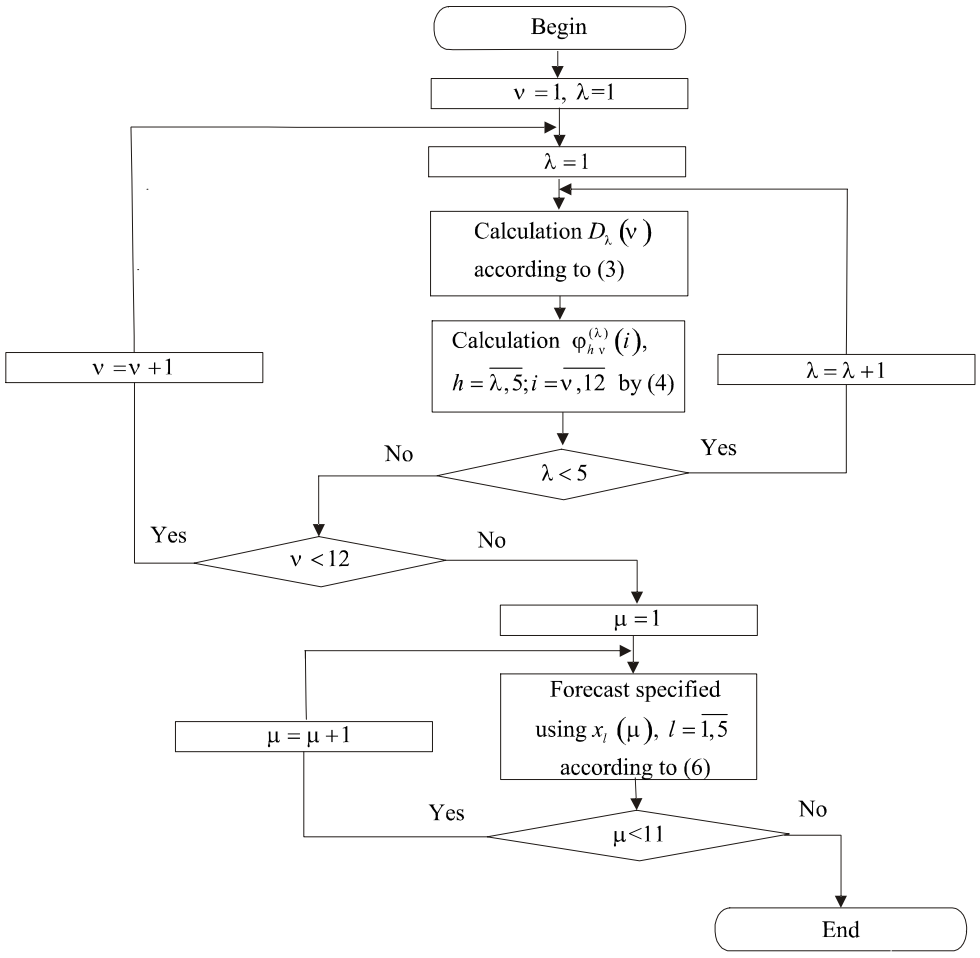
Step 5. The estimation of the future value of the investigated process is specified by the introduction into the calculating process the next value $x_l(\mu),\ l=\overline{1,5}$ (initially $\mu=1$). For $l=1$ the third expression of the formula (6) is used, for $l=\overline{2,5}$ the second one is used;

Step 6. It is checked whether all values are used for the forecast: $\mu=11$. If the condition is fulfilled, the process of calculations is finished, otherwise the value μ. Is increased by one $\mu=\mu+1$ and the transition to Step 5 is fulfilled.

The block diagram in Figure 1 illustrates the work of the algorithm.

Model (6) gives the possibility to estimate gross profit $x_1(12)$ and gross output $x_2(12)$ for 2013 for a certain enterprise basing on the data $x_h(i), h=\overline{1,5},\ i=\overline{1,11}$ of its work for eleven previous years. The comparison of the prognostic values which are obtained by means of the extrapolation algorithm with the statistic data of the results of agricultural enterprises work of Nikolaev region for 2013 indicates high effectiveness of the developed prognostic model (relative error 2-3%).

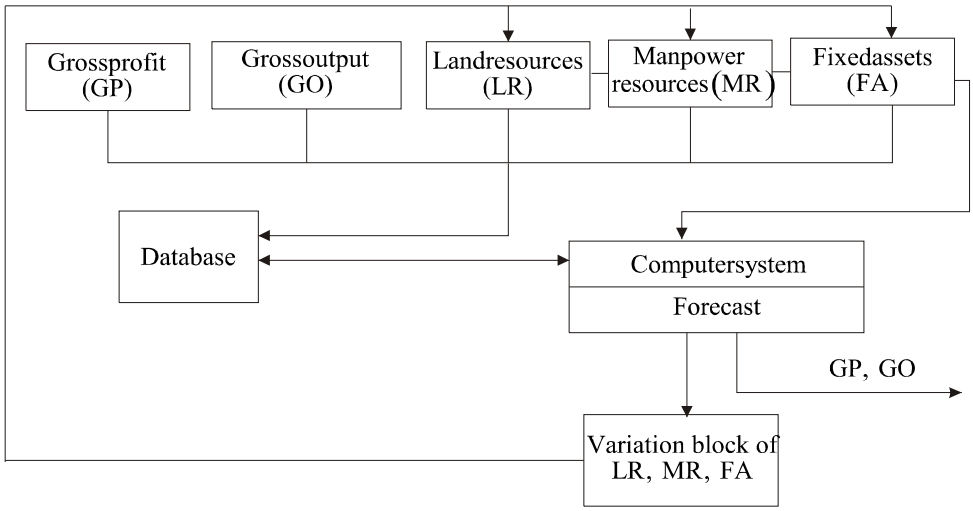
Figure 1. Block diagram of the functioning of the algorithm (6).



New known results of enterprises functioning for 2013 allows to specify the characteristics of the algorithm (6) and the extrapolator can be used for the enterprise management at the level of the parameters $x_3(12)$ - land resources in 2013, $x_4(12)$ - labour resources in 2013, $x_5(12)$ - fixed assets in 2013 for the achievement of the required effect for 2014.

The diagram of the computer system functioning on the basis of the developed technology of management is presented in Figure 2.

Figure 2. Computer system for the forecast and management of an agricultural enterprise.



CONCLUSIONS

The optimal algorithm of the extrapolation of the economic indices of agricultural enterprises which as well as canonical expansion put into its base doesn't impose any essential limitations on the stochastic properties of economic indices is obtained.

The model of the forecast allows to estimate the results of enterprise functioning after its reorganization (the change of land resources, manpower, fixed assets).

The offered method of management may be also realized for nonagricultural enterprises with other set of economic indices.

REFERENCES

Altman E.I., Marco G., Varetto F. (1994) Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks, Journal of Banking and Finance, Vol. 18 – pp. 505-529.

Altman E.I., Narayanan P. (1997) An international survey of business failure classification models, Financial Markets, Institutions and Instruments, Vol. 6, №2 – pp. 81-130.

Atamanyuk I.P. (2005) Algorithm of extrapolation of a nonlinear random process on the basis of its canonical decomposition, Cybernetics and Systems Analysis, №2 – pp. 131-138.

Atamanyuk I.P. (2009) Optimal Polynomial Extrapolation of Realization of a Random Process with a Filtration of Measurement Errors, Journal of Automation and Information Sciences, Volume 41, Issue 8 – pp. 38-48.

Atamanyuk I.P., Kondratenko V.Y., Kozlov O.V., Kondratenko Y.P. (2012) The algorithm of optimal polynomial extrapolation of random processes, Lecture Notes in Business Information Processing, 115 LNBIP – pp. 78-87.

Granger C.W.J., Newbold P. (1986) Forecasting economic time series. Academic Press – 114 p.

Hall S.G. (1994) Applied economic forecasting techniques. Harvester Wheatsheaf – 224 p.

Kudritskiy V.D. (2001) Filtration, extrapolation and identification of the realizations of random functions. – Kyev: FADA ltd. – 176 p.

Połoński M. (2012) Prognozowanie czasu zakończenia inwestycji na podstawie jej bieżącego zaawansowania, Metody ilościowe w badaniach ekonomicznych, Tom XIII/3 – s. 169-179.

Prędki A. (2013) Subsampling approach for statistical inference within stochastic DEA models, Metody ilościowe w badaniach ekonomicznych, Vol. XIV, No. 2 – s. 158-168.

Pugachev V. (1962) Theory of random functions and its implementation. Moscow: Physmathgis – 720 p.

Ryabushkin B.T. (1987) Application of statistical methods in an economic analysis and prognostication: Prakt. guidance. M.: Finances and statistics – 175 p.

Sirenko N.N. (2010) Management of strategy innovative development of agrarian sector of economy of Ukraine. – Mykolaiv – 416 p.

Szmuksta-Zawadzka M., Zawadzki J. (2013) Modele harmoniczne ze złożoną sezonowością w prognozowaniu szeregów czasowych z lukami systematycznymi, Metody ilościowe w badaniach ekonomicznych, Tom XIV/3 – s. 81-90.

Teyl G. (1971) Economy prognosis and making decision. M.: Statistics – 488 p.

Trifonov Yu.V., Plehanova A.F., Yurlov F.F. (1998) Choice of effective decisions in an economy in the conditions of vagueness. Nizhniy Novgorod: Publishing house NNGU – 140 p.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 17 – 26

HERD BEHAVIOR IN INTERNATIONAL MARKET

Milena Balcerzak
Department of Econometrics and Statistics
Warsaw University
e-mail: Milena.balcerzak@gmail.com, mbalcerzak@wne.uw.edu.pl

Abstract: According to literature, herding can be observed not only within single stock market, but also on international level. Poland has entrance to Euro Zone in perspective, so especially interesting is answer to the question whether herding is present on the stock markets of candidate countries to Euro Zone? Moreover, whether this tendency will be more noticeable in periods of higher volatility than during the periods of relative market stability, as suggested by psychological theories.

To find the measure of herd tendency we can refer both to neoclassical and behavioral theories. Basing on models grounded in rationality assumption, there is a need to assume informative efficiency of markets. Thus, it can be expected that aggregated changes in stock markets should be reflected in the process of prices formation. Behavioral literature pointed that if tendency of herding is present on international level, it can be assumed that markets 'in average' will behave in the similar way. Following this idea the size of deviation of the given series from the 'average' can be taken as the measure of herding activity in aggregated market. Methodology created by Christie, Huang and Chang, Cheng, Khoran will be used as a starting point for the consideration of this problem. To propose more perfect measure of herding, in the paper this approach will be developed by using alternative methodology for average and measure of deviations construction. Those improvements will cause that, theoretical models will come closer to reality and will take into account both sociological and behavioral tendencies present in investors activity.

Keywords: herding, international stock market, measure of herd behavior

INTRODUCTION

Herding is defined as activities of decision-makers that are caused by interpersonal interaction and rely on making decisions basing on observation of others participants of the given community and imitating their behavior [Szyszko 2009]. In financial markets such activity is present if in a group of investors some of them proceed as others and invest or not in the given asset in the given period of time (in case of financial markets this definition is rarely related to investments in a single and strictly defined time point) [Sias 2002].

It was shown that behavior of subjects is determined by individuals who from their point of view have relatively bigger stores of knowledge, which they personally do not own. The tendency of behavior imitation will be the larger, the smaller is precision and amount of available information. It is expected that herding does not appear if credibility of private information is bigger than quality of information coming from activity of others market players. The tendency of herding will be the bigger, the more difficult is to gain and process the information.

Literature shows, that herding is present in the international market. In research performed for Germany, Great Britain, United States of America, Mexico, Japan, Spain and France, Blasco and Ferreruela [2008] identified this tendency just in Spain market. Chang, Cheng and Khoran's [2000] study indicated on presence of meaningfully significant herd tendency in South Korea and Taiwan markets in period 1963-1997. Hwang and Salmon [2004] used daily stock returns and also found some circumstances that herding was present in South Korea market in period 1993-2002. In 2004, Demirer and Kutan [2004] applied CSSD (Cross-Sectional Standard Deviation) methodology to Chinese individual and sector markets and did not find any reason to say that behavior imitation tendency appears there.

Herding can be observed not only in stock market. Presence of this propensity was tested also in government stock market [Gleason et al., 2004], in future market for fuel oil and petroleum [Weiner et al., 2004], where symptoms of collective activity were identified.

Although the literature contains a number of examples of herd behavior presence tests on many different markets and for plenty of varying assets, there is noticeable shortage of papers, which investigate this phenomenon in global scale. The example of such work is Demirer, Gubo and Kutan's [2007] paper, where such research was performed on many stock markets grouped into six regions: developed markets (West Europe and USA), Asian, Central and East Europe, Latin America, Mid-East and Africa.

Literature suggests, that imitation of behavior is present not only inside of single stock market, but also this phenomenon can be observed on international level. Thus, it can be expected that the Euro Zone members (having relatively broader knowledge) determine the behaviour of candidate countries to this

international community, especially in periods of greater market volatility. Confirmation of such dependence will introduce new implications in the area of stock market behaviour modelling as allows to assume that in the markets of members of this international community important leading factors for candidate countries to Euro Zone exist and grants to make further use of theory of leading factors in analyses on international level.

Attempting to point herd tendency measure, we can refer both to neoclassical and behavioral theories. It can be expected that aggregated changes in the market will be reflected in the process of prices development. On the basis of behavioral literature, in case of herding presence in international market, it can be assumed that markets in average should behave in a similar way. Following this intuition, as a measure of herding activity in the aggregated market, the size of deviation of the given time series from 'average' can be considered. Such methodology was proposed by Christie, Huang [1995] and Chang, Cheng, Khoran [2000]. To create closer to reality measure of herd behavior, alternative approach in mean construction and deviation measure was used to develop those theories. Thanks to introduced extensions, statistics reflect behavioral tendencies in investors activity in the stock markets.

HERD BEHAVIOR MEASURE

Referring to behavioral literature, if herding is present in international market, it can be assumed that stock market indexes, which are considered to be representant of capital markets and are regarded to reflect aggregated changes in prices development, in average should behave in the similar way. This intuition suggests, that the value of deviation of the given return series from 'average' may be considered as herd measure in aggregated market.

Related literature refers to two measures of herd behavior. First was proposed by Christie and Huang [1995]. In it construction standard deviation is used – it's called Cross-Sectional Standard Deviation (CSSD). The second approach was introduced by Chan, Cheng and Khoran [2000]. This indicator based on absolute deviation from the mean – it's so-called Cross-Sectional Absolute Deviation (CSAD). Construction of CSSD and CSAD assumes that relation between average cross sectional and actual return of asset in interest is symmetric.

Researches depict that investor reactions on decreases and increases of rates in the markets are not the same. Behavioral economy proves that people in suspense have tendency to too quick jumps to conclusions, make decisions just basing on single facts. On the other hand, sometimes investor's reactions are delayed in respect to signals that came to them or downright ignored. On the basis of psychological tendencies, it seems to be reasonable to include in herding measure construction both propensity to overreactions and for ignoring information/ waiting for signals confirmation. Such phenomenon's take place on the level of activity of individual investors, but it can be expected that to some

extent, they will be reflected also in aggregated market. Owing to those facts asymmetric measures can be proposed: Mean Mixed Errors Undervaluation (MMEU) and Mean Mixed Errors Overvaluation (MMEO):

$$MMEU_t = \frac{1}{N}\left[\sum_{i=1}^{N}|R_{ti} - R_t|K_{ti} + \sum_{i=1}^{N}\sqrt{|R_{ti} - R_t|L_{ti}}\right] \tag{1.1}$$

$$MMEO_t = \frac{1}{N}\left[\sum_{i=1}^{N}\sqrt{|R_{ti} - R_t|K_{ti}} + \sum_{i=1}^{N}|R_{ti} - R_t|L_{ti}\right] \tag{1.2}$$

where:

$$K_{ti} = \begin{cases} 1 & if \quad R_{ti} > R_t \\ 0 & if \quad R_{ti} \le R_t \end{cases} \qquad L_{ti} = \begin{cases} 1 & if \quad R_{ti} \le R_t \\ 0 & if \quad R_{ti} > R_t \end{cases}$$

R_{ti} - daily return from market index for country i in period t,

R_t - daily average cross sectional return from market indexes for N countries in period t / indicator of market behavior.

The idea of those measures based on emphasizing (by using root square function) value of deviation of the given return series from aggregated market index in direction which is 'assumed' by measure.

MODELS DEFINITION

It is assumed that in normal conditions investors act in rational way, taking into account all available information when making decisions. Extreme conditions cause extreme emotions and reality shows that investors 'feel more comfortable' acting as other participants of market. This suggests, that it can be expected that herding should be clearly noticeable especially in periods when bigger uncertainty is in the market. On the basis of this presumption, 'normal' and 'extreme' periods can be distinguish in the market behavior. Periods with returns in first (extremely low) and in fourth (extremely high) quartile of distribution are considered as 'extreme'. Moreover, it can be expected that investors will modify the way of their behavior depending on the phase in which the market is.

To verify this hypothesis, model with zero-one variables marking market phases (D^U =1 if the return in day t is in the fourth quartile of distribution, 0 otherwise; D^L =1 if the return in day t is in the first quartile of distribution of market returns, 0 otherwise) can be proposed:

$$\mathrm{MMEU_t}\left[\mathrm{MMEO_t}\right] = \beta_0 + \beta_1^L D_{t1}^L + \beta_2^U D_{t2}^U + \varepsilon_t \tag{2}$$

This linear model allows to investigate the direction of changes in the market. According to classical theories of capital market, developed on rationality fundaments, estimated coefficients should be significantly positive as assets differ in the level of sensitivity. It is expected that herding will cause decrease of dispersion of returns around the mean. Thus, statistically significant and negative

value of at least one of estimated coefficients will indicate on presence of this tendency in the market.

Empirical researches show [Prosad et al., 2012] that relation between measure of herd behavior and market indicator may not be linear. Thus, following alternative nonlinear regression model can be proposed:

$$\mathrm{MMEU}_t[\mathrm{MMEO}_t] = \beta_0 + \beta_1|R_t| + \beta_2 R_t^2 + \varepsilon_t \tag{3}$$

where symbols description as in (1).

If collective behavior is not present in the stock market, relation between MMEU [MMEO] and R_t will be linear and positive. Statistically significant and negative coefficient β_2 implies presence of this tendency in the market.

According to behavioral literature, relation between dispersion measure and market returns may be asymmetric, so two separate models for extreme reactions can be proposed:

$$\mathrm{MMEU}_t^U\left[\mathrm{MMEO}_t^U\right] = \beta_0^U + \beta_1^U\left|R_t^U\right| + \beta_2^U\left(R_t^U\right)^2 + \varepsilon_t \quad \textit{if} \quad R_t > 0 \tag{4.1}$$

$$\mathrm{MMEU}_t^D\left[\mathrm{MMEO}_t^D\right] = \beta_0^D + \beta_1^D\left|R_t^D\right| + \beta_2^D\left(R_t^D\right)^2 + \varepsilon_t \quad \textit{if} \quad R_t < 0 \tag{4.2}$$

where symbols description as in (1).

EMPIRICAL ANALYSIS

Data and analysis period

The research was performed for period from 1st July 2006 to 1st July 2008[1]. July 2007[2] is consider to be the start point of global economic crisis. For analysis purposes two sub periods were determined: period before the crisis start: 1st July 2006 - 1st July 2007 and period after the crisis start: 2nd July 2007 - 1st July 2008. This allows to compare behavior of market in different economic conditions.

Herd tendencies measures presented in literature assume the equal realization each of the series in creation of the average R_t. If the research is performed within single market there are no objections to such approach. However, if the study concerns different countries, using weighted average market capitalization seems to be more appropriate way. Such approach find also methodological justification, as size of the market has very important impact on stock market behavior.

1 This article is part of larger study and that was main determinant of choosing such period of analysis.

2 Appointment of crisis start date is subjective task. In this article decision was made to take more or less a moment of bankruptcy of two hedging funds of Bear Stearns bank – the time when alarming information started appearing in the market.

Daily closing index prices coming from web-sites[3]: http://analytics.tradingeconomics.com, http://www.nasdaqomxbaltic.com/market, http://www.borzamalta.com.mt (for Lithuania and Malta respectively) were used in analysis and for calculation weighted average market capitalization index (R_t) for countries belonging to Euro Zone.

Information about markets capitalization was taken from web-site: http://databank.worldbank.org (annual data were converted into daily data using PROC EXPAND procedure (with options: OBSERVED=total and METHOD=join) available in SAS 9.1®).

Results of investigation

To verify presence of herd behavior in the aggregated market of candidate countries to Euro Zone, models (2), (3) and (4) were estimated. The research was performed both for S&P 500 (as American market still has great impact on world economy) and for weighted average index return of Euro Zone members as a markers of market phases.

Table 1. Regression coefficients for: $\mathrm{MMEU_t}\ [\mathrm{MMEO_t}] = \beta_0 + \beta_1^L D_{t1}^L + \beta_2^U D_{t2}^U + \varepsilon_t$

	1st July 2006 - 1st July 2007		2nd July 2007 - 1st July 2008	
Parameter	Estimates (p-value) #	Estimates (p-value) ##	Estimates (p-value) #	Estimates (p-value) ##
MMEO				
β_0	0.0526 (<.0001)*	0.0408 (<.0001)*	0.0609 (<.0001)*	0.0512 (<.0001)*
β_1^L	-0.0003 (0.0006)*	-0.0117 (0.9288)	-0.0023 (0.5047)	-0.0180 (0.0002)*
β_2^U	-0.0042 (0.0009)*	0.0222 (<.0001)*	0.0049 (0.0091)*	0.0440 (<.0001)*
MMEU				
β_0	0.0421 (<.0001)*	0.0432 (<.0001)*	0.0418 (<.0001)*	0.0474 (<.0001)*
β_1^L	0.0105 (0.0006)*	0.0322 (<.0001)*	0.0066 (0.0390)*	0.0423 (<.0001)*
β_2^U	-0.0002 (0.0064)*	-0.0099 (0.9004)	-0.0011 (0.0027)*	-0.0128 (0.5043)

for S&P 500 as marker of market phases; ## for weighted average index return of Euro Zone members as marker of market phases
* denotes significance at 5%

Source: Author's calculations, performed in SAS 9.1®

Estimated coefficients β_2^U of MMEU regression and β_1^L of MMEO regression are negative in each case. Also, for pre crisis time β_2^U is below zero in the model where S&P 500 as marker of market phases and MMEO as dependent

3 In the research were included all countries belonging to Euro Zone in the analyzed period and the members of European Union that candidate to this international community.

variable are used. Results of analysis do not allow to reject hypothesis about herding presence in the market.

Literature shows that all economies are very sensitive on the situation in USA. It is worth to notice that when American index as marker of market phases is used, estimator's values are closer to zero. Such results suggest that greater impact on behavior of countries that would like to enter to Euro Zone has behavior of members of Euro Zone than S&P 500 index, which is traditionally used for determination of tendencies in the world markets.

MMEO construction highlights issues of overestimation variability. Too strength reaction are characteristic for stressful periods: with higher variability, with very low returns and decreases in the market.

Contrastively MMEU construction highlights issues of underestimation variability. Such tendencies in investors behavior are characteristic especially for 'calm' periods and characterized by increasing trend.

Table 2. Regression coefficients for: $\mathrm{MMEU}_t[\mathrm{MMEO}_t] = \beta_0 + \beta_1|R_t| + \beta_2 R_t^2 + \varepsilon_t$

	1st July 2006 - 1st July 2007		2nd July 2007 - 1st July 2008	
Parameter	Estimates (p-value) #	Estimates (p-value) ##	Estimates (p-value) #	Estimates (p-value) ##
MMEO				
β_0	0.0343 (<.0001)*	0.0345 (<.0001)*	0.0525 (<.0001)*	0.0471 (<.0001)*
β_1	0.2397 (0.0414)*	0.2271 (0.0109)*	-0.0887 (0.4348)	-0.1815 (0.0125)*
β_2	13.0570 (0.0004)*	7.2686 (0.0088)*	5.6789 (0.0561)*	6.1352 (0.0013)*
MMEU				
β_0	0.0427 (<.0001)*	0.0396 (<.0001)*	0.0467 (<.0001)*	0.0412 (<.0001)*
β_1	-0.1149 (0.3236)	-0.2690 (0.0025)*	-0.1603 (0.1770)	-0.0853 (0.2527)
β_2	6.3826 (0.0784)*	13.5097 (<.0001)*	17.1034 (<.0001)*	10.7802 (<.0001)*

where symbols description as in Table 1

Source: as in Table 1

Results of $\mathrm{MMEU}_t[\mathrm{MMEO}_t] = \beta_0 + \beta_1|R_t| + \beta_2 R_t^2 + \varepsilon_t$ estimation suggest that herding is not present in the market as in each model β_2 coefficient is greater than zero. Also, it can be observed that R_t changes have more than proportional impact on the herding behavior measure.

Table 3. Regression coefficients for: $\mathrm{MMEU}_t^x \left[\mathrm{MMEO}_t^x\right] = \beta_0^x + \beta_1^x \left|R_t^x\right| + \beta_2^x \left(R_t^x\right)^2 + \varepsilon_t$, $x \in \{U, D\}$

	1st July 2006 - 1st July 2007		2nd July 2007 - 1st July 2008	
Parameter	Estimates (p-value) #	Estimates (p-value) ##	Estimates (p-value) #	Estimates (p-value) ##
MMEO where $R_t > 0$				
β_0	0.0352 (<.0001)*	0.0358 (<.0001)*	0.0450 (<.0001)*	0.0445 (<.0001)*
β_1	0.6468 (<.0001)*	0.4915 (<.0001)*	0.5346 (0.0006)*	0.2676 (0.0082)*
β_2	5.4493 (0.0799)	10.1682 (0.0152)*	1.4433 (0.5745)	1.5037 (0.7250)
MMEU where $R_t > 0$				
β_0	0.0411 (<.0001)*	0.0385 (<.0001)*	0.0532 (<.0001)*	0.0430 (<.0001)*
β_1	-0.6405 (<.0001)*	-0.6057 (<.0001)*	-0.7233 (<.0001)*	-0.5280 (<.0001)*
β_2	15.1782 (<.0001)*	18.2479 (<.0001)*	18.7647 (<.0001)*	18.0368 (<.0001)*
MMEO where $R_t < 0$				
β_0	0.0353 (<.0001)*	0.0344 (<.0001)*	0.0543 (<.0001)*	0.0461 (<.0001)*
β_1	-0.8458 (<.0001)*	-0.4980 (<.0001)*	-0.4979 (0.0008)*	-0.4708 (<.0001)*
β_2	-16.5644 (<.0001)*	-27.0667 (<.0001)*	-11.2251 (0.0021)*	-11.7262 (<.0001)*
MMEU where $R_t < 0$				
β_0	0.0416 (<.0001)*	0.0391 (<.0001)*	0.0470 (<.0001)*	0.0412 (<.0001)*
β_1	1.1762 (<.0001)*	0.5823 (<.0001)*	0.1664 (0.3292)	0.2392 (0.0238)*
β_2	-0.1154 (0.9776)	-15.9737 (0.0024)*	-3.7381 (0.1662)	-13.4507 (0.0023)*

where symbols description as in Table 1

Source: as in Table 1

Different conclusions provide analysis performed separately for decreasing and increasing periods in the market than those flowing from the results of estimation $\mathrm{MMEU}_t \left[\mathrm{MMEO}_t\right] = \beta_0 + \beta_1 |R_t| + \beta_2 R_t^2 + \varepsilon_t$.

Results for bull market allow to reject hypothesis about existence of herding in the market of candidate countries to Euro Zone. In periods characterized by $R_t < 0$ models for both MMEO and MMEU and for both markers of market phases indicate on herding presence in the analyzed market. This draws conclusions that decreasing trend itself invokes anxiety and increases stress among investors and thus increases probability that herd behavior appears in the stock market. Furthermore, received differences in the results for periods characterized by $R_t < 0$ and $R_t > 0$ confirm theory of behavioral finance about different reaction of

investors for increasing and decreasing trends in the market and show that this tendency is also noticeable on the aggregated level.

As in the first model (2), in most cases values of received estimators indicate relatively stronger (as to absolute values) reaction of herd behavior measures on the changes of weighted average of indexes returns of Euro Zone members than on changes of American index returns.

To complete empirical analysis, short summary of models diagnostic need to be added. All presented models are statistically significant (Pr > F: <.0001). Adjusted R square belongs to interval: [0.0200, 0.4779] for models estimated for pre crisis period and to: [0.0259, 0.3864] for crisis period. For both periods and both market indicators the worst fitted linear models are. Slightly better is model (3). Significantly the best fitted are nonlinear models estimated separately for periods characterized by $R_t < 0$ and $R_t > 0$.

SUMMARY AND CONCLUSIONS

In the paper attemption to create better than known until now measure of herd behavior on the basic of methodology of deviations from mean was made. Theory proposed by Christie, Huang and Chang, Cheng, Khoran was developed for alternative approach of average construction and deviation measure. Thanks to that, it includes predispositions of investors behavior and becomes more close to reality. Referring to literature, it seems that introduction of mentioned elements do not affect significantly the obtained results. One of the reasons may be dilution (by averaging) of the effect on aggregated level. For further methodology verification it is recommended to use data with higher frequency.

Moreover, in the paper the problem of herding presence on the markets of countries that would like to enter to Euro Zone in the period before economic crisis (1st July 2006 - 1st July 2007) and after it beginning (2nd July 2007 - 1st July 2008) was analyzed. Received results do not allow to answer unambiguously for in the introduction questions raised, however seem to be consistent with those described in literature. Even then the models do not indicate on presence of herd tendency among countries that would like to enter to Euro Zone unquestionably, but also do not allow to reject this hypothesis. Also, it should be emphasized that any market can not be fully free from herding effect – it with smaller or greater strength will be reflected in aggregated market.

Obtained results suggest that dependence between the behavior of countries that would like to enter Euro Zone and members of this international community exists. If in the aggregated market imitation tendencies are present, received results indicate that candidates to Euro Zone will copy the way of activity not only in periods of higher uncertainty. Also, it should be emphasized that it is more probable that countries that would like to enter to Euro Zone will be more

sensitive to changes appearing in the markets of members of Euro Zone than to those indicated by American index, which traditionally according to literature is used as marker of market behavior. This outcome supports availability heuristic defined in psychology. Also, this draws very important conclusion for modeling behavior of candidate countries to Euro Zone markets. It can be expected that important leading factors for countries that would like to enter this international structure can be found in markets of members of Euro Zone. Furthermore, it provides a clue for governments about directions in which information systems should be developed to provide appropriate level of information for investors.

REFERENCES

Blasco N., Ferreruela S. (2008) Testing Intentional Herding in Familiar Stocks: An Experiment in an International Context, The Journal of Behavioral Finance.

Boyer B. H., Kumagai T., Yuan K. (2006) How Do Crises Spread? Evidence from Accessible and Inaccessible Stock Indices, The Journal of Finance, Vol. 61(2).

Brown N. C., Wei K. D., Wermers R. (2012) Analyst Recommendations, Mutual Fund Herding, and Overreaction in Stock Prices, Working paper, University of Maryland.

Chang E. C., Cheng J. W., Khoran, A. (2000) An examination of herd behavior in equity markets: An international perspective, Journal of Banking and Finance, Vol. 24.

Christie W. G., Huang R. D. (1995) Following the pied piper: Do individual returns herd around the market?, Financial Analyst Journal, Vol. 51 (4).

Chiang T. C., Zheng D. (2010) An empirical analysis of herd behavior in global stock markets, Journal of Banking and Finance, Vol. 34(8).

Demirer R., Gubo G, Kutan A. M., (2007) An Analysis of Cross-Country Herd Behavior in Stock Markets: A Regional Perspective, Working Paper.

Demirer R., Kutan A. M. (2004) Does herding behavior exist in Chinese stock market?, Journal of International Financial Markets, Institutions and Money.

Demirer R., Kutan A. M., Chen C. (2010) Do investors herd in emerging stock markets?: Evidence from the Taiwanese market, Journal of Economic Behavior and Organization, Vol. 76 (2).

Gleason K. C., Mathur I., Peterson M. A. (2004) Analysis of intraday herding behavior among the sector ETFs, Journal of Empirical Finance, Vol. 11.

Hwang S., Salmon M. (2004) Market stress and herding, Journal of Empirical Finance, Vol. 11.

Lakonishok, J., Shleiferand, Vishny (1992) The Impact of Institutional Trading on Stock Prices, Journal of Financial Economics.

Sias R. (2002) Institutional Herding, Working Paper.

Szyszko A. (2009) Finanse Behawioralne. Nowe podejście do inwestowania na rynku kapitałowym.

Weiner R. J., Green M. A. (2004) Do birds of a feather flock together? Speculator herding in derivatives markets, Working Paper George Washington University.

Wermers R. (1995) Herding, Trade Reversals, and Cascading by Institutional Investors, University of Colorado, Boulder.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 27 – 36

DEMAND FORECAST WITH BUSINESS CLIMATE INDEX FOR A STEEL AND IRON INDUSTRY REPRESENTATIVE

Magdalena Barska
Collegium of Economic Analysis
Warsaw School of Economics
e-mail: magda.barska@gmail.com

Abstract: The steel and iron industry production is dedicated to serve other industries mainly. This makes the exercise of demand forecasting different than for consumer goods. The common sense says that demand fluctuations are influenced by general economic soundness. An attempt was made to address the question of improving forecast's accuracy by adding a business cycle indicator as an input variable. The SARIMAX model was applied. Including a business climate indicator improved model's performance, however no co-integration is observed between the two series.

Keywords: demand forecasting, SARIMAX, business climate indicator

INTRODUCTION

Demand planning is a very important matter in companies' operations. It is an integral part of goods manufacturing and distribution process. Both excessive production that cannot be sold and demand exceeding production levels – despite having adequate production capabilities – are disadvantageous. Estimation of future demand for goods and services is also crucial when determining resources that are used in the production process.

Demand forecasting requires specific tools. Intuitive predictions may turn out to be insufficient due to the number and complexity of influence factors. A belief that the intuitive approach is more advantageous happens to win in some cases as it is based on experience and familiarity with the industry. Not always is this approach advisable as companies' external environment may be unstable.

A lot of attention was paid to demand forecasting for the fast moving consumer goods and energy industries where the data of a high frequency are analysed. In this article the focus is on iron and steel products. The industry is

characterized by the following features: long production process, long planning horizon, high value and weight of produced goods. This makes the exercise of demand forecasting different than for consumer goods. Iron and steel products are dedicated to serve other industries rather than individual agents. The common sense says that demand fluctuations are related to general economic soundness. Moreover, the price and quality of goods are believed to impact demand to a greater extent than individual preferences or fashion do. Political and environmental factors are important as companies need to respect restrictions on technologies in use, minimising the harmful impact on the environment. This results in low elasticity in adjusting to demand fluctuations in a short-term as well as in additional costs of technology improvement. Demand for the heavy industry was analysed among others by [Rippe et al. 1976]. [Lallement and Briffaut 2010] analyse demand volatility in the construction industry using the French construction cost index. To analyse demand in the automotive industry, [Klug 2011] employs the Monte Carlo simulation that controls for uncertainty when measuring random demand levels.

The goal of this article is to verify if a demand forecast for the iron and steel industry in Poland can be improved by adding a business climate index. Empirical analysis was carried out for a representative agent. The company is one of the key players in this market. In Section "FACTORS SHAPING DEMAND …" the factors influencing demand levels in the iron and steel industry as well as sector's characteristics are discussed. Section "RESEARCH METHODS" describes the research approach. In Section "EMPIRICAL ANALYSIS" the SARIMAX model with a business climate indicator added as an input variable is applied to the sample data. Results are compared with SARIMA. Section "CONCLUDING REMARKS" summarizes the main findings.

FACTORS SHAPING DEMAND IN THE STEEL AND IRON INDUSTRY IN POLAND

The environment in which enterprises operate is volatile. Demand levels are impacted by many factors. Economic transformation in Poland posed a challenge for state-owned companies. In the 1990´s they had to undergo a long process of restructuring to adjust to the new market conditions. Nowadays they need to obey some political and environmental restrictions on technologies in use or accept constraints imposed on production levels to reduce emissions of greenhouse gases. Infrastructural investments and the EU policies promoting renewable energy are both an opportunity for the sector. The latter triggers an interest in wind and nuclear energy or ship transport while the industry is a supplier of windmill and ship engines components. At the same time the sector is affected by constantly increasing energy costs. Entrepreneurs face strong competition from Asia and

Europe. Some unfair actions as dumping or tax settlement fraud on steel sales are observed [Forbes 2013].

RESEARCH METHODS

General Considerations

[Cieślak et al. 1997] present an overview of forecasting methods and its applications in the economic context. There are several statistical techniques that can be used to build a forecast based on macroeconomic variables. Traditional methods are criticised for basing solely on past observations. It is justified providing that the external environment is stable and customers' behaviour is not a subject to significant changes. It is rarely the case though. It would be advisable to include some external factors that shape demand. One difficulty here would be that they may not be easy to identify or their value remains unknown. To address that [Crane and Crotty 1967] use exponential smoothing. The authors analyse past trends in banking data. They measure the demand for financial products using outputs of exponential smoothing in a multiple regression. They take benefit from both approaches, building a prediction based on historical data and other explanatory variables. Alternatively, ARIMA described by [Box et al. 1994] can be a foundation for the ARIMAX approach. ARIMA specifies a relation between a variable in the period *t* and its values in previous periods. In ARIMAX, another time series is added as an input variable. Hence, it allows for including a business cycle indicator that provides with early information on agents´ future activity. [Ďurka and Pastoreková 2012] compare results of ARIMA and ARIMAX to model the GDP. In this case ARIMA turns out to be more accurate than ARIMAX where an unemployment rate is added as an input variable. [Bielak 2010] compares ARIMA and ARIMAX when forecasting unemployment. Adding the economic mood index improves model´s accuracy in terms of errors and information criteria. The improvement is slight but applying the index can be justified by its property to signal changes in advance as well as by its strong correlation with fluctuations observed in the market.

Business Climate Indicators

Macroeconomic indicators show trends in social and economic development of a country. In our model a leading or a business climate index would be required as they measure future activity. For price or production indices this assumption is not fulfilled as they reflect past values of a variable (i.e. how much was sold, at what price). Moreover, prices may adjust in the long term; therefore some trends will be presented with a delay. Leading indices show changes in economic performance a few months ahead of the GDP. Using them in a model may increase

the forecasting horizon as well as help obtain better parameters [Szeplewicz 2011]. A variable reflecting general economic conditions serves the following purposes:

- Controls for the variability of factors influencing demand;
- Controls for performance of various sectors that purchase steel goods;
- Increases model's reliability and extends the forecasting horizon;
- Allows for early identification of fluctuations owing to its leading properties.

A relevant indicator was sought for among those that measure the activity of various industries. Both domestic and international databases were considered. Given the nature of demand in the steel and iron industry, business climate index for the processing industry published by the Central Statistical Office (GUS - Główny Urząd Statystyczny) was selected for further research. It refers to the Polish market where the major part of transactions for the sample data is conducted. The indicator comprises of respondents´ current assessment on: order backlog, financial soundness, expected delays in settling payments by contractors, selling prices and employment. Such tests are successfully used to diagnose economic conditions in a short term and are distributed as questionnaires [GUS 2013]. They should be considered as CEO´s subjective opinions. They enrich the knowledge about markets. Subjectivity can be regarded as a weakness but we assume that predictions of entrepreneurs will be translated into action, i.e. they will be more prone to invest when economic conditions are favourable. Selecting the index for the processing industry is explained by the fact that it measures performance of industries producing metals, metal products, machinery and equipment while the industry´s production is addressed to entities operating in the iron and steel or metallurgical industry. On the top of that, the indicator is easily accessible and published with a monthly delay. It replaces subjective opinions of experts in a company. That covers the question of a trade-off between statistical methods and intuition when determining future demand levels. We expect that experts´ assessment is based on monitoring such indicators or tracking information in trade magazines or through participation in trade fairs. Business climate indicators measure the condition of a number of related industries and may provide additional information that an expert does not possess. There are some interesting findings from [Franses and Legerstee 2013] who consider the inclusion of a variable reflecting opinions of people involved in sales and marketing activities. They assume that the final forecast is a weighted average of an expert´s and model´s forecasts. Adding the variable did not improve the forecast significantly. Its accuracy increased only in the event of an expert having very specific knowledge.

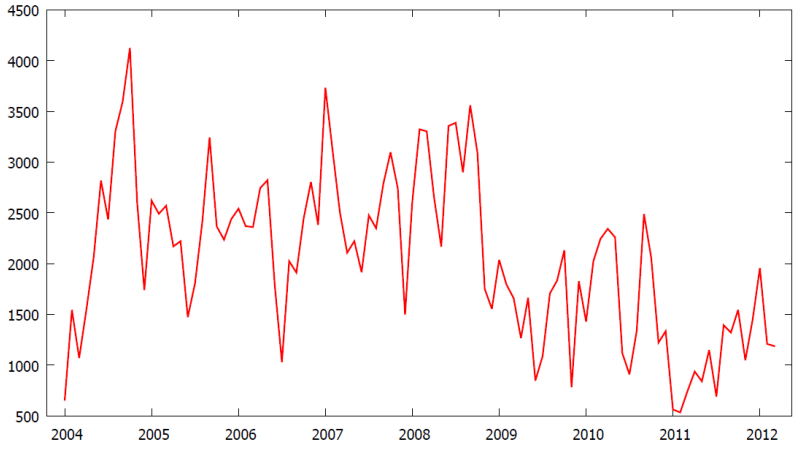
In this sense adding the business climate indicator is a way to include an average opinion of various agents acting in the market instead of CEO's predictions solely.

EMPIRICAL ANALYSIS

Data Description

Demand level data of a Polish enterprise operating in the steel and iron industry constitute a monthly time series for the period 2004-2012. The company is among the leading and mid-range 50 players of the industry. The chart presents significant demand fluctuations (Figure 1). A decrease can be observed in the average demand level starting from 2009. The production is purchased by several industries like metallurgy and machinery and targeted at the Polish market mainly.

Figure 1. Monthly steel products sales of a sample enterprise in 2004 – 2012 (in tons)



Source: author's calculations

The seasonality test was carried out in Demetra[1]. The general evaluation of the model is considered *severe*, which means that results of the seasonality check cannot be trusted. In such instances it is advised to verify the series case by case and modify parameters in the specifications window of Demetra to improve the quality of the adjustment if needed [Grudkowska 2013]. There is a significant drop in the demand level if we compare the period 2009 - 2012 with 2004 - 2008. Therefore the earlier observations may not be relevant. But a time series needs to contain data for at least 5 years so that the procedures used in Demetra can produce a reliable outcome. Eventually an analysis is carried out for the period 1/2007 – 12/2012, which fulfils this requirement. The model turns out to be of a good quality and the analysis carried out in Demetra gives an evidence of seasonality. This

[1] Demetra - a software for seasonal adjustments.

conclusion is based on the results of the F and Kruskal-Wallis tests for stable seasonality. For both tests the null hypothesis of non-seasonality is rejected with *p* value equal to 0.02 and 0.04 respectively. The null hypothesis of the F test assuming the lack of moving seasonality cannot be rejected with *p* value = 0.1. We conclude on the stable seasonality of the narrowed time series.

The Augmented Dickey-Fuller unit root test indicates non-stationarity of the time series. The null hypothesis is not rejected with *p* value equal to 0.84. The same test carried out for variable's first differences rejects the null hypothesis of non-stationarity, giving an evidence of the first degree integration ($p = 0.003$).

The SARIMAX model

ARIMA (p,d,q) is first built for the period 2007 - 2012. It now has 72 observations, which is sufficient to apply the model. To find the optimal parameters the ACF and PACF are examined. Basing on that and on the results of the unit root test, ARIMA (1,1,1) is chosen. Modification of *p* and *q* does not improve the model, indicating statistical insignificance of both parameters. Residuals correlogram is generated showing autocorrelation at lag 10. We would wish to find a better model. In the next step SARIMA (1,1,1)(0,1,1) is built. *P, D, Q* parameters are chosen based on backward selection. Table 1 presents the specification of the model and information criteria.

Table 1. The SARIMA (1,1,1)(0,1,1) model

	Coefficient	Std. error	z	p-value
phi_1	0.5090	0.1559	3.265	0.0011
theta_1	-0.9232	0.1169	-7.899	0.0000
Theta_1	-0.9999	0.3033	-3.298	0.0010

Log-likelihood	-564.937	Akaike criterion	1137.874
Schwarz criterion	1146.981	Hannan-Quinn	1141.500

Source: author's calculations

The SARIMA forecast is entirely based on past values, which calls for a variable controlling for the future. The model does not provide with explanation of demand volatility, except for seasonal influences. To address that, the SARIMAX approach is introduced. The processing industry business climate indicator is added as an input variable. It is not seasonally adjusted. The impact of seasonal factors was examined in Demetra with the TRAMO/SEATS procedure. The seasonally adjusted time series was produced to be used in the model with overall evaluation of the model defined as good. The Augmented Dickey-Fuller unit root test indicates non-stationarity of the series. The null hypothesis is not

rejected with p value = 0.61. The same test carried out for variable's first differences rejects the null hypothesis of non-stationarity, giving an evidence of the first degree integration (p = 0.03).

The starting point is the model with 12 lags for the independent variable. Backward selection is used to determine the longest lag. The lag order 4, 5 and 6 are significant and the lowest values of information criteria are observed for the lag order 6, hence it is chosen solely. The Doornik-Hansen normality test does not allow for rejecting the null hypothesis that residuals are multivariate normal with p value equal to 0.50. Table 2 shows parameters of the model. SARIMAX turns out to be better when log-likelihood, Akaike and Hannan-Quinn criteria are compared.

Table 2. The SARIMAX (1,1,1)(0,1,1) model with business climate index

	Coefficient	Std. error	z	p-value
phi_1	0.4794	0.1060	4.524	0.0000
pheta_1	-1.0000	0.0751	-13.32	0.0000
Theta_1	-0.9999	0.2944	-3.396	0.0007
index_6	-1.4273	0.6955	-2.052	0.0401

Log-likelihood	-563.3111	Akaike criterion	1136.622
Schwarz criterion	1148.006	Hannan-Quinn	1141.154

Source: author's calculations

Co-integration between the two series was tested using the Engle and Granger approach [Engle and Granger 1987]. The sample series shows no evidence of co-integration with the business climate indicator, therefore we cannot conclude on their long-term relationship. This result is confirmed by the Johansen approach. For the Johansen test we fail to reject the null hypothesis that there is no co-integration vector with p = 0.58 for the eigenvalue test and p = 0.66 for the trace test. We expect such a result as due to a high volatility of market conditions. Factors influencing demand change constantly which does not allow for a long-term relationship to prevail. This could be verified with longer time series, both are relatively short however. There are two reasons for that: gathering information on business climate indices is quite a recent activity and it is not advisable to analyse the company´s sales in the transition period, in the 1990´s, as market conditions were extremely different.

Results Comparison

Table 3 presents a comparison of forecast accuracy measures for SARIMA (1,1,1)(0,1,1) and SARIMAX (1,1,1)(0,1,1) with the business climate

indicator for the last 9 observations excluded[2]. Better *ex post* accuracy measures are observed for SARIMAX. An improvement from SARIMA to SARIMAX is slight but noticeable when comparing mean errors and Theil's U which are lower for the latter.

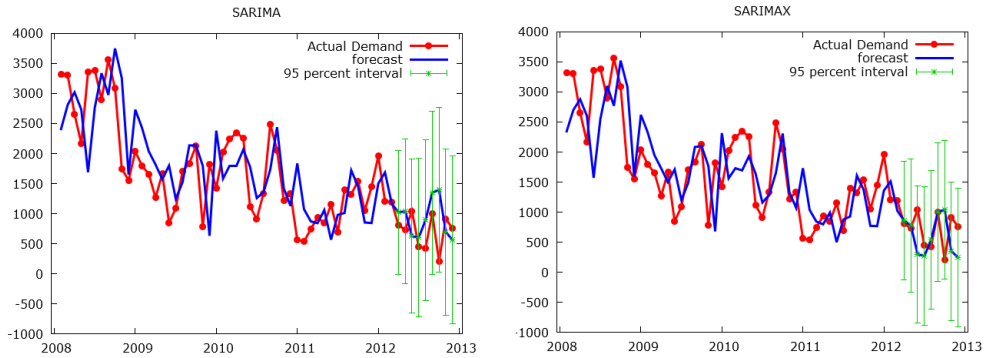
Table 3. *Ex post* forecast accuracy measures

Forecast Evaluation Statistics	**SARIMA**	**SARIMAX**
Mean Error	-204.95t	106.29t
Mean Squared Error	2.41e+005	2.08e+005
Root Mean Squared Error	491t	456.36t
Mean Absolute Error	389.87t	338.14t
Mean Percentage Error	-80.163	-21.52
Mean Absolute Percentage Error	100.13	75.04
Theil's U	0.84	0.70
- Bias Proportion, UM	0.50	0.17
- Regression Proportion, UR	0.36	0.59
- Disturbance Proportion, UD	0.14	0.24

Source: author´s calculations

Figure 2 presents forecasts based on SARIMA and SARIMAX[3]. For ARIMAX the accuracy of an interval forecast is increased. The average confidence interval set at the 95% level shrank by about 13% for SARIMAX (Table 4).

Figure 2. Forecast demand in tons for the period 4/2012 – 12/2012



Source: author´s calculations

[2] This number was chosen to build a forecast for at least 10% of observation and to keep enough cases in the learning set for the relatively short sample time series. On the other hand, the forecast must be for at least 6 months ahead for the company to be useful.

[3] Forecast carried out in a continuous manner, without re-estimating model's parameters.

Table 4. Comparison of SARIMA and SARIMAX average interval forecast

	SARIMA	SARIMAX	% Difference
Average Confidence Interval Width	2598.38	2254.22	-13.25%

Source: author´s calculations

The comparison of mean errors and the confidence interval reduction prove that SARIMAX is a slightly better fit. Choosing SARIMAX serves the purpose of building a relatively simple model, but of a good quality and with good predictive properties.

CONCLUDING REMARKS

The achieved result and verification of SARIMA and SARIMAX quality measures lead to the conclusion that the business climate indicator can be a good predictor of future demand levels. SARIMAX' forecast accuracy slightly improved when compared to SARIMA. The indicator controls for general economic conditions in companies´ external environment therefore not only is the forecast based on past values but a component informing about the future is added. It proves correct to forecast the demand for products dedicated for the metallurgy, engineering and steel industry. As this is a business climate indicator, it may provide with a subjective evaluation, but we assume that entrepreneurs' opinions will be translated into action. As no co-integration with the analysed series is observed, we cannot conclude on the long-term relationship.

REFERENCES

Briffaut, J.P., Lallement, P. (2010) Volatility Forecasting of Market Demand as Aids for Planning Manufacturing Activities. Service Science & Management, 3, pp. 383-389.

Bielak, J. (2010) Prognozowanie ryku pracy woj. Lubelskiego z wykorzystaniem modeli ARIMA i ARIMAX. Barometr Regionalny, nr 1(19), pp. 27-44.

Box, G. E. P., Jenkins, G., Reinsel, G. C. (1994) Time Series Analysis: Forecasting and Control, third edition. Prentice-Hall.

Cieślak, M. et al. (1997) Prognozowanie gospodarcze. Metody i zastosowania. Wydawnictwo Naukowe PWN, Warszawa.

Crane, D. B., Crotty, J. R. (1967) A two-stage forecasting model: Exponential smoothing and multiple regression. Management Science, Vol. 13 (8), Series B, pp. 501-507.

Ďurka, P., Pastoreková, S. (2012) ARIMA vs. ARIMAX – which approach is better to analyze and forecast macroeconomic time series? Proceedings of 30th International Conference Mathematical Methods in Economics, Karviná, Czech Repubic.

Engle, R. F., Granger, C. W. J. (1987) Co-integration and error correction: Representation, estimation and testing. Econometrica, 55(2), pp. 251–276.

Franses, P.H., Legerstee, R. (2013) Do statistical models for SKU level data benefit from including past expert knowledge? International Journal of Forecasting 29, pp. 80–87.

Forbes (2013) W Europie znika stal. Dowód na wyłudzanie VAT? http://www.forbes.pl/artykuly/sekcje/wydarzenia/w-europie-znika-stal--dowod-na-wyludzanie-vat,26826,1

Główny Urząd Statystyczny (2013) Wskaźniki makroekonomiczne. http://www.stat.gov.pl/gus/wskazniki_makroekon_PLK_HTML.htm

Główny Urząd Statystyczny (2013) Uwagi metodyczne. http://www.stat.gov.pl/gus/5840_2794_PLK_HTML.htm

Grudkowska, S. (2013) JDemetra+. User Manual. National Bank of Poland.

Klug, F. (2011) Automotive Supply Chain Logistics: Container Demand Planning using Monte Carlo Simulation. International Journal Automotive Technology and Management, Vol. 11, No. 3, pp. 254-268.

Rippe, R., Wilkinson, W., Morrison, D. (1976) Industrial Market Forecasting with Anticipations Data. Management Science, Vol. 22, No. 6 (Feb., 1976), pp. 639-651.

Szeplewicz, K. (2011) Wskaźniki wyprzedzające koniunktury - analiza ekonometryczna. Prace i materiały Instytutu Rozwoju Gospodarczego, IRG SGH, Warszawa, nr 87, pp. 33-63.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 37 – 46

COMPUTER-ASSISTED CHOICE OF SMOOTHING PARAMETER IN KERNEL METHODS APPLIED IN ECONOMIC ANALYSES

Aleksandra Baszczyńska
Department of Statistical Methods
University of Lodz
e-mail: albasz@uni.lodz.pl

Abstract: In the kernel method, it is necessary to determine the value of the smoothing parameter. Not without significance is the fact of using the objectivity in the selection of this parameter and a certain automation of the selection procedure, which is important especially for novice users of kernel methods in the process of statistical inference. In the paper some methods of choice of the smoothing parameter are presented with the results of the simulation study that indicate these methods of selecting the smoothing parameter as handy tool when kernel methods are used in economic analyses.

Keywords: kernel method, smoothing parameter, Silverman's practical rule, SiZer map

INTRODUCTION

Kernel method is widely used in the estimation procedures of functional and numerical characteristics as well as in the hypothesis verification procedures concerning, for example, symmetry, goodness-of-fit or independence of random variables.

In all mentioned procedures it is necessary to determine the form of the weighting function known as the kernel function and the value of smoothing parameter which specifies the amount of smoothing in the kernel method. The parameters of kernel methods (kernel function and smoothing parameter) greatly influence the results of kernel methods.

In the case of kernel function basic attention is paid to the order of the kernel function which is closely connected with the number of vanishing moments and the number of existing derivatives of unknown density function [cf. Horová et al. 2012].

Classical kernel function, that is second order function, is unimodal and symmetric around zero density function [Gajek, Kałuszka 1996], [Wand, Jones 1995], [Silverman 1996]. It satisfies the following conditions [cf. Domański et al. 2014]:

$$\int_{-\infty}^{+\infty} K(u)du = 1, \tag{1}$$

$$\int_{-\infty}^{+\infty} uK(u)du = 0, \tag{2}$$

$$\int_{-\infty}^{+\infty} u^2 K(u)du = \mu_2(K) > 0. \tag{3}$$

Gaussian kernel function:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{u^2}{2}\right)}, \tag{4}$$

which is the kernel function with the conditions (1)-(3) is mostly used in practical applications.

Kernel function of k-th order, where k is even number, fulfils the following:

$$\int_{-\infty}^{+\infty} K(u)du = 1, \tag{5}$$

$$\int_{-\infty}^{+\infty} u^l K(u)du = \mu_l(K) = 0 \quad \text{for } l = 1,...,k-1, \tag{6}$$

$$\int_{-\infty}^{+\infty} u^k K(u)du = \mu_k(K) \neq 0. \tag{7}$$

A large variety of kernel functions are presented in literature [e.g. Domański, Pruska 2000].

For $k = 2$ and $K(u) \geq 0$ kernel function is a density function. The kernel estimator of density function with such kernel function is also the density. For $k > 2$ kernel function may be negative what can cause that the density estimator may be negative.

Let $X_1, X_2, ..., X_n$ be continuous random variables with density function f. In parametric approach we assume that observations belong to one of the known density and the procedure of estimation means only estimation of parameters. In nonparametric approach we can use e.g. histogram or kernel Rosenblatt-Parzen estimator:

$$\hat{f}(x)=\frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h}\right), \tag{8}$$

where $K(u)$ denotes kernel function, h is a smoothing parameter (bandwidth). Smoothing parameter, depending on the number of the sample n, $h=h(n)$: $\{h(n)\}$, is a sequence of a non-random positive numbers.

CHOICE OF THE SMOOTHING PARAMETER

Choice of the smoothing parameter is a crucial issue in the kernel inference procedures. The subjective method is the simplest but time consuming one. In the case of density estimation, various kernel density estimators with different values of smoothing parameters are constructed and that parameter is chosen as the proper one which is used to construct "the best" estimator. This method is simply the judging by user's eye. Sometimes this method is treated as the pilot method of choosing the smoothing parameter in more complicated methods of selection of parameter. In the practical applications, not without significance is the fact of existing the objectivity in the selection of the smoothing parameter and a certain automation of the selection procedure, which is important especially for novice users of kernel methods in the process of statistical inference.

Silverman's rule of thumb is one of the mostly used objective method of smoothing parameter choice. When asymptotic mean integrated square error is used as the measure of closeness of the estimator $\hat{f}$ to the true density f, smoothing parameter which minimize this measure with the assumption that the unknown density is normal $N\left(0,\sigma^2\right)$ is the following:

$$\hat{h}_{SRT}=1.06\,\sigma n^{-\frac{1}{5}}, \tag{9}$$

where σ can be estimated by $\hat{\sigma}=\sqrt{\frac{\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2}{n-1}}$ or $\hat{\sigma}_{IQR}=\frac{X_{\left\lceil\frac{3n}{4}\right\rceil}-X_{\left\lceil\frac{n}{4}\right\rceil}}{\Phi^{-1}\left(\frac{3}{4}\right)-\Phi^{-1}\left(\frac{1}{4}\right)}$,

for Φ^{-1} standard normal quantile function.

Terrell and Scott maximal smoothing method uses upper bound for the parameter calculated on the base of the asymptotic mean integrated square error. For specific value of scale, for example variance, this bound is attained by family of beta distributions $B(4,4)$ with variance σ^2 and it minimizes $\int_{-\infty}^{+\infty}\left(f^{(2)}(x)\right)^2dx$.

The unknown $\int_{-\infty}^{+\infty}\left(f^{(2)}(x)\right)^2 dx$ is estimated by its upper bound. The smoothing parameter is the following:

$$\hat{h}_{MSTS} = \frac{3}{35^5}\hat{\sigma}\left(\frac{R(K)}{\mu_2^2(K)}\right)^{\frac{1}{5}} n^{-\frac{1}{5}}, \tag{10}$$

where: $R(K) = \int_{-\infty}^{+\infty} K(u)^2\,du$, $\mu_2(K) = \int_{-\infty}^{+\infty} u^2 K(u)du$.

Smoothing parameter selected by least squares cross-validation is the following:

$$\hat{h}_{LSCV} = \arg\min_{h\in H_n} LSCV(h), \tag{11}$$

where: $LSCV(h) = \int_{-\infty}^{+\infty} \hat{f}_h^2(x)dx - 2\frac{1}{n}\sum_{i=1}^{n} \hat{f}_{h,-i}(X_i)$ and

$\hat{f}_{h,-i}(X_i) = \frac{1}{(n-1)h}\sum_{j\neq i}^{n} K\left(\frac{X_i - X_j}{h}\right)$.

Smoothing parameter in the biased cross-validation method of selection has the form:

$$\hat{h}_{BCV} = \arg\min_{h} BCV(h), \tag{12}$$

where $BCV(h) = \frac{R(K)}{nh} + h^4\left(\frac{\mu_2(K)}{2}\right)^2 R\left(\tilde{\tilde{\tilde{f}}}^{(2)}\right)$ and

$R\left(\tilde{\tilde{\tilde{f}}}^{(2)}\right) = \frac{1}{n^2}\sum\sum_{i\neq j}(K'' * K'')(X_i - X_j)$.

In plug-in methods the pilot estimator with initial smoothing parameter is plugged in the formula of mean integrated square error. In this way the estimates of the unknown quantities are used. Smoothing parameter has the form:

$$\hat{h}_{DPI} = \left(\frac{R(k)}{\mu_2^2(K)\hat{R}\left(f^{(2)}\right)}\right)^{\frac{1}{5}} n^{-\frac{1}{5}}, \tag{13}$$

where $\hat{R}\left(f^{(2)}\right) == \frac{1}{n^2 g}\sum_{i,j=1}^{n} L^{(4)}\left(\frac{X_i - X_j}{g}\right)$ is the estimator with the kernel function L and smoothing parameter g, where L and g may be different from K and h.

In the iterative method the smoothing parameter is the following:

$$\hat{h}_{ITE} = \arg\min_{h} \hat{A}MISE(h), \tag{14}$$

where $\hat{A}MISE(h)$ is the suitable estimator of asymptotic mean integrated square error.

SIZER MAP

SiZer map is another method of assessing the smoothing parameter, quite different from methods mentioned earlier. It presents not only one value of smoothing parameter used in kernel density estimation but the range of values. It is a graphical device which shows different structures of curves with different levels of smoothing. In this way it is possible to indicate significant features of estimator with respect to smoothing parameter used in the construction of density estimator.

SiZer map is used in the situation where we are interested in assessing the estimator of density function, especially which local maximums are true and which are false ones. It is a very good method to separate noise from the signal.

The family of kernel estimators for different values of smoothing parameter [cf. Baszczyńska 2014]:

$$\left\{\hat{f}_h(x): h \in [h_{\min}, h_{\max}]\right\} \tag{15}$$

are constructed, where $h_{\min} = 2B$, B is the binwidth, $h_{\max} = x_{\max} - x_{\min}$.

All estimators in the family (15) are constructed with Gaussian kernel function for which number of points where derivative is zero decreases monotonically when the values of smoothing parameter is bigger.

A density estimator has derivatives equal to 0 at points of minimum, maximum and points of inflection. Before a point of minimum (maximum) the sign of the derivative $\frac{\partial^m E\left(\hat{f}_h(x)\right)}{\partial x^m}$ is positive (or negative), and after it, the derivative is negative (or positive). It is possible to identify structure of the estimator by zero crossings of the m*th* order of the derivative.

The hypothesis, verified in SiZer map, are the following [Chaudhuri, Marron 1999]:

$$H_0^{h,x} : \frac{\partial^m E\left(\hat{f}_h(x)\right)}{\partial x^m} = 0, \tag{16}$$

and

$$H_1^{h,x} : \frac{\partial^m E\left(\hat{f}_h(x)\right)}{\partial x^m} \neq 0. \tag{17}$$

The SiZer map is constructed in the following way: on the horizontal axis, values of x are presented and on the vertical there are values of smoothing parameters. Different coloured regions are presented, where each colour has the special meaning. The following colour code is used: blue region means that $\hat{f}_h(x)$ is significantly increasing; red regions – $\hat{f}_h(x)$ is decreasing; purple regions – $\hat{f}_h(x)$ is not significantly increasing or decreasing; grey regions mean that the data are too sparse to make statements about the significance of increasing or decreasing of density estimator.

SIMULATION STUDY

A simulation study was conducted to indicate some of the properties of the kernel density estimator, particularly the dependence of the density kernel estimation results on the parameters of the kernel method (kernel function and smoothing parameter) and, additionally, on the size of the sample. Seven sets of observations were regarded in the study:

1. random sample generated from population of normal distribution $N(0,1)$; $n=10$.
2. random sample generated from population of normal distribution $N(0,1)$; $n=30$.
3. random sample generated from population of normal distribution $N(0,1)$; $n=100$.
4. random sample generated from population with density of a mixture of two normal distributions $N(0,1)$ and $N(10,1)$ with equal weights; $n=10$.
5. random sample generated from population with density of a mixture of two normal distributions $N(0,1)$ and $N(10,1)$ with equal weights; $n=30$.
6. random sample generated from population with density of a mixture of two normal distributions $N(0,1)$ and $N(10,1)$ with equal weights; $n=100$.
7. 107 observations of the futures contract of price for units of carbon dioxide (CO2) in USD per tonne (current price from 2.01.2014 to 30.05.2014). Source of the data is the following: http://finanse.wp.pl/notowania-surowce-online.html.

In this way, different (small, medium and large) number of observations and in addition unimodal populations (sets (1)-(3)) and bimodal populations (sets (4)-(6)) were taken into account.

For all sets of observations the kernel density estimators were calculated with various kernel functions (e.g. classical kernels: Gaussian, Epanechnikov, uniform, quartic, triangular and kernel functions of higher order) and smoothing

parameter chosen in various way. The methods of choosing the smoothing parameter were the following: Silverman's rule of thumb SRT, Terrell and Scott maximal smoothing method MSTS, least squares cross-validation method LSCV, biased cross-validation BCV, direct plug-in method DPI, iterated method ITE.

The implementation of regarded methods was done using software with functions and toolboxes in Matlab (e.g. http://math.muni.cz/english/science-and-research/developed-software/232-matlab-toolbox.html; http://www.unc.edu/~marron/marron_software.html).

Values of smoothing parameters in kernel density estimation, calculated for all regarded sets of observations with Gaussian kernel function are presented in Table 1.

Table 1. Values of smoothing parameter in kernel density estimation with Gaussian kernel function

Observations	Method of smoothing parameter selection					
	SRT	MSTS	LSCV	BCV	DPI	ITE
n=10 N(0,1)	0.72008	0.77764	1.0504	1.5487	0.78788	1.1764
n=30 N(0,1)	0.60328	0.65151	0.7187	0.90829	0.72125	0.74501
n=100 N(0,1)	0.35267	0.38086	0.44758	0.4378	0.47666	0.41665
n=10 N(0,1) and (10,1)	3.6576	3.9500	1.4854	7.8638	2.8882	6.3832
n=30 N(0,1) and (10,1)	2.6542	2.8664	0.69337	5.7036	2.0929	0.87573
n=100 N(0,1) and (10,1)	2.0755	2.2414	0.50585	4.459	1.6562	0.54419
2.01.14-30.05.14 CO_2 price	0.21279	0.2298	0.026502	0.21587	0.26041	0.18841

Source: own calculations

Comparing results of the kernel density estimations for samples from unimodal and bimodal populations gave the conclusion that in this second case the smoothing parameters are bigger. While the parameter controls the amount of smoothing, the issue of global or local smoothing parameter should be regarded. When a bandwidth is local, the amount of smoothing varies at each location. Around the modes the smoothing parameter should be smaller while in the tail of a distribution one can use much smoothing.

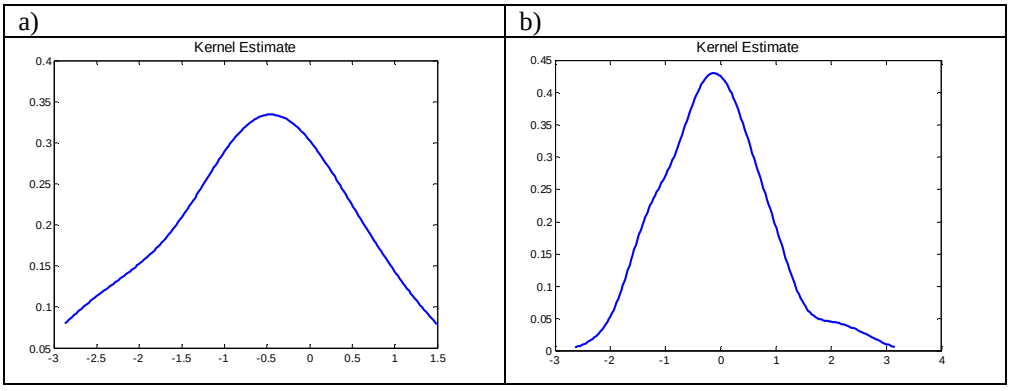
When biased cross-validation is used, the smoothing parameters for all sets of observations are bigger than in other methods of smoothing parameter selection.

It can be noticed that in most cases, the bigger the sample is, the bigger parameter is used in density estimation.

For the data of CO_2 price, the values are rather similar with exception of LSCV method (assuming independence of the observations). Similar results were obtained for other kernel functions.

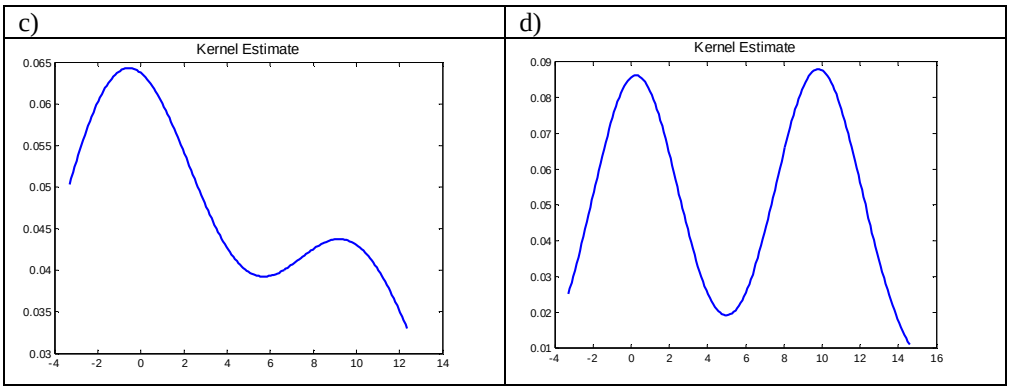
To deepen the analysis of kernel density estimation, including the significance of the sample size in the procedure of estimation, for all sets of observations kernel density estimators were calculated. Figures 1-3 present kernel density estimator where Gaussian kernel function and Silverman's rule of thumb were used for small and big sample sizes and for observations of the futures contract of price for units of carbon dioxide.

Figure 1. Kernel density estimator (Gaussian kernel function, Silverman's rule of thumb), data from unimodal populations a) n=10; b) n=100



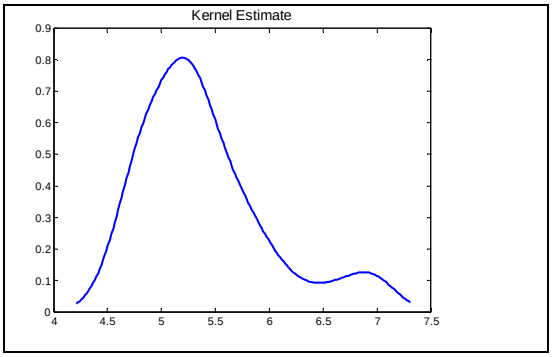
Source: own calculations

Figure 2. Kernel density estimator (Gaussian kernel function, Silverman's rule of thumb), data from bimodal populations c) n=10; d) n=100



Source: own calculations

Figure 3. Kernel density estimator (Gaussian kernel function, Silverman's rule of thumb), data of the futures contract of price for units of carbon dioxide

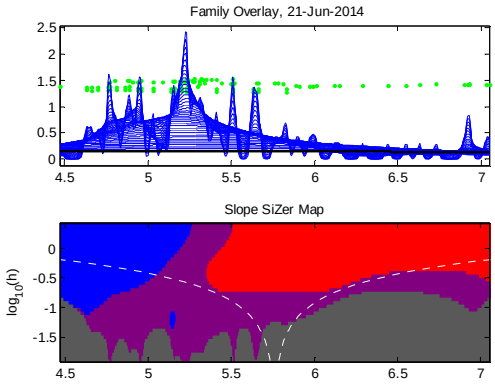


Source: own calculations

In the case of unimodal distribution, kernel density estimator is also unimodal, even for small sample. But bigger value of smoothing parameter should be used in this process of density estimation. The smaller size of the sample, the bigger smoothing parameter is used (cf. Table 1). When sample is chosen from bimodal distribution, only big sample ensures clearly bimodal density estimator, with smaller value of smoothing parameter.

Kernel density estimator for data of the futures contract of price for units of carbon dioxide indicates asymmetry and unimodality of the distribution. Additionally, for this set of observation the Sizer map was used. Figure 4 presents this SiZer map.

Figure 4. SiZer map, data of the futures contract of price for units of carbon dioxide



Source: own calculations

In SiZer map wide range of values of smoothing parameter is considered. For small values of smoothing parameter it is impossible to determine which maximums are true and which are false (for values up to approximately 0.0316, $\log_{10}(h) = -1.5$, grey region is observed). It means that the kernel density estimator is undersmoothed. For values of smoothing parameter bigger than 0.1 it is easy to inference that kernel estimator is positive or negative (blue regions and red regions). This result is consistent with previous result about values of smoothing parameter (cf. Table 1). The bigger value of smoothing parameter, the more smooth the estimator is.

SUMMARY

Value of smoothing parameter in kernel density estimator is an issue of great importance. For users of kernel methods, it is very important that they can use some objective and automatic way of finding "the best" value of this parameter in practical applications. This condition is fulfilled by methods mentioned above with the help of specialised software. The presented results encourage not only scientists but also unexperienced users to apply them. But deeper analysis of "optimal" choice of smoothing parameter and kernel function is necessary in future research.

REFERENCES

Baszczyńska A. (2014) Testing Significance of Peaks in Kernel Density Estimator by SiZer Map, Proceedings of the 8th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena, Foundation of the Cracow University of Economics, Kraków.

Chaudhuri P., Marron J.S. (1999) SiZer for Exploration of Structure in Curves, Journal of the American Statistical Association, 94, 447, 807 – 823.

Domański C., Pekasiewicz D., Baszczyńska A., Witaszczyk A. (2014) Testy statystyczne w procesie podejmowania decyzji, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Domański C., Pruska K. (2000) Nieklasyczne metody statystyczne, PWE, Warszawa.

Gajek L., Kałuszka M. (1996) Wnioskowanie statystyczne. Modele i metody, Wydawnictwa Naukowo-Techniczne. Warszawa.

Horová I., Koláček J., Zelinka J. (2012) Kernel Smoothing in MATLAB. Theory and Practice of Kernel Smoothing, World Scientific, New Jersey.

Silverman B.W. (1996) Density Estimation for Statistics and Data Analysis, Chapman & Hall, London.

Wand M. P., Jones M.C. (1995) Kernel Smoothing, Chapman & Hall, London.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 47 - 63

OKUN'S LAW IN OECD COUNTRIES IN 1990 – 2013

Dariusz J. Błaszczuk
Department of Economics, Vistula University
e-mail: d.blaszczuk@vistula@edu.pl

Abstract: Politicians try to reach targets that usually contradict. It leads to lack of optimal long-term equilibrium.
The main target should be the maximal growth rate of GDP and other targets should remain within predetermined limits. Politicians should influence relations between every two targets.
Empirical study of relations between unemployment and GDP growth rates is done for every OECD country basing on quarterly data for 20 years. According to the results of the investigation countries are put into "strategic groups". Paper ends with preliminary proposals for economic policy makers.

Keywords: economic policy targets, long-term equilibrium, Okun's law, OECD countries, GDP growth rate, unemployment rate

INTRODUCTION

The aim of the overall macro-economic policy is to set a specific combination of individual targets. Depending on the situation in the economy in a given period there are, usually at the same time, for example: growth rate of GDP higher than a certain figure, level of the unemployment rate lower than a specified figure, inflation rate lower than a given value, assumed level of exchange rate, sustainability of public finances. They almost always are at least partly contradictory[1]. Striving to achieve them at the same time leads in a long run to political instability[2], especially in times of economic crises[3].

[1] Although for example, higher economic growth is generally associated with a lower level of unemployment, the two phenomena are accompanied by higher inflation and the stability of domestic prices usually interfere with the stability of the exchange rate.

[2] A classic example of such a situation was the collapse of the centrally planned economies, where the aim was to optimise the objectives of individual economic policies. Of course, the collapse was caused by the whole convolution of reasons and mentioned above was not the only important.

The solution of this problem is probably palpable intuitively by most of economic policymakers. It boils down to determine the maximum GDP growth rate that enables the fastest possible improvements in the standards of living of the citizens of the country[4] while maintaining the levels of other objectives within the predetermined limits. The way to achieve this objective is to make respective changes of the relationships within every pair of specified economic policy targets[5]. Therefore the relationships between every two specified targets of economic policy for the given economy should be earlier determined.

In the paper there are discussed the results of the empirical verification of the modified Okun's law, namely of the relationships between unemployment rates and the GDP growth rates[6]. There is, obviously, a large number of other empirical research studies directed at verification of the Okun's law[7]. The comparison of the received results with the figures provided by other authors is, however, somewhat difficult because of the different analytical forms of functions used in the respective studies. In general one can say, however, that in the presented study there is an expected sign of relationship between unemployment rate and GDP growth one only in case of 13 out of 33 analysed countries[8] and the slope varies from -1.3 to -0.1 (1% increase of unemployment rate causes decrease of GDP growth rate from 0.8% to 10.0%).

The results of the analysis are then used to build three maps of "strategic groups" of the analysed countries. Every map is based on two criteria: one associated with the level of unemployment and the second based on the levels of GDP. Analyses of these maps provide formulation of initial recommendations for economic policy makers of the relevant countries.

Further research will encompass empirical and theoretical long-term equilibrium points for individual countries and further recommendations for economic policy makers of the individual countries.

---

[3] Social unrests in Spain during the last economic crisis have been a striking example of such an instability.

[4] This statement is different from the one represented by eg. P. A. Samuelson and W. D. Nordhaus who claim that the respective aim should be the assumed level of inflation rate. See: Samuelson and Nordhaus (2004).

[5] For discussion of this issue see, for instance, Błaszczuk (2014a) and Błaszczuk (2014b).

[6] The relationship between the levels of inflation and unemployment rates virtually for the same group of countries and in virtually the same period have been presented to the IV Nationwide Scientific Conference in Poznan, Poland on April 25, 2014 [Błaszczuk D. J. (2014a)] and between rates of inflation and GDP growth rates, also for a similar group of countries and for a similar period, were presented to the III International Scientific Conference in Lodz, Poland on June 10, 2014 [Błaszczuk D. J. (2014b)].

[7] A comprehensive list of such studies is given, for instance, in R. Durech, A. Minea, L. T. Mustea, L. Slusna (2014).

[8] Expected and unexpected results have been also received by other authors, for instance, P. Klimczyk, G. Wronowska (2010).

ASSUMPTIONS OF THE INVESTIGATION, STATISTICAL DATA SOURCES AND RESEARCH METHOD

Assumptions apply to the subject, scope and period of the investigation. According to the statements given in the introduction, the subject of this research is the relationship between unemployment rates and the growth rates of GDP in selected countries. The research was assumed to cover all OECD countries in the period 1990Q1-2013Q4, so during the past two Juglar's business cycles.

Additionally, in order to obtain comparable results, it was assumed that the data on the harmonised unemployment rate (HUR) and the rate of GDP would be taken from the OECD statistical sources. However, data on the harmonized unemployment rate[9] for 13 countries[10] are available for shorter periods, and sometimes even much shorter ones than assumed above. Moreover, data on GDP growth rates[11] are not available for Greece and for 20 other countries[12] are available for shorter periods, sometimes much shorter from the assumed above. Unfortunately, these periods usually do not comply with the periods for which data are available on HUR. On top of that, due to the lack of other equally reliable data, in the investigation it has been ignored that, in some cases, only *estimated* figures were available, and in other there were changes of data collection methods (*break*).

Having regard to these considerations, the study of the relation between HUR and GDP growth rates was made for 33 countries. The number of observations for vast majority of these countries were above 60[13] (see column 4 of table 1).

Next it has been assumed that the GDP growth rate of each country separately [$\mathbf{r_{jt}(GDP)}$] can be expressed as a logarithmic function with constant and respective hyperbolic one of the levels of its harmonised unemployment rate ($\mathbf{HUR_{jt}}$):

$$\mathbf{r^L_{jt}(GDP)} = b^L_{0j} + b^L_{1j}\,\mathbf{(\ln HUR_{jt})} + \boldsymbol{\xi}^{\mathbf{L}}_{\mathbf{jt}} \quad (1)$$

$$\mathbf{r^H_{jt}(GDP)} = b^H_{0j} + b^H_{1j}\,\mathbf{(1/HUR_{jt})} + \boldsymbol{\xi}^{\mathbf{H}}_{\mathbf{jt}} \quad (2)$$

where: j = 1, 2, …, 33 – country number and t = 1, 2, …, t' – quarter number.

[9] Harmonised unemployment rate: all persons, seasonally adjusted, http://stats.oecd.org/index.aspx?DatasetCode=KEI [28 Apr 2014].

[10] Namely for Austria, Czech Republic, Estonia, Germany, Greece, Hungary, Iceland, Israel, Poland, Slovak Republic, Slovenia, Switzerland and Turkey.

[11] Gross domestic product - expenditure approach, growth rate compared to previous quarter, seasonally adjusted, http://stats.oecd.org/WBOS/index.aspx [28 Apr 2014].

[12] Namely for: Belgium, Chile, Czech Republic, Denmark, Estonia, Finland, Germany, Hungary, Iceland, Ireland, Israel, Japan, Luxembourg, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden and Turkey.

[13] Exceptions are: Chile (43), Iceland (44), Ireland (55), Switzerland (16) and Turkey (36).

Lastly it has been assumed that OLS method may be used for estimation of every of the 33 + 33 = 66 equations taking into account, inter alia, that functions (1) and (2) are linear after appropriate transformations.

RESULTS OF THE INVESTIGATION

The results of the research are partially in accordance with expectations. Above all, do not dismay relatively low, and sometimes even very low, values of R^2 because dispersions of points on the vast majority of the 33 charts do not allow to assume in advance any trends (see Annex 1).

Therefore, the estimates of the structural parameter in the logarithmic models of 15 countries[14] are clearly irrelevant statistically (t stat (b^L_{1j}) < 1.00). On the other hand estimates of constant are clearly irrelevant statistically (t stat (b^L_{0j}) < 1.00) in 14 cases, however sometimes for different countries[15] (see columns 6 and 8 in Annex 2).

Slightly better results in this respect has been obtained in case of the hyperbolic models (see 10 and 12 columns of Annex 2). Namely, from the above given list of countries with clearly statistically irrelevant estimates of structural parameters have disappeared Korea and the United Kingdom, and from the list of countries with clearly statistically irrelevant estimates of constant have disappeared as many as 11 countries[16], but 5 new[17] popped up on it.

It is worth noticing that, in general, the type of function does not affect the statistical significance of estimates of the structural parameters while it is not exactly true in case of estimates of constant[18].

Surprising is, however, a significant convergence of both theoretical lines practically for all countries (see Annex 1), despite the fact that the values of the independent variable are clearly (and for some countries by far) different from unity[19]. This is connected with a high negative correlation between estimates of the structural parameters of both functions (see Figure 1).

---

[14] Australia, Canada, Chile, Denmark, Estonia, Germany, Israel, Japan, Korea, New Zealand, Poland, Slovakia, Slovenia, Turkey and the United Kingdom.

[15] From the list disappeared: Israel, Korea, New Zealand, Poland and the United Kingdom and appeared on it: Austria, the Netherlands, Norway and Sweden.

[16] All except for Estonia, Japan and Turkey.

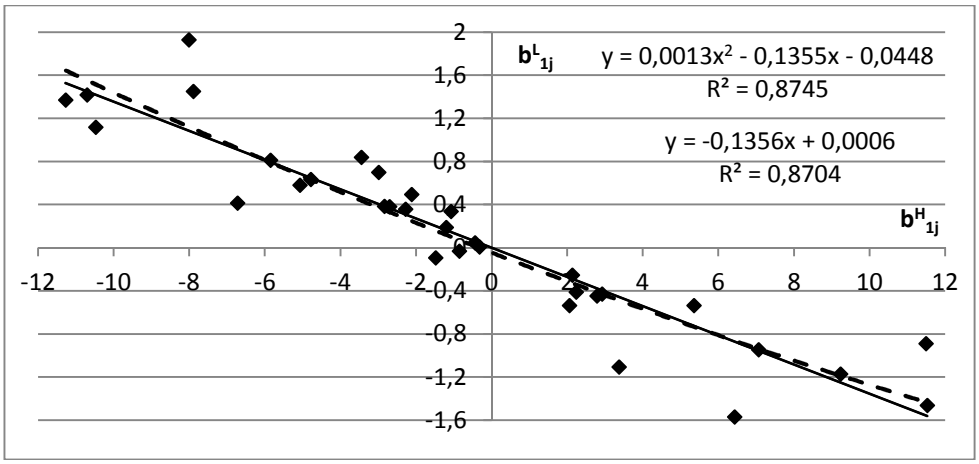
[17] Namely: Iceland, Luxembourg, Mexico, the United Kingdom and the United States.

[18] They are statistically significant in case of both functions for Czech Republic, Finland, France, Hungary and Portugal and almost for Belgium. On the other hand they are statistically relevant in case of hyperbolic functions only for Australia, the Netherlands, New Zealand, Norway, Poland, Sweden and Switzerland and almost for Austria, and in case of logarithmic models only for Ireland, Korea, Luxembourg, Mexico, Spain and the United States and almost for the United Kingdom.

[19] For prove of this statement see, for instance, Kuratowski (1971).

In this context, it should be noted that theoretical lines are compatible with the expected regardless of the type of function only in the case of twelve countries (**Hungary**, Iceland, **Ireland**, Israel, Korea, Luxembourg, Mexico, **Portugal**, **Spain**, Turkey, the United Kingdom and the United States), but only in the case of countries highlighted in bold letters the estimates of structural coefficients are statistically significant[20]. In addition, relevant economically are estimates of the structural parameters in case of logarithmic models for Poland and Slovenia.

Figure 1. Values of b^{L}_{1j} and b^{H}_{1j}



Source: own computations

At the same time, "strategic groups" of countries can be clearly noted, at least for three pairs of the criteria:

a) angle of the (linear) theoretical line r(GDP)^ {∂[r(GDP)^]/(∂[HUR]} and the variability of HUR (the difference between the largest and the smallest values) that is the average change of GDP growth rate, other unchanged, related to the 1 p. p. change in the unemployment rate;
b) the average level of GDP growth rate [r^{av}(GDP)], taking into account the sign of the first derivative of the theoretical curve, and the variation of HUR, as well as
c) the average level of GDP growth rate, taking into account the sign of the first derivative of the theoretical curve, and the average level of HUR [r^{av}(HUR)].

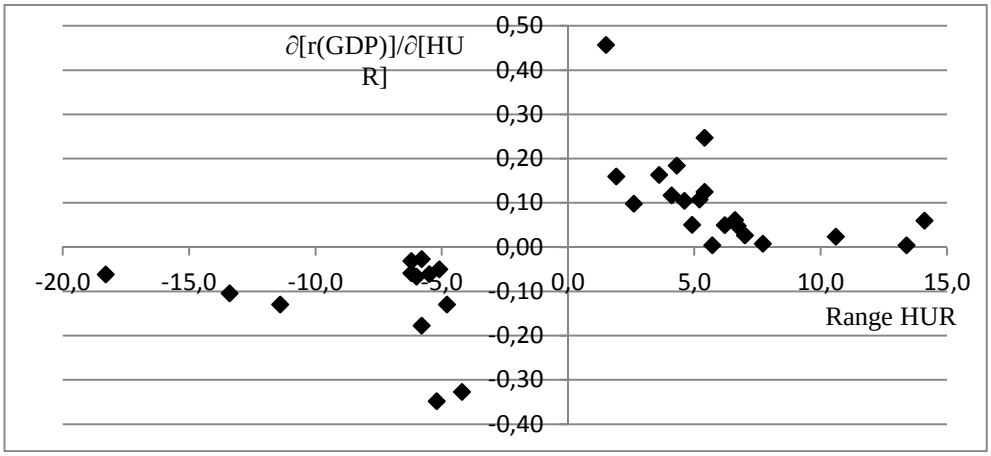
According to the first pair of criteria, countries can be divided, as was mentioned earlier, into two mega-groups (see Figure 2), namely the group of countries with:

[20] By the way it should be noted that three of them (except for Hungary) belong to the so called GIPSI group subject to huge problems during the last economic crisis.

a) a negative first derivative, where the relationships between the levels of unemployment and the rate of GDP levels are in accordance with the Okun's law and
b) a positive first derivative, where these relationships are not in accordance with the Okun's law.

Within the first group covering 13 countries one can distinguish countries with medium (4.2 - 6.2 p. p.) and very large (11.4 - 18.3 p. p.) variations of unemployment rates. On the other hand in the second group one can see countries with small (1.5 - 2.6), medium (3.6 - 7.7) and very large (10.6 - 14.1) variations of these rates. The examined countries can also be divided according to the angle (both negative and positive) into three groups, namely, countries with large, medium and small slopes.[21] Next, combining these two criteria, the examined countries can be divided into nine "strategic groups" (see Annex 3).

Figure 2. Values of the slope of the linear theoretical lines r(GDP)^and the ranges of HUR



Source: own computations

The first "group" consists of countries with medium-sized areas of variation of HUR and theoretical line negative slopes:
a) small: United States, Israel, Slovenia, Turkey, Korea and United Kingdom;
b) medium: Hungary and Mexico;
c) big: Iceland and Luxembourg.

To this group belong also countries with vast areas of HUR variation and:
a) an average negative theoretical line slopes: Ireland and Portugal and
b) small negative theoretical line slope – Spain creating, as the only one, a one-element "group".

[21] On the Figure 2 one can also easily notice the expected inverse relationship between a range of variation of the HUR and the slope of the theoretical line.

On the other hand small positive slope of the theoretical line (0.00 - 0.06) have countries with HUR variation:

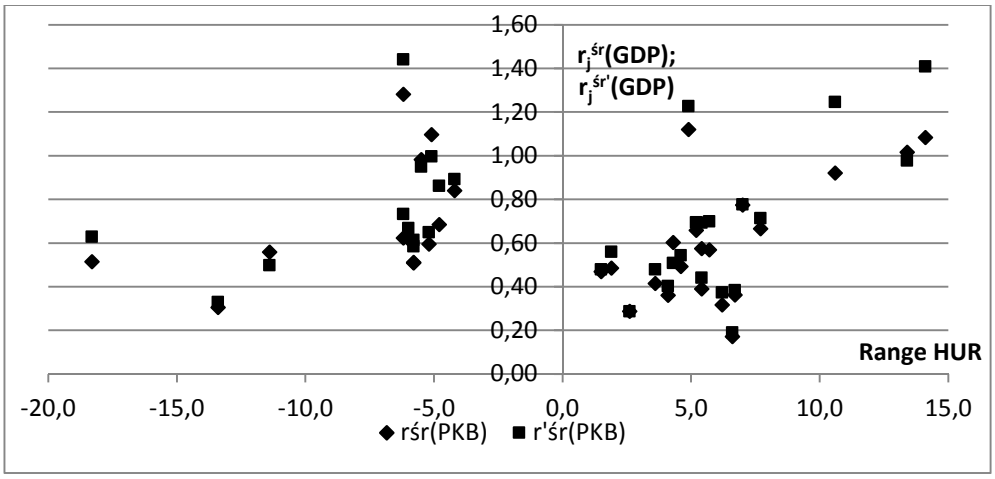
a) very vast: Estonia, Poland and Slovak Republic and
b) medium: Canada, Chile, New Zealand, Australia, Denmark, Germany and Italy.

The last two "strategic groups" according to the criteria in question consist of countries with small or medium ranges of variation of HUR and positive slopes of the theoretical lines:

a) medium (0.10 – 0.18): Japan, Austria, France, Belgium, the Netherlands, Sweden, Finland, Norway and
b) big: Czech Republic (0.25) and Switzerland (0.36).

Use of the second pair of criteria required computation of the average levels of GDP growth rates, $[r_j^{śr}(GDP)]$. However, as one can see in the Annex 1, some extreme, especially negative, GDP growth rates significantly deviated from the respective values directly adjacent to them. In view of this, also the revised average values of GDP growth rates, $[r_j^{śr'}(GDP)]$ have been computed where GDP growth rates significantly varying from the values directly adjacent to them were not included. In total, there were excluded slightly above 5.5% of the all observations, namely: 95 negative values for all 33 countries and 49 positive values for 25 countries. As a result, the adjusted value of the medium GDP growth rates are, on the whole, slightly higher than their uncorrected counterparts (see Figure 3).

Figure 3. Values of ranges of HUR and unadjusted and adjusted average GDP growth rates of OECD countries



Source: own computations

As one can see on the Figure 3, the OECD countries can be classified according to the average (unadjusted and adjusted) levels of GDP growth rates into five classes, namely: countries with a very low, low, medium, high and very high levels of this indicator.

Combining the average GDP rate levels and the ranges of HUR variation, and taking into account the sign of the first derivative of linear theoretical line, the OECD countries can be divided into nine "strategic groups", of which three (Luxembourg, Chile and Korea) are one-element (see Annex 4).

In the case of the uncorrected average GDP growth rates in the first mega-group are four "strategic groups" consisting of the countries with:

a) very large volatilities of unemployment rates and low (Portugal) and medium average GDP growth rates (Spain and Ireland);
b) medium volatility of unemployment rates and medium average GDP growth rates (United States, Mexico, Iceland, United Kingdom, Hungary and Slovenia);
c) medium volatility of unemployment rates and high average GDP growth rates (Luxembourg, Israel and Turkey) and
d) medium volatility of unemployment rates and a very high average GDP growth rate (Korea).

On the other hand in the second mega-group there are six "strategic groups" consisting respectively of countries with:

a) very small volatilities of unemployment rates and low (Japan) and medium levels of the average growth rate of GDP (Switzerland and Austria);
b) small volatilities of unemployment rates and low (Belgium, France and Finland) and medium levels of the average growth rate of GDP (Canada, The Netherlands, Norway, Sweden and Czech Republic);
c) small volatilities of unemployment rates and high level of the average growth rate of GDP (Chile);
d) medium volatilities of unemployment rates and very low (Italy) and low (Denmark and Germany) levels of the average growth rate of GDP;
e) medium volatilities of unemployment rates and medium (New Zealand) and high (Australia) levels of the average growth rate of GDP;
f) high volatilities of unemployment rates and high levels of the average growth rate of GDP (Slovak Republic, Poland and Estonia).

The results in case of the adjusted average GDP growth rates are basically similar, taking into account, of course, changes in their ranges and the overall increase in their levels (due to the predominance of missed negative values). However, Israel and Turkey on one hand and Poland and Slovakia on the other have changed their positions within the frameworks of their "strategic groups", while Mexico and Estonia "improved" their positions, moving to the next "strategic groups"' relative to the levels of the average growth rates of GDP. As a result, Mexico (as well as Israel and Turkey) joined Luxembourg and Estonia has become one-element group.

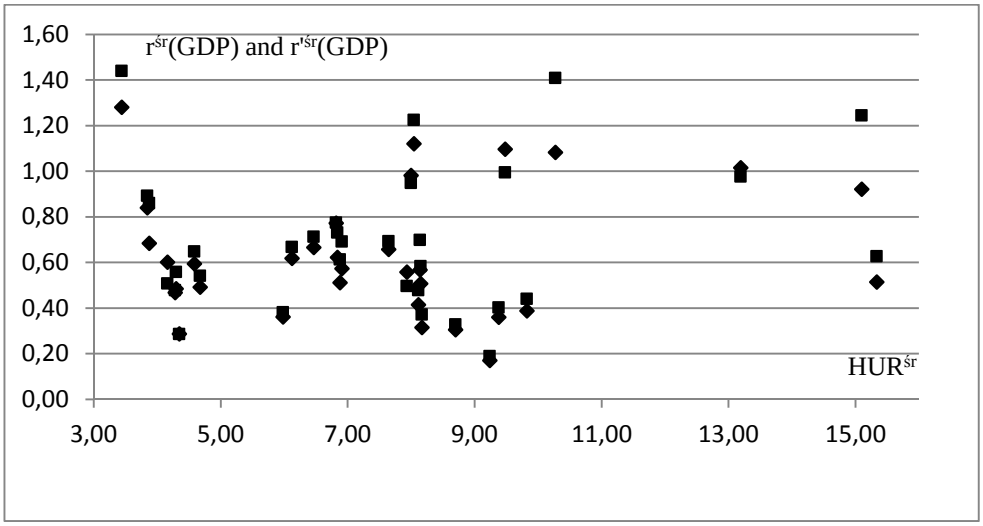
As already mentioned the last pair of the analysed criteria constitute the average levels of GDP and average levels of unemployment[22], ($HUR^{śr}$). According

[22] In the case of unemployment rates there is no sense to take into account the adjusted average rates, because the relative differences between adjusted and unadjusted rates are

to the average levels of unemployment the analysed countries can be divided into five classes, namely: countries with a very low (3.44), low (3.84 – 4,67), medium-5,98 6,91), high (7,64-10,27) and very high (13,19-15,33) levels of this indicator (see Figure 4). In this context it is worth noting that the changes both of the average as well as of the adjusted average GDP growth rates do not depend upon the changes of the average rates of unemployment[23].

Combining these two criteria, the analysed OECD countries can be divided into ten "strategic groups". Among them there is only one one-element group (see Annex 5).

Figure 4. Values of the average unemployment rates and unadjusted and adjusted average GDP growth rates of OECD countries



Source: own computations

A class for everyone is Korea that has a very high average growth rate of GDP and a very low level of the average unemployment rate. Slightly lower levels of the average GDP growth rates had also Chile and Israel, but at high average rates of unemployment, as well as Estonia, Poland and Turkey that had very high levels of unemployment rates.

High levels of average GDP growth rates were accompanied by small, medium and very high levels of unemployment in Luxembourg, Australia, Slovakia, respectively.

negligible (only in two cases these differences exceed the level of 3% (for Spain, the difference is about 3.3%, and for Turkey it is approx. -4.5%).

[23] For example, assuming a linear relationship, increase of $HUR^{śr}$ by 1 p. p. is associated with an about 0.01 p. p. increase in $r^{śr}(GDP)$ ($R^2 = 0.01$) and an about 0.02 p. p. increase in $r^{śr}$'(GDP) (R2 = 0.03).

On the other hand, the average levels of the GDP growth rates corresponded to the different average interest levels of unemployment:

a) low: Switzerland, Austria, the Netherlands, Norway, Iceland and Mexico;
b) medium: United Kingdom, The Czech Republic, Slovenia, the United States and New Zealand;
c) high: Hungary, Sweden, Ireland and Canada and
d) very high: Spain.

Several countries have low levels of the average GDP growth rates. In Japan it was accompanied by low, in Denmark by medium and in Germany, Belgium, France, Portugal and Finland by high average unemployment rates.

Very low average growth rate of GDP accompanied by a high level the average unemployment rate was reached by Italy.

Unfortunately, it is difficult to distinguish other criteria, except for the above given formal ones, for the classification of the individual countries to the given and not to the other "strategic group". In other words, it is difficult to determine the common substantive features that have the countries classified to the given "strategic group".

It should be noted, however, that in the same "groups" or in very close proximity on all three "maps" there are, for example:

a) Australia and New Zealand;
b) Austria, Switzerland, the Netherlands and Norway;
c) Belgium, France and Finland;
d) Estonia, Poland and Slovak Republic;
e) Spain, Ireland, Portugal and Hungary;
f) Israel and Turkey;
g) Luxembourg, Iceland and Mexico;
h) Slovenia and Hungary and United Kingdom, United States and Mexico;
i) Italy, Germany, Denmark and Belgium.

In most of the above cases one can trace to, in addition to the above-mentioned formal criteria also economic, political, geographical, historical or other conditions effecting in "likeness" of countries classified into the same "strategic group" or neighbouring ones. Analysis, both of "similarities" according to different criteria of the countries classified to the same "strategic group", as well as the differences between them, as well as of the differences between the various "strategic groups" will be carried out in the nearest future. It should contribute to the implementation of the mentioned in the second paragraph of this paper, main aim of the whole research study, i.e. formulation (separately for each country) of the accurate recommendations with respect to the direction of movement of the experiential long-term equilibrium point towards the "optimal" one, or in other words of the mid-term and final recommendations for macro-economic policy makers of the individual countries.

PRELIMINARY RECOMMENDATIONS FOR THE ECONOMIC POLICYMAKERS

Recommendations for the economic (and social) policy makers that can be derived on the basis of the results of the above discussed research study are limited because of the assumptions and the research method, as well as because of the availability of the statistical data. Namely, they affected the levels of statistical measures of the results of the survey and, finally, the classification of individual countries into the relevant "strategic groups".

Nevertheless, a thorough analysis of the values of every of the classification criteria in relation to the countries included in the respective "strategic groups" allows to make the following preliminary recommendations[24]:

1. Switzerland, Czech Republic, Norway, Belgium, Austria, Finland, France, Sweden, the Netherlands, Japan, Italy, Estonia, Chile, Germany, Denmark, Australia, Slovakia, New Zealand and Canada, as well as possibly Poland and Slovenia should determine:
   a) list factors affecting relationships between the levels of unemployment and the growth rates of GDP;
   b) the relationships between these factors and the shape and location of the Okun's curve;
   c) list of actions that will result in changes in the shape (the slope) and location of the Okun's curve.
2. the United Kingdom, Korea, Turkey, Israel, the United States and possibly Spain and also Poland and Slovenia should determine:
   a) list of factors influencing volatility of the unemployment rates;
   b) the relationships between these factors and volatility of the unemployment rates;
   c) list of actions that would result in the reduction of this volatility, and simultaneously, the change in shape (increase of the slope) and location of the Okun's curve.
3. all countries, perhaps with the exception for Korea, Luxembourg, Mexico, Norway, Switzerland, Austria, Japan, Iceland and the Netherlands should determine:
   a) list of factors influencing unemployment rates;
   b) the relationships between these factors and unemployment rates;
   c) list of actions that would result in reduction of the unemployment rates, and simultaneously change of the position of the (lowering) and the shape of the Okun's curve.

[24] The orders of listing the countries in this and in the next point indicate the scale and perhaps the urgency of respective actions and in third point countries are appearing in the reverse order.

These recommendations are preliminary. Some more detailed recommendations for macroeconomic policymakers of the respective countries will be formulated after solution, for every country separately, of the equation (5), as indicated in paragraph 3 of this paper and considerations, separately for every country, on the location of the "optimal" long term equilibrium point.

REFERENCES

Abel A. B., Bernanke B. S., Macroeconomics (5th ed.). Pearson Addison Wesley 2005.

Błaszczuk D. J., Cele polityki gospodarczej: zależność między stopą inflacji a stopą bezrobocia w krajach OECD w latach 1990 – 2013, IV Ogólnopolska Konferencja im. Zbigniewa Czerwińskiego „Matematyka i informatyka na usługach ekonomii", Uniwersytet Ekonomiczny w Poznaniu, Poznań 25 kwietnia 2014.

Błaszczuk D. J., Economic Policy Targets: Relationships between Inflation and GDP Growth Rates in OECD Countries in 1990 - 2013, III Konferencja Naukowa „Spatial Econometrics and Regional Economic Analysis, Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Łódź, 09-10 czerwca 2014.

Durech R., Minea A., Mustea L. T., Slusna L., Regional evidence on Okun's Law in Czech Republic and Slovakia, Economic Modelling 2014,v. 42, pp. 57-65.

http://stats.oecd.org/index.aspx?DatasetCode=KEI [28.04.2014].

http://stats.oecd.org/WBOS/index.aspx [28.04.2014].

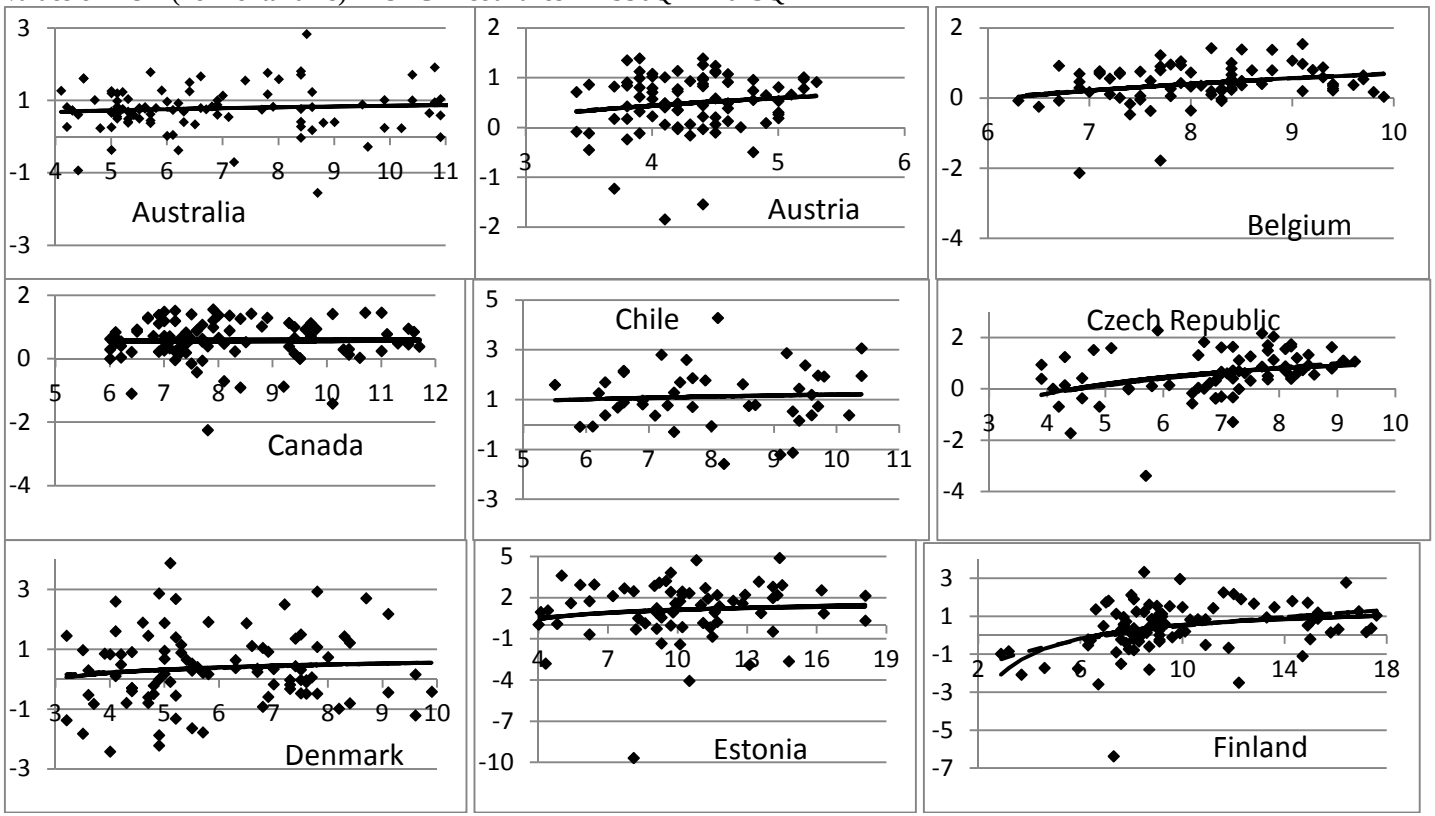
Klimczyk P., Wronowska G., Prawo Okuna w warunkach kryzysu gospodarczego, Wzrost gospodarczy i polityka makroekonomiczna (eds. W. Kwiatkowska, E. Kwiatkowski,

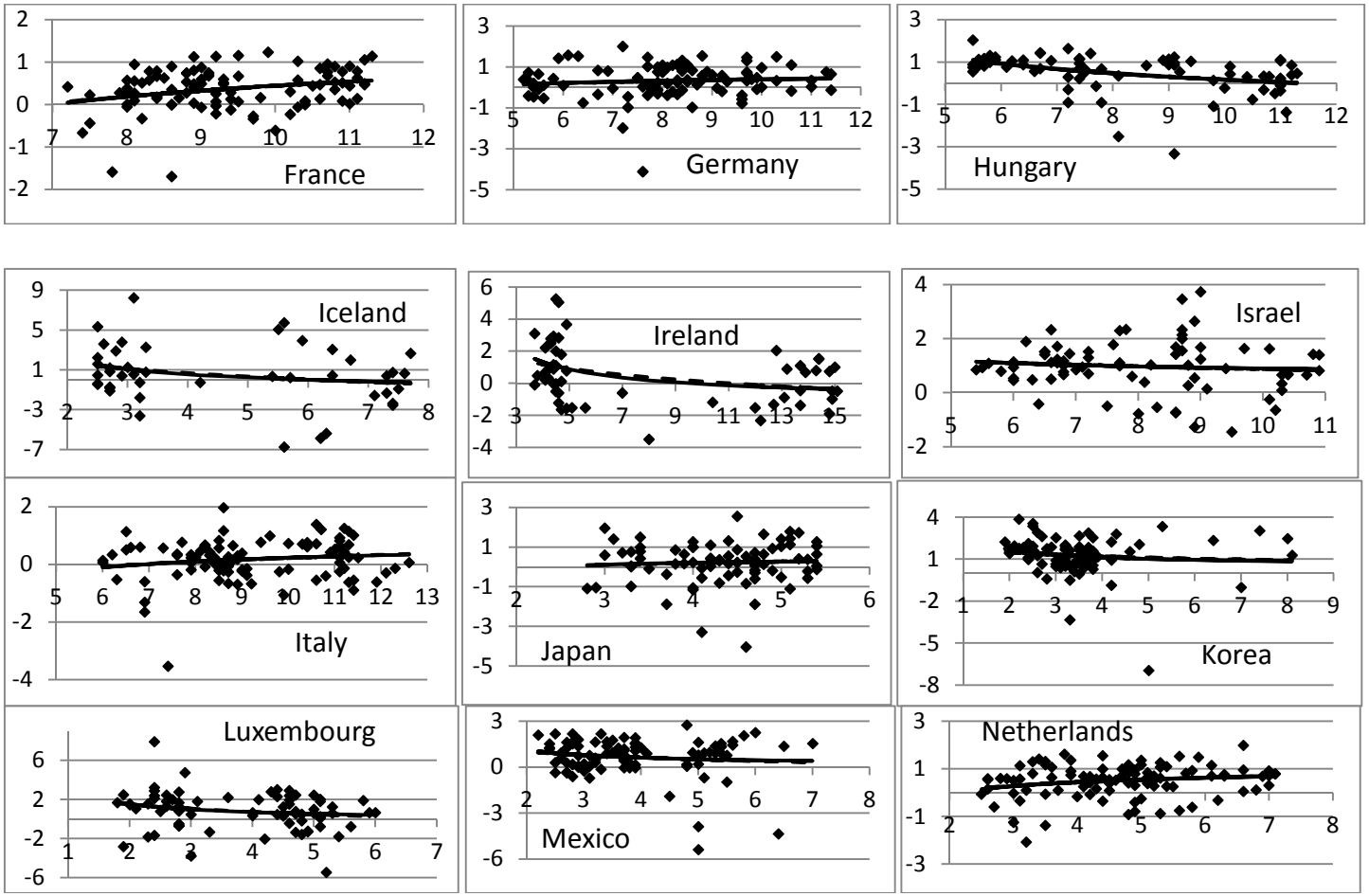
Uniwersytet Łódzki, Łódź 2010, pp. 263-272.

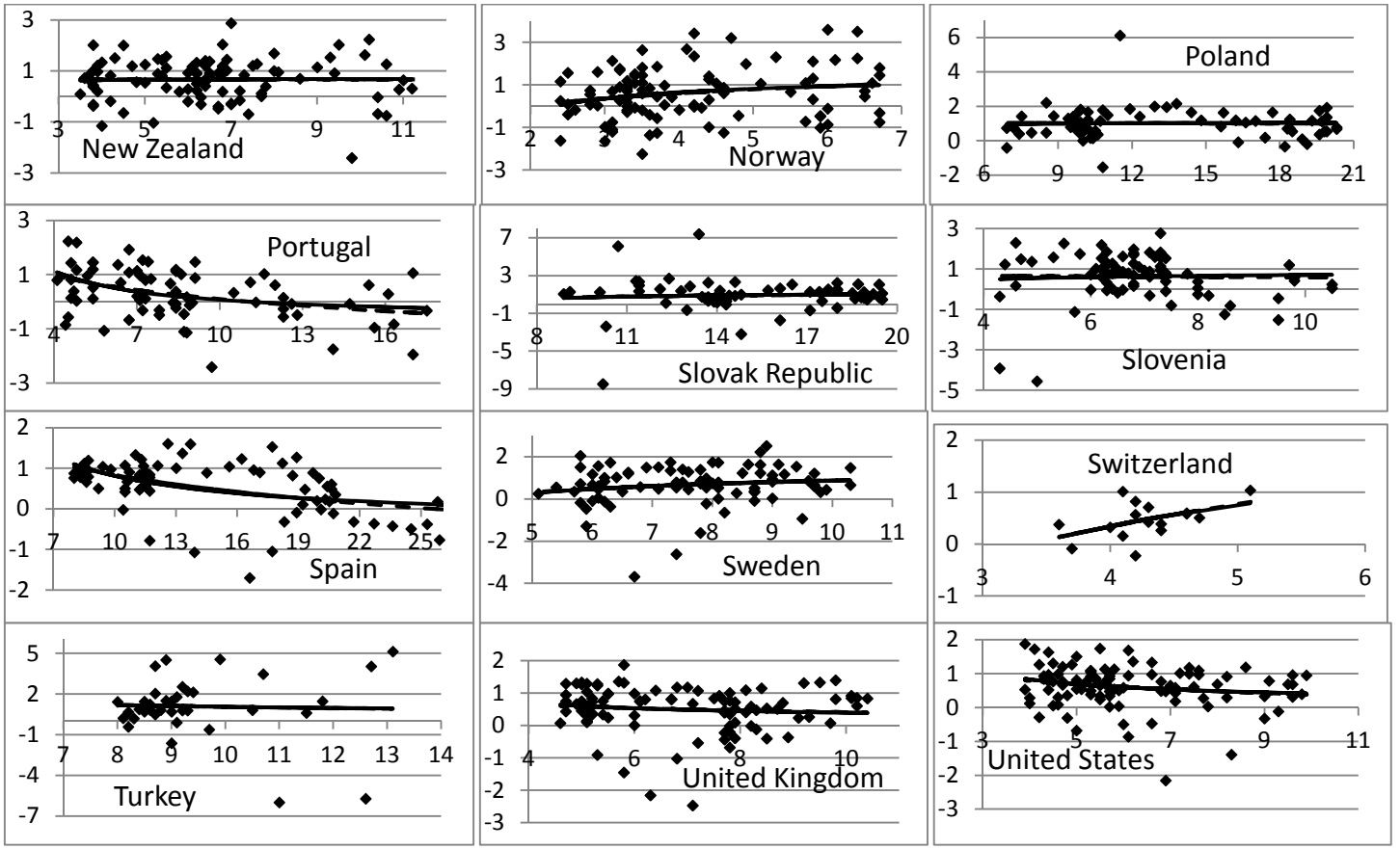
Kuratowski K., Rachunek różniczkowy i całkowy, PWN, Warszawa 1971.

Okun A., Potential GNP: Its Measurement and Significance, American Statistical Association, Proceedings of the Business and Economics Statistics Section 1962, pp. 98-104.

Annex 1. Empirical and theoretical (logarithmic – solid lines, hiperbolic – broken lines) values of r(GDP) (verical axis) and empirical values of HUR (horizontal axis) in OECD countries in 1990Q1 – 2013Q4




France
Germany
Hungary
Iceland
Ireland
Israel
Italy
Japan
Korea
Luxembourg
Mexico
Netherlands



Source: own computations

Annex 2. Results of estimation

No	Country	Observations		r(GNP) =f(ln(HUR)				r(GNP) =f(1/HUR)			
		from	T	b^L_{1j}	t Stat	b^L_{0j}	t Stat	b^H_{1j}	t Stat	b^H_{0j}	t Stat
1	Australia (AU)	Q1-1990	96	0,188	0,81	0,419	0,94	-1,204	0,78	0,964	3,80
2	Austria (A)	Q1-1993	84	0,697	1,14	-0,527	0,59	-2,990	1,15	1,189	1,94
3	Belgium (B)	Q2-1995	75	1,371	2,25	-2,448	1,92	-11,271	2,31	1,821	2,98
4	Canada (CA)	Q1-1990	96	0,045	0,12	0,476	0,61	-0,442	0,15	0,625	1,59
5	Chile (CHI)	Q2-2003	43	0,383	0,37	0,328	0,15	-2,835	0,35	1,484	1,41
6	Czech Republic (CZ)	Q2-1996	71	1,449	3,12	-2,191	2,46	-7,895	2,87	1,776	4,11
7	Denmark (DE)	Q2-1991	91	0,356	0,78	-0,260	0,32	-2,281	0,92	0,775	1,66
8	Estonia (EST)	Q1-1997	68	0,632	0,84	-0,354	0,20	-4,788	0,82	-4,788	0,82
9	Finland (SF)	Q2-1990	95	1,417	3,77	-2,770	3,27	-10,707	3,85	1,615	4,70
10	France (F)	Q1-1990	96	1,117	2,63	-2,131	2,24	-10,479	2,67	1,494	3,49
11	Germany (G)	Q2-1995	91	0,382	0,91	-0,479	0,54	-2,702	0,86	0,662	1,62
12	Hungary (H)	Q1-1996	72	-1,465	3,68	3,539	4,27	11,525	3,75	-0,991	2,41
13	Iceland (IC)	Q1-2003	44	-1,571	1,53	2,842	1,86	6,426	1,51	-1,086	0,91
14	Ireland (IR)	Q2-2000	55	-1,173	2,73	2,809	3,28	9,233	2,99	-0,985	1,74
15	Israel (IS)	Q2-1995	75	-0,431	0,74	1,872	1,55	2,919	0,65	0,604	1,01
16	Italy (IT)	Q1-1990	96	0,579	1,46	-1,107	1,26	-5,076	1,47	0,739	1,88
17	Japan (J)	Q2-1994	79	0,335	0,47	-0,269	0,26	-1,076	0,38	0,472	0,69
18	Korea (K)	Q1-1990	96	-0,412	0,89	1,769	3,11	2,237	1,41	0,568	1,08
19	Luxembourg (L)	Q2-1995	75	-1,108	1,74	2,270	2,67	3,371	1,62	-0,145	0,22
20	Mexico (M)	Q1-1990	96	-0,538	1,17	1,392	2,25	2,054	1,20	0,112	0,23
21	Netherlands (NE)	Q1-1990	96	0,495	1,89	-0,252	0,63	-2,113	1,95	0,979	3,78
22	New Zealand (NZ)	Q1-1990	96	0,006	0,02	0,655	1,30	-0,321	0,20	0,720	2,54
23	Norway (N)	Q1-1990	96	0,837	1,99	-0,555	0,93	-3,447	2,04	1,503	3,29
24	Poland (PL)	Q1-1997	68	-0,032	0,09	1,098	1,25	-0,855	0,20	1,089	2,92
25	Portugal (PT)	Q1-1996	72	-0,947	3,61	2,281	4,10	7,059	3,40	-0,637	2,16
26	Slovak Republic (SL)	Q1-1998	64	0,414	0,36	-0,193	0,06	-6,728	0,42	1,386	1,23
27	Slovenia (SV)	Q1-1996	72	-0,094	0,13	0,802	0,56	-1,478	0,31	0,848	1,15
28	Spain (ES)	Q2-1995	75	-0,889	4,71	2,885	5,68	11,490	4,32	-0,338	1,61
29	Sweden (S)	Q2-1993	83	0,811	1,36	-0,980	0,81	-5,853	1,34	1,448	2,42
30	Switzerland (CH)	Q1-2010	16	1,928	2,04	-2,328	1,70	-8,013	2,00	2,354	2,49
31	Turkey (TQ)	Q1-2005	36	-0,537	0,18	2,300	0,35	5,354	0,18	0,522	0,16
32	United Kingdom (UK)	Q1-1990	96	-0,256	0,84	0,998	1,72	2,131	1,06	0,184	0,58
33	United States (US)	Q1-1990	96	-0,448	1,75	1,415	3,08	2,787	1,81	0,133	0,48

Source: own computations

Annex 3. Map of "strategic groups" of OECD countries according to the slope of the linear theoretical line between the GDP growth rate and HUR and HUR variation areas

∂[r(GDP)]/∂[HUR)]	-0,35 – -0,33	-0,18 – -0,10	-0,07 – -0,03	0,00 - 0,06	0,10 - 0,18	0,25 - 0,46
-18,3			ES			
-13,4 - -11,4		IR PT				
-6,2 – -4,2	IC L	H M	US[a)] IS SV TQ K UK			
1,5 - 2,6					J A	CH
3,6 - 5,7				CA CHI	NE S F SF B N	CZ
6,2 - 7,7				NZ AU DE G IT		
10,6				SL		
13,4 - 14,1				PL EST		

Range HUR

Source: own computations

Legend: a) The yearly slope ∂[r(GDP)]/∂[HUR)] ≈ 0.25 p. p. is evidently lower than estimated by M. F. Prachovny and even by A. B. Abel and B. S. Bernanke.

Annex 4. Map of "strategic groups" of OECD countries according to the range of variation of unemployment and the average (unadjusted and adjusted[a)]) levels of GDP growth rates

$r^{śr}$(GDP)	0,17	0,29 - 0,41	0,47 - 0,68	0,77 - 0,92	0,98 - 1,12	1,28
$r^{śr'}$ (GDP)	0,19	0,29 - 0,44	0,48 - 0,73	0,78 - 1,00	1,23 - 1,25	1,41 - 1,44
-18,3			ES			
-13,4 - -11,4		PT	IR			
-6,2 - -4,2			SV US UK H IC M	IS' TQ' M' L	IS TQ	K
1,5 - 2,6		J	CH A			
3,6 - 5,7		B F SF	N NE S CZ CA		CHI	
6,2 - 7,7	IT	G DE	NZ	AU		
10,6				SL	SL'	
13,4 - 14,1				PL'	PL EST	EST'

Range HUR

Source: own computations

Legend: a) countries classified according to adjusted index to other group than according to the unadjusted one are marked by the sign '.

Annex 5. Map of "strategic groups" of OECD countries according to the levels of $r^{śr}$(GDP) and $r^{śr}$'(GDP) and $HUR^{śr}$

$r^{śr}$(GDP)	0,17	0,29 - 0,41	0,47 - 0,68	0,77 - 0,92	0,98 - 1,12	1,28
$r'^{śr}$(GDP)	0,19	0,29 - 0,44	0,48 - 0,73	0,78 - 1,00	1,23 - 1,25	1,41 - 1,44
3,44						K
3,84 - 4,67		J	CH A NE IC N M	L M'		
5,98 - 6,91		DE	UK CZ US SV NZ	AU		
7,64 - 8,17		G B	H IR CA S		IS CHI	
8,70		PT				
9,24 - 10,27	IT	F SF		TQ'	EST TQ	EST'
13,19 - 15,33			ES	SL PL'	PL SL'	

$HUR^{śr}$

Source: own computations

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 64 – 73

APPLICATION OF REGRESSION MODELS IN ANALYSIS OF THE ADVERTISEMENT'S IMPACT ON CONSUMER'S DECISION MAKING PROCESS

Mariola Chrzanowska[1], Joanna Chudzian[2]
[1]Department of Econometrics and Statistics
[2] Department of European Policy, Public Finance and Marketing
Warsaw University of Life Sciences – SGGW
e-mail: mariola_chrzanowska@sggw.pl; joanna_chudzian@sggw.pl

Abstract Consumer's behaviour in the market is a widely studied and analysed problem. Complexity of social, economic and psychological determinants that influence consumer's decision process is a reason for multilevel and multi-factor approaches to analyse this problem. Therefore the aim of this paper is to describe application of parametric regression model for the effectiveness of advertising. The study described is based on a survey covering 550 consumers of dairy product, all of age over 15 and living in one of the nine biggest Polish agglomerations. Built models were examined and verified statistically. Obtained results clearly show that the approach chosen to describe AIDA model is an appropriate method for analysing impact of advertisement on consumer's decision making process.

Keywords: regression, consumer's decision making process, advertising, AIDA model

INTRODUCTION

Advertisement is the basic form of the company's communication with the market. It is the information connected with the persuasive message. Advertisement not only informs the consumers about the goods and services, but it also aims to persuade them to purchase the advertised goods or to use the specified services [Kotler 2005]. The impact of advertisement on the consumers' purchase decisions is complex and difficult to describe in full. What is particularly difficult is the analysis of the impact of the consumers' decision-making process on the effectiveness of advertising. Literature research [Świątkowska, Berger 2001,

Szwacka-Salmonowicz 2003] confirms that this issue is poorly researched in relation to individual product groups. No indication of the direct relevance of the shape of the process in advertisement creation invokes its limited use in practice. This is due to the descriptive nature of the decision-making process, and thus the difficulty in identifying its relationship with the effectiveness of the advertisement [Maison, 2007].

Studies on the influence of advertisement on the decision-making process resulted in many models. One of them is the AIDA model developed by L. Strong. This model assumes the hierarchical and cause and effect impact stages of the advertisement on a consumer. The model assumes that in each subsequent step the number of recipients of advertisement is reduced [Barry, Howard 1990]. Widely known in the theory, the model is not practical , hence rarely in use. Selected works [Gharibi et all, 2012] show that AIDA model is useful in the effectiveness of advertising research[1].

During the decision-making process it is possible to distinguish the following stages [Engel, Blackwell, Kollat 1968]: problem recognition, information search, evaluation of alternatives, decision implementation, resigning from the purchase, post-purchase processes and post-purchase evaluation. Some authors limit the role of the advertisement only to the stage of making people aware of the need and raising interest [Mazurek-Łopacińska 2003]. However, most of the respondents agree on the existence of a different meaning and effect of the advertisement in particular stages of the decision-making process [Łodziana–Grabowska 1996, Jachnis 2007]. It seems that the appropriately selected set of features and elements of the advertisement may have an impact both at the time of the formation of the need to purchase and during the consideration of available possibilities to meet this need or after the purchase. In addition, it seems that the effectiveness of the advertisement message is different depending on the current consumer's stage of making the purchase decision.

From the advertiser’s point of view, the advertisement should be directly addressed to the current consumer's stage. In practice, it is not always feasible. It is also impossible to identify groups of consumers passing through various stages. On the other hand, the impact on all stages would require huge financial investments, and the results of the studies [Łodziana-Grabowska, 1996, Pociecha, 1996] do not clearly indicate which advertisement stages may have the greatest efficiency. The extraction of these stages would allow the advertisement impact only on these stages of the decision-making process which are important from the sender’s point of view. Such an action would permit to minimise the incurred financial expenditures, as well as to actively enhance and suppress individual stages through the marketing communication. The effect of such action would therefore be the creation of such advertisement which strengthens the stages of great importance,

[1] In marketing, the effectiveness of advertising is defined as the degree of advertisement of goal achievement [Bendixen 1993].

and weakens the stages of low importance for the advertiser. Therefore, it is reasonable to conduct research in this field.

MATERIALS AND METHODS

The aim of this study was to analyse the impact and significance of the advertisement in shaping the consumers' behaviour on the example of the dairy market. To verify that goal the dairy products are divided into two categories: traditional products (e.g. butter, milk well known to the market) and new generation products (e.g. dessert or dairy drinks, new to the market), In the next step of the research, consumers' reaction was also examined.

The study covered dairy products' consumers aged over 15, who declare intake of at least, one of the tested sub-category of dairy products (traditional or new generation) and have access to different forms of advertisement. The advertisement accessibility increases with the size of the city. Hence, the respondents for the study were selected in 9 Polish cities.

Results were collected using two different surveys. One of the surveys contained questions concerning the advertisement's impact described in the AIDA model. The second one included a list of statements described in the Likert scale, which determine the intensity of occurrence of the various stages of making a purchasing decision. Each stage is tested separately and qualification for the various stages were decided through the analysis of variation from the average results for the different stages of decision making process.

The data was collected by using the CAWI method and 630 questionnaires were obtained during the study. During the initial data analysis, some questionnaires were rejected on the basis of database cleanup from vague, arbitrary and contradictory information. Finally, 550 surveys. The analysis was performed using the parametric regression methods. The SPSS and R programs were used for the study[2].

In order to determine the pattern of consumers' conduct on the dairy market, the answers were averaged (obtained in the scale of 1−5) from statements attributed to particular stages. The results were considered as independent variables for each consumer's individual phases In order to determine the stages of the greatest importance for the effectiveness of the advertisement, the regression model was built. The weighed AIDA index[3] was adopted as a dependent variable. The

[2] The article uses the studies' results included in the thesis of one of the article's authors (available but not published). It is held in accordance with the Law on Copyright and Related Rights Act of 4 February 1994, published in OJ 1994 No. 24, item. 83. The contribution of the two authors of the article equals 50%.

[3] The dependent variable was defined in the models as the advertisement effectiveness. It is expressed by the weighed mean developed based on the theoretical AIDA model. The attention was operationalized by the declaration of absorbing advertisement of dairy products, the interest – as the declaration of increased interest as a result of the advertisement, the desire to purchase – as the

averaged values of the answers for subsequent stages of the decision-making process were adopted as the independent variables [Engel, Blackwell, Kollat 1968)]. The defined decision making stages are: problem recognition (I), information search (II). evaluation of alternatives (III), decision implementation (IV), resigning from the purchase (V), post-purchase processes (VI). post-purchase evaluation (VII).

RESULTS

The relationship of advertising and decision-making on the dairy market

Model 1. determining the impact of the stages of the decision-making process on the growth of the advertisement effectiveness of dairy products adopted the linear form. Details of the estimation are presented in table 1.

Table1.Results of estimation of model 1.

	Non-standardized coefficients		Standardized coefficients	*t* Statistic	*p* value
	B	Standard error	*Beta*		
Constans	-0,046	0,043		-1,071	0,285
Stage II Information Search	0,166	0,018	0,405	9,374	0,000
Stage III Evaluation of Alternatives	-0,034	0,008	-0,170	-4,453	0,000
Stage IV Decision Implementation	0,036	0,014	0,109	2,523	0,012

Source: own calculations

Model 1. explains 22% of the variable variance AIDA (R^2= 0,224, R_s^2= 0,220). In addition, based on the analysis of both measures the validity of the reference to the model's results can be confirmed to the whole population from which the test sample came[4]. The value of the *F*-Snedecor statistics ($F_{3;546}$= 52,517)[5] which allows to reject the hypothesis of the zero value of coefficients β (H0: $\beta 1 = \beta 2 = 0$). Such result proves the correctness of the estimated model. In the further part of the analysis the correctness of the model

positive attitude towards the advertisement, and action – through the declaration of purchase under the influence of the advertisement. This way to operationalize the AIDA variable was not used so far in the literature, and it represents an innovative authors' approach. This part of the study with the detailed description was presented in the paper [Chudzian, Chrzanowska 2014].

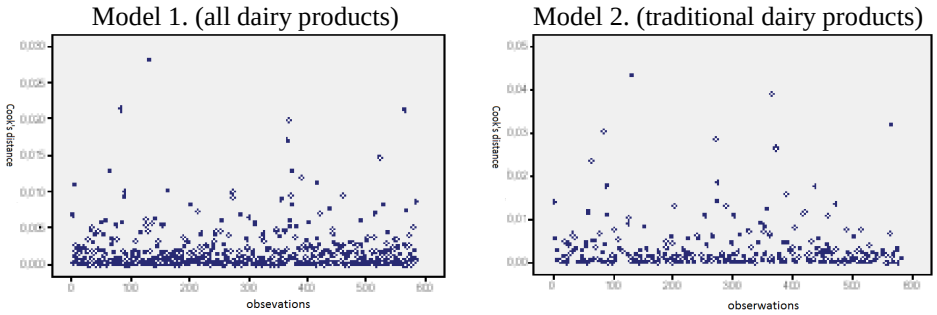
[4] The difference below 0,01 between R2 and Rs2 means the random sampling and the possibility to generalise the model results on the population (Field, 2005).

5 *p* value < 2,2e-16.

was verified through the additional analysis of residues. Detailed results are given in figure 1. All regression coefficients with the independent variables in the presented model are statistically significant. The full analysis of variance model is presented in table 2.

From the advertisement effectiveness point of view, there appeared three major purchase processes in the constructed model (Table 1.). The most important of them, which has the highest standardised regression coefficient Beta, is the stage of information search about the product. Another important stage refers to the circumstances of the purchase decision. Both of them received the positive coefficient, which means that the increase in the occurrence of these stages in consumers' behaviour which causes the advertisement effectiveness increase. Thus, the impact and formation of the advertisement composition which strengthens the occurrence of these stages causes greater effectiveness of the advertisement. In turn, the occurrence of the stage of choice of alternatives weakens the effectiveness of the advertisement (negative standardised Beta coefficient), so the sender's goal should be to persuade the consumers to resign from product comparison.

Figure 1. Residual analysis (Cook's distance)



Source: own elaboration

Table 2. Analysis of variance model 1 for dairy products

	Sum of squares	*df*	mean square	*F* Statistics	*p* value
Regression	7,161	3	2,387	52,517	0,000
Rest	24,815	547	0,045		
Total	31,975	550			

Source: own calculations

Comparing the obtained results to the actual nature of the decision-making process on the dairy product market, it can be stated that it seems reasonable to lower the stage of the assessment of alternatives, which is actually observed for the given market. It seems that due to the stated high degree of the post-purchase evaluation, highlighting in the advertisement differences between the competing products may bring about unnecessary

consumer's interest in other available brands. In turn, the absence of stages: information search and the purchase decisions, prevents the active acquisition of new clients through their attainment, among others, to press advertisement, Internet advertisement or promotions in the sale locations. Supporting in the advertisement communication the active search for information thus causes the increase of advertisement effectiveness, that is the increase in purchase. What is also interesting is the comparison of values of the standardized Beta coefficients, which shows that the strongest effect (in terms of the absolute value) is found in the stage of information search, and the weakest – the decision to purchase. This information may suggest the possible division of means in the advertisement budget.

It seems that the impact on other stages, already existing in the decision-making process (problem recognition and post-purchase evaluation), is irrelevant from the point of view of the effect on the size of their purchases. This would suggest, therefore, that these stages occur on their own and investing funds in their support is not advisable.

The relationship of advertising and decision-making process in the market of traditional products

The linear regression model specifying the impact of individual stages of the decision-making process on the increase of the effectiveness of the traditional dairy product advertisement (model 2.) also adopted the linear form. Detailed information is provided in table 3.

Table 3. Results of estimation of model 2.

	Non-standardized coefficients		Standardized coefficients	*t* Statistic	*p* value
	β	Standard error	*Beta*		
const	-0,067	0,059		-1,131	0,259
Stage II Information Search	0,207	0,023	0,480	8,856	0,000
Stage III Evaluation of Alternatives	-0,031	0,011	-0,149	-2,744	0,006

Source: own calculations

Model 2 explains more than 22% of the variance of the variable AIDA (R^2 = 0.227, R_s^2 = 0.221). The additional diagnosis confirming the correctness of the model was shown in figure 1[6]. Regression coefficients with independent variables are statistically significant. The full analysis of variance model is presented in Table 4. The value of F-Snedecor statistics ($F_{2, 272}$ = 39,922) allows to reject the hypothesis of zero *Beta* coefficients with significance *p* value = 6.28e-16.

[6] For more details about outliers see Cook (1982).

Table 4. Analysis of variance model 2 for traditional products

	Sum of squares	*df*	mean square	*F* Statistics	*p* value
Regression	3,605	2	1,803	39,922	0,000
Rest	12,281	272	0,045		
Total	15,886	274			

Source: own calculations

In case of traditional products a very simplified decision-making process was observed, limited to the feeling of the need, occurrence of post-purchase processes and post-purchase evaluation. In comparison with the results obtained for this model, it can be concluded that the possible effects on the purchase by the structure of the advertisement based on the decision-making process are also simplified. They are limited to the advertisement actions supporting the active search for information about the product and not admitting the assessment of the competing products. Post-purchase evaluation of traditional products is in fact very common, so the possible change of the brand may cause a permanent change of buying habits.

Table 5. Results of estimation of model 3.

	Non-standardized coefficients		Standardized coefficients	*t* Statistic	*p* value
	β	Standard error	*Beta*		
const	-0,161	0,087		-1,844	0,066
Stage II Information Search	0,135	0,025	0,345	5,418	0,000
Stage III Evaluation of Alternatives	-0,037	0,011	-0,185	-3,443	0,001
Stage VI Post-purchase processes	0,044	0,018	0,134	2,440	0,015
Stage IV Decision Implementation	0,044	0,021	0,130	2,098	0,037

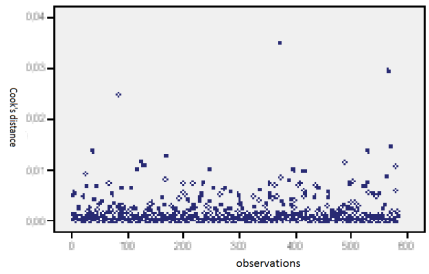
Source: own calculations

The relationship of advertising and decision-making process in the product market of new generation

The linear regression model (model 3.) presents the dependency between individual stages of the decision-making process and the increase of the advertisement effectiveness of the new generation products. The results of the estimation are shown in Table 5.

Model 3. explains more than 23% of the variable variance AIDA (R^2 = 0,234, R_s^2 = 0,223). The regression coefficients with all independent variables in the presented model are statistically important on the level of $p < 0,05$. With the additional diagnostics (figure 2), it allows to confirm the correctness of the model. The full analysis of the variance is shown in table 6. The value of the *F*-Snedecor statistics ($F_{4;270}$ = 20,664) means the ability to reject the hypothesis about the zero value of the Beta coefficients (H_0: $\beta_1 = \beta_2 = 0$) with the significance p = 7,17e-15.

Figure 2. Residual analysis for model 3.



Source: own elaboration

Table 6. Analysis of variance of model 3 for the new generation products

	Sum of squares	*df*	mean square	*F* Statistics	*p* value
Regression	3,763	4	0,941	20,664	0,000
Rest	12,290	270	0,046		
Total	16,053	274			

Source: own calculations

Compared with model 1. (for all dairy products) the model limited to the new generation products (model 3.) is enriched with the stage of post-purchase processes. This stage is important from the advertisement effectiveness point of view. Based on the analysis of the model's coefficients it can be concluded that the relationship between searching for information, decision to purchase and post-purchase evaluation is positive, and the impacts on these stages are caused by the increase of the advertisement effectiveness. As the occurrence of choice of alternatives phase weakens the advertisement effectiveness (negative value of the coefficient with the variable), the aim of the advertisement's sender should be to urge the consumers to resign from this stage.

In case of new generation products, the decision-making process is more elongated than in case of traditional products (in model 3. it has an additional phase of the post-purchase evaluation). The results for this stage indicate that there is a need to communicate in the advertisement the post-purchase evaluation of the same brand, as the loyalty level on the new generation product market is lower than in

case of other dairy products. In addition, this stage is less stable – its level depends on the marketing communication.

Active actions encouraging consumers to reject the conscious analysis of brands and post-purchase evaluation may be based on the consumer's emotional attachment to the given brand. It is important to evoke in the advertisement the positive attitude and close relation of the consumer to the advertised product while using the recommended formation of the mood through music. In comparison with the model recommendation towards supporting the search for information and the purchase decision, it is worth noting that the nature of advertisement actions in this market should be more intense and multifaceted. This is made available by numerous recommended and advised advertisement measures, and perhaps also the non-standard forms of promotion which were not included in the study, and which may explain the additional part of the observed variation of the AIDA rate. As new generation products are more diverse, and as it results from the previously presented results, consumers approach their purchase more emotionally.

CONCLUSIONS

The need to differentiate the advertisement actions according to the product category in the dairy market was indirectly confirmed in the regression models. In case of traditional dairy products, it seems effective to conduct such advertising activities which encourage the consumer to actively search for information about the product, and at the same time weaken his or her need to compare the advertised product to others available on the market. Such actions appear to be more important than making the consumers aware of the need to purchase products or encouraging them to post-purchase evaluation. In addition to supporting consumers in their search for information and discouraging the analysis of the brands available on the market, it seems reasonable to strengthen the already existing loyalty post-purchase evaluation and inducing in the consumer the emotional attachment to the brand. Such activities allow to increase the effectiveness of advertising actions in case of products such as yoghurts or dairy desserts.

REFERENCES:

Barry T. E., Howard D. J. (1990) Review and Critique of the Hierarchy of Effects in Advertising," International Journal of Advertising 9(2), 1990, pp. 121–135

Bendixen M.T (1993). Advertising Effect and effectiveness. European Journal Marketing, vol.127 No 90.

Chudzian J., Chrzanowska M. (2014) Parametric and non-parametric regression methods in identifying impact of advertising on dairy market Econometrics 3(45)

Cook, R. Dennis; and Weisberg, Sanford (1982); Residuals and influence in regression, New York, NY: Chapman & Hal

Engel, J., Blackwell, R., Kollat, D. (1968). Consumer behavior. New York: Holt, Reinehart and Winston.

Field, A. (2005) Discovering Statistics Using IBM SPSS Statistics, SAGE Publications Ltd.

Gharibi S., Yahyah S., Daneshb S., Shahrodi K. (2012) Explain the effectiveness of advertising using the AIDA model. Interdisciplinary Journal of Contemporary Research in Business, Vol 4., No 2, p. 926-940.

Jachnis, A. (2007). Psychologia konsumenta. Psychologiczne i socjologiczne uwarunkowania zachowań konsumenckich. Bydgoszcz-Warszawa: Oficyna Wyd. Bratna.

Kotler, P. (2005) Marketing management , wyd. 11. Prentice Hall, Harlow.

Łodziana-Grabowska, J. (1996) Efektywność reklamy. Warszawa: PWE.

Maison, D. (2007) O badaniach reklamy, czyli jak na podstawie badań przewidzieć skuteczność reklamy i ocenić skuteczność przeprowadzonej kampanii. W: D. Maison, A. Noga-Bogomilski. (red.). Badania marketingowe. Od teorii do praktyki Gdańsk: GWP s.153–177.

Mazurek-Łopacińska J. [2003] Zachowania nabywców i ich konsekwencje marketingowe, Warszawa PWE.

Pociecha, J. (1996) Metody statystyczne w badaniach marketingowych. PWN, Warszawa.

Świątkowska, M., Berger S. (2001). Reklama żywności jako narzędzie kształtowania postaw konsumentów na rynku produktów spożywczych. Warszawa: Wyd. SGGW.

Szwacka–Salmonowicz, J. (2003). Zmiany zachowań nabywców jako determinanta kształtowania strategii segmentacyjnych przedsiębiorstw przemysłu spożywczego w Polsce, Warszawa: Wyd. SGGW.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 74 – 83

ECONOMIES OF SCALES IN EU HOUSEHOLD CONSUMPTION – SOME REMARKS BASED ON A COUNTRY-LEVEL ANALYSIS

Hanna Dudek
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
e-mail: hanna_dudek@sggw.pl

Abstract: Economies of scale in household consumption generally occur as a result of joint consumption of public goods. In order to analyze this phenomenon expenditure shares on housing, which can be treated as a representative of the public good, and expenditure shares on food - representing private goods - are examined. The data used in this study come from the Eurostat database and cover the period between 2004 and 2012. Estimation of panel data models reveals that a large drop in food shares in post-communist countries was mainly due to rising household incomes. It is also found that an increase in housing shares was affected by the rising price of housing relative to other consumer prices in the EU-countries. Reducing differences in the considered components of expenditure structures make use in EU common equivalence scale in 2012 more reasonable than in 2004.

Keywords: economies of scales, households, the EU countries, panel data,

INTRODUCTION

People live in households of different size and composition. Taking into account their consumption behaviour one can state, that in order to attain the same standard of living two individuals living together generally require less money than two individuals living single. The reduction is not proportional because they benefit from economies of scales due to the jointness of consumption. For example, sharing the accommodation, heating and so on and so forth generates economies of scale. In order to take into account economies of scales in comparison of well-being across households of different sizes and compositions the so-called equivalence scales are usually applied. They may be interpreted as parameters informing how much money more/less a household of a given type needs to reach

the well-being of a household of another type [Szulc 2004]. A wide range of equivalence scales exist, many of which are described in [Buhmann et al. 1988; Schröder 2004; Dudek 2011]. The choice of a particular equivalence scale depends on assumptions about economies of scale in consumption as well as on assessments of the needs of different individuals such as children and adults. To the most commonly used scales belong the so called OECD scales. The original OECD scale (also called 70/50 scale or 'Oxford scale') was recommended in the 1980s for possible use in countries which had not established their own equivalence scale. This assigns a weight of 1 to the first household member, of 0.7 to each additional adult and of 0.5 to each child. The Statistical Office of the European Union (Eurostat) adopted in the late 1990s the so-called 'OECD-modified' equivalence scale. This scale[1] (also called 50/30 scale) states that the first adult should be assigned a weight of one, subsequent adults are assigned a weight of 0.5 and children 0.3. It should be mentioned that with the accession of 10 new countries to the European Union in 2004 some experts reported an inadequacy of using a common scale for all Member States [Dennis, Guio 2004; Szulc 2004]. Reported concerns were related to differences in the country-specific structures of consumption expenditure.

Households consume a variety of goods, which can be broadly classified as public and private goods. Economies of scale are generated by the presence of household public goods [Perali 2003]. Such goods can be consumed jointly by several individuals within the household where the satisfaction derived by one person does not reduce that obtained by another[2]. In the opposite to a public good, a private good is defined if it cannot be shared or consumed jointly by more than one person [Dunbar, Lewbel, Pendakur 2013]. If all goods are private, cost of living rises in proportion to the number of people in the household, while if all goods are public, such costs are unaffected by the size of households. This arguments support the intuitive notion that, in very poor economies with a high share of the budget devoted to food (which is almost entirely private) the scope for economies of scale is likely to be small. In other settings where housing (which has a large public component) is important, economies of scale are likely to be larger [Deaton, Zaidi 2002].

The economies of scales depend on the proportion of public versus private goods in the household. This proportion can vary over time and across countries. Therefore, there is a need for empirical researches about this phenomenon. They allow for a deeper insight into the issue of economies of scale in the EU household consumption.

[1] OECD-modified equivalence scale was first proposed by Haagenars et al. [1994].

[2] In reality many goods are partly shared, e.g., an automobile may be used by a single household member part of the time, and by multiple members at other times [Dunbar, Lewbel, Pendakur 2013].

In the paper country-level analysis using Eurostat database is undertaken. In order to study such a complex phenomenon as economies of scale the expenditure shares on food (representing private goods) and the expenditure shares on housing (representing public goods) are considered. To the potential determinants of households behaviour belong: incomes, demographic characteristics and price indices [Deaton, Muellbauer 1980; Rusnak 2007]. The objective of this study is an empirical verification of influence of these factors on households expenditure shares on food and housing.

DATA

The data for the analysis are taken from the European Union statistics office – Eurostat. They are the shares of expenditure on different goods in the household final consumption expenditure. In Eurostat database household consumption expenditures are broken down into twelve main categories by a system known as COICOP[3] classification. In the paper we focus on the percentage of total spending that households in each Member State dedicated to each of two items. The first one relates to food and non-alcoholic beverages and the second one – to housing, water, electricity, gas and other fuels. In short, we call this ratio 'the food share' and 'the housing share' respectively[4]. For the analysis the panel data analysis is used. Such data refer to data containing time series observations of a number of individuals. In our analysis the time span covered is from 2004 to 2012 and the study encompasses the 27 EU Member States.

The analysis includes the following Member States: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, the United Kingdom, Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, Slovenia, Bulgaria and Romania. The first 15 of the above countries (shortly named EU-15) formed the EU before 2004, the subsequent 12 countries mainly from Central and Eastern Europe (CEEC) joined in 2004 and in 2007.

As potential determinants of the expenditure shares following variables are taken into account:

- median of equivalised[5] net income in PPS (Purchasing Power Standard),
- average household size,
- annual average price indices with the base period in 2004: index for food and non-alcoholic beverages, index for housing, water, electricity, gas and other fuels; index for all items HICP (Harmonised Index of Consumer Prices).

[3] COICOP stands for Classification of Individual Consumption by Purpose.

[4] Both shares relate to the spending at current prices as % of total household consumption expenditure.

[5] 50/30 equivalence scale was applied.

There are some gaps in the Eurostat dataset, especially concerning Bulgaria, Greece, Lithuania and Romania. In such cases missing values were interpolated using linear trend.

METHODS

Using panel data on the EU countries we specify the following equation:

$$y_{it} = \alpha + \mathbf{x}_{it}\boldsymbol{\beta} + u_i + \varepsilon_{it} \quad (1)$$

where y_{it} is dependent variable,
$\mathbf{x}_{it}$ – row vector of the time-varying explanatory variables,
i indexes country and t indexes year,
α – intercept,
$\boldsymbol{\beta}$ – column vector of slope parameters,
u_i – is an individual country-specific effect[6], $u_i \sim IID(0, \sigma_u)$, $i=1, 2,\ldots,N$,
ε_{it} – idiosyncratic error term[7], $\varepsilon_{it} \sim IID(0, \sigma_\varepsilon)$, $i=1, 2,\ldots,N$, $t=1, 2,\ldots,T$.

The idiosyncratic error term ε_{it} is assumed to be uncorrelated with the explanatory variables and with the individual country-specific effect. The assumptions about u_i help to determine what kind of panel model should be estimated. In the absence of the individual effect pooled OLS estimator can be applied.

The fundamental distinction is between models random and fixed effects (shortly named RE and FE respectively). The key consideration in choosing between these approaches is whether u_i and $\mathbf{x}_{it}$ are uncorrelated which is an assumption of the RE model [Wooldridge 2002]. To test this assumption Hausman specification test is applied. Since FE is consistent when u_i and $\mathbf{x}_{it}$ are correlated, but RE is inconsistent, a statistically significant difference is interpreted as evidence against the random effects assumption. If the null hypothesis that the individual effects are uncorrelated with the other regressors is rejected, a fixed effect model is usually favoured by applied researchers over its random counterpart[8].

[6] We can think of u_i as representing the effects of all the time invariant variables that have not been included in the model.

[7] In panel data regression the error term $w_{it} = u_i + \varepsilon_{it}$ consists of two components: an 'unobserved heterogeneity' component u_i and an 'idiosyncratic' component ε_{it} denoting the remainder disturbance [Baltagi 2005].

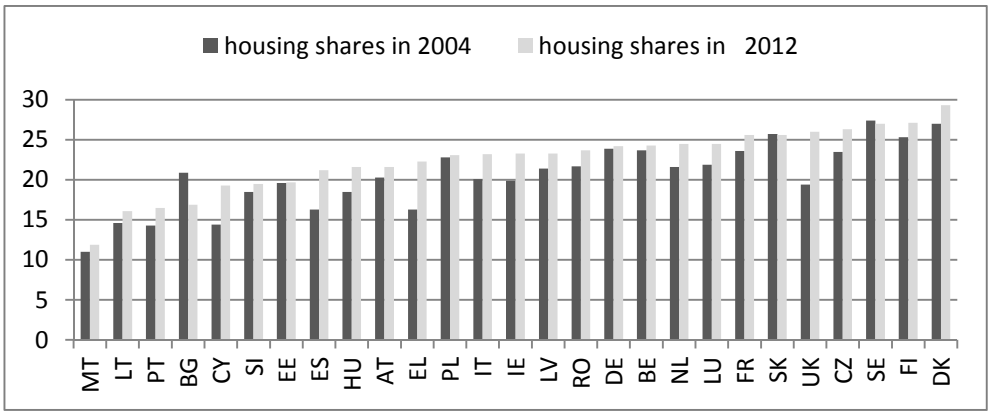
[8] It should be noted that the differences in the estimates of fixed effects and random effects models in finite samples can originate from different sources, therefore results of Hausman test should be interpreted with caution. For example, FE estimator may also be inconsistent due dependence of time-varying explanatory variables and idiosyncratic error term (see for example [Ahn, Low 1996]).

Fixed effects are compared with pooled OLS regression by the F test, while random effects are examined by Breusch and Pagan test (LM test). If the null hypothesis is not rejected in either test[9], the pooled OLS regression is favoured.

RESULTS

There were great disparities in patterns of households' expenditures across countries and time. Housing, water, electricity, gas, other fuels were amongst the most important consumption items for the majority of EU households in years 2004-2012. In many countries shares of the budget devoted to them exceed 20% and in almost all countries they increased during the period in question, as Figure 1 shows.

Figure 1. Housing shares in 2004 and 2012 (% of total expenditures)



Source*:* own elaboration based on Eurostat data

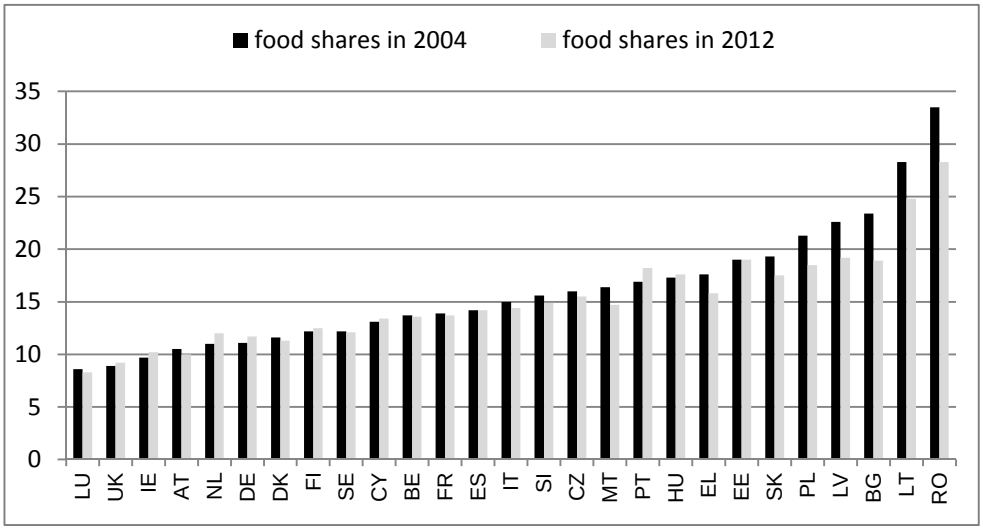
Country codes: AT: Austria; BE: Belgium; BG: Bulgaria; CZ: the Czech Republic; DK: Denmark; DE: Germany; EE: Estonia; IE: Ireland; EL: Greece; ES: Spain; FR: France; IT: Italy; CY: Cyprus; LV: Latvia; LT: Lithuania; LU: Luxembourg; HU: Hungary; MT: Malta; NL: the Netherlands; PL: Poland; PT: Portugal; RO: Romania; SI: Slovenia; SK: Slovakia; FI: Finland; SE: Sweden; UK: the United Kingdom.

Countries with high shares of expenditures on housing were not only affluent countries of Western and Northern Europe but also certain post-communist countries such as the Czech Republic, Slovakia, Poland, Romania and Latvia. At the opposite pole, especially at the beginning of the period in question, were mainly countries of Southern Europe such as Malta, Portugal, Cyprus, Spain, Greece and Slovenia.

[9] The null hyphothesis in Breusch-Pagan test is that variances of individual effects are zero.

Other significant components of expenditure were food and non-alcoholic beverages. There were significant differences across countries and time with respect to them. For example, in 2004 the average households in Romania devoted more than 30% of total consumption to food, while in the most affluent countries, such as Luxembourg and the United Kingdom, the average food shares did not exceed 10% in 2012. This is illustrated in Figure 2.

Figure 2. Food shares in 2004 and 2012 (% of total expenditures)



Source*:* own elaboration based on Eurostat data

Figure 2 highlights a few important facts. It should be noted that the countries of Central and Eastern Europe generally spent a higher proportion on food and non-alcoholic beverages than the 'old' EU Member States. During the period 2004-2012 the decrease in the share of food expenditures could be observed in almost all EU countries. Among the EU-15 countries, these changes are minor, while in the CEEC – more dynamic. Turning to an analysis of country-specific data, one can observe that the southern countries such as Portugal, Spain, Greece and Italy exhibited higher food shares then other EU-15 countries.

The drop of the expenditure shares of food (representing private goods) and growth of the expenditure shares of housing (representing public goods) indicate increasing economies of scale in household consumption. Moreover, as it was proved in [Dudek 2014], the growth rates of these two expenditure shares are inversely correlated with their appropriate initial levels. It means that during the period 2004–2012 the 'catching up' process in the field of economy scale took place. It was also found that standard deviations of food shares and housing shares decreased during the period in question (from 5.9 percent points in 2004 to 4.6 in

2012 and from 4.1 percent points in 2004 to 3.9 in 2012, respectively). This means a reduction of differences in the behaviour of household consumption. The above results give rise to the conclusion that the use of the common equivalence scale in the EU was more reasonable in 2012 than in 2004.

The main objective of this study is identification of determinants of households expenditure shares on food and housing. Many models with different sets of explanatory variables are estimated[10]. Selection of the final models is based on information criteria. Table 1 shows the panel estimation results for the countries analyzed. The results of the tests are also reported. In the second column we indicate the results using as dependent variable the food share; the third column refers to the housing share as dependent variable.

Table 1. Results of estimation of panel data models (1)

Variables and statistics	Food shares	Housing shares
Log of income*d_1	-2.70 (0.33) ***	-
Ratio of price indices for food and for all items	3.54 (1.67) **	-
Ratio of price indices for housing and for all items	-	10.26 (0.84)***
Constant	13.41 (1.62)***	10.04 (0.75)***
R^2	0.98	0.95
F test	397.91***	164.82***
LM test (Breusch and Pagan test)	819.76 ***	787.78 ***
Hausman test	45.46 ***	5.46 **

Source: own calculations.

Standard errors in parentheses. * indicates statistical significance at 0.1, ** at 0.05, and *** at 0.01, d_1 is a dummy variable is that equals 1 if for CEEC and 0 otherwise.

We first test whether or not the panel estimation is preferred to the pooled OLS estimation. Usual F and LM tests confirm that a specification considering individual effects is more appropriate. As can be seen from Table 1, the FE model is probably a better specification than RE, since the Hausman test is significant. As a consequence of this, in table 1 only fixed effects model estimates are presented.

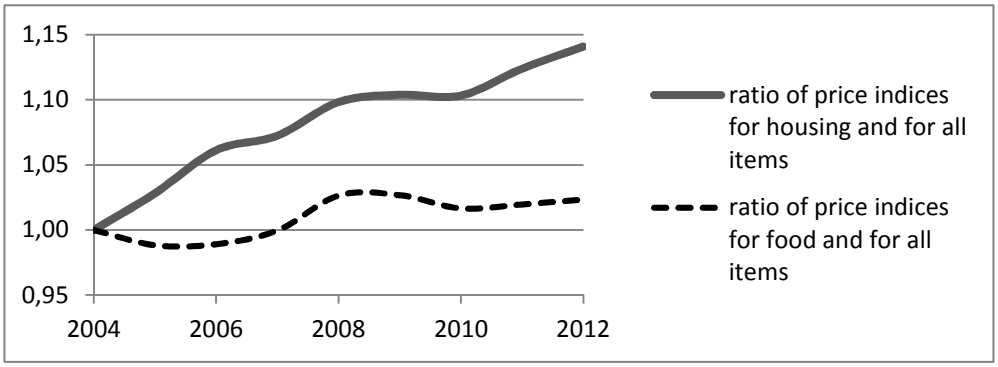
Incomes are found to have a significantly negative effect on food shares only for Central and Eastern Europe countries. It must be emphasized that in almost all EU countries a huge growth of households' equivalised incomes was observed in time in question[11]. However, increment of incomes in the EU-15 did not have a significant impact on food shares. Therefore, Engel law, according to which an increased income leads to a reduced proportion of income spent on food, is

[10]It was found that the basic demographic characteristics - household size - was almost time-invariant variable, and thus it was excluded.

[11] This situation also applies to the Baltic countries, where after 2009 there was drop such incomes.

confirmed only for the poorer countries of the EU, which were the CEEC. Since the logarithm of income is included as an explanatory variable, it denotes a decreasing marginal decline in food shares with respect to an increase in incomes in those countries. As it was expected, a ratio of price indices for food and for all consumed items has a positive effect. It should be noticed, however, that changes in this area were rather small in the years 2004-2012, as Figure 3 shows.

Figure 3. The average values of ratios of price indices in all the EU countries



Source: own elaboration based on Eurostat data

Similarly to food shares, it was found that the ratio of price indices for housing and for all consumed items has a positive and significant effect on the housing share. Except Bulgaria, this ratio was greater than unity in all periods and countries. It was increasing in 2004-2012, as it is presented in Figure 3, causing an increase in housing shares. The 'new' Member States were countries with a particularly high growth of housing prices relatively to all consumer items, except Poland and Slovakia, where the increase was moderate.

CONCLUSIONS

There were big differences across countries with respect to economies of scales in the EU household consumption in 2004-2012. In countries such as Luxembourg, the United Kingdom, Denmark, Ireland, Sweden, Finland and Germany expenditures on housing, water, electricity, gas, other fuels were twice higher than expenditures on food and non-alcoholic beverages. One can say that in the years 2004-2012 these countries were characterized by the highest economies of scales among Member States. The opposite situation was in Bulgaria, Romania, Lithuania, Malta and Portugal, where in the whole period in question average households devoted more of their budgets to food and non-alcoholic beverages than to housing, water, electricity, gas and other fuels. Such a large disparities in the ratio of spending on typical public and private goods raises doubts as to the application of the common equivalence scale for all Member States.

During the period 2004-2012 differences in expenditure shares decreased. The drop in the food shares in the CEEC was caused mainly by increase in households' incomes. Changes in these shares in most other EU-countries were rather small. It is also found that an increase in housing shares was affected by the rising price of housing relative to other consumer prices in the EU-countries.

The decrease of the expenditure shares of food and growth of the expenditure shares of housing indicate increasing economies of scale in EU household consumption.

The aim of this study is to provide an overall picture of the changes in two important components of expenditure structures. The future directions of the research will include a microeconometric analysis using individual household-level data. Such an analysis would enable a deeper insight into the issue of economies of scale in the European Union.

REFERENCES

Ahn S. C., Low S. (1996) A refolmulation of the Hausman test for regression models with pooled cross-section time-series data, Journal of Econometrics, Vol. 71, pp. 309-319.

Baltagi B. (2005) Econometric analysis of panel data, 3rd ed., John Wiley & Sons, West Sussex.

Buhmann B., Rainwater L., Schmaus G., Smeeding T.M. (1988) Equivalence scales well-being, inequality, and poverty: sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) Database", Review of Income and Wealth, Vol. 34, pp. 115-142.

Deaton A., Muellbauer J. (1980) Economics and consumer behavior, Cambridge University Press, Cambridge.

Deaton A., Zaidi S. (2002) Guidelines for constructing consumption aggregates for welfare analysis, World Bank Publication nr 14101, Washington DC.

Dennis I., Guio A.C. (2004) Monetary poverty in new Member states and candidate countries, Statistics in Focus, Vol. 12

Dudek H. (2014) Zmiany w zakresie udziałów wybranych wydatków konsumpcyjnych w UE a problem skal ekwiwalentności, Roczniki Kolegium Analiz Ekonomicznych, (in press).

Dudek H. (2011) Skale ekwiwalentności – estymacja na podstawie kompletnych modeli popytu, Wydawnictwo SGGW, Warszawa.

Dunbar G., Lewbel A., Pendakur K. (2013) Children's resources in collective households: identification, estimation, and an application to child poverty in Malawi, American Economic Review, Vol. 103(1), pp. 438-71.

Hagenaars A., de Vos K., Zaidi M.A. (1994) Poverty statistics in the late 1980s: Research based on micro-data, Office for Official Publications of the European Communities, Luxembourg.

Perali F. (2003) The behavioral and welfare analysis of consumption, Kluwer Academic Publisher, Dordrecht.

Rusnak Z. (2007) Statystyczna analiza dobrobytu ekonomicznego gospodarstw domowych, Wydawnictwo AE im. O. Langego we Wrocławiu, Wrocław.

Schröder C. (2009) The construction and estimation of equivalence scales and their uses, in: D. Slottje (ed.), Quantifying consumer preferences, contributions to economic analysis, Vol. 288, Emerald, Bingley, pp. 349-362.

Szulc A. (2004) Poverty measurement in transition countries joining the European Union: a Polish perspective, paper presented at the conference „Aligning the EU Social Inclusion Process and the Milennium Development Goals", Vilnius 26-27.04.2004.

Wooldridge J. (2002) Econometric analysis of cross section and panel data, The MIT Press, Massachusetts.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 84 –93

MODIFICATIONS OF THE MAXIMIN JOY CRITERION FOR DECISION MAKING UNDER UNCERTAINTY

Helena Gaspars-Wieloch
Department of Operation Research
Poznań University of Economics
e-mail: helena.gaspars@ue.poznan.pl

Abstract: Decision making under uncertainty (DMU) occurs when the decision maker (DM) has to choose an appropriate alternative on the basis of a set of decisions and a set of scenarios (with an unknown probability distribution). The author suggests two modifications of the maximin joy criterion (MJC) - one of the classical decision rules used in DMU by pessimists searching for an optimal pure strategy. The goal of the alterations for MJC is to accentuate more effectively the position of particular outcomes in comparison with other outcomes connected with a given scenario.

Keywords: decision making under uncertainty, optimal pure strategy, maximin joy criterion, payoff matrix

INTRODUCTION

According to the Knightian definition [Knight 1921], decision making under uncertainty (DMU), in contrast to decision making under certainty (DMC) or risk (DMR), is characterized by a situation where future factors are neither deterministic nor probabilistic at the time of the decision. Actually the decision maker (DM) has to choose the appropriate alternative (decision, strategy) on the basis of some scenarios (events, states of nature) whose probabilities are not known [Groenewald and Pretorius 2011, Neumman and Morgenstern 1944, Sikora 2008][1]. Usually the DM can describe the problem on the basis of a payoff matrix representing possible states of nature, decisions and outcomes. There are many decision rules for pure and mixed strategy searching. They differ one from another

[1] The forth category, decision making under partial information, is characterized by probabilities not known completely [Cannon and Kmietowicz 1974, Weber 1987].

with respect to DM's attitude towards risk. Additionally, some classical decision rules and a considerable number of extended decision rules take into account how particular outcomes assigned to alternatives are ordered in the payoff matrix and what the position of a given result is in comparison with other payoffs of the same state of nature. This feature is rather an advantage since it enables to obtain different rankings depending on the status of particular scenarios (dominated or not) and depending on the superiority of a given outcome to other values of the same scenario. The maximin joy criterion (MJC) is one of the classical procedures whose rankings vary after transposing the profits related to particular alternatives. In this paper we present the benefits of this rule and suggest two possible modifications whose aim is to accentuate more effectively the rank of a given outcome in comparison with other results connected with the same scenario. The remainder of this paper is organized as follows. Section DECISION MAKING UNDER UNCERTAINTY deals with the main features of DMU. Section THE MAXIMIN JOY CRITERION briefly describes the maximin joy criterion. Section POSSIBLE MODIFICATIONS OF THE MAXIMIN JOY CRITERION is devoted to the presentation and illustration of possible modifications of MJC. Conclusions are gathered in the last Section.

DECISION MAKING UNDER UNCERTAINTY

As it was mentioned in the introduction, DMU may be presented by means of a profit matrix where m is the number of mutually exclusive scenarios (S_1, …, S_m), n denotes the number of decisions (D_1, …, D_n) and a_{ij} stands for the profit connected with scenario S_i and alternative D_j. In this paper we assume that the distribution of payoffs related to a given decision is discrete. We only discuss one-criterion decision problems and focus on optimal pure strategy searching. A pure strategy is a solution assuming that the DM chooses only one decision. Meanwhile the mixed strategy allows him or her to select a weighted combination of several alternatives [Gaspars-Wieloch 2014b, Puppe and Schlag 2009, Sikora 2008].

Among classical decision (CD) rules designed for DMU one can enumerate the Wald's criterion, maximax criterion, Hurwicz's criterion, Savage's criterion, maximin joy criterion and Bayes' (Laplace's) criterion [Wald 1950, Hurwicz 1952, Savage 1961, Hayashi 2008]. The literature also offers many extensions of these methods, e.g. [Basili et al. 2008, Basili and Chateauneuf 2011, de Finetti 1974, Ellsberg 2001, Etner et al. 2012, Gaspars 2007, Gaspars-Wieloch 2012, 2013, 2014a, 2014c, 2014d, 2015, Ghirardato et al. 2004, Gilboa 2009, Gilboa and Schmeidler 1989, Marinacci 2002, Piasecki 1990, Schmeidler 1986, Tversky and Kahneman 1992], which can be named extended decision (ED) rules. The majority of them refers to the probability calculus (e.g. expected utility maximization, α-maximin expected utility, restricted Bayes/Hurwicz, cumulative prospect theory, Choquet expected utility), which is rather characteristic of DMR where the likelihood is given – let us remind that according to the Knight's definition the

uncertainty occurs when we do not know the probabilities of particular scenarios [Knight 1921, Chateauneuf et al. 2008, Domurat and Tyszka 2004].

THE MAXIMIN JOY CRITERION

The concept of the maximin joy criterion [Hayashi 2008] is extremely similar to the reasoning characteristic of the Savage's procedure which uses a regret matrix and minimizes the maximal opportunity loss, but here, instead of a regret table, a matrix of relative profits is applied.

The first step consists in computing an index for each alternative according to Equation (1) which represents the worst relative profit connected with D_j:

$$J_j = \min_i \{t_{ij}\} \qquad j = 1,...,n \tag{1}$$

$$t_{ij} = a_{ij} - \min_j \{a_{ij}\} \qquad i = 1,...,m;\ j = 1,...,n \tag{2}$$

where a_{ij} is the payoff related to decision D_j and scenario S_i. Symbol t_{ij} denotes the relative profit which is always a non-negative number.

The second step is to choose the alternative that has the highest index:

$$j^* = \arg\max_j J_j \tag{3}$$

Table 1. Payoff matrices for cases I, II and III

Case	Case I			Case II			Case III		
Scen./Dec.	D1	D2	D3	D1	D2	D3	D1	D2	D3
S1	5	1	3	-6	1	3	2	0	3
S2	2	-4	8	2	-4	2	5	1	8
S3	-6	4	-10	5	4	8	4	4	-10
S4	4	3	2	1	3	-10	-6	-4	2
S5	1	0	5	4	0	5	1	3	5

Source: created by the Author

Table 2. Relative payoff matrices for cases I, II and III

Case	Case I			Case II			Case III		
Scen./Dec.	D1	D2	D3	D1	D2	D3	D1	D2	D3
S1	4	0	2	0	7	9	2	0	3
S2	6	0	12	6	0	6	4	0	7
S3	4	14	0	1	0	4	14	14	0
S4	2	1	0	11	13	0	0	2	8
S5	1	0	5	4	0	5	0	2	4
J_j	1	0	0	0	0	0	0	0	0
Ranking	I	II	II	I	I	I	I	I	I

Source: own calculations

The MJC is a decision rule for pessimists, i.e. for people representing a risk-averse behavior. Notice that in this context we do no treat risk as a situation where the probability distribution of each parameter of the decision problem is known, but we mean the possibility that some bad circumstances will happen (losses or low outcomes). The alternative recommended by MJC may be executed only once (see one-shot decisions [Guo 2011]). Despite the fact that the steps of the Savage's rule are very similar to the steps of MJC, there are cases where rankings generated by both methods are different [Gaspars-Wieloch 2015]. The goal of MJC is to show the superiority of particular outcomes connected with a given scenario to its worst result, while within the framework of the Savage's rule the target is to demonstrate the inferiority of particular payoffs related to a state of nature to its best result.

The serious weak point of MJC concerns final rankings. That is, imagine that each decision is the worst for at least one state of nature. In such a case one can find at least one zero in each column (for each decision) of the relative profit table, which means that the MJC's index is equal to zero for each alternative. And then, all the strategies are treaded as optimal solutions (they all obtain the first rank in the ranking), which certainly does not facilitate the decision making process. The problem aforementioned is illustrated in Tables 1-2. Table 1 presents three payoff matrices for three decision problems. Note that these cases are very similar. In each example the sets of payoffs for particular alternatives are identical. Outcomes have only different positions. Table 2 demonstrates relative payoff matrices and final indices for each decision. MJC indices are almost always equal to zero, which does not allow the DM to obtain a reliable ranking. When applying MJC, the size of particular relative profits may be totally ignored if each decision is the worst for at least one event. That feature of MJC is quite alarming. Notice that this phenomenon does not occur when the Savage's rule is used. For our three cases the rankings would be as follows: $D_1(10)$, $D_2(12)$, $D_3(14)$; $D_2(6)$, $D_1(9)$, $D_3(13)$ and $D_2(7)$, $D_1(8)$, $D_3(14)$ respectively (numbers in brackets indicate the Savage's indices). Therefore, in the next section, two possible amendments are proposed.

POSSIBLE MODIFICATIONS OF THE MAXIMIN JOY CRITERION

The alterations of MJC may result from the observation that the relative profit should consider the superiority of a given outcome to all remaining outcomes of the same scenario, not merely to the worst one. If so, the modified relative profit will contain a more precise information about the position of particular results in comparison with other payoffs. The following subsections describe two approaches enabling one to include this information in the final index.

Dominance joy criterion (DJC)

In the first procedure (the dominance joy criterion), instead of computing the relative profit, the number of "dominance cases" is calculated, i.e. the number of payoffs lower than a given outcome within the framework of a scenario:

1. Calculate an index for each alternative according to Equation (4) which represents the sum of all "dominance cases" connected with D_j:

$$J_j^D = \sum_{i=1}^{m} t_{ij}^D \qquad j = 1,...,n \tag{4}$$

$$t_{ij}^D = n - \max\left\{p^i\left(a_{ij}\right)\right\} \qquad i = 1,...,m;\ j = 1,...,n \tag{5}$$

where $J^D{}_j$ denotes the dominance joy criterion, $t^D{}_{ij}$ stands for the sum of "dominance cases" related to a_{ij} and $p^i(a_{ij})$ is the position of payoff a_{ij} in the non-increasing sequence of all results connected with scenario S_i (when a_{ij} is equal to at least one other payoff concerning a given event, then it is suggested to choose the farthest position of this payoff in the sequence, Equation 5).

2. Choose the alternative that has the highest index:

$$j^* = \arg\max_j J_j^D \tag{6}$$

Cumulative maximin joy criterion (CMJC)

Now let us have a look at the second possible procedure (the cumulative maximin joy criterion), where instead of computing the number of "dominance cases", the sum of all relative profits concerning a given payoff is calculated and further steps do not differ from those used in the original MJC:

1. Calculate an index for each alternative according to Equation (7) which represents the worst cumulative relative profit (CRP) connected with D_j:

$$J_j^C = \min_i\left\{t_{ij}^C\right\} \qquad j = 1,...,n \tag{7}$$

$$t_{ij}^C = n \cdot a_{ij} - \sum_{j=1}^{n} a_{ij} \qquad i = 1,...,m;\ j = 1,...,n \tag{8}$$

where $J^C{}_j$ denotes the cumulative maximin joy criterion and $t^C{}_{ij}$ stands for the cumulative relative profit.

2. Choose the alternative that has the highest index:

$$j^* = \arg\max_j J_j^C \tag{9}$$

Results and discussion

As it can be observed, in the case of the dominance joy criterion we do not analyze differences between the lowest outcome and the remaining outcomes

belonging to the same state of nature, but we check the rank of each payoff in the non-increasing sequence of all results related to this scenario.

The modified relative payoff matrix for the cumulative maximin joy criterion also allows the DM to take into account the position of a given outcome in comparison with other outcomes of the same scenario, but in this case, just like in the original MJC, the superiority is measured by means of relative profits rather than by dominance. Nevertheless, this time the way the relative profit is calculated differs from that used in MJC, because within the framework of CMJC a given outcome is compared with all outcomes of the same scenario (not only with the lowest one) and those differences are summed up. If the CRPs are computed in such a way, condition (10) is always fulfilled. Additionally, irrespective of how payoffs related to a given decision are distributed in the matrix (see the three cases presented in Table 1), the sum of all cumulative relative payoffs concerning this decision does never change (analyze Equation 11 and Table 4). Thus, for CMJC we can not use an index being the sum of all cumulative relative profits (instead of the minimal cumulative relative profit) as a decision criterion, because such a measure will not vary depending on the position of particular outcomes in the payoff table.

$$\sum_{j=1}^{n} t_{ij}^{C} = \left(a_{i1} - a_{i1} + a_{i1} - a_{i2} + \ldots + a_{i1} - a_{in}\right) + \left(a_{i2} - a_{i1} + a_{i2} - a_{i2} + \ldots + a_{i2} - a_{in}\right) + \\ + \ldots + \left(a_{in} - a_{i1} + a_{in} - a_{i2} + \ldots + a_{in} - a_{in}\right) = 0 \qquad i = 1,\ldots,m \tag{10}$$

$$\sum_{i=1}^{m} t_{ij}^{C} = \left(n \cdot a_{1j} - \sum_{j=1}^{n} a_{1j}\right) + \left(n \cdot a_{2j} - \sum_{j=1}^{n} a_{2j}\right) + \ldots + \left(n \cdot a_{mj} - \sum_{j=1}^{n} a_{mj}\right) \\ = n\sum_{i=1}^{m} a_{ij} - \left(\sum_{j=1}^{n} a_{1j} + \sum_{j=1}^{n} a_{2j} + \ldots + \sum_{j=1}^{n} a_{mj}\right) \qquad j = 1,\ldots,n \tag{11}$$

Table 3. Dominance case matrices for cases I, II and III

Case	Case I			Case II			Case III		
Scen./Dec.	D1	D2	D3	D1	D2	D3	D1	D2	D3
S1	2	0	1	0	1	2	1	0	2
S2	1	0	2	1	0	1	1	0	2
S3	1	2	0	1	0	2	1	1	0
S4	2	1	0	1	2	0	0	1	2
S5	1	0	2	1	0	2	0	1	2
J^{D}_{j}	7	3	5	4	3	7	3	3	8
Ranking	I	III	II	II	III	I	II	II	I

Source: own calculations

Table 4. Cumulative relative payoff matrices for cases I, II and III

Case	Case I			Case II			Case III		
Scen./Dec.	D1	D2	D3	D1	D2	D3	D1	D2	D3
S1	6	-6	0	-16	5	11	1	-5	4
S2	0	-18	18	6	-12	6	1	-11	10
S3	-6	24	-18	-2	-5	7	14	14	-28
S4	3	0	-3	9	15	-24	-10	-4	14
S5	-3	-6	9	3	-9	6	-6	0	6
Sum of t^C_{ij}	0	-6	6	0	-6	6	0	-6	6
J^C_j	-6	-18	-18	-16	-12	-24	-10	-11	-28
Ranking	I	II	II	II	I	III	I	II	III

Source: own calculations

Tables 3 and 4 contain intermediate calculations and final indices for the three cases described in Table 1 according to both recommended procedures. Thanks to the suggested modifications it is possible to obtain varied indices for particular alternatives (compare the measure values in Tables 2, 3 and 4). Rankings generated by these decision rules depend on the structure of the payoff matrix and they do not have the flaw characteristic of the original MJC, i.e. the indices of particular actions are equal to different values (compare with Table 2).

At first glance DJC seems to be logic and easy in use, but it is worth emphasizing that the analysis of "dominance cases" does not reveal the size of this dominance. For instance, we see that outcome "2" dominates "-4" and "-4" dominates "-6" (case III, scenario S4), but these two examples of dominance are totally different. We would say that the first dominance is more significant than the second one. Meanwhile, when the DJC is applied, both dominances receive the same weight equal to 1, which may be a little questionable and controversial.

Therefore, if the size of dominance is a crucial factor for the DM, it is recommended to use CMJC, where "dominance cases" are computed more conscientiously. In the example aforementioned (case III, scenario S4) the superiority (measured by cumulative relative profits) for the outcome "2", "-4" and "-6" equals to (2-(-6))+(2-(-4))=14, (-4-(-6))+(-4-2)=-4 and (-6-(-4))+(-6-2)=-10 respectively. Note that the change of the decision rule (from DJC to CMJC) has an enormous impact on final rankings. For instance, in case II decision D3 has the first rank according to DJC and the last rank when CMJC is applied. In our example CMJC rankings resemble Savage's orders. The choice of the procedure (DJC or CMJC) should depend on the DM's preferences, i.e. whether he/she only analyzes the order of payoffs related to a given state of nature or whether the distance between particular results is of great importance for him/her. Note that, as a matter of fact, the first modification of MJC (DJC) can be reduced to the simplest case of stochastic dominance, i.e. the state-by-state dominance (or statewise dominance).

Table 5. Payoff matrices for cases IV and V.

Case	Case IV				Case V			
Scen./Dec.	D1	D2	D3	D4	D1	D2	D3	D4
S1	11	9	9.5	9	5	9	9.5	8
S2	8	9	7.5	8	8	6	7.5	9
S3	11	10	7.5	8	11	10	7.5	8
S4	5	6	8	8	11	9	8	8

Source: created by the Author

Let us analyze one more example (cases IV and V) where payoff matrices differ one from another only in terms of the structure of profits (Tables 5-6). This time, we present rankings for all classical rules and two new modifications of MJC. The general conclusion that can be made when comparing all results (Tables 1-6) is as follows: DJC is rather devoted to optimists (its rankings are similar to maximax rankings) since it considers only the superiority of particular payoffs to other outcomes, whereas CMJC is rather designed for pessimists (its rankings resemble Wald's rankings) because in this case that superiority is conscientiously measured.

Table 6. Rankings for cases IV and V.

Case	Case IV				Case V			
Crit. / Dec.	D1	D2	D3	D4	D1	D2	D3	D4
Wald criterion	IV	III	II	I	IV	III	II	I
Maximax criterion	I	II	III	IV	I	II	III	IV
Bayes criterion	I	II	IV	III	I	II	IV	III
Hurwicz crit. (α=0.4)	II	III	I	II	II	III	I	II
Savage criterion	II	I	III	II	III	I	II	I
Maximin joy criterion	I	I	I	I	I	I	I	I
Dominance joy criterion	I	II	III	III	I	II	IV	III
Cumulative maximin joy criterion	IV	I	III	II	III	II	II	I

Source: own calculations

CONCLUSIONS

The factor demonstrating the significance of the rank of a payoff in the set of all outcomes in a given matrix is the fact that people consider not only the value they receive, but also the value received by others [Frank 1997]. Some classical rules and many extended procedures designed for DMU take into account this aspect and for that reason they are quite often applied by DMs. In this paper we suggest two possible alterations for the maximin joy criterion. They may interest people who want to know how their outcome compares with other results, not only with the worst one (maximin joy criterion), the best one (Savage's rule) or the

reference point (prospect theory). The proposed methods focus on payoffs, but the concept of dominance or cumulative relative profits can also be used for utilities.

REFERENCES

Basili M., Chateauneuf A., Fontini F. (2008) Precautionary principle as a rule of choice with optimism on windfall gains and pessimism on catastrophic losses, Ecological Economics, 67, pp. 485–491.

Basili M., Chateauneuf A. (2011) Extreme events and entropy: A multiple quantile utility model, International Journal of Approximate Reasoning, 52, pp. 1095–1102.

Cannon C.M., Kmietowicz Z.W. (1974) Decision theory and incomplete knowledge, Journal of Management Studies 11(3), pp. 224–232.

Chateauneuf A., Cohen M., Jaffray J.-Y. (2008) Decision under uncertainty: the classical models, Documents de Travail du Centre d'Economie de la Sorbonne, Maison des Sciences Economiques, Paris.

de Finetti B. (1974) Theory of Probability, Wiley, New York.

Domurat J., Tyszka T. (2004) Historic and contemporary controversies on the concept of utility, Studies in Logic, Grammar and Rhetoric, 7(20), pp. 85–107.

Ellsberg D. (2001) Risk, ambiguity and decision, Garland Publishing, New York.

Etner J., Jeleva M., Tallon J.-M. (2012) Decision theory under ambiguity. Journal of Economic Surveys, 26(2), pp. 234–270.

Frank R.H. (1997) The frame of reference as a public good. The Economic Journal 107, pp. 1832–1847.

Gaspars H. (2007) Alokacja zasobu w warunkach niepewności: modele decyzyjne i procedury obliczeniowe. Badania operacyjne i decyzje 2007(1), pp. 5–27.

Gaspars-Wieloch H. (2012) Ograniczona skuteczność metod optymalizacyjnych w rozwiązywaniu ekonomicznych problemów decyzyjnych, Ekonomista 2012/3, pp. 303–324.

Gaspars-Wieloch H. (2013) On a decision rule supported by a forecasting stage based on the decision maker's risk aversion, In SOR'13 Proceedings, The 12th International Symposium on Operational Research in Slovenia, September 25-27, 2013, ed. L. Zadnik Stirn, J. Zerovnik, J. Povh, S. Drobne and A. Lisec, Slovenian Society INFORMATIKA, Section for Operational Research, pp. 53–59.

Gaspars-Wieloch H. (2014a) Modifications of the Hurwicz's decision rules, Central European Journal of Operations Research 22(4), pp. 779–794, doi: 10.1007/s10100-013-0302-y.

Gaspars-Wieloch H. (2014b) On a decision for mixed strategy searching under uncertainty on the basis of the coefficient of optimism, Procedia - Social and Behavioral Sciences 110(2014), pp. 923–931.

Gaspars-Wieloch H. (2014c) On a decision rule supported by a forecasting stage based on the decision maker's coefficient of optimism, Central European Journal of Operations Research, doi: 10.1007/s10100-014-0364-5**.**

Gaspars-Wieloch H. (2014d) Propozycja hybrydy reguł Hurwicza i Bayesa w podejmowaniu decyzji w warunkach niepewności, In Modelowanie preferencji

a ryzyko‘14, ed. T. Trzaskalik, Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach, 178, pp. 74–92.

Gaspars-Wieloch H. (2015) The impact of the structure of the payoff matrix on the final decision made under uncertainty, Asia Pacific Journal of Operational Research (in print).

Ghirardato P., Maccheroni F., Marinacci M. (2004) Differentiating ambiguity and ambiguity attitude, Journal of Economic Theory 118, pp. 133–173.

Gilboa I. (2009) Theory of decision under uncertainty, Cambridge University Press.

Gilboa I., Schmeidler D. (1989) Maxmin expected utility with non-unique prior, Journal of Mathematical Economics 18, pp. 141–153.

Groenewald M.E., Pretorius P.D. (2011) Comparison of decision making under uncertainty investment strategies with the money market, Journal of Financial Studies and Research, doi: 10.5171/2011.373376.

Guo P. (2011) One-shot decision theory, IEEE Transactions on Systems, Man, and Cybernetics, Part A 41(5): pp. 917–926.

Hayashi T. (2008) Regret aversion and opportunity dependence. Journal of Economic Theory 139(1), pp. 242–268.

Hurwicz L. (1952) A criterion for decision making under uncertainty, Technical Report 355, Cowles Commission.

Kahneman D., Tversky A. (1979) Prospect theory: an analysis of decision under risk. Econometrica 47, pp. 263–291.

Knight F.H. (1921) Risk, uncertainty, profit, Hart. Schaffner & Marx, Houghton Mifflin Co, Boston MA.

Marinacci M. (2002) Probabilistic sophistication and multiple priors, Econometrica 70, pp. 755–764.

Piasecki K. (1990) Decyzje i wiarygodne prognozy, Akademia Ekonomiczna w Poznaniu.

Puppe C., Schlag K. (2009) Choice under complete uncertainty when outcome spaces are state dependent, Theory and Decision 66, pp. 1–16.

Savage L. (1961) The foundations of statistics reconsidered, In Studies in Subjective Probability, Wiley, New York, pp. 173–188.

Schmeidler D. (1986) Integral representation without additivity, Proceedings of the American Mathematical Society 97, pp. 255–261.

Sikora W. (2008) Badania operacyjne, Polskie Wydawnictwo Ekonomiczne, Warszawa.

Tversky A., Kahneman D. (1992) Advances in prospect theory: cumulative representation of uncertainty, Journal of Risk and Uncertainty 5, pp. 297–323.

von Neumann J., Morgenstern O. (1944) Theory of games and economic behavior, Princeton University Press.

Wald A. (1950) Statistical decision functions, Wiley, New York.

Weber M. (1987) Decision making with incomplete information, European Journal of Operational Research 28, pp. 44–57.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 94 –101

FAMILIES OF CLASSIFIERS – APPLICATION IN DATA ENVELOPMENT ANALYSIS

Urszula Grzybowska, Marek Karwański
Department of Informatics
Warsaw University of Life Sciences – SGGW
e-mails: urszula_grzybowska@sggw.pl; marek_karwanski@sggw.pl

Abstract: Economic description of firms and companies is based on a number of indicators. The indicators are related to each other and can be considered only in a specific context. Regression models allow for such approach. Unfortunately, the problems we deal with are usually nonlinear and the choice of relevant information is very difficult. The aim of the paper is to present a method of variable selection based on random forest and gradient boosting approach and its application to companies ranking in DEA method. The results will be compared with the ordering obtained using expert supported approach for variable selection in DEA.

Keywords: random forests, gradient boosting, DEA, rating classes, variable selection, ranking, high rated portfolio

INTRODUCTION

In many economic issues it is essential for a decision maker to obtain a ranking of entities under consideration. So is the application of internal-rating based approach to estimation probabilities of default (PDs) for the bank obligors. One of the obstacles connected with PD estimation is a low number of defaults, especially in high rating grades. High rating categories might experience many years without any default, for example a part of bank assets called Low Default Portfolios (LDP). These portfolios may consist of assets of the same type, e. g. trust funds. Several methods have been proposed to estimation of PD for LDP [Dzidzevičiūtė 2012]. The only key assumption in the method is a correct ordinal rating of borrowers. Therefore we propose a method of rating which is based on efficiency measure given by Data Envelopment Analysis (DEA). We illustrate our research on an example. The example presents rating in a group of companies from the production sector traded

on Warsaw Stock Exchange (WSE). Our approach involves application of financial indicators describing financial standing of considered firms. The most important thing in DEA approach is a proper selection of indicators and their assignment to output and input. In our research we present two approaches. At first we use the set of indicators suggested by experts to obtain DEA rating. Next we apply ensemble classifiers: random forests and gradient boosting, to select indicators that influence the division into classes. We also compare the results and draw conclusions.

METHOD

In our approach we apply Data Envelopment Analysis (DEA), see for example [Cooper et al. 2006], to obtain division of companies into homogeneous groups. DEA is an Operation Research approach for evaluating the performance of a set of peer entities called Decision Making Units (DMU). DEA can be applied to a wide variety of activities. It can be used to evaluate the performance of governmental agencies, hospitals, universities, non-profit organizations, banks, firms. The method gives an efficiency rating, i. e., a score θ for each DMU and an efficiency reference set (a peer group of objects that are efficient), which is a target for the inefficient DMUs. Traditionally, the efficiency is measured as the output to input ratio. In DEA approach the output and input are linear combinations of variables describing performance of the DMU and the efficiency score is obtained by solving linear programming problems in their primal or dual form. The DMUs with the efficiency score equal to 1 are called efficient. The exception is a super-efficiency DEA where the efficiency score can be greater than 1 in input orientation [Andersen et al.1993]. An important advantage of the method is that the inputs and outputs can be measured in various units. Calculation of the efficiency can be helpful in improving productivity and performance of an inefficient DMU. We have however concentrated our efforts not on efficiency measure but on distinguishing groups of similar i. e., homogeneous DMUs.

In order to obtain division into homogeneous groups of companies, we have performed the DEA algorithm to the whole set of DMUs. The efficient units with efficiency score 1 constitute the first group – see for example [Kaczmarska 2010]. After removing all efficient units we applied DEA algorithm to the remaining set. This resulted in distinguishing the next group of units. The procedure was repeated until the number of DMUs in the remaining group was not sufficient to perform further divisions. The most important obstacle is that the results obtained with DEA refer only to the considered set of DMUs and can be neither generalized nor compared with results concerning even slightly differing sets of objects, not to mention sets of different objects. There are many various DEA models. In our calculations we have applied input-oriented BCC model. The model can be formulated in the following way:

Let us assume that we have n DMUs, denoted by DMU_o, $o = 1,2,\dots,n$. We denote by x_{ij}, i=1,2,..,m the inputs and by y_{rj}, r=1,2,…,s the outputs for $j = 1,2,\dots,n$. For each DMU_o, $o = 1,\dots,n$, described by the inputs x_{io}, $i = 1,2,\dots,m$ and outputs y_{ro}, $r = 1,2,\dots,s$, the efficiency measure θ_o is the solution of the following problem:

$\theta_0^* = min\theta_0$ subject to

$$\sum_{j=1}^{n} x_{ij}\lambda_{jo} \le \theta_o x_{io} \quad i = 1,2,\dots,m \tag{1}$$

$$\sum_{j=1}^{n} y_{rj}\lambda_{jo} \ge y_{ro} \quad r = 1,2,\dots,s \tag{2}$$

$$\sum_{j=1}^{n} \lambda_{jo} = 1,\ \lambda_{jo} \ge 0 \quad j = 1,2,\dots,n \tag{3}$$

A very important issue in DEA approach is variable selection that involves also division of variables into inputs and outputs. A variable classified as an output should have a positive correlation with efficiency while a variable classified as an input should have a negative correlation with efficiency (see [Demirova 2010]). Variable selection in DEA is usually based on expert knowledge and is subject to many discussions during scientific conferences. In our calculations we have decided to follow the choice of financial ratios suggested by experts and compare it with a selection of variables obtained with help of ensemble methods: random forests and gradient boosting [Berk 2008, Hastie et al. 2009, Koronacki et al. 2008].

Random forests were introduced in 2001 by L. Breiman as a method ofclassification [Breiman 2001]. In this approach a large number of unpruned trees is constructed with a random sample of predictors taken before each node is split. The object is classified based on a majority vote of the full set of trees [Berk 2008]. One can use random forests to rank the importance of variables in a classification problem. The importance of predictors can be measured in terms of a Gini index or by Breiman's importance measure [Breiman 2001, Berk 2008].

Random forests and gradient boosting [Berk 2008, Hastie et al. 2009, Koronacki et al. 2008] are extensions of regression trees, that is simply the partition of the space *X*, which consists of predictors of target variable y, into disjoint regions R_j. Let f be the prediction function for regression tree (sometimes simply referred to as a tree):

$$x \in R_j \Rightarrow f(x) = \hat{y}_j \tag{4}$$

Thus regression tree can be represented as

$$T(x;\Theta) = \sum_{j=1}^{J} \hat{y}_j I\left(x \in R_j\right), \tag{5}$$

where $\Theta = \left\{R_j, \hat{y}_j\right\}_{j\in\{1,\dots,J\}}$.

The idea behind random forest is to build a large collection of de-correlated trees and then to average prediction functions. Each tree was constructed based on a random selection of the predictor variables. After B such trees $\{T(x,\Theta_b)\}_{b\in\{1,\dots,B\}}$ are grown the random forest predictor is:

$$\hat{f}^{B}_{random\ forest} \Leftrightarrow \frac{1}{B}\sum_{b=1}^{B} T(x, \Theta_b) \tag{6}$$

Gradient boosting prediction function is yield by formula

$$\hat{f}_{gradient\ boosting} \Leftrightarrow T(x, \Theta_g) \tag{7}$$

where the parameters Θ_g should be found by minimizing the loss function L [Hastie et al. 2009]:

$$\widehat{\Theta} = \arg\min_{\Theta} \sum_{j=1}^{J} \sum_{x_i \in R_j} L(y_j, \hat{y}_j) \tag{8}$$

The solution can be constructed in an iterative way. At m-th iteration it is needed to find:

$$\widehat{\Theta}_m = \arg\min_{\Theta_m} \sum_{i=1}^{N} L(y_i, f_{m-1}(x_i) + T(x_i, \Theta_m)) \tag{9}$$

The above equation can be reformulated as numerical optimization task analog to steepest descent method,

$$f_m = f_{m-1} - \rho_m g_m \tag{10}$$

where ρ_m is the *step length* and g_m is gradient vector:

$$\begin{bmatrix} g_{1m} \\ \cdots \\ g_{Nm} \end{bmatrix} = \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}\right]_{f(x_i)=f_{m-1}(x_i)} \tag{11}$$

The difference between stochastic gradient boosting and an ordinary steepest descent is at the points x_i. Gradient boosting should be applied to the new points that are not represented in training set X used by optimization procedure. The simple solution is to induce a tree *f* using square error to get the tree as close as possible to the gradient vector

$$\widehat{\Theta}_m = \arg\min_{\Theta_m} \sum_{i=1}^{N} w_i \left(-g_{im} - T(x_i, \Theta)\right)^2 \tag{12}$$

In our calculations we used weights w_i derived from multinomial distribution, i.e., we use multinomial deviance as a loss function.

The relevant algorithms were implemented in R package randomForest and SAS Miner. The main advantage of random forests and gradient boosting approach is their high performance on a large set of variables. Their application for economic data does not require examining the structure of financial ratios, their interactions or correlations.

RESULTS OF THE RESEARCH

The sets of financial indicators applied in DEA by various authors differ considerably [Feruś 2006, Demirova 2010, Chodakowska et al. 2013]. In our calculations we have decided to follow the expert knowledge and choose Assets Turnover and Total Liabilities/Total Assets (Debt Ratio) as input indicators and Return on Assets (ROA), Return on Equity (ROE), Current Ratio (CR), Operating

profit margin (OPM) as output indicators. Our data for a set of 76 production companies traded on WSE with quarterly financial reports covered two years: 2011 and 2012. The results of our calculations are shown in column DEA1 of Table 2. We have distinguished 6 groups of homogeneous objects. The first group consists of the best companies. One can venture an opinion that for these companies the probability of default is very low. We were not interested in examining the ways of improving efficiency of the remaining companies but in division into groups of similar objects. We were also interested in selecting variables that determine obtained DEA classification. In order to select variables that influence division into DEA groups we have applied two ensemble methods: random forests and gradient boosting. The calculations were done both in SAS (ver. 13.2) and R (ver. 3.1.0). We have used 20 financial indicators, which were divided into four groups: profitability ratios, liquidity ratios, activity ratios and debt ratios. The results are shown in Table 1.

Table 1. Variables importance in various ensemble methods

	R-CRAN randomForest		SAS Miner Random forests		SAS Miner Gradient boosting	
	Ratio	Variable importance	Ratio	Gini coefficient	Ratio	Variable importance
1	RC	3.91	ROA	0.043	RC	1
2	ROA	3.39	GPM	0.025	EBIT	0.983
3	RT	2.96	RC	0.022	ROA	0.935
4	EBIT	2.64	DSR	0.018	GPMoS	0.772
5	GPM	2.47	OPM	0.016	QR1	0.696
6	DR	2.39	NPM	0.015	DSR	0.691
7	GPMoS	2.07	CR	0.012	ROE	0.678
8	QR1	2.07	EBIT	0.009	WC	0.625
9	OPM	2.04	DR	0.009	DR	0.616
10	NPM	2.02	QR2	0.008	GPM	0.613
11	DSR	1.98	ROE	0.008	AR	0.580
12	QR2	1.95	QR1	0.008	IT	0.565
13	ROE	1.72	AR	0.005	CR	0.539
14	AR	1.56	RT	0.003	OPM	0.533
15	CR	1.48	OC	0.003	NPM	0.498
16	CCC	1.44	WC	0.002	CCC	0.491
17	RA	1.37	CCC	0.002	RT	0.474
18	IT	1.31	GPMoS	0.002	RA	0.432
19	OC	1.1	IT	0.002	OC	0.306
20	WC	0.86	RA	0.002	QR2	0.294

Source: own calculations

We have decided to use four indictors that were simultaneously distinguished by at least two of applied ensemble methods: Liabilities Turnover (RC), ROA, Debt to EBITDA (EBIT) and Gross Profit Margin. Two ratios can be regarded as input: Debt to EBITDA and Liabilities turnover (RC). The other ratios, Return on Assets (ROA) and Gross Profit Margin (GPM), can be regarded as output.

Table 2. DEA rating for 76 production companies and their efficiency scores

Company	DEA1	DEA2	Eff.	Company	DEA1	DEA2	Eff.
AC	1	1	1.00	DEBICA	4	6	0.48
APATOR	2	1	1.00	IZOSTAL	3	5	0.47
CIGAMES	2	1	1.00	PATENTUS	4	6	0.43
CITYINTE	2	1	1.00	ZPUE	4	7	0.42
EKO_EXP	1	1	1.00	BIOMAXIM	3	4	0.42
HYDROT.	1	1	1.00	ZUE	6	5	0.41
PANITERE	1	1	1.00	WINDMOB	1	2	0.41
PGE	1	1	1.00	MIESZKO	6	8	0.40
PULAWY	2	1	1.00	ZPC_OTM	6	8	0.38
SONEL	2	1	1.00	ZYWIEC	2	4	0.37
WAWEL	2	1	1.00	MOJ	4	8	0.37
ZELMER	1	1	1.00	POLNA	3	2	0.37
BERLING	1	2	0.95	INTERCAR	3	7	0.37
DUDA	3	2	0.91	INVICO	4	7	0.36
RELPOL	3	3	0.85	SUWARY	6	8	0.36
MEGAR	2	2	0.81	PLASTBOX	5	9	0.35
BSCDRUK	2	2	0.78	ENERGOIN	6	8	0.34
STALPROD	2	3	0.77	AMICA	5	8	0.33
SYNEKTIK	5	3	0.75	PAMAPOL	6	10	0.31
ESSYSTEM	2	3	0.73	FERRO	5	8	0.29
MENNICA	1	3	0.73	WIELTON	6	10	0.28
POLICE	2	2	0.68	MUZA	6	5	0.28
NOVITA	3	4	0.67	POZBUD	4	4	0.28
BUDVAR	3	4	0.67	FASING	5	7	0.27
ALKAL	3	3	0.66	BORYSZ.	4	7	0.27
TAURON	4	3	0.63	INTEGER	4	2	0.27
HUTMEN	3	4	0.61	RAFAMET	5	8	0.27
IZOL_JAR	2	5	0.60	SNIEZKA	6	9	0.26
KETY	3	3	0.60	GROCLIN	5	10	0.23
FORTE	4	4	0.59	VISTULA	6	10	0.22
LOTOS	4	5	0.57	GRAAL	6	10	0.20
STOMIL_S	3	4	0.56	FERRUM	6	10	0.20
ZUK	3	5	0.54	RAFAKO	5	9	0.20

Company	DEA1	DEA2	Eff.	Company	DEA1	DEA2	Eff.
LENTEX	5	6	0.54	WOJAS	6	10	0.17
KPPD	3	6	0.54	KOELNER	6	10	0.14
PROJPRZM	4	5	0.53	RAWLPL.	6	10	0.14
PEPEES	3	5	0.53	GRAJEWO	6	10	0.14
ERG	5	7	0.48	ARMATUR.	6	10	0.12

Source: own calculations

After performing DEA again for selected set of ratios we have obtained 10 groups of companies. The results of the division are shown in column DEA2 of Table 2. The column Eff. contains relevant efficiency measure for each DMU. The first group of efficient objects consists of 12 companies. The second group consists of 8 companies, etc. It has to be noticed, that the ordering given by efficiency measure does not reflect the ranking of companies given by DEA groups (compare [Chodakowska et al. 2013]). For example, firms with quite low efficiency score were assigned to the second or third DEA group. The division into 10 DEA groups is more precise but, with minor exceptions, reflects previous ordering. The correlation coefficient between both assignments to DEA groups is high. It is equal 0.87.

CONCLUSIONS

In the paper we propose a new approach to classification of companies based on DEA. The method can be regarded as an alternative approach to classical statistical classification methods. We have shown on the example that application of random forests and gradient boosting provides a good tool for variable selection. Both methods, random forests and gradient boosting, are particularly well suited to the search for factors that could be used in DEA because of their response to highly local features of the data and possibility of using in cases with small numbers of observations without risk of overfitting.

Application of ensemble methods seems to be a promising approach to variable selection for the needs of DEA. Our calculations repeated on the group of 17 construction companies revealed that the ratios distinguished by ensemble methods differ depending on the companies' profile. Moreover, membership into DEA groups will be violated even if the set of considered DMUs will differ by one object only. Nevertheless, DEA seems to be a promising tool, alternative to traditional scoring models. It enables ranking of agents and it can be used for distinguishing classes of homogeneous object, e.g., rating classes. The support of ensemble methods in variable selections makes DEA approach an universal tool.

Random forests and gradient boosting can be expected to improve the automation of procedures to evaluate the status of companies by banks and other financial institutions.

REFERENCES

Andersen P., Petersen N. C. (1993) A Procedure for Ranking Efficient Units in Data Envelopment Analysis, Management Science, Vol. 39, pp.1261-1264.

Berk R. A. (2008) Statistical learning from a regression perspective, Springer, New York.

Breiman L. (2001) Random Forests, Machine Learning, Vol. 45 (1), pp. 5-32.

Chodakowska E., Wardzińska K. (2013) The attempt to create an internal credit risk rating of production companies with the use of Operational Research method, Quantitative Methods in Economics, Vol. XIV, No. 1, pp. 74-83.

Cooper W. W., Seiford L. M., Tone K. (2006) Introduction to Data Envelopment Analysis and Its Uses with DEA-Solver Software and References, Springer, New York.

Demirova M. (2010) An empirical application of data envelopment analysis in credit rating. Theses and dissertations, Paper 981, Ryerson University, Canada.

Dzidzevičiūtė L. (2012) Estimation of default probability for low default portfolios, Ekonomika 2012, Vol. 91 (1), pp.132-156.

Feruś A. (2006) The Application of the DEA Method to Define the Level of Company Credit Risk, Bank i Kredyt, Vol. 37, No. 7, pp. 44-59.

Hastie T., Tibshirani R., Friedman J. (2009) The elements of statistical learning. Data Mining, Inference and Prediction, Second Edition, Springer, New York.

Kaczmarska B. (2010) The Data Envelopment Analysis Method in Benchmarking of Technological Incubators, Operations Research and Decisions, Vol. 20, No. 1, pp. 79-95.

Koronacki J., Ćwik J. (2008) Statystyczne systemy uczące się, Akademicka Oficyna Wydawnicza EXIT, Warszawa.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 102 –111

EFFICIENCY AND RETURNS TO SCALE – A CONCEPT OF USING DETERMINISTIC APPROACH[1]

Sebastian Jarzębowski
Department of Economics and Organisation of Enterprises
Warsaw University of Life Sciences –SGGW
e-mail: sebastian_jarzebowski@sggw.pl

Abstract: In the paper the author considered estimation of efficiency, which measures the ability of the company to obtain the maximum output from given inputs. The analysis has been carried out by using deterministic method (on the example of the DEA method, Data Envelopment Analysis). The two-step procedure was proposed, namely the returns to scale were specified in the sample and, basing on that, the efficiency of individual enterprises was assessed. In the paper the data from the companies of a key food processing sector in Poland, namely the meat processing, was used. The analysis covered the period 2006–2011, the sample covered from 195 up to 210 enterprises (depending on the analyzed year).

Keywords: efficiency, economies of scale, the DEA method, food processing sector

INTRODUCTION

The aim of the article was to propose an approach to measuring the basic economic category, namely efficiency, by using a non-parametric DEA method. A two-step procedure was applied. Firstly, the returns to scale were specified in the sample, and then the efficiency of individual enterprises was assessed using a model that takes into account the previously determined returns to scale. The discussed method has been applied to the meat processing industry in Poland. This sector was chosen due to the large size of the sample, as well as the strategic

[1] Praca naukowa finansowana ze środków na naukę w latach 2011-2013 jako projekt badawczy nr 2011/01/B/HS4/02612 *Stopień integracji w łańcuchu dostaw a efektywność przedsiębiorstw przetwórstwa rolno-spożywczego.* Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2011/01/B/HS4/02612.

importance and significant contribution to the production of the entire agri-food sector. Furthermore, with respect to the meat processing sector there are no comparative analyzes carried out, which justifies the need for their conduction.

In the literature, there are concepts of economic efficiency that determines the ratio of outputs achieved and inputs used. T.J. Coelli, D.S.P. Rao, Ch.J. O`Donnell and G.E. Battese, that refer to the dual approach to the issue in their researches on the efficiency, argue that the efficiency ratio increases by maximizing outputs with given inputs (an output-oriented approach), or by minimizing inputs with given outputs (an input-oriented approach). A company uses materials, labour and capital (inputs) in order to manufacture the final product (output), on the basis of which the authors define efficiency of companies as their ability to transform inputs into outputs.[2]

In the discussion on the efficiency the returns to scale are an important aspect. A manufacturer gains increasing returns to scale, if he achieves more than a proportional increase in production as a result of an increase in involved production factors. If this increase is proportional, we are dealing with constant returns to scale. In the case of an increase in production that is smaller than an increase in involved production factors one can say about decreasing returns to scale.[3]

THE DEA METHOD FOR ASSESSING THE RETURNS TO SCALE

The purpose of the appliance of the DEA method was to determine the efficiency for individual enterprises. The assumptions of this method were presented in this part of the paper. Due to the adoption of the two-step procedure for the efficiency assessment, an approach for determining the returns to scale using the DEA method was included in the content of this chapter.

The nonparametric DEA method was developed in 1978 by Charnes, Cooper, Rhodes.[4] The DEA is a method which assumes no random component and does not require functional form relating inputs to outputs for each of the analyzed DMUs[5].[6] The efficiency score is calculated by using the Debreu-Farrell measure. According to definition of Debreu and Farrell, the measurement of technical efficiency is the

[2] Coelli T.J., Rao D.S.P., O`Donnell Ch.J., Battese G.E. (2005) An introduction to efficiency and productivity analysis, 2. Edition, Springer, New York.

[3] Rembisz W. (2011) Analityczne właściwości funkcji produkcji rolniczej, Komunikaty, Raporty, Ekspertyzy, nr 544, Wyd. IERiGŻ-PIB, Warszawa, p. 18.

[4] Charnes, A., Cooper, W., Rhodes A. (1978) Measuring the Efficiency of Decision Making Units. „European Journal of Operational Research", 2 (6), p. 429.

[5] Decision making units (DMUs) are known as the objects of analysis, see: Cooper W., Seiford L., Tone K. (2007) Data Envelopment Analysis. A comprehensive text with models, applications, references, 2. Edition, Springer-Verlag, Berlin, p. 6-12.

[6] Cooper W., Seiford L., Tone K. (2007) Data Envelopment Analysis. A comprehensive text with models, applications, references, 2. Edition, Springer-Verlag, Berlin, p. 13.

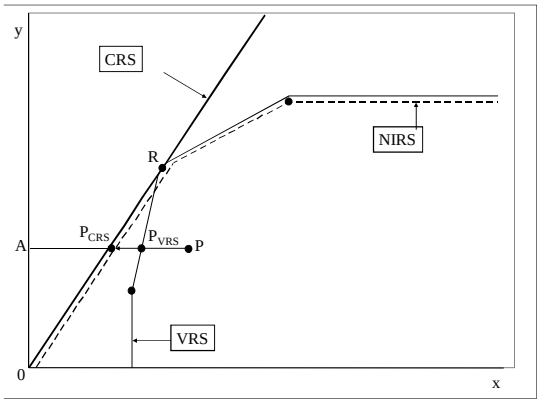
difference between one and the maximal possible reduction of inputs, while production of a certain volume of inputs is technologically possible.[7]

Depending on what is the aim of the analyzed DMUs, minimizing inputs or maximizing outputs, one can calculate the input-oriented technical efficiency or the output-oriented technical efficiency. Besides the input- and output-oriented DEA model, there is a further division including returns to scale. There are following models distinguished:

- CRS – Constant Returns to Scale
- VRS – Variable Returns to Scale
- NIRS – Non-Increasing Returns to Scale[8]

In 1984, Banker, Charnes and Cooper proposed a development of the model assuming constant returns to scale to the model with variable returns to scale. However, it is not possible to determine whether a tested DMU shows increasing or decreasing returns to scale. In 1985, Färe, Grosskopf and Lovell modified the BCC model, complementing it with the additional assumption of convexity,[9] which resulted in the creation of the model assuming non-increasing returns to scale (*NIRS*).

Figure 1. The DEA models (VRS, CRS and NIRS) and returns to scale



Source: own work based on Coelli T.J., Rao D.S.P., O`Donnell Ch.J., Battese G.E. (2005) An introduction to efficiency and productivity analysis, 2. Edition, Springer, New York, p. 174.

[7] Fried H.O., Lovell C.A.K., Schmidt S.S. (1993) The Measurement of Productive Efficiency Techniques and Applications. Oxford University Pres, New York, Oxford, p. 10.

[8] The NIRS includes the decreasing (DRS, Decreasing Returns to Scale) or constant returns to scale (CRS, Constant Returns to Scale). More about the DEA models in: Jarzębowski S. (2011) The efficiency of grain milling companies in Poland and in Germany- application of DEA method and Malmquist index, Universität Bonn-ILB Press, Bonn.

[9] Coelli T.J., Rao D.S.P., O`Donnell Ch.J., Battese G.E. (2005) An introduction..., op. cit.

Under the assumption of constant returns to scale, efficient DMUs create the CRS efficiency frontier, however, assuming that returns to scale are variable, the VRS efficiency frontier is created (see figure 1, where x – set of inputs, y – set of outputs).

Depending on which assumption of returns to scale will be adopted, technical efficiency is equal to[10]:[11]

- the ratio AP_{CRS}/AP for constant returns to scale (the ratio is equal to the technical efficiency at constant returns to scale - TE_{CRS});
- the ratio AP_{VRS}/AP for variable returns to scale (the ratio is equal to the technical efficiency at variable returns to scale - TE_{VRS}).

The ratio of AP_{CRS}/AP_{VRS} determines the efficiency of scale. The efficiency of scale can be (see figure 1) interpreted graphically as the ratio of the mean score of a given DMU in point P_{VRS} to the mean score in point R (point of technically optimal production scale).[12] The value, however, does not indicate whether the given DMU is in the area of increasing or decreasing returns to scale. The nature of returns to scale[13] (caused by both increasing and non-increasing returns to scale) can be determined by analyzing the results of the technical efficiency obtained in the NIRS and VRS model. If these values are not equal, it indicates the presence of increasing returns to scale in the DMU. However, if these values are equal, the DMU is characterized by decreasing returns to scale (if additionally the values obtained in the VRS and CRS models are equal, the DMU shows constant returns to scale).[14]

THE RETURNS TO SCALE AND EFFICIENCY OF ANALYSED ENTERPRISES

The study was carried out on the basis of data collected from meat processing enterprises across Poland, for which financial statements were available (panel data for the period 2006–2011). The sample covers from 195 up to 210 companies, depending on the analyzed year (including micro, small, medium and large enterprises). The production data is reported as revenue/expenditure denominated in PLN in constant prices. The production frontiers are fitted for a single output and two inputs. The inputs are: value of fixed assets (x_1), operating costs (x_2), and the output is net revenues from sales of goods and materials (y).[15]

[10] The technical efficiency was determined on the example of point P. The point is not neither on the CCR efficiency frontier nor on the VRS efficiency frontier, therefore the DMU is considered to be inefficient.

[11] Coelli T.J., Rao D.S.P., O`Donnell Ch.J., Battese G.E. (2005) An introduction..., op. cit., p. 173.

[12] Ibidem, p. 173.

[13] Increasing returns to scale occur when outputs increase faster in relation to growth of used inputs. Decreasing returns to scale occur, when outputs increase slower in relation to growth of used inputs. Constant returns to scale occur when outputs increase proportionally to growth of used inputs. Non-increasing returns to scale occur when outputs increase slower or proportionally to used inputs.

[14] Ibidem, p. 174.

[15] The source of data was the MONITOR POLSKI B where financial statements are published.

Specification of the DEA model

The DEA method (Data Envelopment Analysis) is a relatively modern tool based on a non-parametric approach to the creation the efficiency frontier.[16] In the DEA method, a system of linear equations is being solved, usually using the Simplex method. [17] With this technique, one can simultaneously process a large number of variables, taking into account the internal relations in an enterprise with the assumed objective function.[18] In the case of the DEA method it is possible to use several optimization models. They differ among others in assumptions about the returns to scale in the sample. While choosing a model one bases on the expertise or practices used by other research groups (similarly to the choice of variables), nevertheless the model selection affects the value of the obtained efficiency ratios. As shown in the previous part of the article, the calculation of three models, namely VRS, NIRS and CRS, is needed in order to determine the returns to scale.[19] An algebraic form of the models was followed after J. Zhu and W.D. Cook.[20]

The CRS model is presented in equations 1-4:

$$\max_{\phi_k, \lambda_k} \phi_k \tag{1}$$

$$\phi y_k \leq \sum^{I} \lambda_{ik} y_i \tag{2}$$

$$x_{nk} \geq \sum_{i=1}^{I} \lambda_{ik} x_{ni} \tag{3}$$

$$\lambda_{ik} \geq 0, \tag{4}$$

Equations 5-9 include the NIRS model:

$$\max_{\phi_k, \lambda_k} \phi_k \tag{5}$$

$$\phi y_k \leq \sum_{i=1}^{I} \lambda_{ik} y_i \tag{6}$$

[16] Rembisz W., Sielska A., Bezat A. (2011): Popytowo uwarunkowany model wzrostu produkcji rolno-żywnościowej, Wyd. IERiGŻ-PIB, Warszawa, p. 108.

[17] The linear programming has been commonly used since the 60's, when the Simplex algorithm was applied for problems of farmers, see: Zapf R. (1965) Zur Anwendung der linearen Optimierung in der landwirtschaftlichen Betriebsplanung, Berichte über Landwirtschaft..

[18] Steffen G., Born D. (1987) Betriebs- und Unternehmensführung in der Landwirtschaft, UTB für Wissenschaft, Stuttgart.

[19] No assumptions regarding returns to scale were made in advance. It was assumed that the functions are homogeneous of degree +1 for $\lambda > 0$. A set of observations is convex if for any two points in the set, all weighted average of these two points are also points in the same set.

[20] Zhu J., Cook W.D. (2007) Rank Order Data in DEA, Interval and Ordinal Data, [w:] Zhu J., Cook W.D. (red.): Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis. Springer, New York, p. 13-62.

$$x_{nk} \geq \sum_{i=1}^{I} \lambda_{ik} x_{ni} \tag{7}$$

$$\lambda_{ik} \geq 0, \tag{8}$$

$$\sum_{i=1}^{I} \lambda_{ik} \leq 1 \tag{9}$$

The VRS model is presented in equations 10-14:

$$\max_{\phi_k, \lambda_{ik}} \phi_k \tag{10}$$

$$\phi y_k \leq \sum_{i=1}^{I} \lambda_{ik} y_i \tag{11}$$

$$x_{nk} \geq \sum_{i=1}^{I} \lambda_{ik} x_{ni} \tag{12}$$

$$\lambda_{ik} \geq 0, \tag{13}$$

$$\sum_{i=1}^{I} \lambda_{ik} = 1 \tag{14}$$

where:

k – index indicating an analyzed object,

ϕ_k – multiplier of output level for an object k,[21]

i – index indicating next object $i=1,\ldots,I$, where I is a number of objects in the sample,

y_i – output of an object i,

n – index of next input,

x_{ni} – input n used by an object i,

λ_{ik} – coefficients of linear combination between objects i and k.

For all models the output-orientation was adopted, which was expressed in the objective function of the optimization problem (equations: 1, 5, 10). The output-orientation means that by a given level of input higher and higher level of output may be achieved. On the basis of the literature review it was stated that the purpose of the business activity is the profit maximization (output of the activity), which confirms the correctness of the use of the DEA method with output-oriented models. The evaluation was conducted by using the EMS software.[22]

[21] This is the inverse of the efficiency coefficient.

[22] EMS, Efficiency Measurement System, Scheel H. (2000): Software Version 1.3, University Dortmund [July 2013].

Evaluation of the returns to scale and efficiency in the analyzed enterprises

The DEA method was used for the means of the article, as it provides a lot of detailed information on the individual objects in the sample, among others on returns to scale for individual objects.

Table 1 includes the synthetic results of the analysis of the returns to scale for individual enterprises (calculated using the procedure presented in the paper). The percentage share of the companies characterized by decreasing, increasing and constant returns to scale was determined. The results are summarized by year. The "DRS" (Decreasing Returns to Scale) indicates that companies are characterized by decreasing returns to scale, the "IRS" (Increasing Returns to Scale) – companies with the increasing returns to scale, and the "CRS" (Constant Returns to Scale) – companies characterized by the constant returns to scale. The statement about returns to scale was made basing on the Fig. 1.

On the basis of the conducted analysis it was observed that the majority of the companies operating in the meat processing sector is characterized by decreasing returns to scale (within the period 2006-2011).

Table 1. Companies of meat processing sector with the returns to scale within the period 2006-2011

Year / returns to scale	2006	2007	2008	2009	2010	2011
DRS	86%	86%	86%	85%	85%	60%
IRS	12,5%	13%	13%	13%	13%	38,5%
CRS	1,5%	1%	1%	2%	2%	1,5%

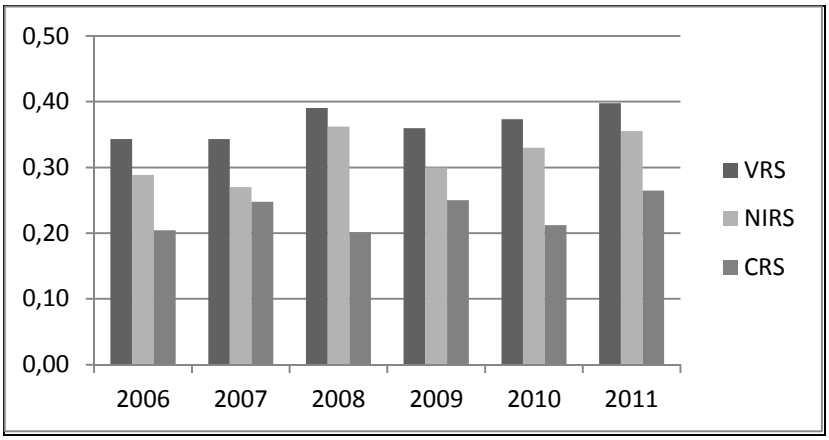
Source: own calculations using EMS software

Hence, most of the analyzed companies does not achieve the returns to scale, so the obtained increase in production is less than the increase in involvement of the production factors. According to R.D. Banker, W.W. Cooper, L.M. Seiford, R.M. Thrall, J. Zhu,[23] the presence of non-increasing returns to scale was assumed. This was the basis for the model specification.

In order to determine the efficiency on the basis of the returns to scale, the NIRS model (presented in equations 6-10) was adapted for the analyzed sample. The average efficiency ratios obtained using the NIRS model were presented in Figure 2. For a comparison, the efficiency ratios that would be obtained using the VRS and CRS models were also included in the analysis. The achieved efficiency ratios were presented by year (Figure 2).

[23] Banker R.D., Cooper W.W., Seiford L.M., Thrall R. M., Zhu J. (2004) Returns to scale in different DEA models, European Journal of Operational Research, Vol. 154, p. 359.

Figure 2. The average efficiency ratio calculated with use of the DEA method within the period 2006-2011[24]



Source: own calculations using EMS software

On the basis of the results showed in Figure 2 it can be stated that the average value of the efficiency ratio in the analyzed period was different depending on the model used. For the NIRS model with non-increasing returns to scale, the average efficiency ratio varied from 0,27 in 2007 to 0,36 in 2008 and 2011. The ratios obtained using the VRS model (Variable Returns to Scale) would be on average 15% higher compared to the results of the NIRS model. While using the CRS model (Constant Returns to Scale), the efficiency ratios would be on average 25% lower than the results of the NIRS model. The differences in the results are determined by assumptions of the DEA method, which was explained in the privies part of the paper. Technical efficiency score for constant returns to scale is not equal to efficiency score for variable returns to scale (see figure 1).

In the context of the conducted analysis, the proposed approach, according to which the returns to scale dominated in the sample are determined and then the DEA model specification is prepared for the efficiency measurement, seems to be appropriate. The proposed way of the model selection allows making a choice of the adequate model for a given sample with regards to the returns to scale and as a result it allows to a proper interpretation of the results. Basing on the conducted analysis it was claimed that in the analyzed sector in each year there is a place and necessity for efficiency improvement through effectively used manufacturing techniques reflecting use of inputs in order to manufacture the output.

[24] Due to the fact that while using the DEA method the relative efficiency is determined, there is no possibility to compare the results between years, see: Bezat A. (2012) Efficiency of Polish grain trade companies: an integrated application of SFA and DEA methods, Universität Bonn-ILB Press, Bonn.

SUMMARY

A company uses inputs in order to manufacture the output, on the basis of which the authors define efficiency of companies as their ability to transform inputs into outputs. A manufacturer gains increasing returns to scale, if he achieves more than a proportional increase in production as a result of an increase in involved production factors. If this increase is proportional, we are dealing with constant returns to scale. In the case of an increase in production lower than an increase in involved production factors one can say about decreasing returns to scale.

The deterministic tools, which analytical background is an optimization problem, (e.g. the DEA method, Data Envelopment Analysis) are methods for assessing the returns to scale, as well as efficiency. The methods require all decision making units to have comparable inputs and outputs and can handle multiple input and multiple output models.

The DEA method was applied for the assessment of the returns to scale of enterprises operating in the meat processing sector. The proposed approach involves determining the returns to scale that dominate in the sample, then the specification of the DEA model was prepared in order to assess the enterprises' efficiency. As demonstrated within the framework of the conducted analysis, the proposed way of model selection allows making a proper model specification and achieving right results and formulate reliable conclusions. It was also indicated in the article that basing – while determining the returns to scale in a given sample – on the expertise or practices used by other research groups affects the value of the obtained efficiency ratios. Hence the article contributes to the discussion on the methodological considerations about the key economic issues of efficiency measurement.

REFERENCES

Banker R.D., Cooper W.W., Seiford L.M., Thrall R. M., Zhu J. (2004) Returns to scale in different DEA models, European Journal of Operational Research, Vol. 154, p. 345–362.

Bezat A. (2012) Efficiency of Polish grain trade companies: an integrated application of SFA and DEA methods, Universität Bonn-ILB Press, Bonn.

Charnes, A., Cooper, W., Rhodes A. (1978) Measuring the Efficiency of Decision Making Units. „European Journal of Operational Research", 2 (6), p. 429-444

Coelli T.J., Rao D.S.P., O`Donnell Ch.J., Battese G.E. (2005) An introduction to efficiency and productivity analysis, 2. Edition, Springer, New York.

Cooper W., Seiford L., Tone K. (2007) Data Envelopment Analysis. A comprehensive text with models, applications, references, 2. Edition, Springer-Verlag, Berlin, p. 6-13.

EMS, Efficiency Measurement System, Scheel H. (2000): Software Version 1.3, University Dortmund [July 2013].

Fried H.O., Lovell C.A.K., Schmidt S.S. (1993) The Measurement of Productive Efficiency Techniques and Applications. Oxford University Pres, New York, Oxford.

Jarzębowski S. (2011) The efficiency of grain milling companies in Poland and in Germany- application of DEA method and Malmquist index, Universität Bonn-ILB Press, Bonn.

Rembisz W. (2011) Analityczne właściwości funkcji produkcji rolniczej, Komunikaty, Raporty, Ekspertyzy, nr 544, Wyd. IERiGŻ-PIB, Warszawa.

Rembisz W., Sielska A., Bezat A. (2011): Popytowo uwarunkowany model wzrostu produkcji rolno-żywnościowej, Wyd. IERiGŻ-PIB, Warszawa.

Steffen G., Born D. (1987) Betriebs- und Unternehmensführung in der Landwirtschaft, UTB für Wissenschaft, Stuttgart.

Zapf R. (1965) Zur Anwendung der linearen Optimierung in der landwirtschaftlichen Betriebsplanung, Berichte über Landwirtschaft.

Zhu J., Cook W.D. (2007) Rank Order Data in DEA, Interval and Ordinal Data, [w:] Zhu J., Cook W.D. (red.): Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis. Springer, New York, p. 13-62.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 112 – 124

THE RATING SCALE MODEL IN THE CONSTRUCTION OF FUZZY TOPSIS METHOD

Bartłomiej Jefmański
Department of Econometrics and Computer Science
Wrocław University of Economics
e-mail: bartlomiej.jefmanski@ue.wroc.pl

Abstract: Fuzzy TOPSIS method enables linear ordering of objects characterized by linguistic variables, which values constitute expressions emerging from natural language. Crucial, however, often neglected phase of this method is a selection of the way of introducing linguistic expressions by fuzzy numbers. Therefore, in this article one suggested a modification of fuzzy TOPSIS method using Rating Scale Model (RSM) to establish triangular fuzzy numbers. A suggested method enables establishing the rank of objects on the basis of objective criteria and subjective weights expressed in the form of triangular fuzzy numbers. Usability of the suggested method was confirmed by an empirical example, concerning linear ordering of selected smartphones models.

Keywords: IRT models, rating scale model, TOPSIS, fuzzy number

INTRODUCTION

TOPSIS method belongs to the linear ordering group allowing for synthetic assessment of multidimensional objects. In an original version of the method suggested by Hwang and Yoon [1981], synthetic measure is evaluated on the basis of Euclidean distance from positive-ideal solution and negative-ideal solution. Fuzzy modification of TOPSIS method suggested by Chen [2000], enables for a synthetic assessment of multidimensional objects with the implication of linguistic variables and triangular fuzzy numbers. However, this method does not suggest the way of expressing linguistic variables by triangular fuzzy numbers. The way of estimating fuzzy numbers' parameters plays an important role in a final linear ordering of objects according to synthetic criterion. Therefore, the purpose of this article is to present a proposition of fuzzy TOPSIS method modification, which

is based on the implementation of the Rating Scale Model in estimating the parameters of triangular fuzzy numbers. A suggested method assumes that criteria's values are shown on a metric scale, while criteria's weights are introduced in the form of linguistic values. Usability of a suggested method was presented on the basis of empirical example.

FUZZY TOPSIS METHOD

Let's assume that a certain set of objects $A=\{A_i|i=1,...,n\}$ and a set of criteria $C=\{C_j|j=1,...,m\}$, where $\tilde{X}=\{\tilde{x}_{ij}|i=1,...,n;\ j=1,...,m\}$ stand for a set of fuzzy evaluation criterion and $\tilde{W}=\{\tilde{w}_j|j=1,...,m\}$ a set of fuzzy weights. Linear ordering of objects with the above outlined assumptions is possible among others through the application of fuzzy TOPSIS method. An example of applying this method can be found among others in studies of: Uyun and Riadi [2011], Yayla and in. [2012], Madi and Tap [2011], Matin et al. [2011], Chang and Tseng [2008], Erdoğan et al. [2013], Ataei [2013].

Application of fuzzy TOPSIS method requires the accomplishment of the following steps [Chen 2000]:

Step 1. Calculation of normalized fuzzy evaluation criteria:

$$\tilde{z}_{ij}(x)=\frac{\tilde{x}_{ij}}{\sqrt{\sum_{i=1}^{n}\tilde{x}_{ij}^2}},\ i=1,...,n;\ j=1,...,m. \tag{1}$$

Step 2. Calculation of weighted normalized fuzzy evaluation criteria:

$$\tilde{v}_{ij}(x)=\tilde{w}_j\tilde{z}_{ij}(x). \tag{2}$$

Step 3. Appointing positive-ideal solution A^+ and negative-ideal solution A^- development:

$$\tilde{A}^+=\{\tilde{v}_1^+(x),\tilde{v}_2^+(x),...,\tilde{v}_j^+(x),\tilde{v}_m^+(x)\}$$
$$\{(\max_i\tilde{v}_{ij}(x)|j\in J_1),(\min_i\tilde{v}_{ij}(x)|j\in J_2)|i=1,...,n\}, \tag{3}$$
$$\tilde{A}^-=\{\tilde{v}_1^-(x),\tilde{v}_2^-(x),...,\tilde{v}_j^-(x),\tilde{v}_m^-(x)\}$$
$$\{(\min_i\tilde{v}_{ij}(x)|j\in J_1),(\max_i\tilde{v}_{ij}(x)|j\in J_2)|i=1,...,n\}, \tag{4}$$

where J_1 and J_2 are respectively the benefit criterion and the cost criterion.

Step 4. Calculation for each object a distance from positive-ideal solution d_i^+ and negative-ideal solution d_i^- (in an original work it is an Euclidean distance).

Step 5. Calculation of a synthetic measure:

$$C_i^+ = \frac{d_i^-}{d_i^+ + d_i^-},\ i = (1,...,n). \tag{5}$$

Measure values (5) are normalized in an interval $<0;1>$. The smaller the distance of an object is from a positive-ideal solution and the bigger from a negative-ideal solution, the closer the value of a synthetic measure is to cohesion.

Step 6. Establishing the objects ranking. The best object owns the biggest value of a synthetic measure.

FUZZY TOPSIS METHOD BASED ON THE RATING SCALE MODEL

RSM model

RSM model is one of the best known IRT (*Item Response Model*) models. It was suggested by Georg Rasch [1960] and then extended by David Andrich [1978]. This model enables to estimate the choice probability by a respondent of a certain category, in the assessment of a selected item scale. Probability depends on the level of "difficulty" of item scale, "the ability" of the individual and threshold for a certain category. In accordance with RSM model, the choice probability by the n-th respondent of the category x on i-th item scale is expressed by the equation [Andrich 1978]:

$$\pi_{nix} = \frac{\exp\sum_{j=0}^{x}\left[\beta_n - \left(\delta_i + \tau_j\right)\right]}{\sum_{k=0}^{m}\exp\sum_{j=0}^{k}\left[\beta_n - \left(\delta_i + \tau_j\right)\right]} \tag{6}$$

where: β_n – the level of the n-th respondent's ability to give a correct answer for i-th item scale,

δ_i – the level of difficulty of i-th item scale,

τ_j – threshold for j-th category within i-th item scale.

A presented model allows to convert measure results from ordinal scale into interval scale. However, it does not find an application and answer in the analysis of "extreme" patterns on the item scale (e.g. when a respondent chooses extreme categories like "definitely unimportant" or "definitely important" within all items scale). An advantage of this model is a fact that parameters concerning a respondent and item scale are expressed by a common measure unit (described as logit) on the same continuum. It is also necessary to emphasise that its application requires to accept the assumption about one-dimensionality scale (all items scale measure only one latent variable) and local independence of item scale (an answer for certain item scale is independent from the answer for other suggestions).

The most important parameters of RSM model, from the view of conversion of verbal categories to the form of triangular fuzzy numbers, are threshold values for these categories. Threshold values are appointed on continuum of latent variable in the point of characteristic curves intersection of characteristic and adjacent with each other categories. Therefore, a threshold value constitutes a point in which a choice probability of a respondent one of the two adjacent categories is the same and comes to 50%. A detailed characteristics of threshold values in IRT models together with a graphical presentation includes Linacre study [2010].

Characteristics of fuzzy RSM-TOPSIS method

Let's assume that a measurement of criteria's values takes place in a metrical measure scale. Variables weights are assigned in a direct way (by experts, respondents etc.) through linguistic values (e.g. very important, unimportant etc.). In such a case, linear ordering of objects is possible through application of fuzzy RSM-TOPSIS method. It constitutes the hybrid linking fuzzy TOPSIS method and RSM model. This method assumes that, criterion weights expressed in the form of linguistic values are transformed to the form of triangular fuzzy numbers with the application of RSM model. The accomplishment of this method takes place in five distinguished stages outlined below:

Stage 1. Selection of criteria for objects' assessment;

Stage 2. Normalization of criteria's assessments (there is a need to apply normalization formula appropriate for metric scales);

Stage 3. Assessment of the criteria's importance by linguistic expressions;

Stage 4. Conversion of linguistic values to the form of fuzzy numbers.

A suggested approach assumes at this stage the use of RSM model procedure. Support for triangular fuzzy numbers for ordered linguistic values are determine in accordance with threshold values, which are assigned to certain categories. Table 1 illustrates formulas needed to establish triangular fuzzy numbers parameters for each of the categories distinguished within j-th criterion. An example concerns rating scale with five ordered verbal criteria, which are properly appointed as follows: definitely unimportant (DUI), unimportant (UI), medium important (MI), important (I), definitely important (DI).

Table 1. Formulas for estimating triangular fuzzy numbers parameters for linguistic values

Category	Fuzzy number parameters		
	a	b	c
DUI	-4	-4	τ_{i1}
UI	τ_{i1}	$\frac{\tau_{i1}+\tau_{i2}}{2}$	τ_{i2}
MI	τ_{i2}	$\frac{\tau_{i2}+\tau_{i3}}{2}$	τ_{i3}
I	τ_{i3}	$\frac{\tau_{i3}+\tau_{i4}}{2}$	τ_{i4}
DI	τ_{i4}	4	4

Source: own study

In the case of parameters b and c of a fuzzy number assigned to the most beneficial category it is determined on level 4. A characteristic feature of RSM model is a fact that, threshold values for particular categories can differ within criteria of objects' assessment. It means that, suggested in this article approach requires estimation of fuzzy numbers parameters separately for each criterion.

Stage 5. Averaging the assessment of significance through calculating arithmetic mean of fuzzy numbers in accordance with an equation:

$$\tilde{w}_j = \sum_{i=1}^{n} \tilde{w}_{ij} \tag{7}$$

where: $\tilde{w}_{ij} = (a_{ij}, b_{ij}, c_{ij})$ - weight j-th criterion assigned by i-th respondent.

Stage 6. Rating normalized weighted criteria's assessments

Normalized weighted assessment i-th object according to j-th criterion is estimated in the following pattern:

$$\tilde{v}_{ij}(x) = \tilde{w}_j z_{ij}(x) \tag{8}$$

According to the principles of arithmetic of fuzzy numbers described among others in the study of Iron [1998] equation's result (8) is also a fuzzy number.

Stage 7. Determining a positive-ideal solution and negative-ideal solution of development.

According to the fact that, the obtained results in stage 4 have triangular form of fuzzy numbers, there is a problem of determining criteria's values for positive-ideal solution and a negative-ideal solution. Therefore, there is an issue of comparing fuzzy numbers and then a choice of maximal and minimal fuzzy number for each of these criterion. In order to do that, defuzzification of normalized weighted assessments is suggested, to show for each of criterion the best and worst value. Next, they will constitute a positive-ideal solution's coordinates, depending on the fact if a criterion influences benefit criterion or cost criterion on synthetic criterion.

Formulas of defuzzification fuzzy numbers were included among others in the study of Opricovic and Tzeng [2003].

Stage 8. Calculation the distance of assessed objects from a positive-ideal solution and negative-ideal solution.

It is suggested to apply Euclidean distance in estimating the distance of i-th object from positive-ideal solution and negative-ideal solution. This distance for two triangular fuzzy numbers $\tilde{A}$ and $\tilde{B}$ is expressed by an equation:

$$d(\tilde{A},\tilde{B}) = \sqrt{\frac{1}{3}(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2} \tag{9}$$

Stage 9. Determining the ranking of objects on the basis of an equation:

$$C_i^+ = \frac{d_i^-}{d_i^+ + d_i^-} \tag{10}$$

The higher the value of synthetic measure is, the higher object's position in a ranking is.

EMPIRICAL EXAMPLE

Fuzzy RSM-TOPSIS method was applied to linear ordering of 10 selected smartphones, available on Polish market. Selected models according to report[1] of skąpiec.pl, were the most wanted ones in January 2014 r. The models were characterized by six criteria:

x_1 – screen size (inches),

x_{2a} – screen resolution horizontally (px),

x_{2b} – screen resolution vertically (px),

x_3 – resolution of built-in digital camera (Mpx),

x_4 – quantity of built-in memory (GB),

x_5 – RAM memory (GB),

x_6 – maximal time of conversations (h).

The importance of particulary criteria in using smartphones was assessed on the basis of survey research results (internet survey), which was conducted among smartphones' users in June 2014 r. This attempt was of purposeful character and its numerical amount came to 47 respondents.

Criteria were normalized in accordance with a formula of linear scale transformation [Shih et al. 2007]. Normalized criteria's values were distinguished in Table 2.

[1] Special report: Telephones, Servis Skąpiec.pl, January 2014.

Table 2. Normalized data matrix

Model	x_1	x_{2a}	x_{2b}	x_3	x_4	x_5	x_6
Samsung Galaxy S4 I9505	0,11	0,15	0,15	0,16	0,14	0,16	0,14
Samsung Galaxy S3 i9300	0,11	0,10	0,10	0,10	0,14	0,08	0,17
myPhone Next	0,10	0,08	0,08	0,10	0,03	0,08	0,04
Samsung Galaxy S III mini I8190	0,09	0,07	0,06	0,06	0,07	0,08	0,11
Samsung Galaxy S DUOS S7562	0,09	0,07	0,06	0,06	0,02	0,06	0,10
Samsung Galaxy Note III N9005	0,13	0,15	0,15	0,16	0,28	0,23	0,17
Sony Xperia Z	0,11	0,15	0,15	0,16	0,14	0,16	0,11
Goclever Quantum 4	0,09	0,07	0,06	0,02	0,03	0,04	0,04
Apple iPhone 5 16 GB	0,09	0,09	0,09	0,10	0,14	0,08	0,06
Sony Xperia J	0,09	0,07	0,07	0,06	0,02	0,04	0,06

Source: own calculations

The importance of particulary criteria, respondents assessed through rating scale, which points constituted linguistic value: “definitely unimportant”, “unimportant”, “medium important”, “important”, “definitely important.” None of the respondents assessed criterion as “definitely unimportant” or “unimportant”, therefore in a further analysis three other linguistic values were used.

On the basis of the results of assessments‘ importance and in accordance with RSM model, characteristic curves were estimated for all criteria. Graphical picture of curves was distinguished in Figure 1.

In accordance with a suggested in this article method a basis of conversion of the results of assessments’ importance to the form of fuzzy numbers are the points of intersection of particulary characteristic curves. These points’ values for each of criteria were distinguished in Table 3.

In accordance with formulas included in Table 1, the conversion results is an expression of linguistic values through triangular fuzzy numbers. Estimating within each criterion an arithmetic mean from fuzzy numbers allowed to obtain averaged weights for each of criterion. Parameters of triangular fuzzy numbers being criteria’s weights were distinguished in table 3.

Table 3. Threshold values and weights for criteria

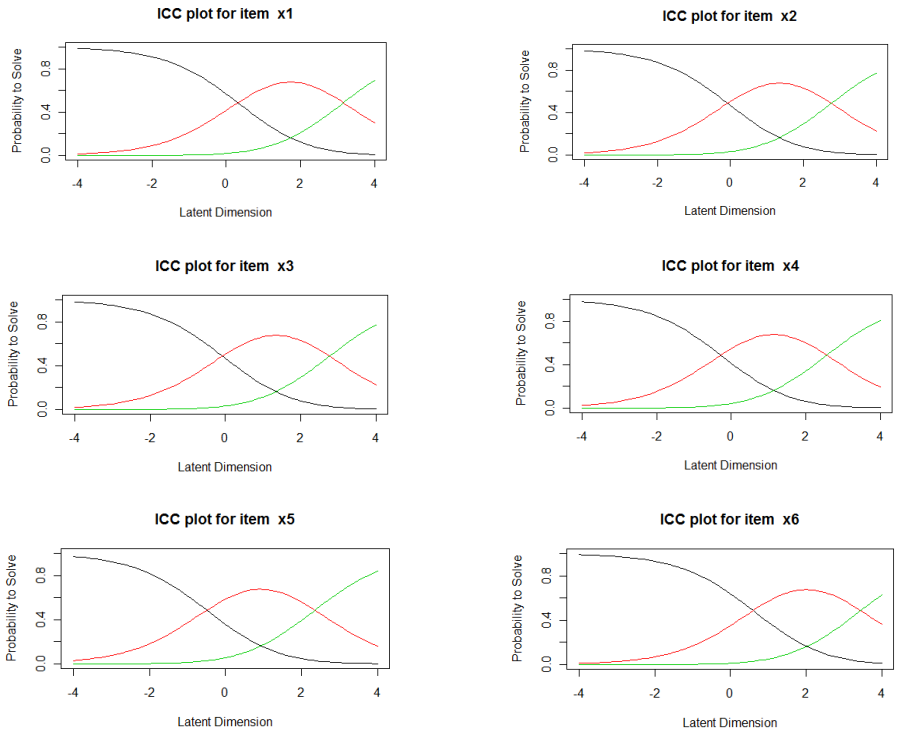
Criteria	Thresholds		Weights		
	τ_1	τ_2	a	b	c
x_1	0,32	3,16	1,06	2,21	3,30
x_{2a}	-0,08	2,77	0,99	2,25	3,17
x_{2b}	-0,08	2,77	0,99	2,25	3,17
x_3	-0,28	2,57	0,96	2,18	3,13

Criteria	Thresholds		Weights		
	τ_1	τ_2	a	b	c
x_4	-0,49	2,36	0,97	2,36	3,17
x_5	0,60	3,44	0,92	2,84	3,15
x_6	0,32	3,16	0,83	1,59	3,08

Source: own calculations with the application of eRm package of R programme

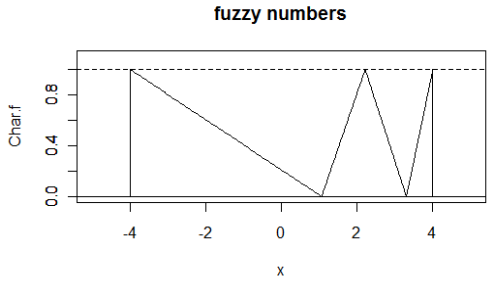
A graphical form of the obtained average assessments' importance in the forms of triangular fuzzy numbers was presented on the basis of criterion x_1 (see figure 2).

Figure 1. Characteristic curves for criteria



Source: own calculation with the application of eRm package of R programme

Figure 2. Average assessment's importance for criterion x_1



Source: own study with the application of fuzzyOP package of R programme

Normalized weighted assessments of criteria were estimated in accordance with arithmetic principles for fuzzy numbers and equation 8. The results in the form of parameters of triangular fuzzy numbers are introduced in table 4.

Establishing positive-ideal solution and negative-ideal solution requires recommendation among normalized weighted assessments maximal and minimal value. In order to do that through the defuzzification method presented in the studies of Ding and Liang [2005] and Wysocki [2010] 10 fuzzy numbers were compared within each criterion:

$$q_i = (a_i + 4b_i + c_i)/6 \tag{11}$$

Because all criteria influence in a benefit way on synthetic criterion, it was assumed in ideal-point solution for each criterion maximal weighted assessments, however in a negative-ideal solution minimal assessments. Values for fuzzy ideal-point solution and negative-ideal solution were distinguished in table 5.

The distance of particulary objects from a positive-ideal solution and negative-ideal solution was estimated in accordance with an equation (9). The distance of objects together with the values of synthetic measure (10) was introduced in table 6.

In a ranking gained on the basis of a suggested method, there are objects very close to a positive-ideal solution as well as the ones of a very low value of a synthetic criterion. The highest value of a synthetic measure was obtained for Samsung Galaxy Note III N9005 model. There are also on a high place in a ranking, other models like Galaxy S4 I9505 and Sony Xperia Z. The lowest value of a synthetic measure gained Goclever Quantum 4 and Sony Xperia J models.

Table 4. Values of normalized weighted criteria's assessments

No.	Model	Criteria																				
		x_1			x_{2a}			x_{2b}			x_3			x_4			x_5			x_6		
		a	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*c*	*b*	*c*
1	Samsung Galaxy S4 I9505	0,12	0,25	0,37	0,34	0,34	0,48	0,15	0,35	0,49	0,16	0,35	0,51	0,13	0,33	0,44	0,14	0,44	0,49	0,11	0,21	0,42
2	Samsung Galaxy S3 i9300	0,11	0,24	0,35	0,23	0,23	0,32	0,10	0,23	0,33	0,10	0,22	0,31	0,13	0,33	0,44	0,07	0,22	0,25	0,14	0,27	0,52
3	myPhone Next	0,11	0,22	0,33	0,17	0,17	0,24	0,08	0,17	0,25	0,10	0,22	0,31	0,03	0,08	0,11	0,07	0,22	0,25	0,03	0,06	0,12
4	Samsung Galaxy S III mini I8190	0,09	0,20	0,29	0,15	0,15	0,22	0,06	0,15	0,20	0,06	0,14	0,20	0,07	0,16	0,22	0,07	0,22	0,25	0,09	0,18	0,35
5	Samsung Galaxy S DUOS S7562	0,09	0,20	0,29	0,15	0,15	0,22	0,06	0,15	0,20	0,06	0,14	0,20	0,02	0,04	0,05	0,06	0,17	0,19	0,09	0,16	0,32
6	Samsung Galaxy Note III N9005	0,13	0,28	0,42	0,34	0,34	0,48	0,15	0,35	0,49	0,16	0,35	0,51	0,27	0,65	0,88	0,22	0,67	0,74	0,14	0,27	0,51
7	Sony Xperia Z	0,12	0,25	0,37	0,34	0,34	0,48	0,15	0,35	0,49	0,16	0,36	0,51	0,13	0,33	0,44	0,14	0,44	0,49	0,09	0,18	0,34
8	Goclever Quantum 4	0,09	0,20	0,29	0,15	0,15	0,22	0,06	0,15	0,20	0,02	0,05	0,08	0,03	0,08	0,11	0,04	0,11	0,13	0,03	0,06	0,12
9	Apple iPhone 5 16 GB	0,09	0,20	0,29	0,20	0,20	0,29	0,09	0,21	0,29	0,10	0,22	0,31	0,13	0,33	0,44	0,07	0,22	0,25	0,05	0,10	0,20
10	Sony Xperia J	0,09	0,20	0,29	0,15	0,15	0,22	0,07	0,16	0,22	0,06	0,14	0,20	0,02	0,04	0,05	0,04	0,11	0,13	0,05	0,09	0,18

Source: own calculations

Table 5. Criteria's values for positive-ideal solution and negative-ideal solution

Criteria	Positive-ideal solution			Negative-ideal solution		
	a	b	c	a	b	c
x_1	0,12	0,25	0,37	0,09	0,20	0,29
x_{2a}	0,34	0,34	0,48	0,15	0,15	0,22
x_{2b}	0,15	0,35	0,49	0,06	0,15	0,20
x_3	0,16	0,35	0,51	0,02	0,05	0,08
x_4	0,27	0,65	0,88	0,02	0,04	0,05
x_5	0,22	0,67	0,74	0,04	0,11	0,13
x_6	0,14	0,27	0,52	0,03	0,06	0,12

Source: own calculations

Table 6. Objects' distance from a positive-ideal solution and negative-ideal solution together with the values of synthetic values

No.	Model	d^-	d^+	C_i	Ranking place
1	Samsung Galaxy S4 I9505	2,71	1,03	0,73	2
2	Samsung Galaxy S3 i9300	1,80	1,94	0,48	4
3	myPhone Next	0,67	3,07	0,18	7
4	Samsung Galaxy S III mini I8190	0,78	2,96	0,21	6
5	Samsung Galaxy S DUOS S7562	0,47	3,29	0,12	8
6	Samsung Galaxy Note III N9005	3,79	0,08	0,98	1
7	Sony Xperia Z	2,63	1,12	0,70	3
8	Goclever Quantum 4	0,07	3,67	0,02	10
9	Apple iPhone 5 16 GB	1,25	2,49	0,33	5
10	Sony Xperia J	0,23	3,51	0,06	9

Source: own calculations

SUMMARY

TOPSIS method belongs to the group of most often applied methods in a linear ordering of multidimensional objects. Its fuzzy modification enables to conduct analyses in the fuzzy information conditions, when objects' assessments and/or criteria's weights are introduced in the form of linguistic values. It allows respondents to formulate assessment in a more natural way than through numbers, but at the same time it causes that this description is less precise and subjective. A theory of fuzzy sets seems to be helpful and allows for among others to express out of vague and ambiguous terms thanks to fuzzy numbers. Taking into consideration a fact that the results of a linear ordering of objects can depend on

parameters describing fuzzy numbers at their estimation there is a need to apply appropriate methods.

This article presents a suggestion of fuzzy TOPSIS method modification based on RSM model. It allows for conversion of the assessments of criteria's importance expressed by linguistic values to the form of triangular fuzzy numbers. As it results from the algorithm of a suggested in this article method, it can be sensitive for several parameters subjectively chosen by a researcher: normalization formula of variable, distance measure between fuzzy numbers and the way of comparing fuzzy numbers by defuzzification methods. It is also worth emphasizing that, RSM model is one of many IRT models, which can be applied in a fuzzy TOPSIS method. Therefore, in further research for this method, it is planned to conduct a conversion of linguistic values to the form of triangular fuzzy numbers, also thanks to such models like: Partial Credit Model, Generalised Partial Credit Model and Graded Response Model.

REFERENCES

Andrich D. (1978) A rating formulation for ordered response categories, "Psychometrika", vol. 43, pp. 561-573.

Ataei E. (2013) Application of TOPSIS and Fuzzy TOPSIS Methods for Plant Layout Design, "World Applied Sciences Journal", vol. 24, iss. 7, pp. 908-913.

Chang S.-H., Tseng H.-E. (2008) Fuzzy Topsis Decision Method for Configuration Management, "International Journal of Industrial Engineering", vol. 15, iss. 3, pp. 304-313.

Chen C.-T. (2000) Extensions of the TOPSIS for group decision-making under fuzzy environment, "Fuzzy Sets and Systems", no. 114, pp. 1-9.

Ding J.F., Liang G.S. (2005) Using fuzzy MCDM to select partners of strategic alliances for linear shipping, "Information Sciences", vol 1-3, pp. 197-225.

Erdoğan M., Bilişik Ö.N., Kaya İ., Baraçh H. (2013) A customer satisfaction model based on fuzzy TOPSIS and SERVQUAL methods, "Lecture Notes in Management Science", vol. 5, pp. 74-83.

Hwang C.L., Yoon K. (1981) Multiple Attributes Decision Making Methods and Applications, Springer, Berlin Heidelberg.

Iron A. (1998) Fuzzy rules and fuzzy functions: A combination of logic and arithmetic operations for fuzzy numbers, "Fuzzy Sets and Systems", vol. 99, iss. 1, pp. 49-56.

Linacre J.M. (2010) Transitional categories and usefully disordered thresholds, "Online Educational Research Journal", pp. 1–10.

Madi E.N., A.O.M Tap (2011) Fuzzy TOPSIS Method in the Selection of Investment Boards by Incorporating Operational Risks, "Proceedings of the World Congress on Engineering", vol. 1, pp. 291-295.

Opricovic S., Tzeng G. (2003) Defuzzification within a multicriteria decision model, "International Journal of Uncertainty Fuzziness and Knowledge-Based Systems", vol. 11, no. 5, pp. 635-652.

Special report: Telephones, Skąpiec.pl, January 2014.

Rasch G. (1960) Probabilistic Models for Some Intelligence and Attainment Tests, Danish Institute for Educational Research, Copenhagen (Expanded edition, University of Chicago Press, 1980).

Shih H.-S., Shyur H.-J., Lee E.S. (2007) An extension of TOPSIS for group decision making, "Mathematical and Computer Modelling", vol. 45, no. 7, pp. 801-813.

Uyun S., Riadi I. (2011) A Fuzzy Topsis Multiple-Attribute Decision Making for Scholarship Selection, "Telkomnika", vol.9, no.1, pp. 37-46.

Wysocki F. (2010) Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich, Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu, Poznań.

Yayla A.Y., Yildiz A., Özbek A. (2012) Fuzzy TOPSIS Method in Supplier Selection and Application in the Garment Industry, "FIBRES & TEXTILES in Eastern Europe", vol. 20, no. 4, pp. 20-23.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 125 – 134

THE ASSESSMENT OF DISEQUILIBRIUM OF THE PODLASKIE VOIVODESHIP LABOUR MARKET USING SYNTHETIC INDEX

Jarosław Kilon
Chair of Business Informatics and Logistics
Faculty of Management, Bialystok University of Technology
e-mail: j.kilon@pb.edu.pl

Jacek Marcinkiewicz
Department of Econometrics and Statistics
Faculty of Economics and Management, University of Bialystok
e-mail: marcinkiewicz_j@o2.pl

Abstract: One of the key tasks of the provincial labour offices is the regular monitoring of the situation on the labour market. It includes assessment of the equilibrium in each profession, in order to determine a deficit, balance or excess in professions. In this paper the results of a study designed to determine the condition of groups of occupations have been presented. The study was based on a multidimensional analysis of the data collected in regard to demand and supply of labour in terms of the individual data, which allowed for a relative assessment of the situation in the various professions.

Keywords: multivariate data analysis, synthetic index, labour market balance

INTRODUCTION

Regular monitoring of the situation on the labour market, including the development of in-depth analysis at the regional level (which covers monitoring of deficit and surplus in professions) is one of the key statutory tasks of the Polish provincial labour offices [Ustawa o promocji… 2004]. In the Podlaskie region, starting from 2009, this task was accomplished partly through the initiatives of the

Podlaskie Labour Market and Economic Forecasting Observatory[1]. The idea of this project was to develop a regional system for obtaining and analyzing information on current trends and projected changes in the economy (including the labour market), and then to develop a system of dissemination and exchange of collected data. One of the measurable results of this project was the regular publication of a study entitled Podlaskie Map of Occupations and Qualifications (Podlaska Mapa Zawodów i Kwalifikacji, PMZiK). The annual reports from the study covered, among others, a thorough assessment of supply and demand in occupations and groups of occupations in the region. The assessment of the situation (balance) in the professions and specialties, in order to determine the deficit, balanced or surplus professions was included in this monitoring. The analysis of demand and supply of labour was performed in spatial (at county level) and time terms (quarterly) [Podlaska Mapa 2014]. The study presented in this paper (in the original and narrower version) was one of the sub analysis of Podlaskie Map of Occupations and Qualifications 2014 published by the Voivodeship Labour Office in Bialystok. The survey was conducted by VIVADE Institute of Research and Analysis. The authors of this paper were members of the research team.

The relationship between the number of job seekers (labour supply) and the number of employment positions offered in a given time (labour demand) is one of the key analytical issues in relation to the labour market. Although the unemployment rate (amounting in Podlaskie 15.1% in 2013[2]) undoubtedly is the basic economic category associated with the phenomenon of imbalance in the labour market, it does not reflect the nature of the imbalance. The unemployment rate provides us only with the overall scale of the imbalance. The authors of a number of theories relating to the labour market perceive the causes of such imbalance in three main areas: structural (infrastructural inadequacy, for example theory of segmentation), the long-term nature of the adjustment process in the labour market (such as theory of human capital and job search) and wage rigidity (theories of effective wage, insiders-outsiders theory) [Jarmołowicz, Knapińska 2011]. Apart from the overall of imbalances, it should be emphasized that the processes taking place in today's labour market are very complex. The market is evolving in a dynamic manner and is influenced – in addition to structural changes – by seasonal fluctuations and frequent changes. All the direct market participants, both employers and potential employees, try to adapt to the changes in a flexible way. As a result, the frequent change of jobs, working part-time or self-employment are common phenomena [Strawiński 2013]. The analysis of the nature of the imbalance as well as monitoring of deficit and surplus in professions and

[1] The idea and tasks of Podlaskie Labour Market and Economic Forecasting Observatory is described in details at: http://www.obserwatorium.up.podlasie.pl/ [15.04.2014 r.].
[2] Data source: Local Data Bank, Central Statistical Office of Poland: [http://stat.gov.pl/bdl/; 20.07.2014 r.].

specialties becomes therefore particularly important. Such a monitoring allows for the acquisition of relevant information on the structure of unemployment, and the staffing needs of employers. In a wider perspective, it should also serve to coordinate the training of the unemployed and provide a basis for the development of the educational offer (with particular emphasis on vocational training).

In this paper the methodology and results of a study aiming at a broader recognition of the problem of determining the chances of finding a job were presented. The analysis of disequilibrium was performed at the level of the groups of professions. The study was based on a multidimensional data analysis of labour demand and supply data, obtained in terms of the individual job offer or unemployed (including jobs in relation to specific positions and information on unemployment registered by the district labour offices). Using a set of eight variables a synthetic index of the situation in the group of professions was constructed, which allowed for the relative assessment of the chances of finding work in groups of professions. The main aim of this study was to rank the groups of professions and specialties with the use of proposed synthetic index.

RESEARCH METHODOLOGY

Source data

In this paper the databases developed for the PMZiK were used. All the datasets covered the year 2013. The labour supply data consisted of information on the number and structure of the unemployed individuals registered at the local labour offices of Podlaskie voivodeship (nearly 71 thousand records). The database was made available for the study by Voivodeship Labour Office in Białystok and contained of half-yearly information on the registered unemployment (at 30.06.2013 and 31.12.2013 as well).The data used in the study were originally subject to public employment services reporting, and consisted of basic information about the unemployed with regard to professions and specialties according to the Polish Classification of Occupations and Specialties 2010 (Klasyfikacja Zawodów i Specjalności, KZiS)[3].

The analysis of the labour demand was based on the public information on the jobs offered in the Podlaskie region in 2013. The database of job opportunities was developed by the EU-CONSULT Ltd. on the basis of local labour offices internal databases and a wide range of offers from the other sources as well

[3] Polish Classification of Ocupations and Specialties (Polska klasyfikacja zawodów i specjalności, KZiS) is based on the ISCO-08 International Standard of Classification of Occupations developed by the International Labour Organisation and recommended by EUROSTAT for use in the European Union (Commission Recommendation of 29 October 2009. concerning the use of the International Standard for Qualifications (ISCO-08) Journal. OJ L 292, 10/11/2009 P. 0031-0047).

[Podlaska Mapa ... 2014]. Therefore a relatively broad spectrum of job offers was included in this study as the analysis of the demand was based on more than 86 thousand records.

It is important to consider a number of restrictions arising from the adopted methodology of collecting data on job vacancies in the Podlaskie in 2013, especially when the interpreting of the presented results is concerned. The restrictions relate in particular to the issues such as the scope of the sources included in the process of collecting data on job vacancies and the methodology adopted to eliminate repeated records (vacancies). With regard to the first problem, the analyzed database contained information gathered through the cataloguing of the jobs in the period 01.01.2013 - 31.12.2013. A wide range of data sources was used in that process, such as public employment services, academic labour offices, regional newspapers, and Internet portals (national and regional). Nevertheless, the adopted directory, though undoubtedly very wide, does not cover all sources, in which potential employers publish their offers. When the second restriction is concerned, the elimination of repeated offers included a two-step selection of advertisements on vacancies. In the first stage all the ads that appeared again in a short period of time in the source were skipped (all of the offers published in the same source in less than 14 days were treated as a repeated offer). In the second stage the elimination over a variety of sources was made. Although such a procedure was effective, it has to be assumed that in the case of certain professions (such as salesman or sales representative), in which employers conduct continuous recruitment (there is a large turnover of employees), the number of jobs identified in job advertisements may, however, differ from the real vacancies.

A continuous identification of deficits and surpluses in the occupations and specialties (performed by the Voivodeship Labour Office) is conducted on a very detailed data. The lowest level of classification of occupations (meaning the most detailed data – professions and specialties) used in these studies causes, however, that analyzes are significantly hindered. It precludes or significantly restricts the substantive interpretation of the obtained results (both in terms of supply and demand for labour). Moreover, a detailed analysis of the quality of data made by authors, suggests that there are number of errors in regard to determining the codes of professions in the available collections of information (both for the unemployed and job offers).

Therefore, aiming for a broader interpretation of the results obtained in this study, the assessment of the situation on the labour market in the Podlaskie voivodeship was conducted at the level of large groups of KZiS 2010 (double-digit symbol of occupational groups).

Synthetic index development

In order to determine imbalances in groups of occupations in the Podlaskie region, and thus answer the question about the chances of finding work in

occupations belonging to the groups, an analysis including the following variables was made:

X_1 – the number of jobs other than internships offered in the occupational group in which the Podlaskie voivodeship was indicated as a place of work;

X_2 – the number of jobs offered in the occupational group in which the whole Poland or the voivodeship other than Podlaskie was indicated as a place of work;

X_3 – the number of internships in the occupational group;

X_4 – the number of abroad jobs offered in the occupational group;

X_5 – the inflow of unemployed to the Podlaskie voivodeship labour offices, classified in the group of occupations;

X_6 – total number of unemployed people in the group of occupations, registered in the public labour offices (at the end of the year);

X_7 – the number of unemployed in the occupational group who were unemployed for a period longer than 12 months, registered in the public labour offices (at the end of the year);

X_8 – the number of unemployed in the occupational group who completed their education within previous 12 months (having graduate status), registered in the public labour offices (at the end of the year).

The values of X_1-X_8 were computed for each of the occupational groups. In the second stage the standardization of the obtained variables was made, using the following formula [Panek 2009]:

$$u_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \tag{1}$$

where:

x_{ij} – obserwation i of X_j variable, i = 1, 2, …, n; j = 1, 2,…, 8;

$\bar{x}_j$ – mean of X_j variable;

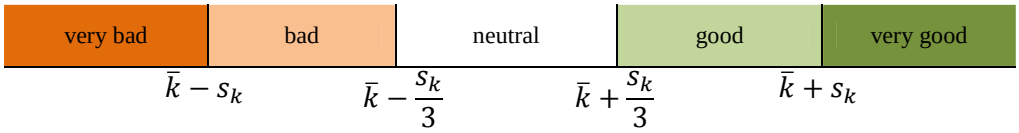
s_j – standard deviation of X_j variable.

All of the standardized variables of the demand side (U_1-U_4) were treated as stimulants, while supply-side variables (U_5-U_8) as destimulants of the analysed phenomena (situation in the group of occupations). The variables were then used to compose a synthetic variable, that descibed a relative condition of each of the occupational groups. In the construction of synthetic variable the not-pattern approach was used [Gatnar, Walesiak 2004]. The value of synthetic variable (index of a relevant condition of the group of occupations) was obtained according to the formula:

$$K_i = \sum_{j=1}^{4} u_{ij} w_j - \sum_{j=5}^{8} u_{ij} w_j \tag{2}$$

where w_j stands for weights relevant to the share of X_j in the sum of X_1-X_4 (for the demand side variables) or X_5-X_8 (for the supply). The obtained values of a synthetic variable were the basis of assignment of the groups of occupations to one of five clusters (Figure 1).

Figure 1. The classification of occupational groups in accordance to the the synthetic index value ('condition' of the occupational group)



$\bar{k}$ – mean of synthetic index in the analysed occupational groups;

s_k – standard deviation of the synthetic index.

Source: own study

The method adopted in this paper allowed us for identifying occupations in which finding a job was very difficult (there was a significant oversupply of job seekers compared to the number of jobs offered in the occupational group). On the other hand, the groups of professions in which finding a job was relatively easy (the number of jobs in the occupational group significantly exceeded the number of potential employees) were identified as well.

RESEARCH RESULTS

According to the official data, the total number of unemployed in the Podlaskie voivodeship in 2013 was by nearly 25,000 higher than the number of jobs offered (assuming that the Podlaskie voivodeship was declared as a place of work) [Podlaska Mapa ... 2014]. The level of discrepancy between the number of job seekers and the number of available positions in specific occupational groups, however, was significantly differentiated. According to the study there were five groups of occupations that could be described as those with the smallest difficulties of finding employment. The first cluster covered the following groups of occupations: administration associate professionals, information and communications technology professionals, building and related trades workers (excluding electricians), business and administration professionals and administrative and commercial managers (Table 1). Such a diagnosis could be confirmed by the structure of jobs and the unemployed registered in labour offices (Table 2). There were approx. 15% of the unemployed remained in the groups of occupations with a very good situation (taking into account both the number of unemployed on 31.12.2013 (X_5), and the inflow of the new unemployed in the period (X_6)). At the same time nearly 37% of jobs in which Podlaskie voivodeship

was the place of work (X_1), and nearly half of the jobs out of the region (X_2) concerned occupations classified in the first cluster.

Table 1. Occupational groups with a good or very good relative situation on the labour market in the Podlaskie region in 2013

Group ID	Group name	Synthetic index	Class
33	Business and administration associate professionals	1,4211	Very good
25	Information and communications technology professionals	1,3191	Very good
71	Building and related trades workers, excluding electricians	1,2083	Very good
24	Business and administration professionals	1,0663	Very good
12	Administrative and commercial managers	0,9209	Very good
96	Refuse workers and other elementary workers	0,5994	Good
13	Production and specialised services managers	0,4495	Good
83	Drivers and mobile plant operators	0,4474	Good
92	Agricultural, forestry and fishery labourers	0,3017	Good
41	General and keyboard clerks	0,2915	Good
11	Chief executives, senior officials and legislators	0,2614	Good

Source: own calculations

According to the research results, there were only approx. 5% of the unemployed registered in groups of occupations with good situation (X_5, Table 2). At the same time almost 10% of jobs in which Podlaskie was the place of work, and more than 11% of jobs outside the province concerned occupations classified in these groups. Good situation of people representing these groups of occupations (understood as a relatively high chance of finding employment) was partly caused by a relatively high percentage of internships and job offers abroad (respectively 27.3% (X_3) and 26% (X_4) of the total numbers).

Table 2. The structure of vacancies and unemployment in the Podlaskie region in 2013 in the scope of the analyzed variables

Class of occupational group	Labour demand				Labour supply			
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
very good	36,2%	46,6%	9,9%	23,2%	15,9%	15,6%	14,8%	31,2%
good	9,9%	11,4%	27,3%	26,0%	5,3%	5,4%	6,0%	12,8%
neutral	25,3%	20,6%	34,0%	20,8%	21,6%	21,0%	21,5%	28,5%
bad	21,3%	14,1%	22,4%	15,2%	31,2%	29,9%	29,4%	21,4%
very bad	7,4%	7,3%	6,4%	14,9%	26,0%	28,1%	28,3%	6,1%
TOTAL	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Source: own calculations

The situation described as neutral involved approx. 21% of the total number of persons registered in labour offices in the Podlaskie region. There were approx. 25% of vacancies addressed to these job-seekers in the region, and as many as 34% of total internship offers. Detailed analysis of the occupational groups with a neutral situation leads to the conclusion that there was a relatively high staff turnover in these professions (such as seller, security staff, customer service staff, etc. (Table 3)). It means that it was relatively easy to find both the job and the employee in the professions belonging to this group (which often did not require special competences).

Table 3. Occupational groups with a neutral situation on the labour market in the Podlaskie region in 2013

Group ID	Group name	Synthetic index	Class
52	Sales workers	0,2477	Neutral
14	Hospitality, retail and other services managers	0,2193	Neutral
94	Food preparation assistants	0,2184	Neutral
95	Street and related sales and service workers	0,2176	Neutral
34	Legal, social, cultural and related associate professionals	0,1771	Neutral
42	Customer services clerks	0,1732	Neutral
44	Other clerical support workers	0,1356	Neutral
54	Protective services workers	0,1349	Neutral
63	Subsistence farmers, fishers, hunters and gatherers	0,1349	Neutral
53	Personal care workers	0,1114	Neutral
22	Health professionals	0,0657	Neutral
62	Market-oriented skilled forestry, fishery and hunting workers	0,0530	Neutral
35	Information and communications technicians	0,0256	Neutral
82	Assemblers	-0,0163	Neutral
43	Numerical and material recording clerks	-0,0612	Neutral
32	Health associate professionals	-0,0698	Neutral
91	Cleaners and helpers	-0,0704	Neutral

Source: own calculations

Among all of the analysed professions the metal, machinery and related trades workers, food processing, wood working, garment and other craft and related trades workers as well as science and engineering associate professionals (Table 4) had the most difficult situation on the regional labour market. This situation is fully reflected in the structure of jobs: there were only about 7% of all offers related to these professions (X_1, Table 2), whereas more than a quarter of the registered unemployed belonged to the groups above.

Table 4. Occupational groups with a bad or very bad situation on the labour market in the Podlaskie region in 2013

Group ID	Group name	Synthetic index	Class
73	Handicraft and printing workers	-0,2702	Bad
74	Electrical and electronic trades workers	-0,2799	Bad
81	Stationary plant and machine operators	-0,2951	Bad
51	Personal service workers	-0,3828	Bad
21	Science and engineering professionals	-0,3959	Bad
61	Market-oriented skilled agricultural workers	-0,4076	Bad
23	Teaching professionals	-0,4546	Bad
93	Labourers in mining, construction, manufacturing and transport	-0,4810	Bad
26	Legal, social and cultural professionals	-0,7595	Bad
72	Metal, machinery and related trades workers	-1,8284	Very bad
75	Food processing, wood working, garment and other craft and related trades workers	-1,9112	Very bad
31	Science and engineering associate professionals	-2,5173	Very bad

Source: own calculation

The classification of occupational groups carried out in this paper gave also a rough estimate on how the vocational schools and educational institutions offer was fitted to the needs of the local labour market and to the expectations of employers as well. According to the study, nearly 31% of all graduates registered in labour offices have completed their education in occupations belonging to the groups with the best situation, and more than 6% – with the worst (Table 2). It leads to the conclusion of relatively good adjustment of educational offer to the labour market needs (there was a small percentage of graduates in the least popular professions). On the other hand, it emphasizes the question on the quality of education (it was difficult to take up employment immediately after completion of education by the graduates).

SUMMARY AND CONCLUSIONS

The method of assessing imbalances in the regional labour market presented in this paper allows for an in-depth analysis of the occupational groups in relative terms. It allows to formulate general conclusions in terms of both the current situation and trends (in case of analysing the data from consecutive periods) in relation to the chances of finding work in groups of occupations. Furthermore, it provides a broader approach to the problem of the imbalance in the labour market than the analysis including only identification of surplus and deficit in occupations or specialties. The level of generality presented in that paper (meaning large groups of KZiS, not professions or specialties) increases the transparency of the obtained

results, and partly eliminates the classification errors occurring within the process of data collection. The assessment of the situation on the labour market in large groups of occupations can be used in the analyses conducted by the public employment services, educational institutions, municipalities and other entities participating in the regulation of the labour market and vocational education.

REFERENCES

Gatnar E, Walesiak M. (2004) Metody statystycznej analizy wielowymiarowej w badaniach marketingowych, Wyd. AE we Wrocławiu, Wrocław, s. 352-355.

Jarmołowicz W., Knapińska M. (2011) Współczesne teorie rynku pracy a mobilność i przepływy pracowników w dobie globalizacji, Zeszyty Naukowe PTE, nr 9, Kraków, s. 123-144.

Panek T. (2009) Statystyczne metody wielowymiarowej analizy porównawczej, SGH w Warszawie, Warszawa, s. 38, 64-65.

Podlaska Mapa Zawodów i Kwalifikacji (2014), Wojewódzki Urząd Pracy w Białymstoku, Białystok.

Strawiński P. (2013) Uzupełniające miary statusu osób na rynku pracy, „Wiadomości statystyczne", nr 3, Główny Urząd Statystyczny w Warszawie, Warszawa, s. 37-52.

Ustawa o promocji zatrudnienia i instytucjach rynku pracy z dnia 20 kwietnia 2004 r. (Dz. U. z 2013 r., poz. 674 i 675).

Zalecenie Komisji z dnia 29 października 2009 r. w sprawie stosowania Międzynarodowego Standardu Kwalifikacji (ISCO-08) Dz. U. L 292, 10/11/2009 P. 0031-0047.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 135 –143

ON LOW-FREQUENCY ESTIMATION OF BID-ASK SPREAD IN THE STOCK MARKET

Marek Andrzej Kociński
Department of Applied Mathematics
Warsaw University of Life Sciences – SGGW
e-mail: marek_kocinski@sggw.pl

Abstract: In the article two popular low-frequency methods od bid-ask spread estimation are presented and applied to the stocks quoted on the Warsaw Stock Exchange (WSE): the Roll method [Roll 1984] and Corwin-Schultz method [Corwin and Schultz 2012]. The widely available data on average spreads published by WSE are used as benchmark and proxy of information, usually received from difficult to access and limited high-frequency financial data

Keywords: high-frequency data, low-frequency data, Roll estimator, Corwin-Schultz estimator

INTRODUCTION

Transaction costs are any payments for an opportunity of exchange of the good from one party of the transaction to the other. They are inherently part of a market microstructure. Estimation and analysis of transaction costs are important for portfolio managers, regulators and scholars. The costs of transaction executions are significant determinant of the investment performance since the net profits can be substantially reduced by these costs. Though a trading cost is often a small fraction of the value of a single transaction, in long-time horizon such expenses can considerably lower the return attained by the implementation of the investment strategy, especially when many purchases or sales are required. This is particularly important in the area of algorithmic and high-frequency trading which has been developing fast in recent years. Whether transaction costs are handled effectively or not by the asset manager can thus make all difference in possibility to outperform a given benchmark, for example a stock index. Moreover the transaction costs are in practice the measure of liquidity which is the essential

concept in financial theory and practice, since it represents the ease with which financial instruments can be traded. Monitoring liquidity risk, which means monitoring the amount of transaction costs, is one of the main tasks of financial regulators, particularly after the crisis of 2008. It was the time that due to the shortfalls of liquidity, some funds were not in able to redeem their units. The funds operating in the money market were particularly affected, because they methods of investing required high degree of liquidity.

Typically, in investment management, three major sources of transaction costs are taken into consideration: commissions and other direct fees, bid-ask spreads and market impact [Elton et al. 2010], [Sharpe et al. 1999]. The expenses such as broker commissions, taxes, and other fees are direct and predictable therefore they usually are not included in the quantitative analysis of transaction costs. Market impact refers to the effect of the order on the price of the traded security. This consists in inducing by the order the movement of the price against the order maker that is upward when buying and downward when selling. Thus, it is a source of transaction costs. The market impact is a complex phenomenon which, despite numerous articles on this subject, is not explained and quantified in a satisfactory manner. However, in practice it influences only very large transactions and can be neutralized by splitting the order and executing it incrementally over a longer period. The most important source of the transaction costs appears to be the bid-ask spread and estimation of this cost based on easy available financial data is the subject of this study. The bid-ask spread is the difference between the highest bid price and the lowest ask price for the stock; it measures the loss from buying a share of stock and then immediately selling it. It is a common measure of market liquidity and costs incurred by investors [Anand and Karagozoglu 2006]. Some authors include the bid-ask spread in market impact, given that it is a market phenomenon, in opposition to commissions and taxes [Grinold and Kahn 2000]. Bid–ask spreads are so important in market microstructure that researchers must have reliable estimators for spreads if they are to have substantial research on financial markets. Additionally, with the development of electronic trading in stock markets, investors, exchanges, and regulators are likely to be interested in precise bid–ask spread estimation methods. An accurate bid–ask estimator can help to compare the transaction costs in two competing markets. Moreover, the reliable spread estimators could help to optimize the investment strategies. The aim of this paper is to examine the possibility of bid-ask spread estimation based on low-frequency data from Warsaw Stock Exchange. The transaction costs in the polish stock market were consider before in the literature, see, for example [Olbryś 2012] or [Gniadkowska 2012] but it seems that there is no study on estimation of bid-ask spread based on easy available data for polish stocks, though the methods of estimation are present in financial theory at least from the work of [Roll 1984]. Thus, in my opinion, this article may be considered as the contribution to fill the gap in studies on the structure of the polish stock market.

LOW-FREQUENCY VERSUS HIGH-FREQUENCY DATA

In recent years the growing popularity of high-frequency data has been observed. It is driven by the rapid development of computer and information technologies which enable to effectively handle with enormous sets of data. The average daily number of number of observations of an actively traded NYSE stock can be higher than 20000 [Zivot 2005]. Transaction or quotes data observed in time periods shorter than once a day are called high-frequency data, in practice the length of time interval between consecutive moments of observations is very short and measured in seconds. The full information on market prices contains tick-by-tick data where tick refers to the change in the price of a security from trade to trade. The tick-by-tick data are the highest possible frequency data. In the extreme case, when the analyst has access to full record of characteristics of each transaction then one can talk about ultra-high-frequency data. The high-frequency financial data sets have been widely used to study various aspects of market microstructure. At first sight, the higher the number of independently measured observations, the higher is the precisions of estimation and the better model of market behavior can be obtained, however the situation is not as simple as it appears. The very detailed financial data can cause many problems. Some of them may result from human input errors, such as typing errors leading to data outliers. Other errors may be computer system errors, such as transmission failures which lead to data gaps, and database bugs causing mis-ordered time series observations [Zivot 2005]. Cleaning and correcting the data can be cumbersome. Furthermore tick-by-tick data are very noisy, because of discrete prices and nonsynchronous reporting which may deform inferences based on standard statistical models. The enormous number of data enforces the need of time-consuming data handling and filtering techniques. The sets of high-frequency data are materially limited. In the US markets transaction data are only available since 1983 and in many countries they are not available at all [Goyenko et al. 2009]. The presented problems with high-frequency data are sufficient to justify the need of the reliable bid-ask spread estimator based on low-frequency data by which I mean easy accessible daily characteristics of stocks such as opening price, closing price, low price, high price and trading volume. Researchers have an access to such data over a long price history and in many markets. The US daily stock returns and volume data are available from the Center for Research in Security Prices (CRSP) for NYSE/AMEX firms from 1926 to the present and for NASDAQ firms from 1983 to the present. A wide variety of services provide daily stock returns and volume data for international equity markets. For example, Thomson Financial's Data-stream provides the considered data for more than 60 countries from 1994 to the present and daily stock returns for several developed markets from the early 1970s [Goyenko et al. 2009]. Thus, reliable low-frequency spread proxies enable research of the bid-ask spread for a long time period and many markets. High-performing

low-frequency spread measures would be very helpful to the theory of market efficiency and in corporate finance. Such estimators are needed to verify whether described in literature trading strategies that appear to generate significant abnormal returns are truly profitable net of a relatively precise measure of transaction costs. Moreover, spread estimators across countries would greatly extend the potential diversity of international corporate finance environments to analyze [Holden 2009]. Recent articles provide indications that there is a chance of useful estimation of high-frequency benchmarks by means of low-frequency spread proxies in stock market [Goyenko et al. 2009], foreign exchange market [Karnaukh et al. 2014] and it seems that there is a hope for a success of low-frequency spread estimation in other types of markets where bid and ask prices appear. In fact, most studies published in financial literature deal with low frequency, regularly spaced data.

ESTIMATORS AND THEIR EVALUATION

In the financial literature there exist quite a number of methods of low-frequency spread estimation, a good overview of such estimators one can find, for example, in [Anand and Karagozoglu 2006], [Corwin and Schultz 2012] and [Goyenko et al. 2009]. In this paper I test the applicability of two of them on polish stock market: Roll estimator [Roll 1984] and Corwin-Schultz estimator[1] [Corwin and Schultz 2012]. The Roll method appeared as the first tool for measuring the bid-ask spread by means of easily available financial daily data. Although nowadays intraday data are widely available, researchers still frequently use this estimator or its extensions in applications such as, for example, asset pricing and testing market efficiency [Corwin and Schultz 2012]. The popularity of the Roll estimator may result from the fact that it enables the rapid measurement of transaction costs solely on the basis of the observed prices [Doman 2011]. As describes [Roll 1984], the method requires only the prices, so it is very cheap. It requires two major assumptions [Roll 1984]:

1. The asset is traded in an informationally efficient market.
2. The probability distribution of observed price changes is stationary (at least for short intervals).

The Roll estimate of the spread understood as the relative proportion of the actual or theoretical price is given as follows:

$$S = 2\sqrt{-\operatorname{cov}\left(R_t, R_{t+1}\right)} \tag{1}$$

where R_t and R_{t+1} are rate of returns over a day t and $t+1$ respectively.

[1] also known as high-low estimator

If the returns are arithmetic then the formula holds approximately [Roll 1984], it holds exactly for logarithmic returns [Doman 2011]. The elegant derivation of the Roll formula can be found in [Zhang and Hodges 2012].

The method of Roll was later modified and extended. [Bleaney and Li 2014] even classify the bid-ask estimators as the Roll family of estimators and the other ones. Although simple and based on reasonable assumptions, the Roll method is strongly impaired by the fact that in about half the cases the autocovariance is positive The common ad hoc way to deal with this problem is to set Roll spreads to zero in these cases [Corwin and Schultz 2012].

[Corwin and Schultz 2012] spread estimator uses the daily high and low prices to estimate the spread. It assumes that:

1. the daily high price is a buyer-initiated trade and the daily low price is a seller-initiated trade
2. the percentage spread is constant over the 2-day estimation period
3. the true, unobserved price follows a geometric Brownian motion with zero drift

Especially the first assumption is a brilliant idea which seems to be very well adjusted to market reality.

Let H_t^0 and L_t^0 denote the observed daily high and low price, respectively, on day t. Moreover, let $H_{t,t+1}^0$ and $L_{t,t+1}^0$ be observed high and low, respectively, over the 2 days t and $t+1$. Then, the parameters β and γ are defined as follows:

$$\beta = E\left(\left[\ln\left(\frac{H_t^0}{L_t^0}\right)\right]^2 + \left[\ln\left(\frac{H_{t+1}^0}{L_{t+1}^0}\right)\right]^2\right) \text{ and } \gamma = \left[\ln\left(\frac{H_{t,t+1}^0}{L_{t,t+1}^0}\right)\right]^2. \tag{2}$$

The value S of the Corwin-Schultz spread estimator in its original form is determined by the equations [Corwin and Schultz 2012]:

$$S = \frac{2\left(e^{\alpha}-1\right)}{1+e^{\alpha}} \tag{3}$$

$$\alpha = \frac{\sqrt{2\beta}-\sqrt{\beta}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma}{3-2\sqrt{2}}}\,. \tag{4}$$

The last equation can be, by means of elementary transformations, significantly simplify and present the formula for α in the following, more appealing form [Karnaukh et al. 2014]:

$$\alpha = \left(1+\sqrt{2}\right)\left(\sqrt{\beta}-\sqrt{\gamma}\right). \tag{5}$$

In practice, β can be calculated as the sum over two days of the squared daily natural logarithms of the proper high/low ratios [Karnaukh et al. 2014]. The results of [Corwin and Schultz 2012] suggest that such approach produces more

accurate estimates of average monthly spreads than approximation of the expected value by averaging β over the month.

There are a number of implicit assumptions underlying the high–low spread estimator which are usually not true in practice. In order to benefit effectively from the considered method it is good to make the following adjustments which are described in detail in [Corwin and Schultz 2012]:

- adjustment for overnight price changes
- adjustment for the same price for all daily trades
- adjustment for the negative values of the estimates

To evaluate the performance of bid–ask spread estimators it is natural to examine two issues:

1. how much the estimates differ from the real bid-ask spread
2. to what extent they exhibit the same behavior as the spread in different trading environments and time periods.

With regard to the first issue, one uses mean absolute error (MAE) or root mean squared error. Both measures capture both the bias in the estimate and its variability. Lower values of the considered errors indicate that the estimated bid-ask spreads are closer to the actual spread. Second issue can be studied by calculation the correlations between the estimated and actual spreads. The higher the correlation the closer similarity of behavior od estimates and true spreads. [Corwin and Schultz 2012] carefully study the second issue and find that the accuracy of their model in terms of correlation is very good. Interestingly, they devote very little attention to the question of error of estimation. They write in the footnote that calculated mean absolute errors based on the difference between monthly spread estimates for each of considered by them estimators and monthly effective spreads from TAQ (trades and quotes database) across all sample months, and present the results of these calculations. For their high-low method, mean absolute error was 0,9%. Taking into account that the simple average effective spread from TAQ across all stock-months was 2.38%, one can easily compute that the ratio of these values which is the relative error is about 38%. It is not dramatically bad but the simulations of [Bleaney and Li 2014] reveal that the Corwin-Schultz estimator may be yet seriously biased. When it comes to the Roll method, it is rather commonly known that is far from satisfactory with respect to both aspects of performance assessment.

EVALUATION OF SPREADS OF THE STOCKS LISTED ON THE WARSAW STOCK EXCHANGE

The estimation concerns the stocks quoted on the Warsaw Stock Exchange (WSE) in 2013. From 460 stocks of which average annual spreads have been published in WSE Statistic Bulletin for the year 2013 I took a random sample of 30 stocks. The average annual spread from the bulletin is a benchmark by which the

accuracy of low-frequency estimation is measured. The requirements from a share to be taken into account in drawing were very low. They were the following:

1. the full record of daily data (247 days) and non-ambiguous name, in the quotations archive in the portal http://www.gpwinfostrefa.pl
2. non-zero spread in 2013 WSE Statistic Bulletin

This approach quite strongly differs from standard methods of stock selection described in academic literature where usually the criteria are much more demanding. But in my opinion such high requirements are designed for the comfort of researchers rather than for managers and analysts of real financial markets. Of course, applying the given estimator to some set of data may be theoretically unjustified, though numerically possible. But also then the obtained result is the indicator of the method accuracy. The complete results are presented in table 1.

Table 1. Annual average spreads and Roll and Corwin-Schultz estimations

Stock name	Average spread	Roll estimation	Corwin-Schultz estimation
ALCHEMIA	0,87%	0,78%	0,94%
AVIASG	1,59%	0,00%	0,77%
BSCDRUK	3,56%	2,09%	1,45%
BUDIMEX	0,58%	0,00%	0,64%
BYTOM	2,26%	3,95%	1,65%
CASHFLOW	3,41%	0,00%	1,40%
COALENERG	1,58%	0,00%	2,11%
ERBUD	1,61%	0,00%	1,26%
FERRUM	1,82%	2,09%	1,05%
GETINOBLE	0,60%	1,40%	0,72%
INPRO	3,59%	1,57%	2,59%
INVISTA	3,35%	2,48%	2,28%
IQP	2,72%	2,10%	1,55%
IVMX	2,87%	2,03%	1,95%
KDMSHIPNG	4,18%	0,46%	1,90%
KOMPAP	0,91%	3,35%	1,35%
LENTEX	0,72%	0,00%	0,42%
MAKRUM	2,15%	0,00%	1,31%
MILLENNIUM	0,41%	0,00%	0,81%
MIT	1,92%	1,32%	2,01%
PCCINTER	2,91%	1,04%	0,89%
MOSTALZAB	1,06%	0,00%	1,10%
PLAZACNTR	2,53%	4,81%	2,90%
PRAGMAFA	1,25%	1,74%	0,29%
SANWIL	4,45%	3,77%	3,23%
SKOTAN	0,81%	0,00%	0,97%
TESGAS	1,93%	0,00%	1,14%
TRITON	4,01%	0,00%	1,82%
VARIANT	2,68%	0,00%	1,72%
VINDEXUS	1,89%	0,00%	1,03%

Source: WSE Statistic Bulletin 2013 and own calculations based on data from www.gpw.infostrefa

In 14 cases the Roll estimates is 0 which means that the autocovariance is positive and it is consistent with common knowledge that positive autocovariance happens in about half cases. also calculated the correlation coefficients r_{Roll} and r_{CS} between the averages spreads from 2013 WSE Bulletin and the Roll and the Corwin-Schultz estimators, respectively and verified their significance by the standard significance test with test statistics $t = r\sqrt{\frac{n-2}{1-r^2}}$ where r is the correlation coefficient and n is a number of observations. The following results were obtained (p is the p-value):

$$r_{Roll} = 0{,}290;\ p = 0{,}12006$$

$$r_{CS} = 0{,}790,\ p = 0{,}00001$$

It is apparent from the carried out calculations that the Corwin-Schultz estimator is much better than the Roll one with respect to the similarity if behavior to the actual spread. The results do not even confirm the relationship between the Roll estimates and actual spreads. One may suspect that this is due to small sample but on the other hand it was large enough to prove strong interdependence between the Corwin-Schultz estimator and true values of spread.

Then, the mean absolute errors MAE_{Roll} and MAE_{CS} of the Roll and the Corwin-Schultz estimators were computed:

$$MAE_{Roll} = 1{,}51\%$$

$$MAE_{CS} = 0{,}85\%\,.$$

The results provide clear evidence that the Corwin- Schultz method outperforms the Roll estimation with regard to error of measurement. The ratio of MAE_{CS} and the average value of actual spread in the sample is below 40% which is not small value but seeming acceptable from the point of view of the bid-ask spread assessment in practice.

CONCLUSION AND PERSPECTIVE FOR RESEARCH

The paper by demonstrating the difficulties with high-frequency data argues that low-frequency estimation is important. Two popular estimators based on daily data are applied to the polish shares. The obtained results show that the Corwin-Schultz estimator is an effective tool to measure the bid-ask spread for stocks quoted on WSE on the basis of daily data and it is unambiguously better than the popular Roll method, in this task. The aim was realized by confirmation of the usefulness of the Corwin-Schultz method in the polish stock market. Moreover, the performed calculations for the polish market, confirmed the properties of both

estimators known from the previous literature.. The study was done with the small sample size and concerned only one year. The natural extension of this article could be considering larger sample and longer time period. Moreover, it interesting to see how the situation changes when not annual but, as it is in [Corwin and Schultz 2012], monthly average spreads are considered.

REFERENCES

Anand A., Karagozoglu A. K. (2006) Relative performance of bid–ask spread estimators: Futures market evidence, International Financial Markets, Institutions and Money, 16, pp. 231 – 245.

Bleaney M., Li Z. (2013) The performance of bid-ask spread estimators under less than ideal condition, http://www.nottingham.ac.uk/economics/documents/discussion-papers/13-05.pdf

Bleaney M., Li Z. (2014) A new spread estimator, http://www.nottingham.ac.uk/economics/documents/discussion-papers/14-01.pdf

Corwin S. A., Schultz P. (2012) A Simple Way to Estimate Bid-Ask Spreads from daily High and Low Prices, Journal of Finance, 67, pp. 719 – 759.

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

Elton E. J, Gruber M. J, Brown S. J., Goetzmann W. N. (2010) Modern Portfolio Theory and Investment Analysis, John Wiley & Sons, Hoboken.

Gniadkowska A. (2012) Wpływ płynności obrotu na kształtowanie się stopy zwrotu z akcji notowanych na Giełdzie Papierów Wartościowych w Warszawie, Zarządzanie i Finanse, 10, pp. 563 – 570.

Goyenko R. Y., Holden C. W., Trzcinka C. A. (2009) Do liquidity measures measure liquidity?, Journal of Financial Economics, 92, pp. 153 – 181.

Grinold R. C., Kahn R. N. (2000) Active Portfolio Management, McGraw-Hill, New York.

Holden C. (2009) New low-frequency spread measures, Journal of Financial Markets, 12, pp. 778 – 813.

Karnaukh N., Ranaldo A., Söderlind Paul (2014) Understanding FX Liquidity, http://www1.vwa.unisg.ch/RePEc/usg/sfwpfi/WPF-1315.pdf

Olbryś J. (2012) Tarcie w procesach transakcyjnych i jego konsekwencje, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, 254, pp. 181–189.

Roll R. (1984) A simple implicit measure of the effective bid-ask spread in an efficient market, Journal of Finance, 39, pp. 1127–1139.

Sharpe W. F., Alexander G., J., Bailey J. V. (1999) Investments Prentice Hall, Upper Saddle River.

WSE Statistics Bulletin, http://www.gpw.pl/pub/statystyki_roczne/2013_GPW.pdf

Zhang H., Hodges S. (2012) An Extended Model of Effective Bid-ask Spread, http://www.cass.city.ac.uk/__data/assets/pdf_file/0007/128068/H.Zhang.pdf

Zivot E. (2005) Analysis of High Frequency Financial Data: Models, Methods and Software. Part I: Descriptive Analysis of High Frequency Financial Data with S-PLUS, http://faculty.washington.edu/ezivot/research/hflectures.pdf

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 144 –156

THE REFERENCE POINT METHOD APPLIED TO DECISION SELECTION IN THE PROCESS OF BILATERAL NEGOTIATIONS

Andrzej Łodziński
Department of Econometrics and Statistics
Warsaw University of Life Sciences - SGGW
e-mail: andrzej_lodzinski@sggw.pl

Abstract: The paper presents a method of supporting the decision selection in the process of bilateral negotiations. The negotiation process is modeled as using a multi-criteria optimization. The method of finding solutions is the interactive selection process of some proposals. The parties shall submit their proposals to the subjects of the negotiations. These proposals are parameters of the multi-criteria optimization tasks. Selection of solutions is accomplished by solving the optimization task with parameters that define the aspirations of each party involved in the negotiations. Finally, evaluation of the solutions obtained by the parties is done.

Keywords: negotiation process, multi-criteria optimization, equitably effective decision, scalar function, method of the reference point

INTRODUCTION

The paper presents a method of supporting the decision selection in the process of bilateral negotiations. Negotiations serve to agree the decisions when different interests of participants occur. Negotiations are carried out to reach a more favorable result than that which can be achieved without negotiation. Negotiating parties could benefit, coming to the agreement with each other, in comparison to the situation when they act separately. Well arranged agreement is better for the parties than no agreement at all, and some agreements are more favorable for both parties than others. In the complex negotiations, the parties not only want to reach an agreement, but they are looking for the optimal agreement – i.e. the agreement that would be the best for both parties.

Negotiations are characterized by a lack of clear solutions and the necessity of taking into account the preferences of all parties. The process of bilateral negotiations can be modeled using game theory. The solution is then the Nash cooperative solution or Raiffy-Kalai-Smorodinsky solution [Luce, Raiffa 1966], [Malawski i in. 1997], [Raiffa 1998], [Straffin 2004].

The work is devoted to apply a multicriteria optimization to support decision making in negotiation process. The process of bilateral negotiations is modeled as a multi-criteria optimization tasks. The method of decision selection is based on an interactive selection of some proposals of solutions, i.e. the algorithm requires the reaction of parties during this process. The parties submit their proposals for the subjects of negotiations; these proposals are parameters of the multi-criteria optimization task; this way the task is solved. Then, the parties evaluate the solution: they accept it or reject it. In the second case, the parties shall submit new proposals - the new values of parameters and the problem is solved again for these new parameters. The process of selection of solution is not a one-time process, but an iterative process of learning by parties about the negotiated problem.

MODELING OF THE NEGOTIATION PROCESS

The negotiation process is modeled as an interactive decision-making process. Each party presents its proposals of solutions. The negotiation process is then the process of seeking a common decision, which reconciles the interests of both parties. The parties are trying to find a common compromise solution. Decisions require voluntary consent of both parties and are taken together, not unilaterally. Both parties have to accept those decisions.

During the negotiation process, there are many different purposes, which are implemented using the same set of feasible solutions. The negotiation process is modeled by introducing a decision variable that describes the solution as well as two evaluation functions evaluating the solution from the point of view of each party. During the negotiations, each proposal is evaluated by either party by its evaluation function. Such a function is a measure of satisfaction of a party with a given solution. It evaluates the degree of realization of each subject of negotiations by each party. Higher value of the function means higher satisfaction of a party, so each function is maximized. The basis for evaluation and solution selection are two functions of evaluation - the criteria for both parties.

We assume the following terms:

party 1 and party 2 - parties in negotiations,
n - the number of subjects for negotiation,
$x \in X_0$ – solution - a decision, the parties of which are to agree, belonging to a set of feasible decisions

$X_0 \subset R^n$, $x = (x_1, x_2, ..., x_n)$ - each coordinate $x_i, i = 1,...,n$ defines $i-th$ subject of negotiations,

$f1: X_0 \to R^{m1}$ – the evaluation function of decision x by party 1,

$f1 = (f1_1, f1_2, ..., f1_{m1})$ - vector function ,

which determines the degree of realization of solution by party 1,

$f2: X_0 \to R^{m2}$ – the evaluation function of decision x by party 2,

$f2 = (f2_1, f2_2, ..., f2_{m2})$ vector function ,

which determines the degree of realization of solution by party 2.

The problem of a decision selection has the multi-criteria character. The decision is characterized by a complex evaluation function, wherein the first component is a function of evaluation of the decision by the first party and the second component is a function of evaluation of the decision by the other party. Each party wants to maximize its evaluation function, but it must take into account the existence of the other party. The selection of solution is done by using both evaluation functions.

The negotiation process is considered as a task of multi-criteria optimization with the function of purpose $f = (f1, f2)$:

$$\max_x \{(f1(x), f2(x)): \quad x \in X_0\} \tag{1}$$

where:

$x \in X$ – vector of decision variables

$f = (f1, f2)$ – the vector function which maps the decision space X into evaluation space $Y_0 \subseteq R^{m1+m2}$,

X_0 – the set of feasible decisions.

Task (1) is to find such feasible decision $\hat{x} \in X_0$, for which $m1 + m2$ evaluations takes the best values.

Task (1) is considered in the evaluation space, i.e., the following task is considered:

$$\max_x \{y = (y1, y2): \quad y \in Y_0\} \tag{2}$$

where:

$x \in X$ – vector of decision variables,

$y = (y1, y2) = (y_1, ... y_{m1}, y_{m1+1}, ..., y_{m1+m2})$ – vector quality indicator;

individual coordinates $y_i = f_i(x), \quad i = 1,2,..,m1 + m2$ represent single scalar criteria, first $m1$ coordinates are evaluation criteria of a solution by party 1, the next coordinates are evaluation criteria of a solution by party 2,

$m = m1 + m2$ - dimension of the criteria space,
$Y_0 = (f1, f2)(X_0)$ – a set of achievable vectors of evaluation.

The set of achievable results Y_0 is given in the implicit form - through a set of feasible decisions X_0 and mapping of a model $f = (f1, f2)$. To determine the value y, the simulation of the model is needed: $y = (f1, f2)(x)$ *for* $x \in X_0$.

The purpose of task (1) is to help in the selection of such a decision, which takes into account the best interests of both parties [Lewandowski and Wierzbicki 1989], [Ogryczak 2002], [Wierzbicki 1984].

EQUITABLY EFFICIENT SOLUTION

The solution in the negotiation process should satisfy certain properties that the parties accept as reasonable. The solution should be:

- optimal solution in the sense of Pareto – i.e. such that you can not improve the solution for one party without worsening the solution for the other party,
- symmetric solution – i.e. that it should not depend on the way the parties are numbered, no one is more important, the parties are treated in the same way in the sense that the solution does not depend on the names of the party or other factors specific to a given party,
- equalizing solution - that is, a vector that has less variation of coordinates of evaluation is preferred in comparison to the vector with the same sum of coordinates, but with a greater diversity of coordinates,
- the solution should take into account the strength of the parties in the negotiations.

A decision, which satisfies these conditions is an equitably efficient decision. This is Pareto-optimal decision which satisfies additional conditions – anonymity and the axiom of equalizing solution.

Not dominated results (Pareto - optimal) are defined as follows:

$$\hat{Y}_0 = \{\hat{y} \in Y_0 : (\hat{y} + \tilde{D}) \cap Y_0 = \varnothing\} \tag{3}$$

where:

$\tilde{D} = D \setminus \{0\}$ – positive cone without the top. As a positive cone, it can be adopted $\tilde{D} = R_+^m$ [Górecki 2000], [Lewandowski and Wierzbicki 1989].

In the decision space, the appropriate feasible decisions are specified. The decision $\hat{x} \in X_0$ is called effective decision (Pareto - optimal), if the corresponding vector of evaluation $\hat{y} = f(\hat{x})$ is a not dominated vector.

In the multi-criteria problem (1), which is used to select a decision in the negotiation process, the relation of preferences should satisfy additional properties: anonymity property and property of equalizing solution.

This preference relation is called an anonymous relation if, for every assessments $y=(y_1,y_2,...,y_m)\in R^m$ and for any permutation P of the set $\{1,...,m\}$, the following property holds:

$$(y_{P(1)},y_{P(2)},...,y_{P(m)})\approx(y_1,y_2,...,y_m) \tag{4}$$

No distinction is made between the results that differ in their arrangement. Evaluation vectors having the same coordinates, but in a different manner are identified.

Relation of preferences satisfies the axiom of equalizing transfer if the following condition is satisfied:

for the evaluation vector $y=(y_1,y_2,...,y_m)\in R^m$:

$$y_{i'}>y_{i''}\Rightarrow y-\varepsilon\cdot e_{i'}+\varepsilon\cdot e_{i''}\succ y \text{ for } 0<y_{i'}-y_{i''}<\varepsilon \tag{5}$$

Equalizing transfer is a slight deterioration of a better coordinate of the evaluation vector and simultaneously improvement of a poorer coordinate, giving the evaluation vector strictly preferred in comparison to the initial evaluation vector. This is a structure of equalizing – the evaluation vector with less diversity of coordinates is preferred in relation to the vector with the same sum of coordinates, but with a greater diversity of coordinates.

Not dominated vector satisfying the anonymity property and the axiom of equalizing transfer is called a equitably not-dominated vector. The set of equitably not dominated vectors is denoted by $\hat{Y}_{0w}$. In a decision space, the equitably efficient decisions are specified. The decision $\hat{x}\in X_0$ is called the equitably efficient decision, if the corresponding evaluation vector $\hat{y}=f(\hat{x})$ is an equitably not dominated vector. The set of equitably efficient decisions is denoted by $\hat{X}_{0w}$ [Ogryczak 2002].

The relation of equitable domination can be expressed as the relation of inequality for cumulative, ordered evaluation vectors. This relation can be determined with the use of transformation $\overline{T}:R^m\to R^m$ that accumulates coordinates of decreasing order in the evaluation vector.

The transformation $\overline{T}:R^m\to R^m$ is defined as follows :

$$\overline{T}_i(y)=\sum_{l=1}^{i}T_l(y) \quad \text{for } i=1,2,...,m \tag{6}$$

where:

$T(y)$ is the vector with decreasing ordered coordinates of the vector y, i.e. $T(y) = (T_1(y), T_2(y), ..., T_m(y))$, where $T_1(y) \le T_2(y) \le ... \le T_m(y)$ and
there is a permutation P of the set $\{1,...,m\}$, such that $T_i(y) = y_{P(i)}$ for $i = 1,..,m$.

The relation of equitable domination $\succeq_w$ is a simple vector domination for the evaluation vectors with coordinates which are accumulated values of ordered evaluation vector [Ogryczak 2002].

The evaluation vector y^1 dominates in equitable way the vector y^2 if the following condition is satisfied:

$$y^1 \succeq_w y^2 \Leftrightarrow \overline{T}(y^1) \ge \overline{T}(y^2) \tag{7}$$

Solving the problem of decision selection in the negotiations process consists in determination of equitably efficient decision that satisfies the preferences of parties.

SCALARING THE PROBLEM

For determination of equitably efficient solutions of multi-criteria task (1) a specific multi-criteria task is solved. It is the task with the vector function of the cumulative, ordered evaluation vectors, i.e. the following task:

$$\max_y \{(\overline{T}_1(y), \overline{T}_2(y), ..., \overline{T}_m(y)) : \quad y \in Y_0\} \tag{8}$$

where:
$y = (y_1, y_2, ..., y_m)$ – evaluation vector,
$\overline{T}(y) = (\overline{T}_1(y), \overline{T}_2(y), ..., \overline{T}_m(y))$ cumulative, ordered evaluation vector,
Y_0 – set of achievable evaluation vectors.

An efficient solution of multi-criteria optimization tasks (8) is an equitably efficient solution of the multi-criteria task (1).

To determine the solution of a multi-criteria task (8) the scalaring of this task with the scalaring function $s : Y \times \Omega \to R^1$ is introduced:

$$\max_x \{s(y, \overline{y}) : x \in X_o\} \tag{9}$$

where:
$y = (y_1, y_2, ..., y_m)$ – evaluation vector,
$\overline{y} = (\overline{y}_1, \overline{y}_2, ..., \overline{y}_m)$ – control parameters for individual evaluations.

It is the task of single objective optimization with specially created scalaring function of two variables - the evaluation vector $y \in Y$ and control parameter $\bar{y} \in \Omega \subset R^m$. It is the function $s : Y \times \Omega \to R^1$. The parameter $\bar{y} = (\bar{y}_1, \bar{y}_2, ..., \bar{y}_m)$ is available to the parties, it allows them to review the set of equitably efficient solutions.

A optimal solution of the task (9) should be a solution to the multiple criteria task (8). A scalaring function should satisfy certain properties - property of completeness and property of sufficiency. The property of sufficiency means that for each control parameter $\bar{y}$ the solution of the scalaring task is the equitably efficient solution, i.e. $\hat{y} \in \hat{Y}_{0w}$. The property of completeness means, that by appropriate changes of the parameter $\bar{y}$ it can be achieved any solution $\hat{y} \in \hat{Y}_{0w}$. Such a function completely characterizes the equitably efficient solutions. Inversely, each maximum of such a function is an equitably efficient solution. Each equitably efficient solution can be achieved with some appropriate values of control parameters $\bar{y}$.

Complete and sufficient parameterization of the set of equitably efficient solutions $\hat{Y}_{0w}$ can be achieved, using the method of reference point for the task (8). This method makes use of aspiration levels as control parameters. Aspiration levels are such values of evaluation function that satisfy the decision owner.

The scalaring function in the method of reference point is as follows:

$$s(y, \bar{y}) = \min_{1 \le i \le m} (\bar{T}_i(y) - \bar{T}_i(\bar{y})) + \varepsilon \cdot \sum_{i=1}^{m} (\bar{T}_i(y) - \bar{T}_i(\bar{y})) \tag{10}$$

where:

$y = (y_1, y_2, ..., y_m)$ – evaluation vector,

$\bar{T}(y) = (\bar{T}_1(y), \bar{T}_2(y), ..., \bar{T}_m(y))$ - cumulative, ordered evaluation vector,

$\bar{y} = (\bar{y}_1, \bar{y}_2, ..., \bar{y}_m)$ – vector of aspiration levels,

$T(\bar{y}) = (T_1(\bar{y}), T_2(\bar{y}), ..., T_m(\bar{y}))$ - cumulative, ordered vector of aspiration levels,

ε – arbitrary small, positive adjustment parameter.

Such scalaring function is called function of achievement. The aim is to find a solution that approaches as close as possible the specific requirements – the aspiration levels.

Maximizing this function with respect to y determines equitably efficient solution $\hat{y}$ and the equitably efficient decision $\hat{x}$. Note, the equitably efficient decision $\hat{x}$ depends on the aspiration levels $\bar{y}$ [Lewandowski and Wierzbicki 1989], [Ogryczak 2002], [Wierzbicki 1984].

SET OF NEGOTIATIONS

The aim of the complex negotiations is not only the achievement of an agreement between the parties, even if it is beneficial for both parties, but finding a solution that meets the expectations of parties as much as possible and, if it is not worse than a solution attainable without negotiations.

Before starting the negotiations, parties should consider what is the result they can achieve if negotiations are not successful - the status quo point. This point is the result which can be achieved by each party without negotiation with the other one. If the parties can achieve the result $ys = (y1s, y2s)$ without negotiations - part 1 can achieve the result $y1s$, part 2 - the result $y2s$, then, no one party will agree to the worse result. During negotiations, parties want to improve the solution in relation to this point. The status quo point determines the strength of the parties in the negotiations and, what is their impact on the result.

The set of negotiations is a collection of equitably dominated evaluation values dominating the status quo point.

The set of negotiation is as follows:

$$B(\hat{Y}_{OW}, ys) = \{\hat{y} = (\hat{y}1, \hat{y}2) \in \hat{Y}_{OW} \wedge \hat{y}1 \geq y1s \wedge \hat{y}2 \geq y2s\} \tag{11}$$

where:

$\hat{y} = (\hat{y}1, \hat{y}2) \in \hat{Y}_{0W}$ – equitably not-dominated vector for the set $\hat{Y}_{0W}$,

$ys = (y1s, y2s)$ – status quo point - the result, which can be achieved by both parties without agreement.

A set of negotiations embraces the points from the set of equitably not-dominated results, which give each party at least as much as it can achieve individually (without negotiation).

The parties wish to find such a decision, $\hat{x} \in X_0$, that the corresponding evaluation vector $\hat{y} = (\hat{y}1, \hat{y}2) = (f1(\hat{x}), f2(\hat{x}))$ belongs to the set of negotiations $B(\hat{Y}_{0W}, ys)$ [Luce, Raiffa 1966], [Raiffa 1998].

METHOD OF SOLUTION SELECTION

The solution to multi-criteria optimization task (8) is the set of equitably efficient decisions. In order to resolve the problem there should be selected one solution that will be evaluated by both parties. Since the solution is a whole set, the parties shall select the solution with the help of an interactive computer system. Such a system allows us a controlled overview of the whole set. Each party attending the negotiation determines its proposed solutions as aspiration levels. These are the values of evaluation of individual negotiation issues, that each party would like to achieve. These values are control parameters of the scalaring

function. For these values the system indicates different equitably efficient solutions for analysis; they correspond to current values of the control parameters. The aim is to find solutions which meet, as close as possible, the specific requirements – aspiration levels.

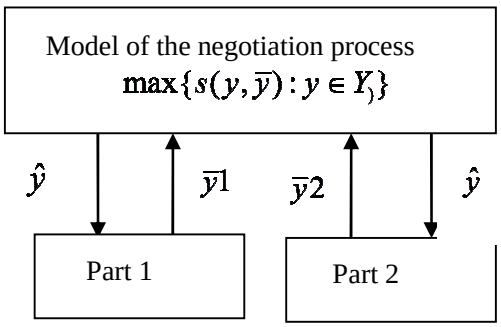
The method of decision selection is as follows:

1. The initial arrangements.
2. Iterative algorithm - proposals for further decisions.
   2.1. The interaction with the system - parties define their proposals for individual subjects of negotiations, as aspirations levels $\bar{y}1$ and $\bar{y}2$.
   2.2. Calculations - giving another solution from the set of negotiations, $\hat{y} = (\hat{y}1, \hat{y}2) \in B(\hat{Y}_{0W}, ys)$.
   2.3. Evaluation of the obtained solutions $\hat{y} = (\hat{y}1, \hat{y}2)$ – the parties may accept the solution or not. In the latter case, the parties shall submit new proposals – they provide new values of their aspiration levels $\bar{y}1$ and $\bar{y}2$ and a new solution is determined (see sec. 2.2).
3. Determination of the decision that meets the requirements of both parties.

A choice is not a single act of optimization, but a dynamic process of searching solutions. That means the parties learn and can change their preferences during the process. Comparing the results of the evaluation $\hat{y}1$ and $\hat{y}2$ to their aspiration points $\bar{y}1$ and $\bar{y}2$, we see that each party has information about what is and what is not achievable, and how far the parties` proposals $\bar{y}1$ and $\bar{y}2$ are from the possible solutions $\hat{y}1$ and $\hat{y}2$. This allows the parties to do appropriate modifications of their proposals: to provide their new aspiration points. These levels of aspiration are determined adaptively during the learning process. The process ends when the parties find such a decision, which allows them to achieve results that meet their aspirations or, in a sense, are as close as possible to these aspirations.

The method of finding a solution is show at Figure 1.

Figure 1. The method of decision selection



Source: own work

This method of decision selection does not impose any rigid scenario on parties and allows them to change their preferences while solving the problem. As we see, parties are learning about the problem during the negotiation. The computer does not replace the parties in selection of solution. It should be witnessed that the entire process of solution selection is controlled by both parties.

EXAMPLE OF BILATERAL NEGOTIATIONS

To illustrate the method of decisions selection in the process of bilateral negotiations the following example is shown [Górecki 2000].
The negotiation problem is as follows:
part 1 and part 2 – the parties attending the negotiations,
$n = 2$ - number of subjects for negotiations,
$x = (x_1, x_2) \in X_0$ - a decision that the parties are to agree and which belongs to set of feasible decisions $X_0 \subset R^2$,
x_1 - the decision concerning the first subject of negotiations,
x_2 - the decision on the second subject of negotiations,
$X_0 = \{(x_1, x_2) \in R^2 : -2 \cdot x_1 + 3 \cdot x_2 \le 27,\ \ 6 \cdot x_1 + 7 \cdot x_2 \le 175,\ \ 0 \le x_1 \le 21,\ \ 0 \le x_2 \le 13\}$
- the set of feasible decisions,
$f1: X \to R^1$ $f1(x) = \frac{20}{21} x_1 - \frac{2}{3} x_2$ – evaluation function of decision x by party 1,
$f2: X \to R^2$ $f2(x) = -\frac{4}{21} x_1 + \frac{2}{3} x_2$ – evaluation function of decision x by party 2,

$ys = (ys1, ys2) = (10, \quad 1)$ - status quo point.

The resulting task is a typical optimization problem with parameters that every party may dispose. It is solved with the use of a standard optimization software (Solver in Excel).

As a first step of the multi-criteria analysis, a single-criterion optimization of evaluation function of each party is done. The result is a matrix of implementation goals, containing the values of all criteria of each party, received during solving two single-criterion problems. This matrix allows for the estimation of the range of changes of each evaluation function on the feasible set, as well as provides some information about the conflictual nature of evaluation function. The matrix of the implementation of goals generates an utopia vector representing the best value of each separate criterion.

Table 1 The matrix of the implementation of goals with the utopia vector

Optimized criterion	Solution	
	$y1$	$y2$
Evaluation by the party 1 $y1$	20	-4
Evaluation by the party 2 $y2$	-2,85	7,52
Utopia vector	20	7,52

Source: own calculations

The Table 1 shows a clear advantage of party 1 in negotiations.

The multi-criteria analysis is presented in Table 2.

Table 2. Interactive analysis of seeking a solution

Iteration	Evaluation of party 1	Evaluation of party 2
	$y1$	$y2$
1. Aspiration point $\bar{y}$	20	7,52
Solution $\hat{y}$	15,33	0,66
2. Aspiration point $\bar{y}$	15	7
Solution $\hat{y}$	14,99	0,83
3. Aspiration point $\bar{y}$	14	6,5
Solution $\hat{y}$	13,33	1,33
4. Aspiration point $\bar{y}$	13	6,5
Solution $\hat{y}$	12,99	1,83
5. Aspiration point $\bar{y}$	12	5
Solution $\hat{y}$	11,99	2,33

Source: own calculations

At the beginning of the analysis, the parties specify their preferences as an aspiration point equal to the utopia vector. The resulting solution clearly prefers the first party and it is not acceptable to the other party – it is worse than its status quo point. To improve the solution, the first party reduces its requirements in the next iteration. The result is a slight deterioration of solution for the first party and a slight improvement of solution for the other party. The solution still does not exceed the status quo point of the other party. In subsequent iterations, the first party still reduces its requirements and the obtained solutions are now better than the status quo point of the other party. They become more rewarding for it. The analysis shows, that the solution depends in a significant way from the first party, which has a stronger position in negotiations and can impose a solution. For iteration 5 the relevant decisions are as follow $\hat{x}^5 = (18{,}81; 8{,}88)$.

Final selection of a specific solution depends on the parties` preferences. The example shows that this method allows the parties to explore the capabilities of decision-making during the interactive analysis and to search a mutually satisfactory solution.

CONCLUSIONS

The paper presents a method of modeling a process of bilateral negotiations in the form of multi-criteria optimization task. It is used to support the decision selection. The model of the negotiation process as a multi-criteria optimization task allows us to create variants of decision and to track their consequences.

The method of interactive analysis, based on the reference point, is applied for multi-criteria task with a cumulative, ordered evaluation vector. It allows us to determine solutions, well-tailored to the parties preferences. The numerical example shows that the proper computational task efficiently can be solved by the standard optimization software.

This procedure does not determine the final solution, but supports and teaches the parties about the specific negotiation problem. The final decision is to be taken by the parties involved in the negotiations.

REFERENCES

Fisher R., Ury W., Patron B. (2002) Getting to YES. Negotiating agreement without giving in. (in polish) PWE. Warsaw.

Górecki H. (2000) Optimization and control of dynamic systems. (in polish) Institutional Educational Science Publishing House of the Academy of Mining and Metallurgy. Kraków.

Lewandowski A. and Wierzbicki A. eds. (1989) Aspiration Based Decision Support Systems. Lecture Notes in Economics and Mathematical Systems. Vol. 331, Springer-Verlag, Berlin-Heidelberg.

Luce D., Raiffa H. (1966) Games and decisions. (in polish) PWN, Warsaw.

Malawski M., Wieczorek A., Sosnowska H. (1997) Competition and Cooperation. Game Theory in Economics and the Social Sciences. (in polish) PWN, Warsaw.

Ogryczak W. (2002) Multicriteria Optimization and Decisions under Risk. Control and Cybernetics, vol. 31 (2002) No. 4.

Raiffa H. (1998) The art. And Science of Negotiations. Harvard University Press, Cambridge Mass.

Straffin Ph., D. (2004) Game Theory. Scolar, (in polish) Warsaw.

Young H. P. (2003) Equity: In Theory and Practice. (in polish) Warsaw.

Wachowicz T. (2006) E-negotiations. modeling, analysis and support. Publisher University of Economics. Karol Adamiecki in Katowice, Katowice.

Wierzbicki A. (1984) Negotiation and mediation in conficts. Plural rationality and interactive decision processes, Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin Heidelberg New York Tokyo.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 157-166

THE CONVERGENCE OF THE ECONOMIC SIZE OF FARMS IN POLAND – THE ECONOMETRIC ANALYSIS

Joanna Muszyńska, Iwona Müller-Frączek
Department of Econometrics and Statistics, Nicolas Copernicus University
e-mail: Joanna.Muszynska@umk.pl, Muller@econ.umk.pl

Abstract: The research discussed in the article concerned the level of development of agriculture in Poland. The main aim of the study was to verify the hypothesis of convergence of the economic size of farms. The analysis confirmed the existence of β-convergence, however, showed no occurrence of σ-convergence and γ-convergence. Based on the results of the analysis of marginal vertical β-convergence, Poland was divided into two clubs with different paths of development. In one of them the process of convergence was confirmed, but in the second one the phenomenon of divergence was revealed.

Keywords: agricultural development, economic size of the farm, β-convergence, σ-convergence, γ-convergence

INTRODUCTION

The subject of research that has been discussed in the article was the level of development of agriculture in Poland, more precisely, the answer to the question whether private farms are able to reach a similar level of agricultural development, regardless of their location. In the paper, the economic size of the farm, that describes its potential production capacity, was adopted as a measure of development.

Since the authors' interest had focused on the equalization of the level of agricultural development between regions, the main aim of the study was therefore to verify the hypothesis about the convergence of economic size of farms in Poland. The average in the province, private farm was the research unit.

During the study, different types of convergence were analysed, both the basic – beta, sigma and gamma, as well as less popular – vertical marginal β-convergence and club convergence.

Economic size of farm does not reflect all aspects of the agricultural development. In other paper, the authors [Muszyńska, Müller-Frączek 2013] presented a broader approach to the analysis of the development of agriculture, in which the economic size of a farm was only one component of the synthetic variable.

THE ECONOMIC SIZE OF FARM[1]

FADN (The Farm Accountancy Data Network) is an European system for accountancy data collection from agricultural farms. It is one of the tools used by the EU for creating Common Agricultural Policy. Farms participating in FADN are classified according to Community Typology for Agricultural Holdings. One of the criterion for this classification is the economic size of the farm.

Economic size of a farm is a sum of all Standard Outputs[2] (SO) for all agricultural activities existing in that farm. It describes the potential production capacity of the farm. Farmers – FADN participants[3] – estimate the economic size of their farms, using standard output coefficients, on voluntary basis.

For the purposes of their study on the regional development of agriculture in Poland, the authors have constructed, for each province, a measure as similar as possible to the economic size, defined by FADN for a single farm[4]. Computations were based on the regional coefficients of standard output and the annual data on major crops and acreage of basic animal husbandry. All empirical data were derived from the Local Data Bank of CSO from the period 2004-2012. Due to the lack of data in public statistics, some products were not taken into account, e.g. mushrooms, flowers and ornamental plants. Then, the economic size of the region was divided by the number of farms in the province. Estimated this

[1] Detailed information on the topics covered in this chapter can be found in: Goraj L., Bocian M., Cholewa I., Nachtman G., Tarasiuk R. (2012) Współczynniki Standardowej Produkcji "2007" dla celów Wspólnotowej Typologii Gospodarstw Rolnych, Institute of Agricultural and Food Economics, National Research Institute, Warsaw.

[2] Standard Output is defined as the average monetary value of the agricultural output of an agricultural product (crop or livestock) over the reference period of 5 years, per 1 ha or 1 head of livestock per a year, in average production conditions in particular regions. Standard Output coefficients, used for the calculation of economic size, are different for the four Polish regions. In this way local conditions of each region are reflected.

[3] In Poland a group of approx. 12100 farms.

[4] Details of the calculations and the values of the economic size of the average farm in each province can be found in: Müller-Frączek I., Muszyńska J. (2014) Zmiany wielkości ekonomicznej polskich gospodarstw rolnych w latach 2004-2012, The Annals of The Polish Association of Agricultural and Agribusiness Economists, vol. XVI, no.3, pp. 205-210.

way the economic size of the average farm in the region formed the basis of the further research.

RESEARCH METHODOLOGY

Verification of the hypothesis about assimilation of the regions in terms of the level of agricultural development proceeded on the basis of analysis of the various types of convergence of the economic size of the average farm. There were beta, sigma and gamma convergence, widely discussed in the literature[5]. The study was extended by the analysis of marginal vertical β-convergence. Based on its results, Poland was divided into two clubs with different paths of development. Then the hypothesis regarding club convergence also was verified.

β-type convergence

The phenomenon of β-convergence occurs when there is a constant over time, negative correlation between the level of the analysed process and its growth rate. It means the regions with initially lower level of the investigated process will catch up the better developed provinces. The analysis of this phenomenon was based on the dynamic panel data model in the form:

$$\ln \frac{Y_{it}}{Y_{it-1}} = \alpha_0 - \alpha_1 \ln Y_{it-1} + \eta_i + u_{it}, \tag{1}$$

where:

- Y – the analysed process,
- i – the number of the region, $i = 1, ..., N$,
- t – number of period $t = 1, ..., T$,
- η_i – group effects,
- u_{it} – error term.

A positive value of the parameter α_1, in equation (1), proves the existence of β-convergence, a negative value means the occurrence of divergence phenomenon. When the β-convergence occurs, the speed of convergence[6] to equilibrium – the point at which all the regions are at the same level of development – can be estimated as follows: $\beta = -\ln(1-\alpha_1)$.

In order to estimate parameters the dynamic panel data model, described by the equation (1), is transformed to the model:

[5] An extensive discussion of these topics can be found, inter alia, in Kusideł E. (2013) Konwergencja gospodarcza w Polsce i jej znaczenie w osiąganiu celów polityki spójności, Wydawnictwo Uniwersytetu Łódzkiego, Łódź, Wolszczak-Derlacz J. (2007) Wspólna Europa, różne ceny – Analiza procesów konwergencji, CeDeWu – Wydawnictwa Fachowe, Warszawa and Łaźniewska E., Górecki T., Chmielewski R. (2011) Konwergencja regionalna, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.

[6] The rate of catching up.

$$y_{it} = \alpha_0 + (1-\alpha_1) y_{it-1} + \eta_i + u_{it}, \quad (2)$$

where $y_{it} = \ln Y_{it}$.

Model parameters can be obtained using the Blundell and Bond System Generalized Method of Moments Estimator (GMM-sys). The quality of the estimated model is verified on the basis of statistical tests[7]. The basic ones are: the Arellano-Bond test for autocorrelation and the Sargan test of over-identifying restrictions. The last one evaluates the correctness of the selection of instrumental variables during estimation stage in the sense of their being uncorrelated with the error terms of the first difference model. The Arellano-Bond test verifies the assumption regarding autocorrelation of the model error term. The model is properly specified if the test provides no grounds for rejecting the null hypothesis about the absence of the second-order autocorrelation of the first difference model error term. Occurrence of the first-order autocorrelation resulting from the model construction is an expected phenomenon.

σ-type convergence

The existence of β-type convergence is not tantamount to the presence of σ-type convergence [Wolszcak-Derlacz 2007]. β-convergence is a necessary but insufficient condition for σ-convergence occurrence. The last one occurs when the diversification of values of analysed variable among regions decreases in time. In practice, this means a significant change in values of measures of dispersion or concentration of variable distribution.

In the study to evaluate changes of the dispersion of economic size distributions the variance equality test was applied [Lichtenberg 1994]. To verify the null hypothesis of the statistical insignificance of changes in the dispersion[8], the Snedecor's F statistic was used. The empirical value of F-statistic was calculated as the ratio of the variance for outermost (2004, 2012) periods of study.

Because the σ-convergence analysis using the variance equality test was based only on the first and the last year of the investigated period, the research had been expanded. To evaluate the changes in consecutive units of time the parameters of linear trend models for the variation coefficient (a measure of dispersion) and Gini coefficient (a measure of concentration) were estimated.

[7] The detailed descriptions of the estimation methods and the statistical tests used to assess the quality of the dynamic panel data models can be found, among others, in: Baltagi B.H. (2005) Econometric Analysis of Panel Data, John Wiley & Sons, Ltd., Chichester and Dańska-Borsiak B. (2011) Dynamiczne modele panelowe w badaniach ekonomicznych, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

[8] The null hypothesis of equality of variance in the examined periods means lack of σ-convergence / σ-divergence.

γ-type convergence

The subject of the next stage of the study was the analysis of γ-convergence. In order to verify the existence of this type of convergence the rankings of objects were created. Each province in each year of the study was assigned a rank in respect of an examined variable. Gamma convergence occurs when the regions have changed their ranking position.

As before, the analysis covered changes of the rankings for outermost years as well as the whole investigated period. In both cases, compatibility of the rankings was assessed using Kendall's coefficient of concordance.

Marginal vertical β-type convergence

In order to extend the analysis, contribution of individual regions to the process of β-convergence was also examined. The concept of marginal vertical β-type convergence, proposed by Batóg [2010], was used for that purpose. Coefficients of marginal vertical β-convergence (β_i) were calculated according to the formula:

$$\beta_i = \beta - \beta_i^{m-1}, \tag{3}$$

where β is a speed of convergence of all m regions and β_i^{m-1} means a speed of convergence of $m-1$ regions (without i province).

This method allowed to point out the regions that had a positive and negative influence on the process of alignment of the level of agricultural development.

Club convergence

The final step of the study was an attempt to divide Polish provinces into two groups (clubs) of a different nature from the point of view of convergence in the level of agriculture. The results of the analysis of the marginal vertical β-convergence were used for that purpose.

Depending on the sign of βi coefficient, calculated in accordance with equation (3), the clubs of negative and positive impact on the convergence process were distinguished. Then, for each club separately, beta, sigma and γ-type convergences were examined.

THE EMPIRICAL RESULTS

β-convergence analysis

The analysis of β-convergence was the first stage in the process of verification of the main hypothesis of work about assimilation of Polish

provinces in terms of agricultural development. The model of β-convergence, described by equation (2), was estimated. It took the following empirical form[9]:

$$\hat{y}_{it} = 0{,}557 + 0{,}949\, y_{it-1}. \quad (4)$$

The correctness of the estimated model was verified using statistical tests. The results are compiled in Table 1.

Table 1. The test results for model described by equations (4)*

test	value of the test statistics	p-value
α_0	0,488	0,6253
$1-\alpha_1$	8,831	0,0000
AR(1)	-2,527	0,0115
AR(2)	-1,412	0,1578
Sargan	14,677	0,9984
Wald	77,984	0,0000

*-verification was conducted at 5% level of significance

Source: own computations

All the tests confirmed the proper specification of the model. The Arellano-Bond test gave no arguments for rejecting the null hypothesis about the lack of the second-order autocorrelation of model error terms. Also the Sargan test provided no grounds for rejecting the null hypothesis. The instruments applied during the estimation process were not correlated with the error terms of the model. Significance of the parameter estimates was proved using the Wald test.

The characteristics of β-convergence are shown in Table 2. A positive value of the coefficient α_1 confirmed the hypothesis of the occurrence of β-convergence. The average speed of convergence was estimated at 5,25% per year. The time required to cover half of the distance[10] between the average level of economic size of farm and the economic size of farm in the steady state was estimated at 13 years.

Table 2. The characteristics of β-type convergence

α_1- model parameter	β – speed of convergence (%)	τ – half-life (in years)
0,051	5,25	13

Source: own computations

σ-convergence analysis

Since the presence of β-convergence is a necessary condition for the existence of σ-type convergence, a positive verification of the hypothesis of the

[9] All computations, presented in the paper, were performed in GRETL.

[10] Half-life was calculated as follows: $\tau = -\ln(0{,}5)/\beta$.

occurrence of that first phenomenon gave grounds to the next stage of study – analysis of σ-convergence.

In the analysis based on the variance equality test, covering the first and the last year of the study, there was no significant decrease in variance, and hence the hypothesis of σ-convergence has not been confirmed[11].

To assess the changes in consecutive units of time, the parameters of linear trend models of the variation coefficient and Gini coefficient were estimated. Table 3 presents the results.

Table 3. The parameters of linear trend models of the variation coefficient (V) and Gini coefficient (G)

measure	trend slope	p-value
variation coefficient (V)	0,0078	0,0003
Gini coefficient (G)	0,0042	0,0004

Source: own computations

Since the parameters of empirical models did not indicate a downward tendency (the slopes of the trends were statistically significant, but positive) it was concluded that there was no σ-convergence. Positive values of the parameter estimates of the time variable showed a slight increase in the dispersion of the examined phenomenon, i.e. the existence of σ-type divergence.

γ-convergence analysis

In order to determine the occurrence of γ-type convergence, Kendall's concordance coefficient was used. The analysis covered changes of the rankings for outermost years as well as the whole investigated period. In all cases, concordance between the rankings of regions was high and statistically significant[12]. That meant the lack of changes in the arrangements of regions – i.e. the lack of γ-type convergence[13].

The values of Kendall's concordance coefficients, describing the compatibility of the arrangements of regions, and the values of test statistics are compiled in Table 4.

[11] The empirical value of F-statistic amounted to 1,86 and provided no arguments for rejecting the null hypothesis about the lack of the sigma-type convergence.

[12] Since in case of small samples (e.g. the sample of 16 Polish provinces), χ^2-statistic, due to the adopted values, can lead to errors of I and II type, to test the significance of Kendall's concordance coefficients t-statistics also were used (see: Kusideł (2013) Konwergencja gospodarcza w Polsce i jej znaczenie w osiąganiu celów polityki spójności, Wydawnictwo Uniwersytetu Łódzkiego, Łódź).

[13] In all cases, the null hypothesis of the presence of γ-type convergence has been rejected.

Table 4. The values of Kendall's concordance coefficients and the test statistics

rankings	Kendall's concordance coefficients	χ^2-statistic	t-statistic
2004 and 2012	0,974	29,206	15,937
all	0,975	131,559	16,253

Source: own computations

Analysis of the marginal vertical β-convergence

Since the results of estimation of the model (2) have confirmed the occurrence of β-convergence of economic size of the average farms the next step of the study was to assess individual province contribution to this phenomenon. According to equation (3), coefficients of marginal vertical β-convergence were calculated. They were used to define a positive or negative impact of the region on the speed of convergence. The values of the coefficients are shown in Table 5.

Table 5. The values of coefficients of marginal, vertical β-convergence (%)

province	β_i	significance	province o	β_i	significance
dolnośląskie	-4,21	+	podkarpackie	-1,12	no
kujawsko-pomorskie	0,01	no	podlaskie	0,14	no
lubelskie	1,28	no	pomorskie	1,56	no
lubuskie	6,94	+	śląskie	0,45	no
łódzkie	-0,55	no	świętokrzyskie	-1,99	no
małopolskie	-2,34	no	warmińsko-mazurskie	-1,21	no
mazowieckie	1,57	no	wielkopolskie	-0,01	no
opolskie	-0,25	no	zachodniopomorskie	2,62	+

Source: own computations

The analysis of the coefficients' significance[14] revealed that lubuskie and zachodniopomorskie have contributed to the convergence of agricultural level of development whereas dolnośląskie province has had a negative impact on the speed of that process.

Analysis of the club convergence

The analysis of vertical β-convergence allowed to identify groups of regions that differed in the influence on the convergence process. The first one consisted of provinces with the negative impact[15]. There were the following regions: dolnośląskie, łódzkie, małopolskie, opolskie, podkarpackie, świętokrzyskie,

[14] The statistical significance of the coefficients was assessed by comparing the coefficients values and their standard deviation. The coefficients with absolute value greater than the standard deviation were considered to be statistically significant.

[15] β_i coefficient was negative.

warmińsko-mazurskie and wielkopolskie. The second group included regions that positively influenced the speed of convergence[16]. The “positive impact club” consisted of provinces: kujawsko-pomorskie, lubelskie, lubuskie, mazowieckie, podlaskie, pomorskie, śląskie and zachodniopomorskie.

It seemed natural to use these results in order to find the different path of development for each group, i.e. to apply the results of vertical convergence analysis as the basis for the analysis of club convergence.

The final step of the study was to analyse the convergence process in each of the suggested clubs. In the "positive impact club" the phenomenon of β-convergence with the rate of 17,6%, was confirmed. Nevertheless, in this club, as for the whole country, sigma or γ-convergence were not demonstrated. In the "negative influence club" the occurrence of β-divergence phenomenon with rate of 0,06% was revealed.

SUMMARY AND CONCLUSIONS

The results of the study, described in the paper, confirmed the occurrence of the phenomenon of β-convergence of the economic size of the average private farm in the regions. The analysis did not demonstrate the existence of σ-convergence or γ-convergence. Therefore the main hypothesis of the work regarding the assimilation process of economic size of farms between regions in Poland has not been unambiguously confirmed. However, the actual results of β-convergence seem to indicate that the processes of equalization of level of agriculture in Poland occur, but maybe due to a short period of the study it was impossible to verify them based on other types of convergence.

In addition, during the analysis of vertical marginal β-convergence, the regions, that have a significant positive impact on the process of aligning the level of agriculture in Poland, were indicated. These were the provinces: lubuskie and zachodniopomorskie. It was also pointed to dolnośląskie as the region that inhibits this process. Furthermore, based on the results of the analysis of the vertical marginal β-convergence, Poland was divided into two clubs with different paths of development. The results for one of the clubs were similar as for the whole country, i.e. the existence of β-convergence was confirmed, the occurrence of σ-convergence and γ-convergence were not demonstrated, whereas in the second club the phenomenon of β-divergence was revealed.

[16] Positive value of β_i coefficient.

REFERENCES

Baltagi B.H. (2005) Econometric Analysis of Panel Data, John Wiley & Sons, Ltd., Chichester.

Batóg J. (2010) Konwergencja dochodowa w krajach Unii Europejskiej. Analiza ekonometryczna, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Dańska-Borsiak B. (2011) Dynamiczne modele panelowe w badaniach ekonomicznych, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Goraj L., Bocian M., Cholewa I., Nachtman G., Tarasiuk R. (2012) Współczynniki Standardowej Produkcji "2007" dla celów Wspólnotowej Typologii Gospodarstw Rolnych, IERiGŻ PIB, Warszawa.

Kusideł E. (2013) Konwergencja gospodarcza w Polsce i jej znaczenie w osiąganiu celów polityki spójności, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Lichtenberg F. (1994) Testing the Convergence Hypothesis, The Review of Economics and Statistics, vol. 76, no. 3, pp.576-579.

Łaźniewska E., Górecki T., Chmielewski R. (2011) Konwergencja regionalna, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.

Müller-Frączek I., Muszyńska J. (2014) Zmiany wielkości ekonomicznej polskich gospodarstw rolnych w latach 2004-2012, The Annals of The Polish Association of Agricultural and Agribusiness Economists, vol. XVI, no.3, pp. 205-210.

Muszyńska J., Müller-Frączek I. (2013) The development of agriculture in Poland in the years 2004-2011 – the taxonomic and econometrics analyses, Quantitative Methods in Economics, vol. XIV, no.2, pp. 118-128.

Wolszczak-Derlacz J. (2007) Wspólna Europa, różne ceny – Analiza procesów konwergencji, CeDeWu – Wydawnictwa Fachowe, Warszawa.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 167-176

THE USE OF CSISZÁR'S DIVERGENCE TO ASSESS DISSIMILARITIES OF INCOME DISTRIBUTIONS OF EU COUNTRIES

Jarosław Oczki[1], Ewa Wędrowska[2]
Nicolaus Copernicus University in Toruń
[1] Department of Human Resource Management
e-mail: joczki@umk.pl
[2] Department of Econometrics and Statistics
e-mail: ewaw@umk.pl

Abstract: Income distributions can be described by measures of central tendency, dispersion, skewness, kurtosis or by indexes of polarization. In numerous studies, Gini coefficient and Lorenz curve have been used to investigate inequality of incomes. Income distributions can also be analysed in comparison to one another. In the article two measures belonging to Csiszár's divergence class have been used to identify the degree of differentiation of income distributions among the EU countries in 2005 and 2012. Similar and dissimilar countries with respect to distribution of income have been identified and the change of divergence of EU countries income distributions between 2005 and 2012 has been assessed. European Union Statistics on Income and Living Conditions (EU-SILC) dataset has been used.

Keywords: income inequalities, income distribution, Csiszár's divergence

INTRODUCTION

Income levels and income distributions draw attention of researchers, especially those analysing labour markets, social policy and poverty. Many of studies concern the processes of income convergence or divergence, while other explore the properties of income distributions, including income inequalities, e.g. (Jędrzejczak 2012) and (Quintano et al. 2009). In the article, Csiszár's divergence measure have been applied to assess income inequality in EU countries, and to identify the degree of differentiation of income distributions of all EU countries in

2005 and 2012. Similar approach to analyzing dissimilarity of distributions of economic variables has been used in (Tomczyk 2011), (Podolec et al. 2011) and (Wędrowska 2011).

MEASURES OF DIVERGENCE

The importance of measures of distance between probability distributions arises because of the role they play in problems of inference and discrimination [Ullah 1996]. Divergence measures based on the concept of information-theoretic entropy were first introduced in communication theory by Shannon in 1948 and later developed by Wiener in 1949. These types of measures describe the degree of similarity between a pair of probability distributions.

One of the most general probability measures which plays a significant role in information theory is the well known Csiszár's *f*-divergence [Csiszár 1967].

Csiszár's *f*-divergence between a pair of discrete probability distributions: $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ is defined as:

$$C_f(P,Q) = \sum_{i=1}^{n} q_i f\left(\frac{p_i}{q_i}\right), \tag{1}$$

where $f : [0,\infty) \to \Re$ is a convex function satisfying $f(1) = 0$, $f'(1) = 0$, $f''(1) > 0$ and at $x = 0$, $0 \cdot f\left(\frac{0}{0}\right) = 0$ and $0 \cdot f\left(\frac{p}{0}\right) = \lim_{x\to\infty} \frac{f(x)}{x}$ (Menéndez et al. 2003).

A number of information theory measures are merely the particular cases of Csiszár's *f*-divergence. A list of *f*-divergence measures is provided in (Taneja 2004), (Taneja 2008) and (Wędrowska 2012). In the article Jeffreys-Kullback-Leibler divergence (*J*-divergence) and Jensen-Shannon divergence (*JS*-divergence) have been used in order to measure the degree of similarity between a pair or multiple income distributions.

J-divergence is a function of Kullback-Leibler divergence, the *I*-divergence, or the relative entropy, which assesses the dissimilarity between a pair of probability distributions. The *I*-divergence is defined as:

$$I(P,Q) = \sum_{i=1}^{n} p_i \log_2\left(\frac{p_i}{q_i}\right), \tag{2}$$

for probability distributions $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$.
It is well known that *I*-divergence is non-negative, additive, but not symmetric. The *I*-divergence is coincident with Csiszár's *f*-divergence for convex function (Wędrowska 2012):

$$f_{KL} = x \log_2 x. \tag{3}$$

The sum of the two mentioned divergences is Kullback's symmetric divergence, also known as the *J*-divergence (Cavanaugh 1998). To obtain a symmetric measure, one can define:

$$J(P,Q) = I(P,Q) + I(Q,P) = \sum_{i=1}^{n}(p_i - q_i)log_2\left(\frac{p_i}{q_i}\right). \quad (4)$$

The *J*-divergence coincides with *f*-divergence for convex function (Reid et al. 2009):

$$f_J = (x-1)log_2 x. \quad (5)$$

The properties of *J*-divergence are discussed in (Seghouane et al. 2004), (Lefebvre et al. 2010) and (Taneja 2013).

Lin introduced an information-theory based divergence measure regarding two or more probability distributions (Lin 1991) known as Jensen-Shannon divergence. It is based on the Shannon entropy and is related to the Kullback-Leibler divergence. The *JS*-divergence is defined as:

$$JS(Q,P) = H\left(\frac{P+Q}{2}\right) - \frac{1}{2}(H(P) + H(Q)), \quad (6)$$

where $H(P) = \sum_{i=1}^{n} p_i log_2 p_i$ is the Shannon entropy.
Jensen-Shannon divergence is the difference between the Shannon entropy of the mean density and the mean value of their entropies.

The Jensen-Shannon divergence is a symmetrized and smoothed version of the Kullback-Leibler divergence:

$$JS(P,Q) = \frac{1}{2}\left[I\left(P,\frac{P+Q}{2}\right) + I\left(Q,\frac{P+Q}{2}\right)\right]. \quad (7)$$

It coincides with Csiszár's *f*-divergence for convex function (Taneja 2005):

$$f_{JS}(x) = \frac{x}{2}log_2 x + \frac{x+1}{2}log_2\left(\frac{2}{x+1}\right). \quad (8)$$

Discussion of properties of JS-divergence can be found in (Menéndez et al. 1997), (Lamberti 2008) and (Grosse 2002).
The generalization of JS-divergence is defined as (Lin 1991):

$$JS(Q,P) = H(\pi_1 P + \pi_2 Q) - \pi_1 H(P) - \pi_2 H(Q), \quad (9)$$

where $\pi_1, \pi_2 \geq 0$, $\pi_1, +\pi_2 = 1$ are arbitrary weights for the probability distributions *P* and *Q*. Since *H* is concave function, *JS*(*P*,*Q*) is nonnegative and equal to zero, when $P = Q$. For an arbitrary set of probability distributions $P_1, P_2, \cdots, P_m$ with weights $\pi_1, \pi_2, \cdots, \pi_m \geq 0$, $\sum_{i=1}^{m}\pi_1 = 1$, the Jensen-Shannon divergence is defined by:

$$JS(P_1, P_2, \cdots, P_m) = H(\sum_{i=1}^{m}\pi_i P_i) - \sum_{i=1}^{m}\pi_i H(P_i). \quad (10)$$

DIVERGENCE OF INCOME DISTRIBUTIONS OF EUROPEAN UNION COUNTRIES

In this part of the article an analysis of divergence of income distributions of the EU countries is carried out. The analysis is preceded by the investigation of income inequalities in the European Union countries in 2005 and 2012. In order to assess countries income inequalities, Gini coefficient – one of the most popular measures of income concentration, as well as Shannon entropy have been presented (table 1). Gini coefficient values have been taken from EU-SILC database (for disposable income after social transfers). Shannon entropy for each country has been calculated for income distributions represented by shares of national disposable income in the relevant decile as percentage of total national disposable income.

The value of disposable income after social transfers is dependent on:

- labour market outcomes, such as: wages of employees or profits of self-employed, which in turn can be a result of labour market institutions (e.g.: minimum wages, flexible employment contracts regulations), dispersion of qualifications, or discrimination, e.g. against immigrants or employees working in flexible employment forms,
- transfers, which are part of countries' tax and social policies.

High level of income inequality can be an effect of increased variation of wages (or profits), and (or) low degree of income redistribution achieved by social transfers, and fiscal policy in general.

Data in Table 1 show that in 2005-2012 period income inequalities decreased in majority of EU countries. Latvia and three Mediterranean countries: Portugal, Spain and Greece can be identified as those with highest income inequalities throughout the whole period. Two least wealthy countries in the EU – Bulgaria and Romania, which joined the block in 2007, both have above-average income inequalities in 2012. High inequalities and low level of average incomes in these countries indicate that there is a threat that substantial groups of their societies, earning incomes in the first several deciles of the income distribution, could be at high risk of poverty. This situation creates a challenge for economic and social policies pursued by the governments of these countries.

Gini coefficients, as well as entropy values also indicate countries with lowest income inequalities throughout the analysed period – two Scandinavian countries: Sweden and Finland, and three Central and Eastern European countries: Slovenia, Slovakia and Czech Republic.

Table 1. Income inequalities in European Union countries in 2005 and 2012

2005			2012		
Country	Entropy	Gini Coefficient	Country	Entropy	Gini Coefficient
Portugal	2.980	38.1	Latvia	3.031	35.7
Lithuania	3.019	36.3	Spain	3.045	35.0

2005			2012		
Country	Entropy	Gini Coefficient	Country	Entropy	Gini Coefficient
Latvia	3.022	36.2	Portugal	3.046	34.5
Poland	3.033	35.6	Greece	3.053	34.3
United Kingdom	3.046	34.6	Bulgaria	3.069	33.6
Estonia	3.057	34.1	United Kingdom	3.074	32.8
Italy	3.071	32.8	Romania	3.078	33.2
Greece	3.074	33.2	Estonia	3.084	32.5
Spain	3.090	32.2	Italy	3.086	31.9
Ireland	3.094	31.9	Lithuania	3.091	32.0
Cyprus	3.136	28.7	Cyprus	3.104	31.0
Belgium	3.145	28.0	Poland	3.106	30.9
Hungary	3.147	27.6	Croatia	3.110	30.5
France	3.149	27.7	France	3.110	30.5
Netherlands	3.157	26.9	Ireland	3.123	29.9
Malta	3.159	27.0	Denmark	3.138	28.1
Germany	3.163	26.1	Germany	3.145	28.3
Luxembourg	3.164	26.5	Luxembourg	3.146	28.0
Austria	3.166	26.2	Austria	3.149	27.6
Czech Republic	3.167	26.0	Hungary	3.158	26.9
Slovakia	3.168	26.2	Malta	3.160	27.1
Finland	3.169	26.0	Belgium	3.163	26.6
Denmark	3.192	23.9	Finland	3.169	25.9
Slovenia	3.192	23.8	Netherlands	3.173	25.4
Sweden	3.199	23.4	Slovakia	3.178	25.3
			Czech Republic	3.181	24.9
			Sweden	3.182	24.8
			Slovenia	3.195	23.7

Source: own calculations based on EU-SILC database

In 2005, Poland, with relatively high value of Ginni coefficient and small Shannon entropy, belonged to the group of the EU countries characterized by the largest income inequalities. After 2005, Gini coefficient in Poland had been falling gradually to reach a level close to EU average in 2012. In the whole period, Poland experienced the largest drop in that index. Also significant decreases in inequalities were observed in Lithuania and Portugal. The downward tendency of the values of Gini coefficient could have been observed in almost all countries with above-average initial levels of income inequalities. On the other hand, the largest increase in inequality between 2005 and 2012 occurred in Denmark, France and Spain.

In the next step of our analysis we identify the degree of divergence between the income decile distributions of EU countries. The Jesnen-Shannon divergence has been calculated for 25 countries in 2005 and for 28 countries in 2012 (27 EU member states in 2012 and Croatia which joined the EU in 2013). Comparison of values of *JS*-divergence (the bottom row of table 2) suggests that in 2005-2012 period the divergence of income distributions of all EU countries decreased, from

JS=0.00521 in 2005 to JS=0.00392 in 2012. The fall in divergence of income distributions between 2005 and 2012 can be attributed to a trend observed in countries with initially high income inequalities towards a more egalitarian distribution of income.

Table 2. Distribution of income deciles and *JS*-divergence for EU countries in 2005 and 2012

2005			2012		
Country	First decile	Tenth decile	Country	First decile	Tenth decile
Portugal	2.5	30.3	Latvia	2.3	27.1
Lithuania	2.2	27.2	Spain	1.5	24.8
Latvia	2.1	27.7	Portugal	2.7	27.3
Poland	2.2	26.9	Greece	1.8	25.1
United Kingdom	2.6	27.1	Bulgaria	2.3	25.4
Estonia	2.4	25.7	United Kingdom	2.7	25.9
Italy	2.5	25.4	Romania	2.1	23.3
Greece	2.5	25.0	Estonia	2.6	24.2
Spain	2.5	23.8	Italy	2.4	24.3
Ireland	3.3	25.2	Lithuania	2.7	23.9
Cyprus	3.5	22.8	Cyprus	3.5	25.1
Belgium	3.8	23.2	Poland	3.1	24.2
Hungary	3.7	23.2	Croatia	2.6	22.8
France	3.8	22.9	France	3.6	25.6
Netherlands	3.2	22.1	Ireland	3.1	23.2
Malta	3.7	21.0	Denmark	2.3	22.2
Germany	3.7	22.1	Germany	3.4	22.4
Luxembourg	3.7	21.6	Luxembourg	3.6	22.2
Austria	3.8	21.9	Austria	3.2	22.1
Czech Republic	4.0	22.2	Hungary	3.7	22.2
Slovakia	3.4	21.5	Malta	3.8	21.8
Finland	4.1	22.1	Belgium	3.5	21.1
Denmark	3.4	19.7	Finland	4.0	21.6
Slovenia	3.9	19.9	Netherlands	3.8	21.3
Sweden	3.9	19.8	Slovakia	3.6	20.3
			Czech Republic	4.1	21.6
			Sweden	3.4	20.0
			Slovenia	3.9	19.6
Jensen-Shannon divergence JS=0.00521			Jensen-Shannon divergence JS=0.00392 (JS*=0.0038901)[1]		

Source: EU-SILC database and own calculations

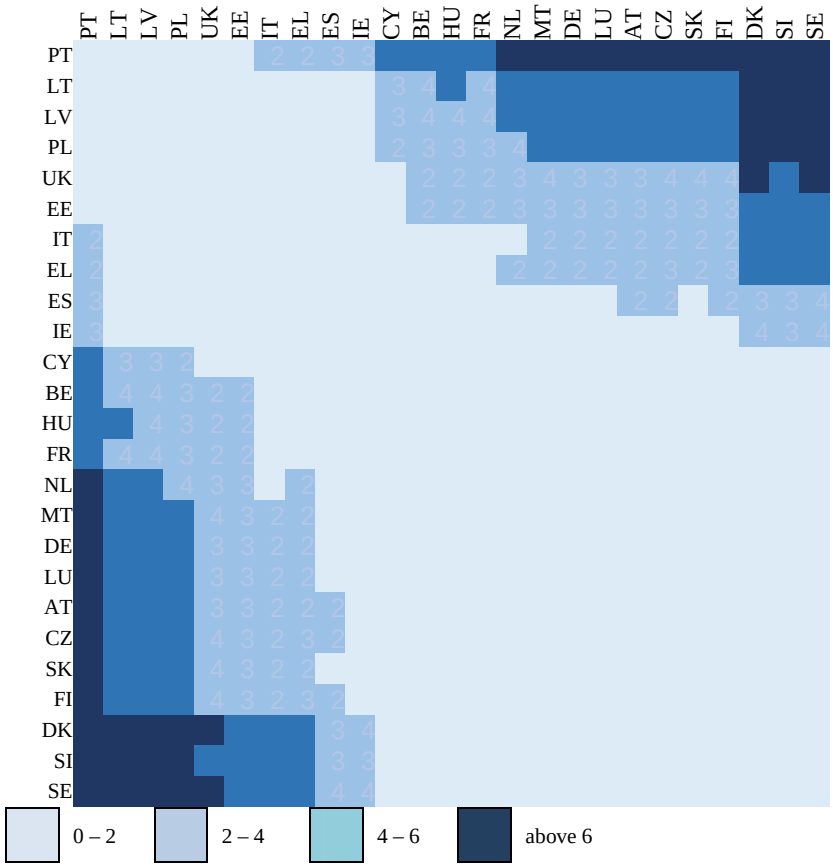
[1] The value of divergence JS* has been calculated for the same group of countries as JS for 2005, i.e. all EU countries, excluding Bulgaria, Croatia and Romania.

Table 2 also presents the shares of population earning first and tenth decile of income. Analysis of the data leads to a conclusion that the most significant decrease in inequalities measured by shares of population earning top and bottom 10 percent of income was observed in Lithuania and Poland. For example, in Lithuania, the share of

population earning bottom 10 percent of income fell from 27.2 percent in 2005 to 23.9 in 2012. In some countries there have been increases of inequalities, especially in Spain where the proportion of population earning top 10 percent of income dropped from 2.5 to 1.5 and Denmark, where the respective proportion fell from 3.4 to 2.3 percent.

In the next step, the degree of dissimilarity between income decile distributions of each pair of countries have been investigated, as measured by Jeffryes-Kullback-Leibler divergence. The results for 2005 and 2012 are presented in Figures 1 and 2, respectively. The darker squares in the figure indicate larger values of divergence between a pair of income distributions of countries representing a particular row and column. The darkest areas are concentrated in top-right and bottom-left corners of the chart simply because income distributions of countries with high income inequalities vary greatly from the distributions of countries with lowest inequalities.

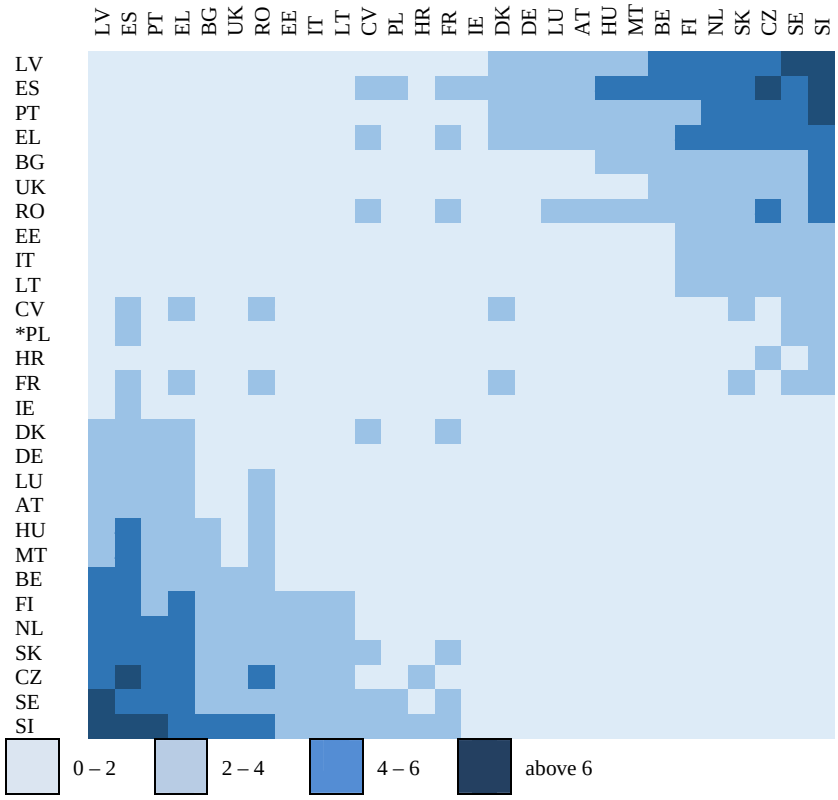
Figure 1. Jeffryes-Kullback-Leibler divergence for pairs of EU countries in 2005



Source: own calculations

The smaller deep-dark areas in top-right and bottom-left corners in Figure 2 in relation to Figure 1, indicate that, in period 2005-2012, EU countries became more similar in their income distributions – in 2012 there were fewer pairs of income distributions for which the value of Jeffryes-Kullback-Leibler divergence exceeded the value of 6. This conclusion confirms the finding mentioned earlier in the article, where Jensen-Shannon divergence values for 2005 and 2012 had been compared. As it has already been discovered earlier, this process of increase in similarity of income distribution patterns is the effect of reduction of income inequalities in countries where they were highest in 2005.

Figure 2. Jeffryes-Kullback-Leibler divergence for pairs of EU countries in 2012



Source: own calculations

CONCLUSIONS

Measures based on entropy can be a useful tool for assessment of income inequalities, as well as divergences between income distributions. Analysis based on Gini coefficient and Shannon entropy concluded that, in period 2005-2012,

income inequalities decreased in majority of EU countries, with Poland, Lithuania and Portugal experiencing especially strong moves towards more egalitarian income distributions.

The use of Jensen-Shannon measure has shown that divergence of income distributions of all EU member countries decreased between 2005 and 2012. Income distributions in the EU became more similar mainly as a result of the decline of income inequalities in countries with initially high inequalities. Since disposable income after social transfers has been used as a measure of income, further research is needed in order to assess, to what extent the decline in divergence of distributions was a result of labour market outcomes, and how it had been influenced by tax and social policies.

REFERENCES

Cavanaugh J. (1999) A Large-Sample Selection Criterion Based on Kullback's Symmetric Divergence, Statistics & Probability Letters, vol. 42, p. 333–343.

Csiszár I. (1967) Information-Type Measures of Difference of Probability Functions and Indirect Observations, Studia Sci. Math. Hungar, 2, p. 299–318.

Grosse I., Bernaola-Galva P., Carpena P., Román-Roldan R., Oliver J., Stanley H.E. (2002) Analysis of Symbolic Sequences Using the Jensen-Shannon Divergence, Physical Review E, vol. 65, p. 041905-1 – 041905-16.

Jędrzejczak A. (2012) Estimation of Concentration Measures and Their Standard Errors for Income Distributions in Poland, International Advances in Economic Research, vol. 18, issue 3, p. 287-297.

Lamberti P.W., Majtey A.P., Borras A., Casas M., Plastino A. (2008) Metric Character of the Quantum Jensen-Shannon Divergence, Physical Review A, no. 77, p. 052311-1 – 052311-6.

Lefebvre G., Steele R., Vandal A. (2010) A Path Sampling Identity for Computing the Kullback–Leibler and J Divergences. Computational Statistics & Data Analysis, vol. 54, p. 1719 – 1731.

Menéndez M.L., Pardo J.A., Pardo L., Pardo M.C. (1997) The Jensen-Shannon Divergence, Journal of the Franklin Institute, vol. 134, no. 2, p. 307-318.

Menéndez M.L., Pardo J.A., Pardo L., Zografos K. (2003) On Tests of Homogeneity Based on Minimum ϕ-divergence Estimator with Constraints, Computational Statistics and Data Analysis, 43, 215–234.

Podolec B., Ulman P., Wałęga A. (2011) Próba oceny zróżnicowania spożycia artykułów żywnościowych w Polsce na podstawie wyników Badań Budżetów Gospodarstw Domowych, Acta Universitatis Lodziensis, Folia Oeconomica, vol. 253, p. 225-237.

Quintano C., Castellano R., Regoli A. (2009) Evolution and Decomposition of Income Inequality in Italy, 1991-2004,Statistical Methods and Applications, vol. 18, issue 3, p. 419-443.

Reid M.D., Williamson R.C. (2009) Generalised Pinsker Inequalities, Proc. of the 22nd Annual Conference on Learning Theory (COLT 2009), Montreal – Quebec.

Seghouane A.K., Bekara M. (2004) A Small Sample Model Selection Criterion Based on Kullback' s Symmetric Divergence. IEEE Transactions on Signal Processing, vol. 52, p. 3314 – 3323.

Shannon, C. E. (1948) The Mathematical Theory of Communications. Bell Syst. Tech. Journal, vol. 27, p. 423-467.

Taneja I.J. (2005) Refinement Inequalities Among Symmetric Divergence Measures, The Australian Journal of Mathematical Analysis and Applications, 2 (1), Art. 8, p. 1-23.

Taneja I.J. (2008) On Mean Divergence Measures, Advances in Inequalities from Probability Theory & Statistics (eds.) N.S. Barnett, S.S. Dragomir, Nova Science Publishers, p. 169–186.

Taneja I.J. (2013) Generalized Symmetric Divergence Measures and the Probability of Error. Communications in Statistics: Theory & Methods, vol. 42, p. 1654 – 1672

Taneja I. J., Kumar P. (2004) Relative Information of Type s, Csiszár's f-Divergence, and Information Inequalities, Information Sciences 166, p. 105-125.

Tomczyk E. (2011) Application of Measures of Entropy, Information Content and Dissimilarity of Structures to Business Tendency Survey Data, Przegląd Statystyczny, Zeszyt 1-2, p. 88-101.

Ullah A. (1996) Entropy, Divergence and Distance Measures with Econometric Applications, Journal of Statistical Planning and Inference, 49, p. 137 – 162.

Wędrowska E. (2011) Application of Kullback-Leibler Relative Entropy for Studies on the Divergence of Household Expenditures Structures, Olsztyn Economic Journal, vol. 6, p. 133-142.

Wędrowska E. (2012) Miary entropii i dywergencji w analizie struktur, Wyd. Uniwersytetu Warmińsko-Mazurskiego, Olsztyn.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 177 – 187

THE INTERNATIONAL COMPETITIVENESS OF EUROPEAN UNION COUNTRIES AND STRATEGIES FOR ITS ENHANCEMENT: A TAXONOMIC ANALYSIS

Magdalena Olczyk
Department of Economic Sciences
Gdańsk University of Technology
e-mail: Magdalena.Olczyk@zie.pg.gda.pl

Abstract: The aim of this article is to identify clusters of countries with similar levels of competitiveness among the EU-27 countries and to identify for each of the 27 EU countries the path of competitiveness growth. The cluster analysis conducted confirms the hypothesis that the European Union is an area with a high differentiation in terms of levels of competitiveness. The analysis shows that the strategy to increase competitiveness should be significantly different for each EU country. It is suggested that in order to increase its international competitiveness each EU country should follow the pattern of the country which stands above it in terms of competitiveness ranking, but at the same time is most similar.

Keywords: global competitiveness index, cluster analysis, competitive strategy, European Union

INTRODUCTION

Competitiveness is one of the most misunderstood concepts in economics [Waheeduzzaman, Ryans, 1996]. The main problem of economists dealing with issues of competitiveness is the lack of a single universally accepted definition of this phenomenon. Even such an expert as Porter in his book "The competitive advantage of nations" does not define competitiveness directly, although the term is used repeatedly [Porter, 1990]. Furthermore, in the literature there is the problem of an excess of definitions, which stems from the fact that the phenomenon is considered at up to four levels, i.e. micro-, mezo-, macro- and mega-competitiveness. Moreover, the concept of competitiveness is derived from at least

three economic theories: those of international trade, of economic growth and microeconomics.

One of the few successful attempts at defining and measuring competitiveness is the methodology proposed by the Word Economics Forum. This approach can be called holistic, and is based on both macro and micro theories and uses both hard and soft data. The Growth Competitiveness Index (GCI) published by the Word Economic Forum aims to quantify crucial pillars of a country's competitiveness [Schwab, 2013]. They are:

- Institutions (I) – the institutional environment, which is determined by the legal and administrative framework;
- Infrastructure (II) – the quality and the extensiveness of infrastructure;
- Macroeconomic environment (III) – the stability of the macroeconomic environment (inflation rate, government spending, public debt);
- Health and primary education (IV) – the quantity and quality of basic education and access to the health care system;
- Higher education and training (V) – the quality of higher education and the intensity of vocational and continuous on-the-job training;
- Goods market efficiency (VI) – the structure of production and its compatibility with national and international demand; the intensity of domestic market competition; the degree of customer orientation and buyer sophistication;
- Labour market efficiency (VII) – the flexibility of the labour force; the transparency of labour law; the adjustment of supply to demand in the labour market;
- Financial market development (VIII) – the level of development and health of the financial system; the trustworthiness and transparency of the banking system;
- Technological readiness (IX) – ICT (Information and Communication Technology) access and use in daily activities and production processes; foreign direct investment intensity (FDI intensity);
- Market size (X) – the size of the domestic market and openness to trade;
- Business sophistication (XI) – the quality of a country's overall business networks and the quality of individual firms' operations and strategies;
- Innovation (XII) – the intensity of technological innovation; R&D expenditure in the private sector.

The GCI calculated for an individual country informs about the level of competitiveness of its economy and allows a comparison of this level with that achieved by other countries.

However, for an integrated group like the EU-27 countries, knowing which country is more competitive than the others does not contribute much to EU growth policy. The previous competitiveness growth strategy for the EU-27, based on a common denominator and employing identical tools for competitiveness support, did not work, which could be evidenced by the failure of the implementation of the

Lisbon Strategy in the period 2000-2010. For this reason, in this paper similarities in the competitiveness of the EU-27 countries will be identified, and then the different path for competitiveness growth will be determined for each EU country. The paths will be determined, based on the principle that to increase the international competitiveness, each EU country should try to follow the path of another EU country which is very similar but stands above it in terms of competitiveness ranking. Rather than advocating the concept of a "two-speed Europe", the paper proposes viewing the strengths of each particular Member State as an engine of growth for the whole EU.

METHODOLOGY

For the analysis, the above-mentioned set of 12 competitiveness indicators (diagnostic variables) is used. First, the usefulness of the diagnostic variables is determined by examining their degree of variation and correlation. The analysis requires variables which have sufficient variation and are not correlated too strongly with each other (Table 1). Only in this case will they be good carriers of information, allowing different processes to be identified [Grabinski at all, 1993]. The desired level of the coefficient of variation and the correlation coefficient, i.e. such that the variables selected can be considered diagnostic variables, is taken from the literature. It is assumed that if the coefficient of variation exceeds 10% the feature has statistically significant variation [Zieliaś, 2000], [Kozłowska 2010]. Due to the very low value of the coefficient of variation for variables IV and VI, they are excluded from the analysis.

To assess whether significant correlations exist between the variables, a matrix of correlation coefficients for pairs of variables is generated. No correlation which could contribute to abnormal results in the analysis (min. r_{xy} = -0.1483, max r_{xy} = 0.913) is found [Nowak, 1990], [GUS, 2012]. However, correlation coefficients do not allow to determine the dependence between variables other than linear. Therefore, the correlation plots for all the pairs of variables are analyzed. It confirms the absence of significant curvilinear correlations.

The next step of analysis should be the data normalization. In our case this is not necessary due to quite similar scales or magnitudes among the variables.

Two taxonomic methods are used to identify similarity in the competitiveness level of the EU-27 countries.

Firstly, to group the countries analyzed into relatively homogeneous groups the cluster analysis is applied. This method allows a determination of the similarity of objects without establishing a hierarchy among them. Classification and separation of the object clusters is carried out by means of a distance matrix. To create this, Ward's method is used. It is based on an analysis of variance to evaluate the distances between clusters, i.e. it attempts to minimize the sum of the squared distances of points from the cluster's centroid.

Table 1. The variables and their basic statistical measures

Name		I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
AT	Austria	5,04	5,8	5,35	6,32	5,48	4,91	4,69	4,65	5,7	4,62	5,52	5,07
BE	Belgium	5	5,68	4,66	6,75	5,81	5,12	4,54	4,68	5,57	4,81	5,32	5,09
BG	Bulgaria	3,39	3,79	5,42	5,92	4,31	4,17	4,54	3,97	4,3	3,82	3,62	2,98
CY	Cyprus	4,59	4,8	3,86	6,5	4,98	4,68	4,57	4,56	4,85	2,81	4,18	3,36
CZ	Czech Republic	3,67	4,81	5,19	5,87	4,87	4,53	4,32	4,25	5,06	4,51	4,45	3,81
DK	Denmark	5,4	5,74	5,4	6,19	5,59	5,03	5,22	4,69	6,17	4,22	5,41	5,08
EE	Estonia	4,94	4,72	6,01	6,21	5,17	4,73	5,11	4,51	5,29	2,98	4,2	3,93
FI	Finland	6,03	5,58	5,7	6,82	6,18	5,05	5	5,5	5,92	4,18	5,49	5,75
FR	France	4,83	6,28	4,64	6,32	5,14	4,47	4,41	4,73	5,72	5,76	5	4,91
GR	Greece	3,37	4,7	2,42	6,04	4,74	3,93	3,56	3,13	4,54	4,38	3,74	3
ES	Spain	4,25	5,92	4,17	6,09	5,02	4,37	3,98	3,9	5,29	5,45	4,51	3,77
IE	Ireland	5,22	5,34	3,44	6,46	5,3	5,24	5	3,6	5,82	4,13	5,09	4,66
LT	Lithuania	4,01	4,74	4,57	6,05	5,15	4,36	4,41	3,86	5	3,53	4,16	3,51
LU	Luxemburg	5,6	5,84	6,18	6,2	4,74	5,32	4,65	5,21	6,21	3,07	4,96	4,82
LV	Latvia	4,01	4,11	5,06	5,99	4,78	4,42	4,78	4,4	4,73	3,11	3,89	3,25
MT	Malta	4,61	4,91	4,6	6,34	4,93	4,62	4,14	5,11	5,59	2,38	4,27	3,43
NL	Netherlands	5,72	6,18	5,2	6,6	5,79	5,29	4,99	4,96	5,98	5,11	5,63	5,31
DE	Germany	5,31	6,36	5,48	6,3	5,8	4,92	4,51	4,66	5,71	6,02	5,71	5,42
PL	Poland	4,11	3,89	4,6	6,03	4,92	4,39	4,48	4,59	4,66	5,12	4,06	3,25
PT	Portugal	4,28	5,5	3,87	6,19	4,98	4,31	3,8	3,71	5,27	4,34	4,17	3,86
RO	Romania	3,33	3,22	4,83	5,51	4,36	3,86	4,01	3,98	4,09	4,41	3,47	2,92
SK	Slovakia	3,44	4,23	4,87	6,03	4,5	4,37	4,2	4,45	4,46	4	4,02	2,98
SI	Slovenia	4,05	4,91	4,94	6,29	5,2	4,42	4,15	3,29	4,96	3,46	4,18	3,85
SE	Sweden	5,73	5,69	6,16	6,46	5,75	5,14	4,81	5,29	6,29	4,62	5,56	5,56
HU	Hungary	3,7	4,39	5,15	5,84	4,67	4,28	4,27	4,05	4,43	4,25	3,74	3,61
GB	United Kingdom	5,41	6,22	4,01	6,39	5,57	5,09	5,42	5,16	6	5,78	5,48	5,17
IT	Italy	3,56	5,19	4,23	5,84	4,73	4,29	3,72	3,57	4,71	5,63	4,75	3,73
	min	**3,33**	**3,22**	**2,42**	**5,51**	**4,31**	**3,86**	**3,56**	**3,13**	**4,09**	**2,38**	**3,47**	**2,92**
	max	**6,03**	**6,36**	**6,18**	**6,82**	**6,18**	**5,32**	**5,42**	**5,5**	**6,29**	**6,02**	**5,71**	**5,75**
	median	**4,59**	**5,19**	**4,87**	**6,2**	**5,02**	**4,53**	**4,51**	**4,51**	**5,29**	**4,34**	**4,45**	**3,85**
	average	**4,55**	**5,11**	**4,78**	**6,20**	**5,13**	**4,63**	**4,49**	**4,39**	**5,27**	**4,31**	**4,61**	**4,15**
	standard deviation	**0,86**	**0,88**	**0,94**	**0,32**	**0,51**	**0,43**	**0,50**	**0,66**	**0,67**	**1,02**	**0,72**	**0,94**
	coefficient of variation	**0,19**	**0,17**	**0,20**	**0,05**	**0,10**	**0,09**	**0,11**	**0,15**	**0,13**	**0,24**	**0,16**	**0,23**

Source: own calculations

The error sum of squares and r^2 values are computed using the following formulae:

$$\text{ESS (error sum of squares)} = \sum_i \sum_j \sum_k \left| x_{ijk} - \bar{x}_{ik} \right|^2, \quad (1)$$

$$\text{TSS (total sum of squares)} = \sum_i \sum_j \sum_k \left| x_{ijk} - \bar{x}_k \right|^2, \quad (2)$$

$$\text{R Squared } (r^2) = TSS - ESS/TSS, \quad (3)$$

where: x_{ijk} denotes the value for variable k in observation j belonging to cluster i,
$\bar{x}_{ik}$ denotes the cluster mean for variable k, and $\bar{x}_k$ denotes the mean for variable k.

Among very different distance (similarity) matrices, Euclidean distance is chosen, as it is the recommended distance measure for Ward's method [see more Kaufman and Rousseeu, 1990 and Everitt, Landau, Leese, 2001].

Secondly, to create the path of competitiveness growth for an object/ country, an object map is used. To create this, two steps are needed. The first of these is to build the above-discussed matrix of distances, indicating the distance of an object relative to the rest. The second step is to rank all the objects (countries) in a ranking procedure. To create the ranking we calculate a synthetic variable for each country. We follow the methodology proposed by Hellwig. In this method, we first choose an "ideal object" which is described by a set of the maximum values of each variable. The variables must be standardized and they should be stimulants. The synthetic variable uses the following formula:

$$d_i = 1 - \frac{c_{i0}}{c_0}, \quad (4)$$

where
d_i is the taxonomic measure of development proposed by Hellwig,
c_{i0} is the Euclidean distance between the country and the "ideal object",
c_0 is the critical distance between objects and the "ideal object", and

$$c_0 = \bar{c}_0 + 2Sd, \quad (5)$$

$$\bar{c}_0 = \frac{1}{n} \sum_{i=1}^{n} c_{i0}, \quad (6)$$

$$Sd = \left[\frac{1}{n} \sum_{i=1}^{n} (c_{i0} - \bar{c}_0)^2 \right]^{\frac{1}{2}} \quad (7)$$

$$c_{i0} = \left[\sum_{i=1}^{n} (x_{ij} - x_{\max})^2 \right]^{\frac{1}{2}} \quad (8)$$

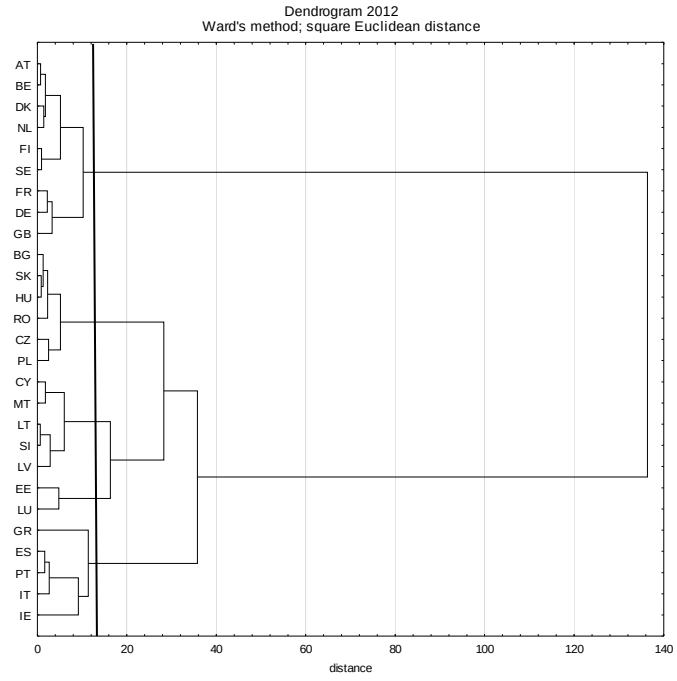
Based on this ranking of the EU-27 countries and on the distance matrix, an object map is created. All the statistical analyses in this article are performed using the statistical software Statistica 11.0, SPSS version 21.0 and R software.

EMPIRICAL ANALYSIS

The above-described similarity matrices of the objects, called distance matrices, are used to create a dendrogram. This shows how many clusters, i.e. homogeneous groups of countries, can be found among the EU-27 countries. The interpretation of the dendrogram, i.e. the identification of the number of clusters, depends on the bond distance chosen . The rule proposed by Mojena, based on the relative size of the different levels of junctions, is chosen to determine the cut-off. The constant in Mojena's inequality has a value of 1.25, which is recommended by Milligan and Cooper [Milligan, Cooper, 1985]. According to the result of this inequality, a cut-off at the level of 12.68 gives a satisfactory division of the 27 countries into clusters.

Using Ward's method, in 2012 five large homogenous groups of countries (clusters) can be distinguished (Figure 1). The basic characteristics of the clusters are presented in table 2.

Figure 1. Dendrogram for 12 competitiveness pillars using square Euclidean distance (27 EU countries)



In cluster 1 we have nine countries, mainly from the EU-15, all with a high level of competitiveness. In this cluster, all the variables not only achieve a better average value compared with the entire EU-27, but 9 of the 10 variables record the

highest average values among the five clusters identified. Because there is no variable with a value clearly differing from the rest, we can regard this cluster as being well-established with a high level of competitiveness.

Table 2. Average values of the variables in each cluster

Cluster number	All countries	1	2	3	4	5
Number of countries in each cluster	27	9	6	5	2	5
Country in each cluster		AT, BE, DK, NL, FI, SE, FR, DE, GB	BG, SK, HU, RO, CZ, PL	CY, MT, LT, SI, LV	EE, LU	GR, ES, PT, IT, IE
Institutions	4,541	5,386	3,607	4,254	5,270	4,136
Infrastructure	5,131	5,948	4,055	4,694	5,280	5,330
Macroeconomic environment	4,815	5,178	5,010	4,606	6,095	3,626
Higher education and training	5,128	5,679	4,605	5,008	4,955	4,954
Labour market efficiency	4,492	4,843	4,303	4,41	4,880	4,012
Financial market development	4,387	4,924	4,215	4,244	4,860	3,582
Technological readiness	5,271	5,896	4,500	5,026	5,750	5,126
Market size	4,315	5,013	4,352	3,058	3,025	4,786
Business sophistication	4,614	5,458	3,893	4,136	4,580	4,452
Innovation	4,151	5,262	3,258	3,480	4,375	3,804

Source: own calculations based on Table 1.

Cluster 2 consists of six countries from Central and Eastern Europe. This group is characterized by a low level of competitiveness, i.e. 6 of the 10 variables have the lowest average level among all the clusters. Any competitive advantage for these countries is based only on above-average development of the size of the domestic market, trade openness and stability of the macroeconomic environment.

Cluster 3 is an interesting case of a group of countries with an average level of competitiveness of their economies. This group comprises Cyprus, Malta, Lithuania, Slovenia and Latvia. The average values of most of the variables differ little from the average value of the competitiveness indicators recorded for the entire group of EU-27 countries. Therefore, this cluster is formed of countries that concentrate on a complex growth of competitiveness in all its aspects (as in cluster 1) rather than on building a competitive advantage based only on two or three selected pillars of competitiveness (as in cluster 2).

Cluster 4 consists of two countries: Estonia and Luxembourg. The main pillar of competitiveness of these countries is an extremely high stability of the macroeconomic environment (inflation rate, government spending and public debt). The level of competitiveness of these two countries can be described as medium-high, because half of the variables analyzed in this cluster have values

greater than the average values of these variables for the entire EU-27 group. Competitive advantage in these countries is also based on an unusually well-developed infrastructure and financial market, a high degree of ICT saturation in the economy and a high intensity of technological innovation.

The last cluster, number 5, consists of four southern European countries (Greece, Spain, Portugal and Italy) and Ireland. These countries are competitive due to a well-developed infrastructure and the large size of the domestic market. Although in this cluster only 2 of the 10 variables have values greater than the average values for the whole EU-27, the majority of the competitiveness indicators vary around the average values for the whole Union. Therefore, the competitiveness level of the countries in this cluster can be described as average.

This cluster analysis confirms the hypothesis that the European Union is an area with high differentiation in terms of levels of competitiveness. We have one cluster of countries with a very high level of economic competitiveness (cluster 1), one of above-average level of competitiveness (cluster 4), two clusters of countries representing the average level of competitiveness, but with different competitiveness bases (clusters 3 and 5), and cluster number 2 consisting of six EU-27 countries with a low level of competitiveness. With such a large differentiation of competitiveness level among the EU economies, it is difficult to build an EU competitiveness strategy based on a single unified growth path.

It is suggested that each of the EU-27 countries, especially in the short term, should follow their individual path to increase competitiveness. In order to determine these paths of competitiveness, all 27 EU countries are ranked according to their competitiveness level. Table 3 shows this ranking of countries based on the value of the synthetic variable, where the "ideal object" is described by a set of the maximum values for each variable.

Table 3. Ranking of 27 EU countries, based on the synthetic variable in 2012

Country		**rank**	**country**		**rank**
NL	0,7733	1	MT	0,2933	15
SE	0,7635	2	PL	0,2885	16
FI	0,7538	3	LT	0,2876	17
DE	0,7062	4	CY	0,2706	18
DK	0,6910	5	PT	0,2658	19
GB	0,6870	6	SI	0,2578	20
AT	0,6422	7	LV	0,2438	21
BE	0,6136	8	IT	0,2201	22
FR	0,5644	9	HU	0,2192	23
LU	0,5097	10	SK	0,1780	24
IE	0,4273	11	BG	0,1203	25
EE	0,4193	12	RO	0,0493	26
ES	0,3469	13	GR	0	27
CZ	0,3406	14			

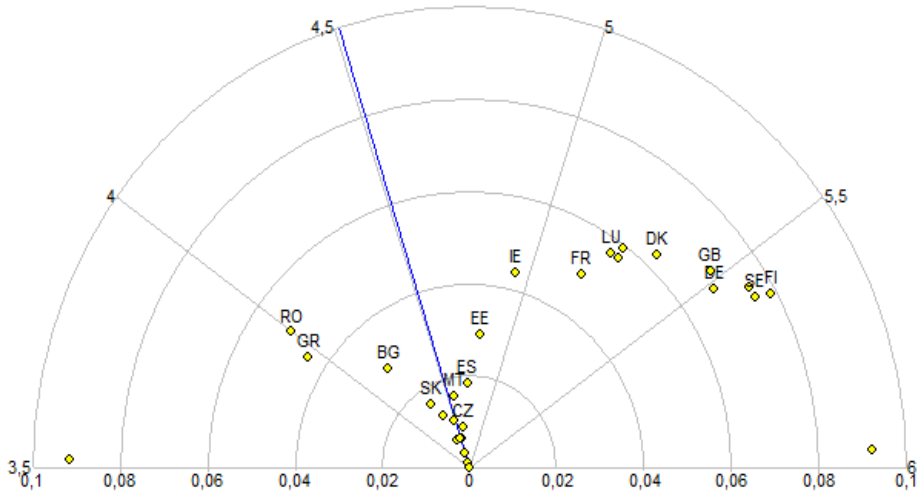
Source: own calculations

The synthetic variable is normalized so that it ranges from 1 (maximum value) to 0 (minimum value). The places of each country in the competitiveness ranking here are slightly different from those in the WEF rankings. This is due to the exclusion from the analysis of variables IV and VI (too low variability). Thus, in the ranking here the following countries gain slightly in competitiveness compared to the WEF rapport: Holland , Denmark, Poland, Malta, Latvia, Slovenia, Cyprus, Bulgaria. On the other hand, these countries lose a little bit: Finland, Sweden, the United Kingdom, the Slovak Rep. and Italy (down by up to 5 places).The best path for increasing the competitiveness of each economy is to follow on the solutions used in countries which are higher in the ranking. However, for example, Luxembourg, which is ranked in 10th place in the competitiveness ranking, does not have to catch up with all the countries ahead of it. The proposal here is to build an easy affordable strategy to improve each country's position in the rankings by adopting the pattern of a country which has a better position in the competitiveness ranking but at the same time is also the most similar. Relying on the experiences of country, which is more competitive but at the same time very similar, ensures easy implementation of the solutions selected. This approach only allows an indication of the countries from which a selected economy should draw patterns. The method does not, however, explain why two countries similar to each other occupy different positions in the competitiveness ranking.

A tool that allows analysis of each country's position in the ranking and allows us to find a better but most similar object, is a map of the objects. The map is a polar diagram, where each point on the map is defined by two values. The first value is the value of the synthetic variable (a measure of angle), with the worst countries on the left and the best objects in the ranking on the right). The second value is the distance matrix for each selected country compared to the other countries (this distance is represented by semi-circles). The country analyzed is always at the bottom of the map, and the bold radius indicates the position of the country analyzed in the ranking. To identify the path to competitiveness growth, we need to pay attention to all the countries on the map lying to the right of the designated radius and at the same time closest to it.

The path of competitiveness growth for the Polish economy is illustrated in Figure 2. The semicircles define the metric distance of all the countries from the country located in the centre at the bottom of the figure (Poland). The radiuses from right to left (anticlockwise) determine the positions of the countries in the ranking. The numbers placed at the end of the last radius of the semi-circle represent the scale of the object values in the rankings. Analysis of this figure shows that to increase the competitiveness of the Polish economy, we should try to implement instruments for competitiveness growth policy based on Czech, Maltese and Estonian experiences.

Figure 2. Paths to competitiveness growth for the Polish economy



Source: own calculations based on calculations from Table 3.

Analogous maps have been drawn for each of the 27 countries. Due to lack of space, only the conclusions which can be drawn from these maps are presented in the table below. Table 4 indicates country to follow for each of the 27 EU countries

Table 4. Paths to competitiveness growth for each of the 27 EU countries

Country analyzed	Country to follow	Country analyzed	Country to follow
NL		MT	Estonia, Luxemburg
SE	Netherlands	PL	Czech Rep., Estonia
FI	Sweden, Netherlands	LT	Poland, Malta
DE	Sweden, Netherlands	CY	Lithuania, Poland
DK	Germany, Netherlands	PT	Spain, Ireland
GB	Netherlands, Germany	SI	Cyprus, Latvia
AT	Denmark, Germany	LV	Slovenia, Lithuania
BE	Austria, Denmark	IT	Portugal, Spain
FR	Germany, United Kingdom	HU	Poland, Czech Rep.
LU	Belgium, Austria	SK	Hungary, Poland
IE	France, Belgium	BG	Hungary, Poland
EE	Luxemburg, France	RO	Slovakia, Hungary
ES	Ireland, Estonia	GR	Italy, Portugal
CZ	Estonia, Luxemburg		

Source: own calculations

CONCLUSIONS

This paper focuses on the identification of similarity in the competitiveness of the EU-27 countries. The cluster analysis has confirmed the hypothesis that the European Union is an area which can be divided into large clusters with very different levels of competitiveness. Therefore, it is suggested that one unified strategy to increase EU competitiveness is not a good solution. A new concept of individual competitiveness growth strategy for each EU country has been proposed, based on the implementation of proven solutions from other EU countries which have a better position in competitiveness ranking but at the same time are the most similar. In order to determine the path to competitiveness growth, the creation of maps based on each country's position in the competitiveness ranking (ranking based on the synthetic variable) and the distance metric of the selected country compared to other countries has been proposed.

REFERENCES

Everitt, B. S., Landau, S., Leese, M. (2001). Cluster Analysis. London, Edward Arnold.

Główny Urząd Statystyczny (2012). Identyfikacja i delimitacja obszarów wzrostu oraz obszarów problemowych w województwie dolnośląskim,
http://www.umwd.dolnyslask.pl/fileadmin/user_upload/Rozwoj_regionalny/SRWD/raport_delimitacja.pdf.

Grabiński, T., Wydmus, S., Zeliaś, A. (1993). Metody prognozowania rozwoju społeczno-gospodarczego, Kraków, Wydawnictwo Akademii Ekonomicznej w Krakowie.

Kaufman, L.& Rousseeuw, P. J. (1990) Finding Groups in Data, New York: John Wiley & Sons.

Kozłowska K. (2010). Zróżnicowanie rozwoju zrównoważonego województwa kujawsko-pomorskiego w latach 2004–2007 w świetle wyników analizy taksonomicznej. Studia Regionalne i Lokalne no 3(41).

Milligan G., Cooper M. (1985). An Examination of Procedures for Determining the Number of Clusters in a Data Set. Psychometrika, no 50.

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

Porter M. (1990). The competitive advantage of nations, Macmillan ,London.

Schwab K. (2013). The Global Competitiveness Report, http://www3.weforum.org/docs/WEF_GlobalCompetitivenessReport_2012-13.pdf

Waheeduzzaman A., Ryans J. (1996) Definition, perspectives and understanding of international competitiveness. Competitiveness Review, no.6.

Zeliaś A. (2000) Metody statystyczne, Warszawa, Polskie Wydawnictwo Ekonomiczne,

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 188 – 197

APPLICATION OF SPATIAL TECHNIQUES FOR PANEL DATA ANALYSIS OF AGRICULTURAL REAL ESTATE MARKET IN THE YEARS 2004 – 2012

Robert Pietrzykowski
KERiMSG, Division of Quantitative Methods
Warsaw University of Life Sciences – SGGW
e-mail: robert_pietrzykowski@sggw.pl

Abstract: In the economic studies models based on panel data are increasingly used. The standard panel models are composed of a cross-section character of the data in the time, but do not include the interaction associated with the location of objects in the geographic space. Spatial panel models are based on the information contained cross-section data in the time with regard to space. The paper proposes a different approach to the changes in time on the basis of spatial matrix weight. The aim of this study was to show the possibility to apply spatial weights matrix with particular consideration of time. Data for the analysis came from the database of the CSO and the ARMA in period 2004-2012. In addition to working methods of spatial statistics classical taxonomic methods were also used to obtain a distance matrix.

Keywords: weight matrix, spatial analysis, cluster analysis, price of agricultural land

INTRODUCTION

In the economic studies models based on panel data are increasingly used. It is probably connected with the general accessibility of computer programs. The panel data have to contain the time and the object in the cross-section dimension. The property can be defined as the observation of phenomena located in space. The standard panel models are composed of a cross-section character of the data in the time, but do not include the interaction associated with the location of objects in the geographic space. Spatial panel models are based on the information containing cross-section data in the time with regard to space. Description of spatial panel models can be found in Elhorst [Elhorst 2003]. The paper proposes a different

approach to the changes in time on the basis of spatial matrix weight. The aim of this study was to show the possibility to apply spatial weights matrix with particular consideration of time. The main aim of the paper was the concluding of time in the weight matrix. In this paper the specific objectives assumed: The first task was to define a links and interactions between neighbours in the time. The second task was to classify regions because of the price of agricultural land based on the modified matrix of weights. The study used Moran scatter plot to classify spatial objects. In addition to working methods of spatial statistics classical taxonomic methods were also used to obtain a distance matrix.

METHODOLOGY AND DATA

Data

Data for the analysis came from the database of the CSO and the ARMA in period 2004-2012. Province was adopted as a unit of spatial. It was due to the availability of data. The observed feature was the price of agricultural land of varying quality.

The matrix of weights - weights based on distances

Undoubtedly fundamental importance for spatial analysis is to define the weight matrix [Suchecka 2014] that reflects the interactions of studied objects. The simplest structure is the weight matrix in which for neighbours we recognize objects which have a common border as follow:

$$\boldsymbol{W} = \begin{cases} w_{ij} = 1, \ \text{the i} - \text{th object is a neighbor of j} - \text{th object} \\ w_{ij} = 0, \ \text{the i} - \text{th object is not a neighbor of j} - \text{th object} \\ w_{ij} = 0, \ when \ i = j \end{cases} \quad (1)$$

where:

W – matrix of weight,

w_{ij} – element of matrix of weight.

Because this approach to the problem may not always correspond to reality, therefore it seems to be natural to select another criterion which may be the selection of the neighbours due to the distance d km according to the formula 2. Take the example of two provinces in Poland: Mazowieckie and Wielkopolskie. However one cannot exclude the relation between these regions.

$$\boldsymbol{W} = \begin{cases} w_{ij} = 1, \ the \ i - th \ object \ is \ away \ an \ object \ j - th \leq \ d \ km \\ w_{ij} = 0, \ the \ i - th \ object \ is \ away \ an \ object \ j - th > \ d \ km \\ w_{ij} = 0, \ when \ i = j \end{cases} \quad (2)$$

where symbols are as in formula 1.

From above considerations we can see that the matrix **W** has a very characteristic structure, composed of zeros and ones, and the diagonal is zero. Taking into account the distance in kilometers to determine the neighbours opens the possibility of using other measures of the position of objects in space. The Doreian and Conley proposition is to include social [Doreian 1980] and economic distance [Conley 1999] which are based on reciprocal trade relations, capital movements and migration between the two spatial units. However another approach is to determine the position of objects in space through the designation of measures by the formula:

$$d_{ik} = \left[\sum_{j=1}^{m} |x_{ij} - x_{kj}|^{p}\right]^{1/p} \tag{3}$$

where:

p – number determining the metrics type,

m – number of variables,

x_{ij}, x_{kj} – determine the accomplishment of j-feature in i-th object and k-th object

After determining the metrics type (p = 1 Manhatan distance, p = 2 Euclidean distance) we can combine the objects in the distance matrix. In the article was assumed p = 2. Note that the distance matrix as well as weight matrix is diagonal zero. In distance matrix instead of zeros and ones specifying the neighbours there are set distance measures for implementation of specific variables. It is therefore necessary to transform distance into the elements of matrix weight. The simplest formula would be written as:

$$w_{ij} = d_{ij} \tag{4}$$

It should be considered that with increasing distance the effect of a neighbor decreases, so the correct formula will be:

$$w_{ij} = \frac{1}{d_{ij}} \tag{5}$$

Other possible transform functions:

$$w_{ij} = e^{-\alpha d_{ij}} \tag{6}$$

$$w_{ij} = \left(d_{ij}\right)^{-\alpha} \tag{7}$$

where α is any positive.

The matrix weights were based on the distance matrix and can be used to analyze the interplay between the two neighbours. However in order to use it in the analysis it is subjected to rows or columns normalization. Due to the fact that the weight matrix should be a symmetric matrix we need a transformation based on sample moments approximation in order to correct estimation [Anselin 2001] according to the formula:

$$\boldsymbol{W} = \frac{(\boldsymbol{W}+\boldsymbol{W}^T)}{2} \tag{8}$$

The prepared matrix **W** can be used for further spatial analysis [Pietrzykowski 2011].

The matrix of weights - to include the time

In the example considered, we observe price per hectare of agricultural land in the period from 1999 to 2012. Observations are on various provinces. Standard weights matrix consists of zeros and ones. When we take into consideration time in the analysis, we typically have been using panel models. The work proposed the effect of time through proper preparation of weights matrix **W**. On the basis of the available data distance matrix was created according to the formula 3. In this study variable prices of agricultural land were observed each year. Distance matrix contained information about provinces (objects) and variables which were the prices of agricultural land in period 2004-2012. Distance matrix defined by:

$$T = \begin{vmatrix} t_{11} & t_{12} & \cdots & t_{1k} \\ t_{21} & t_{22} & \cdots & t_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ t_{i1} & t_{i2} & \cdots & t_{ik} \end{vmatrix} \tag{9}$$

where t_{ij} are distances for the characteristics including the effect of time at the k-th and i-th object.

So the resulting distance matrix **T** was transformed according to equations 5 and 8 giving weight matrix **W**, which was used in further analysis. On the basis of such a weight matrix **W** produced and using the Moran scatter plot [Anselin 1996, Bivand i inn. 2013] provinces were classified. Moran's scatter plot was used to present spatial relationships autocorrelation. Autocorrelation was calculated according to the formula:

$$R_{(M)} = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\boldsymbol{z'Wz}}{\boldsymbol{z'z}} \tag{10}$$

where: **W** – matrix weight, **z** – vector of elements $z_i = x_i - \mu$,

For comparison the classical method of classification using a flexible clustering technique SAHN was applied according to the formula:

$$d_{(ij);k} = \propto_1 d_{i,k} + \propto_2 d_{j,k} + \beta d_{ij} + \vartheta(d_{i,k} - d_{j,k}) \tag{11}$$

which assumed:

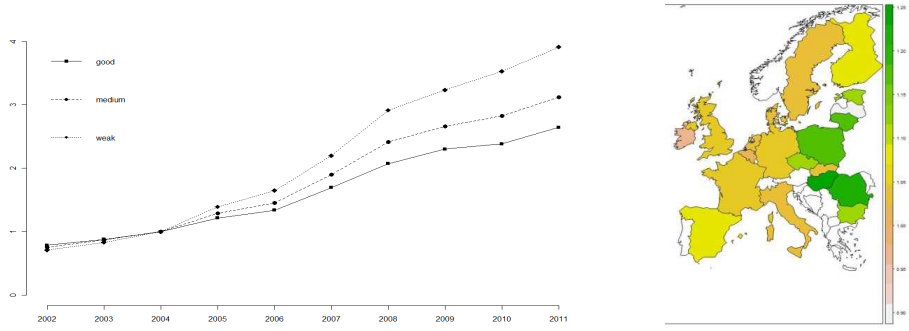
$$\propto_1 + \propto_2 + \beta = 1;\ \propto_1 = \propto_2;\ \beta \leq 1;\ \vartheta = 0$$

This resulted in the division of objects similar to the techniques of furthest neighbor ($where\ \beta = 1$).

RESULTS

Changes in the prices of agricultural land are particularly noticeable in Poland after accession to the EU in 2004. Prior to 2004 agricultural land prices remained at a low level. Price increase from year to year was relatively small, and the largest changes were observed for a good price, then the medium and poor. Since 2004, this situation has been changing. The largest price changes can be observed for the poor land prices (Figure 1) then these changes relate to the medium land prices and good quality. Regarding the situation in Europe, it can be seen that the average variation in the period regard Poland which due to these changes is at the forefront of the countries that joined the EU (Figure 1). A more detailed description of the problem can be found in Pietrzykowski [Pietrzykowski 2012]. In later work hierarchical cluster analysis method was used. The provinces (NUT2) were divided using agricultural land prices in the period from 2004 to 2012 according to the formula 11. Figure 2 shows the distribution obtained by using Euclidean distances and SAHN techniques. In addition dendrogram also contains a visualization of the spatial distribution of prices land on the map of Poland. Analyses were carried out for the land of good, medium and poor. The results have been presented only for price of agricultural land of poor quality due to their specific use [Pietrzykowski 2012].

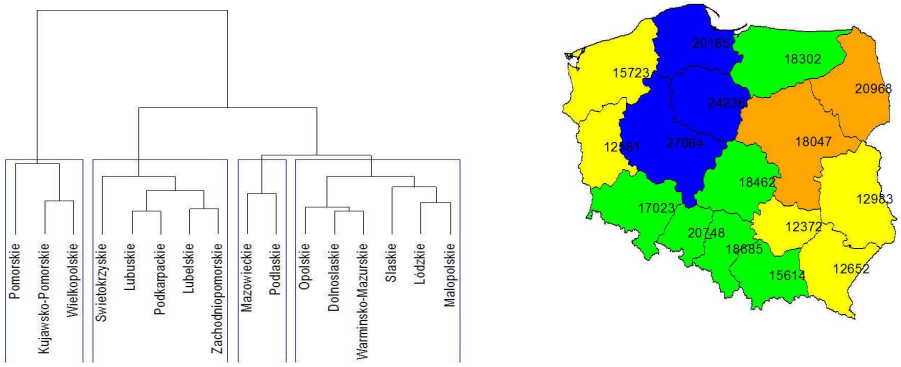
Figure 1. Dynamics of changes in agricultural land prices in the years 2002 – 2011 and the average change in prices in Europe in the period from 2005 – 2011



Source: own calculations [Pietrzykowski 2012]

Figure 2 shows the division of the four clusters. This proposal follows the methodology of work. The Moran scatter plot is a graphic representation that enables a description of the schema of spatial relation. This scatter plot has the standardized value of the variable on the horizontal axis and standardized lagged value on the vertical axis. The Moran scatter plot has four quadrants, where each quadrant corresponds to a specific spatial affiliation that can exist between a region and its neighbours [Anselin 1996, Bivand i inn. 2013]. Because as mentioned earlier it was planned to compare classification of the resulting methods of cluster and spatial analysis divided into four groups.

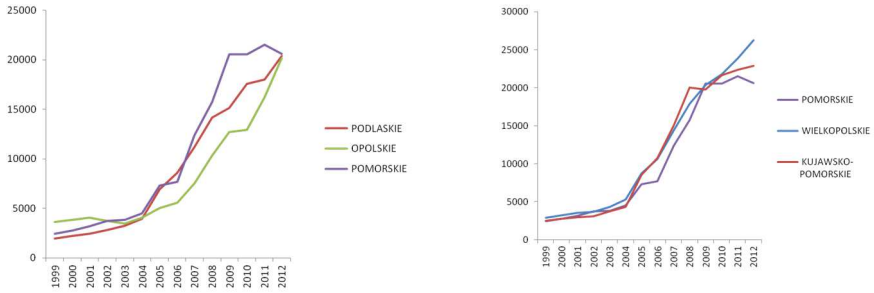
Figure 2. Dendrogram for the provinces due to prices of agricultural land of poor quality in the period from 2004 to 2012 and visualization of four clusters on the spatial map of Poland



Source: own calculations

Map of Poland gives the average prices of agricultural land for the poor quality of the region during the period of time. As you can see the division into four clusters is quite coherent. The first group includes three provinces: Pomorskie, Kujawsko-Pomorskie and Wielkopolskie. In these provinces the average price of land is the highest. One may wonder why Pomorskie (20185 PLN) was also included in this cluster because the average price of land is similar to that in other clusters (Podlaskie 20748 PLN, Opolskie 20968 PLN). Figure 3 is an explanation of this situation. The provinces which are in first cluster had similar prices of land in the time period 1999 - 2012 (Figure 3 right). The prices of agricultural land in provinces Opolskie and Podlaskie differ from price of land in province Pomorskie (see Figure 3, left) in the years 1999 - 2012.

Figure 3. Prices of agricultural land of poor quality for selected regions in the years 1999 – 2012



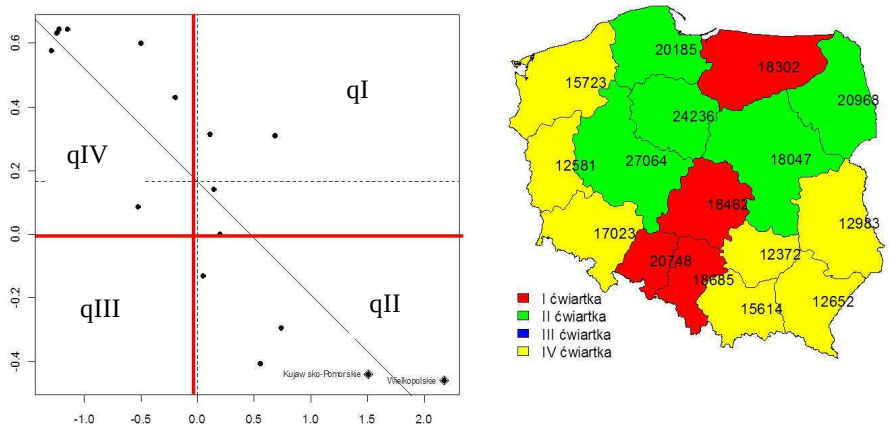
Source: own calculations

The second cluster includes: Świętokrzyskie, Lubuskie, Zachodniopomorskie, Podkarpackie and Lubelskie. The third cluster are two provinces: Mazowieckie and

Podlaskie. The fourth cluster includes six provinces: Opolskie, Dolnośląskie, Warmińsko-Mazurskie, Śląskie, Łódzkie and Małopolskie. Cluster analysis does not account for possible interactions between neighbours, which are the province here.

In the following part of the work on the basis of the obtained distance matrix weight matrix **W** was made. The Moran scatter plot plotted according to the resulting the matrix **W** (Figure 4). The points of Moran scatter plot have been divided into four groups based on four quadrants of this graph. The division due to zero values was made (thick lines in Figure 4, left). As a result of the analysis in the first quarter, four provinces are obtained: Warmińsko-Mazurskie, Łódzkie, Śląskie and Opolskie. In the second quarter: Podlaskie, Mazowieckie, Kujawsko-Pomorskie, Wielkopolskie and Pomorskie. In the third quarter there are no provinces and in the fourth there are seven provinces: Zachodniopomorskie, Lubuskie, Dolnośląskie, Małopolskie, Świętokrzyskie, Podkarpackie and Lubelskie. The Moran scatter plot noted Kujawsko-Pomorskie and Wielkopolskie as points or outliers in the provinces of poor agricultural where land prices obtained the highest value.

Figure 4. The Moran scatter plot for weight matrix based on the distance matrix and their spatial visualization on map of Poland



Source: own calculations

Regarding Moran coefficient it was found to be significant (p-value: 1,181 E-12) was negative -0.3561. It can be concluded that the price of agricultural land of poor quality are not randomly distributed among regions. The matrix weight used instruments that allow the indirect influence of the state on the price of agricultural land. This is due to its construction which takes into account changes variable over time. You can also see that the distribution using the Moran scatter plot allows the identification of regions where there are low prices of agricultural land (fourth quarter) compared to the other provinces (figure 4, right).

Besides it seems that such large provinces like Mazowieckie and Wielkopolskie should interact with each other exactly the price of the land. The

division obtained by the spatial analysis confirms this assumption. Table 1 shows a comparison of divisions using the two methods (differences in bold).

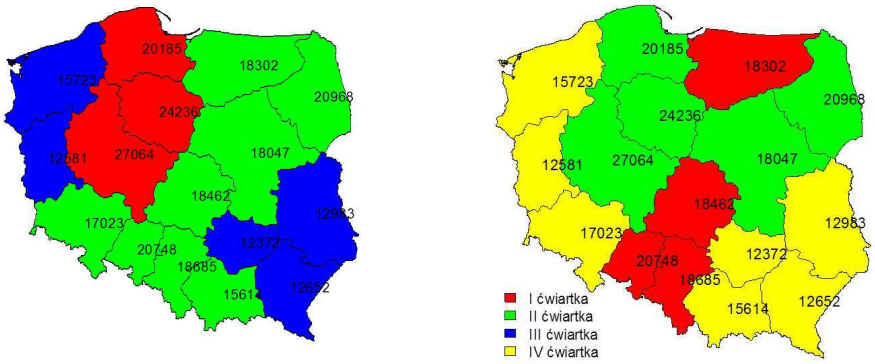
Table 1. Comparison of division the use of cluster analysis and spatial analysis

Cluster (quarter)	Cluster analysis (Figure 2)	Spatial analysis (Figure 4)
First (I)	warmińsko-mazurskie łódzkie, śląskie, opolskie, **dolnośląskie, małopolskie**	warmińsko-mazurskie, łódzkie, śląskie, opolskie
Second (II)	podlaskie, mazowieckie	podlaskie, mazowieckie, kujawsko-pomorskie, wielkopolskie, pomorskie
Third (III)	**kujawsko-pomorskie, wielkopolskie, pomorskie**	
Fourth (IV)	zachodniopomorskie, lubuskie, świętokrzyskie, podkarpackie, lubelskie	zachodniopomorskie, lubuskie, **dolnośląskie**, **małopolskie**, świętokrzyskie, podkarpackie, lubelskie

Source: own calculations

Although at the beginning of the substantive reasons division into four clusters was established, the result of the spatial analysis were obtained for three groups of provinces. Therefore, further analysis examined how it appeared to split into three groups using cluster analysis. Figure 5 compares the resulting division by two methods.

Figure 5. The division by the use of cluster analysis (left) and spatial analysis (right) on the map of Poland.

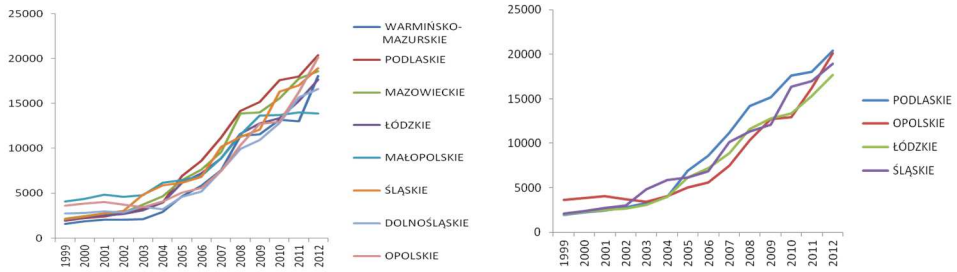


Source: own calculations

In the case of division into three groups using hierarchical methods we obtain the division into eight provinces (Warmińsko-Mazurskie, Podlaskie, Mazowieckie, Łódzkie, Dolnośląskie, Śląskie, Opolskie, Małopolskie), contrary to the spatial

analysis. In this group there are provinces that clearly differ because of the price of land. A thorough analysis of agricultural land prices leads to the conclusion. Comparing the price of land in these provinces, note the differences for provinces Warmińsko-Mazurskie since 2011 and Małopolskie since 2009. It seems that more accurate grouping is achieved by spatial analysis. For example, let us consider the change in the price of land in the years 1999-2012 (Figure 6 right) which is located in the first quadrant (Figure 4).

Figure 6. Prices of agricultural land of poor quality for selected provinces in the years 1999 – 2012.



Source: own calculations

Because of this it seems that the division into clusters, which uses information about reciprocal interactions and changes in prices over time would allow for a better assessment of the market for agricultural land prices.

SUMMARY

The paper proposes a modification of the weight matrix **W** in order to obtain information on changes in the studied phenomenon in time. The use of the weight matrix takes into account the construction aspect of time and information about spatial interactions. The use of spatial analysis allowed us to obtain a division to clusters, which in a more complementary manner presented the phenomenon than the classical cluster analysis did. The paper also shows the possibility of the use of Moran scatter plot and spatial analysis to classify objects. In addition, the work was characterized by changes in the prices of agricultural land of poor quality in Poland in the years 1999-2012.

REFERENCES

Anselin, L. (1996), The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association, in M. Fischer, H. Scholten, and D. Unwin, eds., Spatial Analytical Perspectives on GIS, London: Taylor and Francis, s. 111–125

Anselin L. (2001): Spatial econometrics, Oxford, Basil Blackwell

Bivand R. S., Pebesma E., Gómez-Rubio V. (2013): Applied spatial data analysis with R, Springer, London

Conley T. G. (1999): GMM estimation with cross selection dependence, Journal of Econometrics, 92(1), s. 1 – 45

Doreian P. (1980): Linear models with spatial distributed data. Spatial disturbances or spatial effects, Sociological Methods and Research, 9, 29 – 60

Elhorst J.P. (2003): Specification and estimation of spatial panel data models, International Science Review, nr 26(3), s. 224-268

Pietrzykowski R. (2011): Koncepcja i zastosowanie modyfikacji macierzy wag w przestrzennych badaniach ekonomicznych, Metody ilościowe w badaniach ekonomicznych. Vol. 12, nr 2 / red. Bolesław Borkowski s. 270-278

Pietrzykowski R. (2013): Agricultural land market in Poland and the European Union, Economic Science for Rural Development. Proceedings of the International Scientific Conference 2013, nr 32, s. 76-82

Suchecka J. (2014): Statystyka przestrzenna. Metody analiz struktur przestrzennych, Wydawnictwo C.H. Beck Warszawa

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 198 – 208

ALTERNATIVE WEIGHTING SCHEMES IN SPATIAL ANALYSIS OF GDP PER CAPITA CONVERGENCE[1]

Mariusz Próchniak
Department of Economics II, Warsaw School of Economics
e-mail: mproch@sgh.waw.pl
Bartosz Witkowski
Institute of Econometrics, Warsaw School of Economics
e-mail: bwitko@sgh.waw.pl

Abstract: The article analyzes conditional beta convergence in the EU28 countries with the use of spatial econometrics techniques. We consider alternative structure of the spatial weight matrix based on economic distances. Basing on the spatial Durbin-Watson model, two spatial specifications are tested, which make use of the volume of international trade and the inverted GDP per capita differences between the considered objects. We confirm the existence of GDP convergence and show that the gravity-models-type logic is superior to the approach based on inverted geographic distances.

Keywords: convergence, economic growth, spatial econometrics, Durbin-Watson

INTRODUCTION

The aim of the article is to validate whether there exists conditional β convergence in the EU28 countries. The concept of β convergence means that less developed countries (with lower GDP per capita) grow faster than the more developed ones. Although plenty of studies in this field exist and most of them confirm the existence of convergence among the EU countries [see e.g. European Commission 2009; Rapacki and Próchniak 2009; Kulhánek 2012; Staňisić 2012] but not all of them do [e.g. Monfort et al. 2013], most authors do not consider

[1] The research project has been financed by the National Science Centre in Poland (decision number DEC-2012/07/B/HS4/00367).

spatial dependence between the considered countries and those who do so, apply quite standard techniques of spatial econometrics.

In the spatial growth regression, weight matrices usually refer to the first, second etc. neighbors matrices or the inverted distance matrices [Conley and Topa 2002; Seya et al. 2012; Tian et al. 2010]. Although geography does matter for the evolution of income distribution (as confirmed on the basis of U.S. regions by [Rey 2001]), we consider alternative structure of the spatial weight matrix by incorporating the matrices based on inverted economic distances. This approach is supported by some other authors who suggest that trade flows [Aten 1996], institutional distances [Arbia et al. 2010], socio-economic distances represented by ethnic and occupation distances [Conley and Topa 2002] or demographic and income similarities [Case 1991] are the factors that matter in spatial interactions. In Poland, the first study of this type was carried out by Błaszczuk [1974].

Basing on the spatial Durbin-Watson (SDM) model two non-standard specifications are tested, which make use of the volume of international trade and the GDP per capita differences between the considered objects. These are compared with the usual weights based on inverted distances and the standard non-spatial panel data approach in the EU countries during 1993-2013.

MODEL OF CONVERGENCE WITH SPATIAL INTERACTIONS

The so called "Barro regression" [Barro and Sala-i-Martin 1992] is probably the most frequently found specification in empirical macroeconomics:

$$\Delta lnGDP_{it} = \beta_1 lnGDP_{i,t-1} + x_{it}'\gamma + \alpha_i + \varepsilon_{it}, \quad (1)$$

where GDP_{it} is the gross domestic product of the *i*-th country, $lnGDP_{i,t-1}$ is the lagged value of the GDP of the *i*-th country, x_{it} represents a *kx*1 vector of growth factors for the *i*-th country in period *t*, ε_{it} is the error term, individual effects α_i are introduced in order to represent different steady states while the β_1 and γ are the structural parameters of the model. Usually it is also assumed (though often not mentioned) that $corr(\varepsilon_{it}; \varepsilon_{jt}) = 0$ for every $i \neq j$. This, followed by the specification of the model in which the GDP of *i*-th country is clearly a function of the *i*-th country's growth factors only, constitutes a dynamic panel data model with no spatial dependence in most cases estimated with the use of GMM with the Blundell and Bond's system-GMM being the first choice for most authors [Blundell and Bond 1998]. We also include it in this paper for reference.

The assumption of independency of the data generating processes through the sample is, however, disputable. Globalization of the world economies makes them at least potentially interact and it is possible that the economic situation of one country might have an impact on the rate of growth in another country. In order to overcome this problem, spatial econometrics techniques have been employed in the analysis of growth and also the GDP convergence itself. We consider a model, named spatial Durbin-Watson model (SDM):

$$\Delta lnGDP_t = \beta_1 lnGDP_{t-1} + \rho W \cdot lnGDP_t + X_t\gamma + W \cdot X_t \cdot \theta + \alpha + \varepsilon_t, \quad (2)$$

where $lnGDP_t$ and $lnGDP_{t-1}$ represent the *n*x1 vectors of the GDP of the *n* considered countries in period *t* (and *t*-1, respectively), X_t is the matrix of growth factors for particular countries in period *t*, such that

$$X_t = \begin{bmatrix} x_{1t}{}' \\ \vdots \\ x_{nt}{}' \end{bmatrix}$$

is a *nxk* matrix of growth factors in period *t*, whose *i*-th row represents the growth factors of the *i*-th country in that period, α is a *n*x1 vector of (time invariant) individual effects, ε_t is a *n*x1 vector of error terms in period *t*, θ is a *k*x1 vector of parameters that reflect the influence of growth factors for other countries on the rate of growth of the *i*-th country, ρ points out the importance of growth rates of other countries for the rate of growth of the *i*-th country, W is a spatial weights matrix which emphasizes which of the countries are interrelated stronger than others. If the error term is spherical and further $\theta = 0$ and $\rho = 0$, non-spatial model (1) comes up.

Most authors use either weights based on the inverse of geographic distance or on the fact of being (or not) neighbours, however such an approach seems disputable nowadays. A number of countries form clubs that cooperate between one another, while others do not follow this rule despite geographic closeness with Northern and Southern Koreas being the best example. Also, the issue of geographic distance might be of great importance while the considered partners lay far apart, that is for example on different continents, yet the availability and relatively low transportation costs together with high globalization of contemporary world certainly reduce the importance of this factor in the case of relatively closer countries.

Two alternative specifications of the weighting matrix are then proposed. The first of them is based on the level of the exchange of goods in the considered pair of countries. It is likely that in the case of a pair of countries whose volume of trade is high, also the broadly understood economic changes in one of the partners shall have a serious influence on the other one's growth. A possible example is the CPI. Suppose that it increases vastly in one of the trade partners. In the case of high exchange of goods between the two considered countries, one can expect it to have a serious influence on the level of prices in the partnering countries and thus – indirectly – also on its GDP growth. That is why in the first alternative specification the weighting matrix is constructed with the use of the average volumes of trade between particular countries throughout the period of the analysis.

The other considered alternative stems partly from the logic of the gravity models. Those assume (among others) that the strength of trade relation in a pair of countries, measured by their volume of trade, depends upon their differences in the level of development. The countries that are similar in the sense of their GDP shall cooperate with greater strength, which means that the weak have greater trade

relationships with the weak, while the giants are more likely to cooperate with the giants [Śledziewska and Witkowski 2012]. We transmit the logic of the gravity models to the field of GDP growth in order to test whether the empirically confirmed property in the world of international trade can also be found in the case of growth. We thus construct the second alternative weighting matrix using as weights the inverted absolute differences in the level of GDP per capita measured in PPP. In all the considered cases we use the minmax transformation in order to standardize the weighting matrices [Keleijan and Prucha 2010].

All the spatial models are estimated with the use of maximum likelihood implemented by Belotti et al. [2013]. The consistency of the GMM requires model (1) to be transformed to the

$$lnGDP_{it} = (\beta_1 + 1)lnGDP_{i,t-1} + x_{it}'\gamma + \alpha_i + \varepsilon_{it}, \quad (3)$$

and similarly, model (2) is transformed to

$$lnGDP_t = (\beta_1 + 1)lnGDP_{t-1} + \rho W \cdot lnGDP_t + X_t\gamma + W \cdot X_t \cdot \theta + \alpha + \varepsilon_t \quad (4)$$

which changes nothing in terms of their meaning, but requires subtracting 1 from the estimate of the parameter on $lnGDP_{i,t-1}$ *or* $lnGDP_{t-1}$ in order to attain the convergence parameters.

One can consider limiting the set of growth factors that have a "spatial" influence, that is: that have an influence on the rate of growth of the "neighbors" to just those that are found significant in the sense of $\theta \neq 0$ since it is quite likely that only some of the $\mathbf{x}_{it}$ have an influence not only on the rate of growth of the *i*-th but also some other countries. We thus denote as $\widetilde{\mathrm{X}}_{\mathrm{t}}$ the matrix of those growth factors that are found significant in the sense of $\theta \neq 0$ (not just the $\gamma \neq 0$ and follow the same rule replacing θ with $\hat{\theta}$ and thus the model (4) is converted to:

$$lnGDP_t = (\beta_1 + 1)lnGDP_{t-1} + \rho W \cdot lnGDP_t + X_t\gamma + W \cdot \tilde{X}_t \cdot \tilde{\theta} + \alpha + \varepsilon_t. \quad (5)$$

Last but not least, since a couple of models can be proposed considering the different weighting schemes, a procedure needs to be adopted in order to select one for the analysis. A choice of the best one can be made with the use of information criteria, which is the solution adopted in this paper: specifically the minimization of Schwarz criterion is used in this case, although the difference in the number of degrees of freedom between the particular models is small, thus other criteria would usually yield the same answer. It should be noticed that most researchers treat the weighting matrices as "given" and – in most cases – consider mostly the selection of a functional form of the model itself and to a much lower extent the shape of the *W* matrix. However, as it is shown in the empirical results in this paper, that need not be the always-best approach.

THE DATA

The analysis covers the 1993-2013 period and is based on 3-year subperiods (the first subperiod is 1993-1995 while the last one is 2011-2013). Under such an approach, GDP growth for a given subperiod is calculated as the difference between log GDP per capita levels in the last year of a given subperiod and the last year of a previous subperiod (divided by 3 to express annual changes). The initial GDP per capita is taken as GDP per capita level from the last year of a previous subperiod while explanatory variables are calculated as arithmetic averages for the years covered by a given subperiod (in the case of missing data, the required figures are imputed).

It is necessary to choose the set of explanatory variables which are treated as economic growth determinants in the regression equations. From the theoretical point of view, the appropriate variables are those that characterize different steady-states to which the individual economies are tending. In empirical studies numerous growth factors are tested whose choice is constrained by data availability. Also, it is impossible to include too many variables due to multicollinearity and insufficient degrees of freedom issue. In this study, 21 variables (in addition to the initial GDP per capita) were initially included in theT growth regressions. Those are listed in Table 1.

Table 1. List of considered explanatory variables

Variable name	Variable description	Unit (scale)
Variables included in the final SDW models		
gdp0	Log of initial GDP per capita at PPP	Constant US$
inv	Investment rate	% of GDP
open	Openness rate	(Exports + imports) / GDP
cab	Current account balance	% of GDP
life	Log of life expectancy at birth	Years
econfree_fi	Fraser Institute index of economic freedom	From 0=lowest to 10=highest
wgi	World Bank's worldwide governance indicator	From –2.5=lowest to +2.5=highest
Variables excluded from the final SDW models		
human_cap	Index of human capital*	From 1=lowest to 4=highest
school_tot	Average years of total schooling (age 15+)	Years
school_ter	Population (age 15+) with tertiary schooling	% of total population
edu_exp	Expenditure on education	% of GNI
exp	Exports of goods and services	% of GDP
fdi	Foreign direct investment, net inflows	% of GDP
gov_cons	General government consumption expenditure	% of GDP

Variable name	Variable description	Unit (scale)
infl	Inflation rate (annual)	%
cred	Annual change of the domestic credit provided by banking sector to GDP ratio	% points
dem_fh	Index of democracy (average of civil liberties and political rights according to Freedom House)	From 1=lowest to 7=highest (inverted scale)
fert	Log of fertility rate	Births per woman
pop_15_64	Population ages 15-64	% of total
pop_den	Log of population density	People/km^2
pop_gr	Population growth (annual)	%
pop	Log of total population	Persons

* Index of human capital per person, based on years of schooling and returns to education, taken from Penn World Table 8.0. Note that only a few of the above mentioned variables remain in the final specifications presented in this paper.

Source: Data taken from the Penn World Table 8.0, World Bank *World Development Indicators*, IMF *World Economic Outlook*, Fraser Institute, and Freedom House databases

Initial variants of the econometric model were estimated on the basis of the full set of growth factors. Then, the initial set has been reduced using backward elimination of the least significant variables yielding the final set of six growth factors included in the models discussed in the next section. Also we eliminate (with the use of a standard backward algorithm) the growth factors that are allowed to be significant determinants of other countries' growth.

EMPIRICAL RESULTS

The results of the final models are reported in Table 2. Columns 1-3 refer to SDM models with three different types of including spatial effects while Column 3 concerns the reference Blundell and Bond's GMM system estimator model. For the sake of conciseness, we do not show neither estimates of the initial models based on a greater number of variables.

As regards a given country's growth factors, it turns out that investment rate, the degree of openness, current account surplus, life expectancy, large scope of economic freedom and good quality of governance all lead to a more rapid economic growth of the considered countries. This finding is confirmed both by the reference model as well as by the three alternative SDM approaches with spatial interactions.

Table 2. Spatial Durbin-Watson model for the EU28 countries with (1) inverted geographical distances, (2) inverted GDP per capita differences, (3) volume of trade weights and (4) Blundell and Bond's no spatial effects model (reference model)

regressor	geographic weights (1)	GDP difference weights (2)	volume of trade weights (3)	non-spatial Blundell and Bond (4)
country's growth factors				
gdp0	0.7811***	0.7338***	0.7001***	0.5236***
Inv	0.0088***	0.0069***	0.0086***	0.0143***
Open	0.0005***	0.0003***	0.0004**	0.0007***
Cab	0.0048***	0.0026**	0.0027*	0.0040***
Life	1.0159***	1.2287***	1.1122***	1.8047***
econfree_fi	0.0496***	0.0314***	0.0499***	0.0385***
Wgi	0.0312*	0.0717***	0.0970***	0.2296***
spatial effects (other countries' growth factors)				
gdp[3]	–0.7269	–0.0006***	–0.0026***	-
Open	-	-	0.0000**	-
Inv	1.0210***	0.0002***	-	-
Cab	0.6474***	-	–0.0001***	-
Life	-	-	0.0050***	-
econfree_fi	–2.0812**	-	0.0004*	-
Wgi	-	0.0022***	-	-
selected statistics				
initial GDP per capita in untransformed model (p/a)[1]	–0.0730***	–0.0887***	–0.1000***	–0.1588***
β coefficient[2]	7.6%	9.3%	10.5%	17.3%
BIC	–454.33	–471.41	-432.61	-
R^2 within	0.9035	0.9082	0.8821	-
R^2 between	0.9958	0.9965	0.9962	-
R^2 overall	0.9835	0.9846	0.9800	-

[1]Calculated as: (coefficient on gdp0 minus 1) divided by 3; [2]Calculated as: $\beta = (1/t)\ln(1+\beta_1 t)$, where β_1 is the coefficient on initial GDP per capita in untransformed model (per annum) while t is equal to 1; [3]Other countries' level of GDP per capita in the current period. *** - significant on 1%, ** - significant on 5%, * - significant on 10% significance level. Number of observations: 196 with n=28, T=7

Source: own calculations

It is always an issue how to construct the "best" weighting matrix as long as the assumed criterion is economically sound. Here the assessment of properties of the three SDM models with alternative weights is made on the basis of BIC criterion. It turns out that the model with inverted income distances as weights is found to be best (in terms of the BIC). The SDM model with geographic weights

proves to be notably worse while the model with weights based on the volume of trade is the weakest of the considered ones. Hence, differences in GDP per capita work better as the factor responsible for weighting the other country's growth determinants. This result can be explained by several reasons.

First of all, in the globalized world geographical distances play a diminishing role in determining spatial interactions between countries. This argument is of special importance in the case of one economic and political integration group, like the European Union. The reduction in trade barriers and the facilitation of migration of inputs (labor and capital) all lead to a decreasing role of geographical distances as a factor responsible for the impact of one country's economic performance on its neighbors. Second, it turns out that the volume of international trade is not a strong factor linking the countries either. Indeed, the volume of exports and imports rather depends on the size of a given country. For example, big countries like Germany or Poland record greater volumes of exports and imports in absolute terms while their shares in GDP are usually lower as compared to smaller countries like Slovakia or Estonia. Third, it is the relative level of development which is the most important factor responsible for mutual interactions between countries. This outcome shows that rich EU members are affected by rich EU neighbors while poor member states are influenced by the other poor EU neighbors. This finding is of great importance and it shows the nature of spatial interactions between the considered countries. Hence, the weight matrixes with inverted geographical distances – the approach which dominates in the literature – are not the best way of including spatial effects in the econometric model. Distances in GDP per capita levels seem to be better, at least inside one international organization like the European Union.

All the models confirm the existence of conditional β-convergence. The confirmation of existence of the catching up process among the EU countries is in line with the economic theory and the other empirical studies which is another argument for the economic validity of the considered models. The SDM model with inverted income distances as weights shows that β convergence parameter equals 9.3%. It is a significantly lower estimate than in the case of the reference (non-spatial) model where β amounts to 17.3%. This outcome is likely to show that the standard approach to economic growth modelling might result in artificially high estimates of the pace of convergence. A large part of the pace of convergence reported by standard econometric models with no spatial effects might not be the convergence itself – it is the effect of the impact of the other countries' growth factors (including the level of GDP) on the rate of economic growth of a given country. Indeed, such a supposition is reflected very well by the estimated coefficients for the variables responsible for spatial effects. Column 2 in Table 2 shows that the other countries' level of GDP per capita negatively affects the rate of economic growth of a given country. It means that if neighbor countries become richer the pace of economic growth in a given country is slower. The responsibility

of this interaction for explaining a large part of the pace of convergence is strengthened by the fact that the considered SDM model includes income differences as weights (most weighted are the countries with similar level of economic development). Thus, the SDM model shows that a negative relationship between economic growth and a given country's initial GDP per capita level is not so strong because part of this relationship is caused by a negative impact of GDP in neighboring countries on the rate of economic growth in a given country. However, the standard approach without spatial dependence does not account for this possibility and associates the full aforementioned effect to the classical convergence process.

It is worth to add that a negative relationship between the level of GDP in the other countries and a given country's growth rate refers to the convergence mechanism which is of supply-side in its nature. In contrast, demand-side linkages between the countries theoretically indicate rather the existence of a positive relationship between the other countries' GDP and a given country' economic growth: if neighbors are richer our exports are likely to rise fostering economic growth while a recession or a slowdown in neighbor countries is likely to drop our exports and output growth. These short-term demand-side relationships have not been evidenced in this study – partly because of the fact that the analysis is based on 3-year time spans to exclude the impact of short-term cyclical fluctuations.

The SDM model with inverted income distances (Table 2, column 2) also shows the impact of some other neighboring countries' growth factors on GDP dynamics in a given country. It turns out that the spatial estimate for the investment rate is positive and statistically significantly different than zero. It means that higher investment rate in other countries accelerates the pace of economic growth in a given country. This result points to very beneficial effects of investments in the EU as a whole. It turns out that higher investment rate leads to the acceleration of GDP growth not only inside a given economy but also in the other EU countries. Such positive spillovers can be treated as a novum given the existing literature. While a lot of theoretical and empirical analyses shows beneficial spillover effects of investments among the firms at the microeconomic level, our study widens this perspective as it indicates large beneficial spillover effects among the EU countries at the macroeconomic level. Hence, any increase in investment rate regardless of the country or region has a positive impact on economic growth of the EU as a whole (of course, the strength of this impact is different as reflected by the weight matrix).

Another result of spatial examination is the positive and statistically significant spatial coefficient for wgi variable. It shows that good quality of governance is beneficial not only for a given country but also for the other countries. This result indicates beneficial effects of institutional spillovers among the EU countries. A good quality of institutions in a given country is a significant factor of economic growth for all the EU countries (of course, with different weights as reflected by the weighting matrix). Hence, it is necessary for the EU

politicians and policy makers to enforce institutional reforms in all the EU member states to accelerate economic growth in the whole European Union.

Finally, it is worth to add that all the relationships interpreted here on the basis of regression equations do not formally indicate the existence of causal relationship. Causality is very hard to measure. Although there are formal econometric tests to verify the direction of causal relationship (e.g. Granger tests), such a formal approach has a lot of shortcomings and – as we believe - it is better to analyze causality using narrative-descriptive methods as it is done in this study.

CONCLUDING REMARKS

This study develops the current state of knowledge on real economic convergence by introducing spatial effects to the growth model and assessing the existence of convergence on the basis of spatial relationships. In the paper we propose an alternative structure of weights, based on economic distances. Those are constructed in two ways: (1) with the use of the volume of trade between particular countries, (2) on the basis of inverted differences between the level of GDP per capita of different countries. This procedure can be viewed as robustness check to inverted geographic distances as the main determinant of the strength of cross-sectional correlation. We estimate appropriate spatial Durbin-Watson models for the 28 EU countries over the 1993-2013 period.

We confirm the existence of GDP convergence and show that the model with inverted GDP per capita distances is superior to the approaches based on inverted geographic distances and the volume of trade. The SDM model with inverted GDP per capita distances as weights shows that β convergence parameter equals 9.3%. It is a significantly lower estimate than in the case of the non-spatial model where β amounts to 17.3%. This outcome is likely to show that the standard approach to economic growth modelling might result in artificially high estimates of the pace of convergence.

REFERENCES

Arbia G., Battisti M., Di Vaio G. (2010) Institutions and Geography: Empirical Test of Spatial Growth Models for European Regions, "Economic Modelling", Vol. 27, pp. 12-21.

Aten B. (1996) Evidence of Spatial Autocorrelation in International Prices, "Review of Income and Wealth", Vol. 42, pp. 149-163.

Barro R.J., Sala-i-Martin X. (1992) Convergence, "Journal of Political Economy", Vol. 100, pp. 223-251.

Belotti F., Hughes G., Mortari A.P. (2013) XSMLE – A Command to Estimate Spatial Panel Models in Stata, Material from the 2013 German Stata Users Group Meeting.

Błaszczuk D. (1974) Model ekonometryczny handlu międzynarodowego w 1967 r., "Ekonomista", No. 5, pp. 1095-1106.

Blundell R., Bond S. (1998) Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, "Journal of Econometrics", Vol. 87, pp. 115-143.

Case A.C. (1991) Spatial Patterns in Household Demand, "Econometrica", Vol. 59, pp. 953-965.

Conley T.G., Topa G. (2002) Socio-Economic Distance and Spatial Patterns in Unemployment, "Journal of Applied Econometrics", Vol. 17, pp. 303-327.

European Commission (2009), Five Years of an Enlarged EU. Economic Achievements and Challenges, "European Economy", No. 1/2009, Brussels.

Kelejian H.H., Prucha I.R. (2010) Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances, "Journal of Econometrics", Vol. 157, pp. 53-67.

Kulhánek L. (2012), Real Convergence in Central European EU Member States, in: Proceedings of the 1st International Conference on European Integration 2012, ICEI 2012 (eds. Honová I., Melecký L., Staníčková M.), 1st International Conference on European Integration, Ostrava, VŠB – Technical University of Ostrava, pp. 161-170.

Monfort M., Cuestas J.C., Ordóñez J. (2013), Real Convergence in Europe: A Cluster Analysis, "Economic Modelling", Vol. 33, pp. 689-694.

Rapacki R., Próchniak M. (2009) The EU Enlargement and Economic Growth in the CEE New Member Countries, "European Economy. Economic Papers", No. 367.

Rey S.J. (2001) Spatial Empirics for Economic Growth and Convergence, "Geographical Analysis", Vol. 33, pp. 195-214.

Seya H., Tsutsumi M., Yamagata Y. (2012) Income Convergence in Japan: A Bayesian Spatial Durbin Model Approach, "Economic Modelling", Vol. 29, pp. 60-71.

Staňisić N. (2012) The Effects of the Economic Crisis on Income Convergence in the European Union, "Acta Oeconomica", Vol. 62, pp. 161-182.

Śledziewska K., Witkowski B. (2012) Światowy kryzys gospodarczy a handel międzynarodowy, "Ekonomista", No. 4, pp. 427-448.

Tian L., Wang H.H., Chen Y. (2010) Spatial Externalities in China Regional Economic Growth, "China Economic Review", Vol. 21, pp. S20–S31.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 209 – 222

SECTORAL MONEY SUPPLY EFFECTS IN UKRAINE

Victor Shevchuk
Institute of Economy, Sociology and Philosophy
Cracow University of Technology
e-mail: victorshevchuk@netscape.net

Abstract: This paper examines the effects of anticipated and unanticipated money supply shocks over the 1999─2013 period across several sectors of the Ukraine's economy. It is found that the anticipated money supply shock contributes to output growth in agriculture, food processing and machine-building industries, with no impact for the steel industry. Unanticipated money shock is expansionary for the machine-building industry, while being restrictionary for agriculture. In general, our results reject the Monetary Neutrality Hypothesis (MNH).

Keywords: rational expectations, output, anticipated and unanticipated money supply

INTRODUCTION

Since it had been proposed in the early 1970s [Lucas 1972], the 'Monetary Neutrality Hypothesis' (MNH) suggest that under rational expectations anticipated changes in the money supply have no effect on real output, being translated into proportional changes of the price level. Only unanticipated changes in the money supply have real effects, as economic agents cannot distinguish between current, relative and absolute demand shifts.

Empirical tests of the MNH produced mixed results. Barro [1978] obtain that unanticipated money growth have effects on the U.S. output, while anticipated changes have no effect on output. Such effects are confirmed for the U.S. in several other studies, for instance Ravn and Sola [2004], Uhlig [2005], or Pragidis et al. [2013]. Among other countries, results in line with the theory of rational expectations are obtained for Brazil [Pragidis et al. 2013], Iran [Farahani and Abadi 2012], and Singapore [Maitra 2011]. Evidence in favor of the majority of rational expectations propositions are found for the panel of 41 countries [Apergis and

Miller 2004]. However, many empirical studies do not support assumptions of the rational expectations school. Except the U.S., unanticipated and actual changes in the money supply are found to be about equally poor as explanations of real output growth for other industrial countries [Darby 1980]. Other studies report that money still matters for business fluctuations [Chatterjee 1999]. The non-neutrality of money, at least in the short run, is found for India [Jha and Donde 2001/02], Mexico [Wallace and Shelley 2007], Pakistan [Bilquees *et al.* 2012], and Turkey [Yamak and Küçükkale 1998]. A negative short-run relationship between money supply and output is not ruled out as well, as it is found for Argentina and Brazil [Bae and Ratti 2008].

As mentioned by Devadoss [1991], money neutrality at the aggregate level does not necessarily imply that the hypothesis holds at the disaggregate level as well, assuming that input and/or output price rigidities vary across sectors. Consequently, aggregate level evaluation of monetary shocks can present a distorted picture of the disaggregate level effects.

The purpose of this paper is to review the theoretical underpinnings of monetary policy effectiveness and estimate their relevance for the Ukraine's economy, based on the decomposition of money supply into anticipated and unanticipated components. We test the MNH hypothesis for several sectors of the Ukraine's economy. The empirical results indicate that either anticipated or unanticipated money supply is positively correlated with aggregate output growth, while sectoral effects are quite heterogeneous.

In the next section, the theory underlying important differences between macroeconomic effects of anticipated and unanticipated money supply are discussed in detail. Then statistical methodology is outlined and empirical results are discussed. Finally, a brief summary and suggestions for future research are provided.

THEORETICAL ISSUES

Several very different theories explain effects of anticipated and unanticipated money supply. A seminal paper by Lucas [1972] explains a positive relationship between output and inflation by imperfect information regarding the aggregate price level, despite perfectly flexible prices and wages. The Lucas supply function with rational expectations implies that unanticipated money supply temporarily stimulates output, but eventually only causes inflation, as market participants become fully informed about prices. Anticipated changes in the money supply are neutral in respect to output, affecting only the price level.

Barro [1976] proposed the model that assumes that market participants with imperfect information cannot distinguish between aggregate and market-specific shocks. Because people do not observe the prices of all goods, but are focused on the prices in their sector of the economy, the unanticipated money supply shocks are misinterpreted as market-specific shocks and thus lead to output growth. A shift

in relative demand away from the goods produced by other sectors is instrumental in an increase of output. However, the anticipated money supply shock results in proportional price changes in all sectors, with no effect on output to follow.

In the presence of nominal price and wage rigidities, a Keynesian assumption that the anticipated money supply does affect real macroeconomic variables is restored under the assumption of rational expectations. Fisher [1977] proposed a model with rational expectations that account for existence of long-term contracts, which are valid for period longer than the time it takes the monetary authority to react to macroeconomic shocks. As nominal wage is fixed over the length of the contract, a higher than expected money supply growth leads to a higher inflation and a fall in real wages. This, in turn, induces employers to hire more workers, which raises the output.

There are a few other mechanisms that imply non-neutrality of anticipated money supply: the Tobin effect implying that anticipated inflation reduces capital accumulation [Fischer 1979], myopia of households [Giraud and Tsomocos 2004], imperfect synchronization of price revisions [Caplin and Spulber 1987], a fixed cost of changing firms' pricing plans [Burstein 2006], money functioning in a centralized market [Williamson 2006]. A possibility of the contractionary anticipated increase in the money supply is demonstrated by Rojas-Suarez [1992] under assumptions of the money-based expectations of the exchange rate and financially-constrained producers.

For illustrative purposes, a simplified, stochastic version of an open economy model under rational expectations presents as follows (except interest rates, all variables are expressed in logarithms) [Marston 1985]:

$$y_t = c\left(p_t - E_{t-1}p_t\right) + c_1 E_{t-1}\left(p_t - p_t^* - e_t\right) + c_0, \tag{1}$$

$$y_t = g_p\left(e_t + p_t^* - p_t\right) - g_r\left[r_t^* + E_t e_{t+1} - e_t - \left(E_t i_{t-1} - i_t\right)\right] + g_y y_t^* + g_0, \tag{2}$$

$$m_t - p_t = y_t - k_1\left(r_t^* + E_t e_{t+1} - e_t\right), \tag{3}$$

where y_t, p_t, e_t are domestic output, price and exchange rate, respectively, y_t^* and p_t^* are foreign output and price, respectively, m_t is the aggregate money supply, i_t is the general price level, defined as a weighted average of domestic and foreign prices,

$$i_t = \gamma p_t + (1-\gamma)(e_t + p_t^*). \tag{4}$$

In Equation (1), the aggregate supply is increased by unexpected increases in the domestic price and anticipated appreciation of the real exchange rate. The former refers to a monetary surprise effect, when actual price is above its expected level, while the latter reflects the relative price of imported inputs.

Equation (2) describes the aggregate demand as a positive function of a depreciation of the real exchange rate and an increase in foreign output, with a rise in the expected real interest rate having a contractionary effect.

In Equation (3), the demand for money is increased by output, with income elasticity being set equal to one, while an increase in a nominal interest rate is of an opposite impact. The money supply is assumed exogenous. Although it is a well-established fact that changes in money growth occur prior to changes in output, the real business cycle (RBC) models predict just the opposite [Ahmed 1993]. Money responds to real shocks, such as technological innovations, environmental developments, energy prices, labour market conditions, government spending and taxes, through a change in banking sector deposits. In order to meet the increased transaction service in anticipation of the increase in output, the banking sector reacts by attracting additional funds to increase deposits, which expands the quantity of inside money. Another explanation refers to subjective expectations about future economic activity, which are materialized in asset prices and interest rates. In order to avoid interest rate instability, the central bank could be inclined to accommodate money demand shocks through changes in its monetary base (or outside money).

The aggregate money supply is the sum of deterministic and stochastic terms:

$$m_t = m_0 + \alpha v_{t-1} + u_t^m, \tag{5}$$

where m_0 is the anticipated component of the money supply, v_t and u_t^m are the unanticipated permanent and unanticipated transitory components of the money supply, respectively.

Solutions of the three-equation system (1)–(3) for exchange rate, price and output as functions of the money supply provide with the following expressions:

$$e_t = \bar{e} + m_0 + \frac{\alpha}{A_2} v_{t-1} + \frac{A_1(1+k_1)}{A_2 A_0} u_t^m, \tag{6}$$

$$p_t = \bar{p} + m_0 + \frac{\alpha}{A_2} v_{t-1} + \frac{(g_p + g_r a)(1+k_1)}{A_2 A_0} u_t^m, \tag{7}$$

$$y_t = \bar{y} + \frac{c(g_p + g_r a)(1+k_1)}{A_2 A_0} u_t^m, \tag{8}$$

where $A_0 = ck_1 + (g_p + g_r a)(1 + c + k_1)$,

$A_1 = c + g_p + g_r a > 0$,

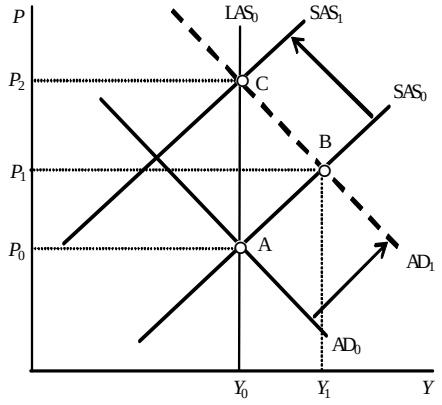
$A_2 = 1 + k_1(1-\alpha) > 0$,

$\bar{e}$, $\bar{p}$, and $\bar{y}$ are constants which are functions of the non-stochastic terms in Eqs. (1)–(3).

Macroeconomic effects of unanticipated and anticipated changes in the money supply are illustrated within the Y–P space on Figure 1. The curve labelled AD describes combinations of income (Y) and the price level (P) that give

equilibrium in the goods and money markets. The aggregate long-run and short-run supply schedules are presented by the curve labelled LAS and SAS, respectively. The temporary increase in the money supply u_t^m brings about a depreciation of the domestic currency, so that the aggregate demand increases (a rightward shift of the demand schedule from AD_0 to AD_1), despite a rise in the price of domestic output. As wages are fixed during the contract period or there is a money illusion among workers due to incomplete information, aggregate supply also increases due to a lower producer's real wage. Consequently, output and price level are increased (p. B). Beyond the current period, the expansionary effect on the economy is lost, because in the absence of further unanticipated money supply shocks macroeconomic variables return to their equilibrium levels (p. A). It is worth noting that such a result is predicted by either the New Keynesian sticky-wage models, or the New Classical models [Ahmed 1993].

Figure 1. Macroeconomic effects of the money supply shocks



Source: based on Marston [1985]

Fully anticipated increase in the money supply $(\alpha = 1)$ is associated with a proportional increase in the exchange rate (equation (6)) and the domestic price (equation (7)), with output unchanged (equation (8)). Following an anticipated increase in the money supply, a rightward shift of the aggregate demand AD curve is combined with an upward shift of the aggregate short-run supply SAS schedule, with a new equilibrium achieved for exchange rate and price level (p. C).

It is not ruled out that the anticipated money supply growth is negatively correlated with output, as it is the case under money-based expectations of the nominal exchange rate in an economy with financial constraints [Rojas-Suarez 1992]. Such a non-conventional outcome is facilitated by a decrease in the real money supply brought about by an increase in the price level in excess of a corresponding increase in the anticipated money supply. The restrictionary

effect is further strengthened by a substantial inflationary pass-through resulting from the exchange rate depreciation (it implies a low value of γ in Equation (4)).

DATA AND STATISTICAL METHODOLOGY

The empirical analysis uses Ukraine's quarterly series for the sample period 1999Q1─2013Q4 on the money supply (m_t), the nominal effective exchange rate (e_t), the real gross domestic product (y_t) and output across several sectors of the Ukraine's economy (agriculture, food processing, machine-building and steel industries), consumer price inflation (cpi_t), the government expenditure (g_t), the world market prices for crude oil (oil_t), metals ($metal_t$) and agricultural raw materials ($praw_t$). All data are taken from the IMF *International Financial Statistics* online database and the database of Ukraine's State Committee of Statistics (www.ukrstat.gov.ua). We use the logarithm of the time series. Except e_t, cpi_t, oil_t, $metal_t$ and $praw_t$, all other time series are seasonally adjusted with the Census X-12 method.

Empirical testing of rational expectations hypotheses are associated with numerous difficulties, including identification of money supply components and choice of appropriate estimators. There are several approaches for testing the MNH: (i) a two-step procedure, (ii) a joint estimation procedure, with the money forecasting equation and output being estimated as a simultaneous system, (iii) the Beveridge─Nelson decomposition. Following Barro [1978], a two-step procedure implies that initially current money supply is regressed on its lagged values, the past unemployment, and a current fiscal variable. Then forecasted values are used as a measure of anticipated money supply, with the difference between actual and anticipated money supply being interpreted as a measure of unanticipated money supply. The difference between actual and expected money growth is viewed as an alternative to estimating a long distributed lag on actual money growth rates [Darby 1980].

Assuming that the aggregate money supply contains unobserved permanent (anticipated) and transitory (unobserved) components, consisting of a random walk (with drift) and stationary autoregressive process with mean zero, the Beveridge─Nelson decomposition [Beveridge and Nelson 1981] is as follows:

$$\phi(L)\left[\Delta m_t - \eta\right] = \theta(L)\varepsilon_t, \tag{9}$$

where the permanent component is defined by

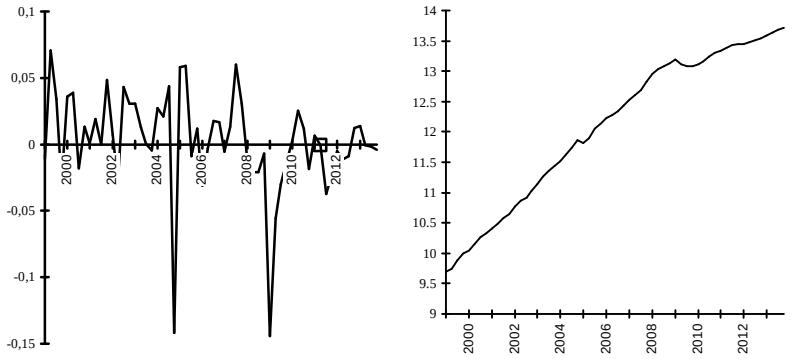
$$\Delta\nu_t = \eta + \psi(1)\varepsilon_t, \quad \psi(L) = \theta(L)\phi(L)^{-1}, \tag{10}$$

and the transitory component is defined by

$$u_t^m = \tilde{\psi}(L)\varepsilon_t, \quad \tilde{\psi}(L) = -\sum_{k=j+1}^{\infty}\psi_k. \tag{11}$$

Applying the Box–Jenkins methodology, it is found that the money supply follows the ARIMA(1,1,3) structure. The forecast residuals represent the unanticipated component of the broad money supply (Fig. 2a). The anticipated money supply component is estimated through in-sample one period ahead forecast (Fig. 2b). The anticipated and the unanticipated components of m_t are denoted by v_t and u_t^m respectively. A strong negative monetary surprise had occurred in 2004Q4 and 2009Q1, in the wake of serious crisis developments in the Ukraine's economy. In both cases, the primary motivation behind a sharp decrease in the money supply was a stabilization of the foreign exchange market, following an outbreak of substantial downward exchange rate pressure.

Figure 2. Unanticipated and anticipated components of the money supply, 1999–2013



a) an unanticipated component; b) an anticipated component;

Source: own calculations

Several criticisms in respect to the rational expectations hypothesis refer to the fact that (i) money surprise last for too long and (ii) conclusions do not hold under reasonable alternative ways of testing [Ahmed 1993]. In order to test an assumption that the inside money (banking deposits) is a better proxy for the money supply effects than the outside money (the monetary base), we used the difference between the money aggregate M2 and the monetary base as a measure of the former.

Also, the bivariate Blanchard–Quah decomposition is used as an alternative way of extracting anticipated and unanticipated components of the money supply [Blanchard and Quah 1989]. The effects of money supply shocks on output are classified as temporary while the effects of aggregate supply shock are assumed to be permanent, both restrictions being consistent with the MNH. The main advantage of this methodology is that the structural VAR approach is closely based on economic theory while allowing the data to determine the short run dynamics. Overall, the process of empirical illustration of economic theories is extended

towards the quantification of structural parameters, though this is not aimed at thorough empirical testing of such theoretical concepts as the MNH [Dunn 2002].

As the Augmented Dickey–Fuller (ADF) and the Phillips–Perron (PP) tests testify that all variables, except unanticipated money supply, are non-stationary in levels and become stationary upon first differencing, output equations for the aggregate (GDP) and disaggregate sectoral data are estimated in first differences.

EMPIRICAL RESULTS

Using the anticipated and unanticipated components of the money aggregate M2 from the Beveridge–Nelson decomposition, estimates for aggregate output growth are as follows (t-statistics in parenthesis):

$$\begin{array}{lcccc} y_t = & 0.116 v_t & +0.205 u_t^m & -0.133 e_{t-1} & +0.150 metal_t . \\ & (3.02^{*}) & (2.41^{**}) & (-2.13^{**}) & (5.13^{*}) \\ & R^2 = 0.41 & DW = 1.82 & ADF = -7.38^{*} & \end{array} \quad (12)$$

According to the coefficient of determination R^2, independent variables explain 41% of the variability of aggregate output growth, measured as the first differences of the logarithm of GDP. The Durbin–Watson statistics (DW) does not indicate the presence of first order serial correlation in residuals. The ADF test rejects the null hypothesis of the unit root in residuals at the 1% level of statistical significance[1].

As predicted by the rational expectations theory, the unanticipated money supply contributes to output growth. Our results indicate that a 1% unexpected increase in the money supply will result in a 0.3% increase in output growth. However, the anticipated money supply has an expansionary effect as well. Despite the fact that the value of the coefficient on v_t is about a half of that of the coefficient on u_t^m, there is no neutrality of anticipated money supply in respect to output growth. Among other factors, GDP growth is stimulated by the exchange rate appreciation and higher world prices for metals. Quite surprisingly, there is no any specific effect of the 2008–2009 financial crisis. A realistic explanation is that the exchange rate depreciation had absorbed all the adverse external shocks associated with the crisis developments.

The outcome from our estimations of sectoral money supply effects is reported in Tables 1–4. Residuals of all equations are white noise, indicating that they are appropriate estimates of sectoral output growth. For both of decomposition methods — Beveridge–Nelson and Blanchard–Quah, column I contains estimates of money supply that is the monetary aggregate M2, while column II refers to the money supply measure based on the 'outside' money, i. e. money aggregate M2 minus monetary base.

[1] * means statistical significance at the 1% level, ** at the 5% level, and *** at the 10% level.

Table 1. Determinants of machine-building output growth

Variables	Beveridge–Nelson		Blanchard–Quah	
	I	II	I	II
v_t	0.411 (2.54**)	0.357 (2.57**)	0.528 (3.28*)	0.448 (3.31*)
u_t^m	1.114 (3.92*)	0.637 (2.71*)	1.082 (3.65*)	0.522 (2.10**)
e_{t-1}	0.405 (1.98***)	0.381 (1.75***)	0.266 (1.26)	0.245 (1.12)
cpi_{t-1}	−0.776 (−2.25**)	−0.650 (−1.94***)	−0.963 (−2.76*)	−0.778 (−2.32**)
R^2	0.26	0.20	0.25	0.19
DW	2.08	2.18	2.02	2.06
ADF	−7.94*	−7.26*	−8.25*	−7.16*

Source: own calculations

Regardless of specification of the money supply measures and the type of its decomposition into anticipated and unanticipated components, the anticipated money supply positively affects output growth in the machine-building industry (Table 1), food processing industry (Table 3), and agriculture (Table 4). The machine-building industry is the only sector of the Ukraine's economy where there is a strong contemporaneous effect of the unanticipated money supply on output growth (the value of coefficient on u_t^m is much larger if compared with the estimate for aggregate output). Money surprise is of negative effect for agriculture (Table 4), while there is no statistically significant impact for the machine-building and food-processing industries. For the steel industry, monetary neutrality is observed for both anticipated and unanticipated money supply components, as no evidence is found, at any lag, of a significant effect of money supply shocks on output growth.

As the unanticipated money supply growth is negatively correlated to the agricultural output growth, it creates a problem for conventional explanations that attribute u_t^m effects to a favorable monetary surprise, even though negative money supply effects are not lacking in the case of other components of the money supply. Devadoss [1991] finds that the anticipated money supply has the adverse impacts on the U.S. agricultural output over the first three quarters, but then it cuts off relatively quickly. Explanation refers to a rapid price/flexible costs squeeze, which causes a decrease in agricultural output. As the adjustments are possible (such as reduced input use), the negative impacts are relatively short lived. However, the effect of unanticipated money supply is positive and prolonged in time, in line with the rational expectations theory. Sounders and Bailey [1986] explain the same negative relationship between the money supply and the U.S. nominal gross farm product by a strong positive impact of money supply on agricultural prices.

Table 2. Determinants of steel output growth

Variables	Beveridge–Nelson		Blanchard–Quah	
	1	2	3	4
v_t	0.120 (1.31)	0.120 (1.31)	0.123 (1.41)	0.119 (1.60)
u_t^m	−0.194 (−0.78)	−0.194 (−0.78)	−0.076 (−0.32)	−0.136 (−0.78)
e_{t-1}	0.441 (2.82*)	0.441 (2.82*)	0.456 (2.98*)	0.452 (2.98*)
$metal_t$	0.175 (2.12**)	0.175 (2.12**)	0.164 (1.96***)	0.163 (1.98***)
crisis	−0.232 (−6.49*)	−0.252 (−6.94*)	−0.228 (−6.79*)	−0.229 (−7.01*)
R^2	0.53	0.53	0.53	0.54
DW	1.91	1.91	1.90	1.91
ADF	−7.35*	−7.18*	−7.45*	−7.42*

Source: own calculations

The result for GDP — that suggests an elasticity of −0.13 between output growth and nominal exchange rate — does not hold for sectoral data. Contrary to the estimates for aggregate output, the exchange rate depreciation contributes to output growth in the machine-building industry. The same positive relationship is obtained for the steel industry, but no exchange rate effect is detected for food processing industry and agriculture. Coefficients on e_{t-1} in specifications for output growth in the food processing industry are negative and large enough, but they lack statistical significance (however, exclusion of the lagged exchange rate leads to a higher degree of serial correlation observed in the residuals). As for agriculture, all the coefficients on e_{t-1} turned out to be small and insignificant ones. When an alternative definition of money supply is used, similar results are obtained.

Inflation is restrictionary in the estimates for machine-building output growth (the coefficient on cpi_{t-1} is negative and significant at the conventional levels for all specifications), but it is neutral in respect to output in other sectors. It is worth mentioning that there is no evidence of any significant effects of inflation on Ukraine's GDP growth (equation (12)).

As suggested by the *crisis* dummy, the steel industry had been heavily affected by the 2008–2009 financial crisis, but other three sectors went unaffected. The result is intuitively appealing as the steel industry had been the largest source of export revenues over the pre-crisis decade, but its importance has been on a decline during the 2013–2014 period, as it has not caught up on the deep slump after the 2008–2009 financial crisis.

The food processing industry benefited from good harvests of 2001 and 2013, as it is revealed by the *harvest1* dummy. Also, there is the reverse relationship between an index of agricultural raw materials prices and food processing output

growth, which implies an improvement of supply conditions for the food processing industry due to a weaker demand for exports of domestically-produced agricultural commodities (wheat, corn, sunflower seeds, sugar) and cheaper imports of such important inputs, as cocoa or palm oil.

Table 3. Determinants of food processing output growth

Variables	Beveridge–Nelson		Blanchard–Quah	
	1	2	3	4
v_t	0.358 (3.08*)	0.207 (2.36**)	0.244 (2.44**)	0.191 (2.22**)
u_t^m	–0.287 (–1.32)	–0.174 (–0.88)	–0.235 (–1.06)	–0.059 (–0.33)
e_{t-1}	–0.274 (–1.29)	–0.057 (–1.27)	–0.274 (–1.34)	–0.213 (–1.05)
$praw_{t-1}$	–0.375 (–2.99*)	–0.396 (–2.85*)	–0.403 (–3.16*)	–0.395 (–3.03*)
harvest1	0.075 (3.04*)	0.074 (2.86*)	0.074 (2.94*)	0.077 (3.04*)
R^2	0.36	0.36	0.34	0.31
DW	2.03	1.98	2.04	2.14
ADF	–7.02*	–6.93*	–7.30*	–7.28*

Source: own calculations

Table 4. Determinants of agricultural output growth

Variables	Beveridge–Nelson		Blanchard–Quah	
	1	2	3	4
v_t	0.203 (3.72*)	0.168 (3.44*)	0.199 (3.59*)	0.174 (3.49*)
u_t^m	–0.285 (–1.94***)	–0.236 (–2.08*)	–0.290 (–1.95***)	–0.228 (–1.97***)
oil_{t-1}	0.067 (2.08**)	0.077 (2.25**)	0.055 (1.72***)	0.058 (1.78***)
g_t	–0.172 (–2.20**)	–0.147 (–1.89***)	–0.149 (–1.92***)	–0.101 (–1.31)
harvest2	–0.119 (–4.87*)	–0.117 (–4.72*)	–0.117 (–4.74*)	–0.113 (–4.53*)
R^2	0.37	0.36	0.37	0.37
DW	2.16	2.19	2.12	2.14
ADF	–8.01*	–8.42*	–8.26*	–8.15*

Source: own calculations

However, the agricultural output growth is not sensitive to changes in the world price of agricultural raw materials. On the other hand, it benefits from higher

oil prices, and this link probably reflects higher demand for biofuels produced from vegetable oil. There is a weak evidence of negative impact on agriculture by the government expenditure (the coefficient on g_t is negative and statistically significant at the 10% level in three out of four specifications). It is quite natural that bad harvests of 2003 and 2007 had slowed the agricultural output growth, as it is indicated by the *harvest2* dummy.

CONCLUSIONS

The results show that money generally plays a statistically important role in explaining GDP and sectoral output growth in Ukraine. It is found that anticipated and unanticipated money supply positively affects GDP growth, with the impact of the latter being almost twice as large. However, empirical estimates indicate that there is a different behaviour regarding the effect of monetary shocks on sectoral output growth. Money supply effects for the machine-building industry mirror those ones for the aggregate level, but it is not the case across other sectors. There are no money supply effects for the steel industry. Though the anticipated money supply contributes to growth in food the processing industry and agriculture, the unanticipated money supply is likely to be neutral in the former and restrictionary in the latter. Generally, our results reject the monetary policy ineffectiveness hypothesis of Rational Expectations─Natural Rate models in that only unanticipated money supply can affect real output.

Although our findings indicate that monetary policy, which is usually related to anticipated money supply, does affect output in Ukraine, it leaves unresolved the problem of substantial sectoral differences of money supply effects. Among possible explanations, different price setting mechanisms, exchange rate effects and industry-specific financial constraints are to be mentioned. Also, there is a puzzle of a strong negative monetary surprise in agriculture. Further exploring of these issues, especially the nature of exchange rate expectations, is left for future research.

REFERENCES

Ahmed S. (1993) Does money affect output? Business Review, No. 3, Federal Reserve Bank of Philadelphia, Philadelphia.

Apergis N., Miller S. (1998) Macroeconomic rationality and Lucas's misperceptions model: Further evidence from forty-one countries, Journal of Economics and Business, Vol. 56, Issue 3, pp. 227–241.

Bae S.-K., Ratti R. (2008) Conservative central banks and nominal growth, exchange rate, and inflation targets, Economica, Vol. 75, Issue 299, pp. 549–568.

Barro R. (1976) Rational expectations and the role of monetary policy, Journal of Monetary Economics, Vol. 2, Issue 1, pp. 1–32.

Barro R. (1978) Unanticipated money, output, and the price level in the United States, Journal of Political Economy, Vol. 86, No. 4, pp. 549–580.

Beveridge S., Nelson C.R. (1981) A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle, Journal of Monetary Economics, Vol. 7, Issue 2, pp. 151–174.

Blanchard O., Quah D. (1989) The dynamic effects of aggregate demand and supply disturbances, American Economic Review, Vol. 79, No. 4, pp. 655–673.

Bilquees F., Mukhtar T., Sohail S. (2012) Dynamic causal interactions of money, prices, interest rate and output in Pakistan, Journal of Economic Cooperation and Development, Vol. 33, No. 3, pp. 37–64.

Burstein A. (2006) Inflation and output dynamics with state-dependent pricing decisions, Journal of Monetary Economics, Vol. 53, Issue 7, pp. 1236–1257.

Caplin A., Spulber D. (1987) Menu costs and the nnutrality of money, The Quarterly Journal of Economics, Vol. 102, No. 4, pp. 703–726.

Chatterjee S. (1999) Real business cycles: a legacy of countercyclical policies? Business Review, Issue 1, Federal Reserve Bank of Philadelphia, Philadelphia, pp. 17–27.

Darby M. (1980) Unanticipated or actual changes in aggregate demand variables: A cross-country analysis, Proceedings, Federal Reserve Bank of San Francisco, San Francisco, pp. 160–198.

Devadoss S. (1991) Tests of monetary neutrality on farm output**,** Western Journal of Agricultural Economics, Vol. 16, No. 1, pp. 163–173.

Dunn S. (2002) A Post Keynesian approach to the theory of the firm, Post Keynesian Econometrics, Microeconomics and the Theory of the Firm, S. Dow, J. Hillard (eds.), Edwin Elgar, Cheltenham, UK, Northampton, USA, pp. 60–80.

Farahani M., M. Abadi (2012) Influence of rise in volume of money on economic growth in Iran during 1367-1387 (1988-2009), Journal of Basic and Applied Scientific Research, Vol. 2, No. 12, pp. 12192–12206.

Fischer S. (1977) Long-term contracts, rational expectations, and the optimal money supply rule, Journal of Political Economy, Vol. 85, No. 1, pp. 191–205.

Fischer S. (1979) Anticipations and the non-neutrality of money, Journal of Political Economy, Vol. 87, No. 2, pp. 211–247.

Giraud G., Tsomocos D. (2004) Global uniqueness and money non-neutrality in a Walrasian dynamics without rational expectations, Economics Series Working Papers No. 2004-FE-15, University of Oxford, Oxford.

Jha R., Donde K. (2001/02) The real effects of anticipated and unanticipated money: A test of the Barro Proposition in the Indian context, The Indian Economic Journal, Vol. 49, No. 1, pp. 21–30.

Lucas R. (1972) Expectations and the neutraliy of money, Journal of Economic Theory, Vol. 4, pp. 103–124.

Marston R. (1985) Stabilization policies in open economies, in: Handbook of International Economics, R. Jones and P. Kenen (eds.), Vol. I, Book 3, North-Holland, Amsterdam; N.Y.; Oxford, pp. 859–915.

Maitra B. (2011) Anticipated money, unanticipated money and output variations in Singapore, Journal of Quantitative Economics, Vol. 9, No. 1, pp. 118–133.

Pragidis I., Gogas P., Tabak B. (2013) Asymmetric effects of monetary policy in the U.S. and Brazil, BCB Working Papers No. 340, Banco Central do Brazil, Brasilia.

Ravn M., Sola M. (2004) Asymmetric effects of monetary policy in the US, Federal Reserve Bank of St. Louis Review, Vol. 86, No. 5, pp. 41–60.

Rojas-Suarez L. (1992) Limited capital markets and the real effects of monetary stabilization policies under alternative exchange rate regimes, Journal of International Money and Finance, Vol. 11, No. 6, pp. 594–613.

Sounders P., Bailey D. (1986) Effects of monetary changes on the price level and output in the U.S. agricultural sector, Agricultural Finance Review, Vol. 46, No. 1, pp. 58–68.

Wallace F., Shelley G. (2007) Long run neutrality of money in Mexico, Economia Mexicana Nueva Época, Vol. XVI, No. 2, pp. 219–236.

Williamson S. (2005) Search, limited participation, and monetary policy, International Economic Review, Vol. 47, No. 1, pp. 107–128.

Uhlig H. (2005) What are the effects of monetary policy on output? Results from an agnostic identification procedure, Journal of Monetary Economics, Vol. 52, No. 2, pp. 381–419.

Yamak R., Küçükkale Y. (1998) Anticipated versus unanticipated money in Turkey, Yapi Kredi Economic Review, Vol. 9, No. 1, pp. 15–25.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 223 – 231

SOFT COMPUTING FOR MODELLING THE VALUE OF SOCIAL CAPITAL AT RED HAT

Julia Siderska
Department of Business Informatics and Logistics
Bialystok University of Technology
e-mail: j.siderska@pb.edu.pl

Abstract: The main objective of this paper is to predict the value of social capital at Red Hat Corporation, using a developed neural network model. Training data were collected and calculated based on balance sheets published for the years 2005 - 2012. Five variables were proposed as an input of the neural network. The output variable presented the value of social capital, calculated by fundamental equation [Walukiewicz 2006]. The artificial neural network had been trained in Statistica Automated Neural Network.

Keywords: artificial neural network, social capital, soft computing, open source software

INTRODUCTION

Social capital, which is the most valuable resource of all IT companies, is defined as all the formal and informal relationships between employees: their ability to cooperate, trust, solidarity, loyalty to each other, etc. These are extremely essential features, since the passion, commitment and enthusiasm of engineers developing software determine the success of such applications and company as well. That is why the competent and reliable measurement of those assets value should therefore be one of the priorities of the IT company. Literature presents and characterises plenty of tools which allow the evaluation of social capital. However, no widely accepted method using soft computing has been proposed so far. In 2013 the author applied an interesting method and conducted the analysis of using an artificial neural network to assess the value of social capital in IT companies [Siderska 2013]. This paper investigates the prediction of the value of social capital at Red Hat in the first quarter of 2013, using proposed model of the neural network.

THE HISTORY OF OPEN SOURCE MOVEMENT

To fully understand an open source development and its potential power it is necessary to assimilate some background information on this movement and its main pioneers and activists: Eric Raymond, Richard M. Stallman and Linus Torvalds. The origins of open source software date back to 1965 when a team of scientists from Bells Labs AT&T, MIT and General Electric Company developed a new operating system - Multics. In 1984 Richard Stallman from MIT started developing a free operating system called GNU [Stallman 1995]. One of the leading members of the GNU project, Eric Raymond, published in 1997 the paper "The Cathedral and the Bazaar", which has been recognised as one of the most fundamental texts in the history of open source. The essay proposed new and innovative models of free software development. Raymond discussed his theories in terms of fundamentally different development styles: the cathedral model and the bazaar model [Silvester 1991]. Raymond compared proprietary and commercial software to monumental cathedrals built in concentration and silence. He was convinced that "the most important software (especially operating systems) needed to be built like cathedrals, carefully crafted by individual wizards or small bands of mages working in splendid isolation, with no beta to be released before its time" [Raymond 1999].

Over 20 years ago a Finnish student - Linus Torvalds, decided to develop the new kernel of the operating system - Linux. The success of the software developed by him convinced Raymond that the bazaar model, because of the massive effort undertaken by a large number of developers, would bring profits to both authors and users. The fundamental statement defining bazaar models was an argument written by Raymond: "given a large enough beta-tester and co-developer base, almost every problem will be characterised quickly and the fix will be obvious to someone [Raymond 1999]".

RED HAT CORPOPRATION

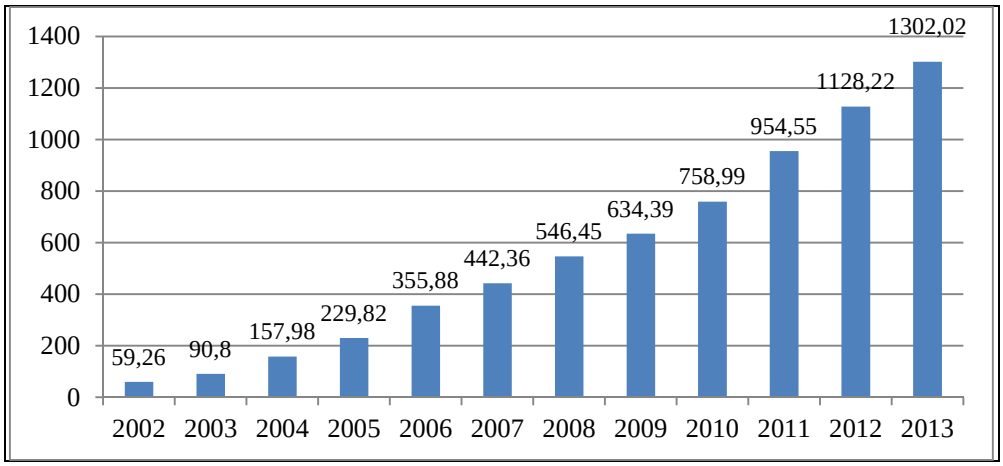
Company background

Nowadays the biggest computer company making open source development a key part of its ongoing software strategy is Red Hat Inc. It has its corporate headquarters in the USA with satellite offices worldwide. The flagship product of the company is Red Hat Linux – one of the most popular distributors of the Linux operating system. As of May 31, 2014 Red Hat had approximately 6500 employees worldwide. However, the most important contribution of the company's success has been its developers - volunteers from all over the world. The number of individual developers working on Linux kernel has been increasing over the different versions of the programme. Over 12% of all the Linux kernel development has been done by developers who are employed by Red Hat. Nearly

18% of developers are doing this work on their own, with no financial contribution from any company [Corbet 2012]. Red Hat is the world leader in providing open source solutions for the enterprise community. Red Hat Linux is being disseminated under the GNU General Public License without any fees. However, the company benefits mostly from support and services [Red Hat 2013].

In 2005 Red Hat became a part of the NASDAQ, an over-the counter, regulated stock market in the United States, under the symbol RHAT. In 2006 the company listed its shares on the New York Stock Exchange (NYSE) under the ticker symbol RHT. It boasts a number of powerful customers including Amazon, DreamWorks, Morgan Stanley etc. [Munga et al. 2009]. The company has achieved significant success in the software market. Over the last four quarters, the company's revenue rose by an average of 16% year-on-year. The largest growth was in the first quarter of 2013, when the revenue climbed by 19%. Notably, Red Hat became the first billion dollar open source company in its fiscal year 2012, reaching $1.13 billion in annual revenue [Babcock 2012]. For the past ten years the gross profit (difference between revenue and the cost of making a product or providing a service) has been constantly growing. These values are shown in Figure 1.

Figure 1. Red Hat's gross profit for years 2002 – 2013 (in millions USD)



Source: own study (on the basis of annual financial reports)

All those mentioned issues prove that the most precious assets of Red Hat are software engineers, programmers and analysts who developed the Linux kernel. Their human capital builds up the social capital of the whole company, therefore it is essential to assess the value of those intangibles.

Business model

The literature lists and characterises many ways of adopting open source software in business models. Frank Hecker suggested some interesting business models and strategies in his paper "Setting up shop: the business of open source software". He summarised the following business models [Hecker 1999]: Service support seller, Accessorising, Widget frosting, Service enablers, Brand Licensing, Software Franchising, Loss - leader, Sell it, Free it.

Red Hat applied *service support seller* model - the most common business model for companies involved with open source software. The revenue comes from media distribution, branding, training, consulting, and post-sales support instead of traditional software licensing fees. This model was originally recommended by Richard M. Stallman in his GNU Manifesto, with Cygnus Solutions being the first company to implement such solutions. Nowadays the best known vendors applying the business model in question are Red Hat and Caldera Software. In the operating system market the company competes with a limited number of large and well-established companies that have significantly greater financial resources and larger development staff: Microsoft, HP, IBM, Oracle, Google [Wikiinvest.com 2013].

Linux is being developed through the joint effort of thousands of volunteers all over the world. They are creative people who share their knowledge and make constructive criticism of the work of other professionals by developing and improving these applications. The idea behind the development of open and free applications is common, parallel and creative work of a team of experts, who communicate with each other via the Internet. Therefore, such a process is a creative process, since such applications are unique, created for the first time. The informal relations between them make the products more perfect, secure, reliable and with less backbiting and less fallible than flag products of ICT giants (such as the Windows operating system developed by Microsoft). The passion, commitment and enthusiasm of the volunteers determine the success of such applications. Here, the Internet plays an important part: efficient communication and the rapid exchange of ideas between programmers have a significant impact on the progress in the development of such projects.

The value of Red Hat

The discussion of value should start with the notion of social capital (SC), according to Walukiewicz [Walukiewicz 2006]:
social capital (SC) is composed of formal and/or informal relations among workers, teams, organisational units, etc. within a firm (internal relations), as well as formal/informal relations with customers, suppliers, banks, regional/central governments, R&D institutions etc. (external relations).

Moving on, **the market value** of a given firm F at a given moment t, denoted by V(F,t), equals the number of issued stocks multiplied by their current

stock price. In other words, the market value is the price investors are ready to pay for firm F at a given moment t of its present (history), taking into account information about the past.

V(RH,t) will denote the value of Red Hat at a given point in time t. Since the entire capital of the firm is conceived of as consisting of four parts (finance capital, physical capital, human capital and social capital) the following formula can be proposed:

$$V(RH,t) = v(FC,t) + v(PC,t) + v(HC,t) + v(SC,t) \quad (1)$$

This is called **fundamental equation** as it forms a base of the accounting model for the social capital analysis. The formula says that in a market economy, under the conditions of equilibrium, when demand equals supply, the value of Red Hat equals the aggregate sum of the four component values of its capital: financial, physical, human and social at any moment t of the firm's past, present and future [Walukiewicz 2006].

The market value of Red Hat at a given moment t, denoted by V(RH,t), equals the number of issued stocks multiplied by their stock price. For instance, in 2009 there were 193 024 423 stocks and the stock price closed at $30.49. The market value of Red Hat was hence $5 964 454 671. In other words, the market value is the price investors are ready to pay for a company at a given moment t of its present (history). **The book value** of Red Hat at a given moment t, denoted by BV(RH,t), is defined as the difference between its total assets and its total liabilities, calculated at the end of a reporting period (year, quarter or month), when balance sheets are prepared. Consequently, the book value is a step-wise function of time. Red Hat's book value in 2009 was $1 304 605 000. The book value is commonly understood as the value of the company calculated in accordance with the accounting principles at the moment t defined as above.

For almost all stock exchange listed companies, the market value is higher – in knowledge-intensive sectors much higher - than the corresponding book value. The author claims that the difference between the market value and the book value of each firm is simply the value of a company's social capital. This means that in traditional accounting a significant share of companies' market value is not included on the company's balance sheets. The difference between the market value and the total registered assets equalled $4 659 849 671 ($4.66 billion) in 2009. This means that 78% of the Red Hat's market value was ignored by traditional accounting and not recorded on its balance sheets. It is assumed that 1/3 of this difference is associated with the human capital and the rest with the social capital (software production is team work). Red Hat is a typical company of the knowledge-based economy in a sense that more than a half of its market value is produced by its social capital. The accounting items: goodwill and intangible assets should be considered as an attempt of Red Hat to account for its social capital, but the value of social capital recorded so far is less than 2% of its market value.

ARTIFICIAL NEURAL NETWORK MODEL

The training data contained information concerning each quarter from 2005 till 2012 and were gathered from the balance sheets published by Red Hat Corporation, Microsoft and IBM. Taking into consideration the data concerning these additional two companies, allowed the enlargement of the whole training set – it contained 96 cases (3 companies, 32 quarters - 4 quarters per each year). Moreover, such procedure resulted in an increase in the ability to generalise the data. For artificial neural network the Statistica 10.0 Automated Neural Networks module was used. For instance, the values of input and output variables for the fourth quarter of 2012 are shown in Table 1. All variables are demonstrated in USD, except employment which is shown in units. For simplicity the values of social capital, assets, liabilities, market value and book value are converted to millions of USD. The volume of employment X_3 is presented in units.

Table 1. Values of input and output variables in the fourth quarter of 2012 (in millions of USD)

Market value	Book value	Employment (in units)	Total asssets	Liabilities	Social capital
X_1	X_2	X_3	X_4	X_5	Y
10 231	1 599	4 500	2 492	892	5 622

Source: own study

80% of the data set was chosen for the training set, 10% of the samples constituted the cross-validation set and another 10% of the samples were used for testing. Measures were determined based on the training set and allowed to assess the capacity for approximation. They point to the precision in determining the output variable value for input vectors presented during learning. The most significant is correctness for input vectors that were not presented in the training process (ability to generalise the data).

There is no doubt that artificial neural network performance depends on learning rather than programming. The main use of a neural network is developing models - the formal structures or mapping a process or a phenomenon. The neural network can be an excellent model of any system. The learning process is an iterative process, repeated many times, step by step, with the fundamental objective to optimise the network parameters, i.e. the weighing factors. Each of the input variables initially gets randomly assigned weight, the strength of its effect on the value of the output variable. The values of the weighing factors are determined in the learning process, by which neurons acquire the knowledge and intelligence. The higher the weight, the more important the variable given [Lula et al. 2007].

In the present case a supervised, learning-with-a-trainer approach was adopted. This type of training is characterised by the fact that the networks receive examples of normal operation that they imitate. Besides the input signals, the expected answers (the output signal) should be determined. The network is trained on the basis of the knowledge of the social capital value [Lula et al. 2007]. The developed network model was trained using back propagation, which is one of the most frequently used and the most effective learning algorithms of multilayer neural network. Its essence was based on minimising the sum of squared errors of the network. Errors that occur at the output of the network are propagated in the opposite direction than the signals passing through the network, i.e., from the output layer to the input [Rutkowska et al. 1997]. In the input layer of the network five neurons are analysed, since it is assumed that this is exactly the number of input variables affecting the output variable. The modelling of the social capital value is a regression task, so only one neuron - characterised by the dependent variable - is presented on the output. The analysis covered the following input, independent variables: *market value* - quantitative variable, calculated by multiplying the number of shares and the price per share at the end of the quarter; *book value* - quantitative variable, calculated as the difference between the sum of total assets and the total liabilities; *total assets* - quantitative variable*; liabilities* - quantitative variable; *employment* - a quantitative variable, defined as the number of employees.

The output (dependent) variable represented the value of social capital. Necessary calculations of the social capital value were made using the fundamental equation, formulated by S. Walukiewicz in 2006. Many parameters of the model had been investigated before the final structure was proposed. The author compared many models with different numbers of hidden neurons, the activation function, the learning algorithm etc. Table 2 presents a summary of the parameters of obtained three-layer perceptron. These parameters allowed to accomplish the best quality in the validation set.

Table 2. Parameters characterising the proposed neural network model

Neural network	Activation function (hidden neurons)	Activation function (output neurons)	Error	Learning algorithm
MLP 5-3-1	Tanh	Linear	SOS – sum of squares	BFGS

Source: own study

The results of social capital value estimated using the fundamental equation and using the constructed model, the MLP 5-3-1 network, were compared. For this purpose, the function: *predictions for new data* was used. The value of the predicted social capital has been provided by the trained neural network.

Table 3 presents the results of the predicted value of social capital (Y*) of Red Hat for the first quarter of 2013.

Table 3. The value of social capital in the first quarter of 2013 (in millions of USD)

Market value	Book value	Employment (in units)	Total asssets	Liabilities	Social capital (Fundamental Equation)	Social capital (MLP 5-3-1)
X_1	X_2	X_3	X_4	X_5	Y	Y*
9 229	1 499	5 500	24 845	986	4 955	5 121

Source: own study

The anticipated value of social capital (Y*) in the first quarter of 2013 was predicted using the proposed three-layer perceptron MLP 5-3-1. The value of social capital (Y) was calculated using Fundamental Equation. The difference between those values is only about 4%. Those values authenticate the proposed method as a reliable tool for credible estimations of the value of social capital.

SUMMARY

Artificial intelligence tools, including neural networks, are increasingly being used not only in the areas of engineering (pattern recognition, signal processing, control, optimisation), but also in economics and management. Neural network models are often used in the forecasting of economic phenomena, such as in predicting trends in the stock market, sales forecasting, pricing, identifying risks of the granting of credit. Artificial neural networks are among the methods of data analysis which are a very powerful application and are therefore increasingly used as an alternative to traditional analytical methods, providing more reliable results [Tadeusiewicz, 1993].

The human factor is the most precious resource of any company from the ICT sector, including Red Hat. Despite numerous methods for estimating the social capital value, no widely accepted tools have yet been proposed. At the same time, no soft computing techniques for modelling the value of social capital as well as for analysing the independent (explanatory) variables have been developed either. This paper presents the draft of an artificial neural network for modelling the value of social capital at Red Hat.

Further research will focus on the feasibility of the application of a neural network model as was built for the world's biggest companies in the information technology industry: Microsoft, IBM, Oracle, Novell, SAP. Moreover, it would be interesting to develop a similar tool (using soft computing) to assess the value of social capital of companies whose shares are not listed on the stock exchanges.

REFERENCES

Babock Ch. (2012) Red Hat: First $1 Billion Open Source Company", InformationWeek, Retrieved 2012-03-29.

Corbet J., Kroah-Hartman G., McPherson A. (2012) Linux Kernel Development: How fast it is going, who is doing it, what they are doing and who is sponsoring it, The Linux Foundation.

Hecker F. (1999) Setting up shop: the business of open source software, IEEE Software, Volume 16, Issue 1, p. 45 – 51.

Lula P., Paliwoda – Pękosz G., Tadeusiewicz R. (2007) Metody sztucznej inteligencji i ich zastosowania w ekonomii i zarządzaniu, Akademicka Ekonomiczna w Krakowie, Kraków, s. 79 - 86.

Munga N., Fogwill T., Williams Q. (2009) The adoption of open source software in business models: a Red Hat and IBM case study, ACM New York, NY, USA, pp.112-121.

Raymond E. (1999) The cathedral and the bazaar: musings on Linux and Open Source by an accidental revolutionary, 1999.

Red Hat Corporation (2013) available at http:/investors.redhat.com 2013, accessed May 2013

Red Hat Corporation (2013) Annual Report 2005 – 2012, available at http://investors.redhat.com/annuals.cfm; accessed February 2013

Reuters.com (2013) available at http://t.reuters.com/finance/stocks/companyInfo?symbol=RHT, accessed May 2013

Rutkowska D., Piliński M., Rutkowski L. (1997) Sztuczne sieci neuronowe, algorytmy genetyczne i systemy rozmyte, Wydawnictwo Naukowe PWN, Warszawa, pp. 34.

Siderska J. (2013) Possibilities of using artificial neural networks to model the value of social capital in IT companies, Economics and Management Volume 5 Issue 1, pp. 84-100.

Silvester P.P. (1991) System operacyjny Unix, Wydawnictwo Naukowo – Techniczne Warszawa, p. 15.

Stallman R. (1985) GNU Manifesto, Dr. Dobb's Journal (DDJ).

Tadeusiewicz R. (1993) Sieci neuronowe, Akademicka Oficyna Wydawnicza, Warszawa, pp. 8-12.

Walukiewicz S. (2006) Systems analysis of social capital at the firm level Working Paper WP-1-2006, System Research Institute of the Polish Academy of Science, Warsaw.

Xu B., Jones D. R. (2010) Volunteers' Participation in Open Source Software Development: A Study from the Social – Relational Perspective, The DATA BASE for Advances in Information Systems, Volume 41, Number 3, pp. 69-84.

Young R., Goldman R. W. (1999) Under the Radar: How Red Hat Changed the Software Business--and Took Microsoft by Surprise, Coriolis Group, LA Vergne, Tennessee, U.S.A.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 232 – 241

INTER-INDUSTRIAL VALUE MIGRATION

Dariusz Siudak
Institute of Social Sciences and Management of Technologies
Lodz University of Technology
e-mail: dariusz.siudak@p.lodz.pl

Abstract: In this paper there is discussed value migration from the perspective of all economic sectors. It was introduced the method for measuring the sectorial value migration and the algorithm for classification with respect to three stages value migration model. The value migration measurement was conducted employing multivariate comparative analyses and in particular linear ordering to construct a synthetic variable of development. On the basis of the proposed measure, the ranking of value migration development and classification of sectors to the particular phases of value migration processes were delivered.

Keywords: value migration, synthetic variable, industry

INTRODUCTION

Value migration is defined as the shift in value-creating forces [Phillips 2012, p.36]. The degree of realization of the companies' goals aimed at value creation for the shareholders causes its migration between individual companies and industries [Szczepankowski 2007, p. 36]. Hence value migration analysis can be carried out in an aggregate way at the level of individual industries.
The analysis of the value migration process can be performed using the three stages of value migration model, proposed by A. Slywotzky in his theoretical framework [Slywotzky 1996, p. 46-59]. The essence of the model is the assumption that every company can be in one of the three stages of value migration [Siudak 2001, p. 195], whose short description is provided in table 1.

Table 1. Description of the individual stages of value migration

Phases of value migration	Description
Inflow stage	Limited competition, high increase in market share, high profitability.
Stability stage	Competitive stability, stable market share, stable margins.
Outflow stage	Competitive intensity, declining sales, low profits, competences, resources, talent, and customers leave at an accelerating rate.

Source: own based on [Slywotzky 1996, s. 50]

The purpose of the article is a classification of the industries based on the presented three stages of value migration model and the value migration analysis in the relation company-industry. The study includes all companies quoted on the Warsaw Stock Exchange in 2007, 270 companies in total. A division into separate industries is based on the industry classification proposed by the Warsaw Stock Exchange and documented in the official bulletin „The Main List of the Warsaw Stock Exchange" [2007]. The number of companies assigned to the individual industries is provided in table 2.

Table 2. The number of companies assigned to the individual industries

Industry		Number of companies
1	Building industry	22
2	Developers	9
3	Power industry	5
4	Finance-other	19
5	Financial industry	16
6	Retail	17
7	Wholesale	21
8	Hotels and restaurants	5
9	Computer science	25
10	Construction materials	12
11	Media	12
12	Chemical industry	22
13	Wood and paper industry	7
14	Electromechanical industry	15
15	Light industry	10
16	Metal industry	14
17	Food industry	18
18	Telecommunications	7
19	Services	14
Total:		270

Source: own work based on The Main List of the Warsaw Stock Exchange [2007]

METHOD OF THE INDUSTRY VALUE MIGRATION ANALYSIS

The measurement of the value migration can be performed by adopting the linear ordering method, constructing an appropriate synthetic variable based on three independent variables acting as stimulant [Siudak 2013b]:

1. Share in the economy migration balance

$$\text{SHARE IN THE MIGRATION BALANCE} = \frac{\Delta MVA_i}{\left|\sum_{i=1}^{n} MVA_i\right|} \qquad \left(\sum_{i=1}^{n} MVA_i \neq 0\right) \qquad (1)$$

where: MVA_i – market value added of i company (i=1, …, n).

2. Share in the industry migration balance

$$\text{SHARE IN THE INDUSTRY MIGRATION BALANCE} = \frac{\Delta MVA_i}{\left|\sum_{i \in I_s} \Delta MVA_i\right|} \qquad \left(\sum_{i \in I_s} \Delta MVA_i \neq 0\right) \qquad (2)$$

where: MVA_i – market value added of i company included in s industry, ($i \in I_s$, i=1, …, s).

3. Change *MVA/K*

$$\Delta(MVA/K) = \left(\frac{MVA}{K}\right)_T - \left(\frac{MVA}{K}\right)_{T-1} \qquad (K \neq 0) \qquad (3)$$

where: K – book value of invested capital.

Market value added (*MVA*) is expressed with the following formula [Steward, 1991]

$$MVA = V - K \qquad (4)$$

where: V – gross market value.

Both categories – market value added and invested capital – on which independent variables are based, are additive. Hence the measurement of the value migration can be carried out among companies as well as in an aggregate way at the level of individual industries.

To measure value migration process at the industry level, market value added and invested capital were aggregated separately for each industry.

The construction of the synthetic variable requires that the following parameters are determined: (1) a system for weighting variables, (2) a variable normalization method, and (3) an aggregation function. The influence of the individual variables on the investigated process was expressed with differentiated weights, whose values were as follows:

- share in the economy migration balance – 25%,
- share in the industry migration balance – 25%,
- change *MVA/K* – 50%.

The variable normalization was carried out with the following equation [Siudak 2013b]:

$$z_{ij} = \frac{x_{ij}}{\max\limits_{i}\{x_{ij}\} - \min\limits_{i}\{x_{ij}\}} \left(\max_{i}\{x_{ij}\} - \min_{i}\{x_{ij}\} \neq 0\right) \quad (5)$$

where: z_{ij} – normalized value of j variable for i company,
x_{ij} – value of j variable for i company.
The aggregation was carried out employing the pattern method which used weighted coefficients and was based on Euclid's distance

$$d_i = \sqrt{\sum_{j=1}^{m} w_j \left(z_{ij} - z_{0j}\right)^2} \quad (6)$$

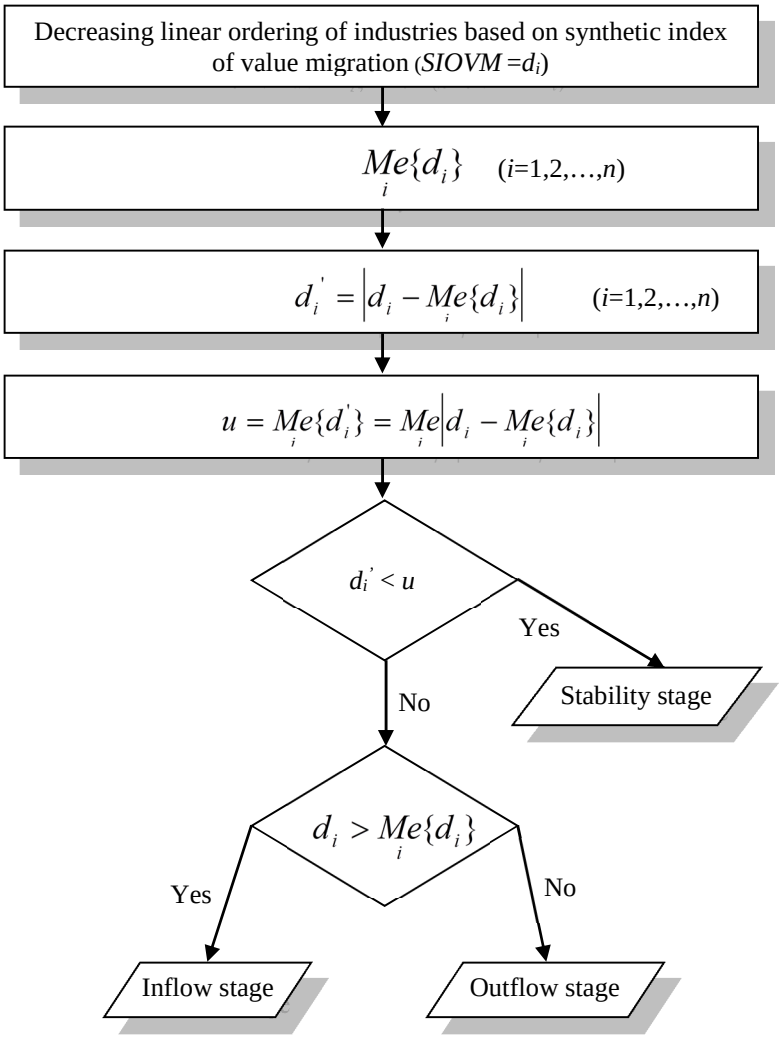
where: d_i – value of the synthetic variable in i company,
w_j – weighted coefficient of j variable (j=1, 2, …, m),
z_{ij} – normalized value of j independent variable in i company (j=1, 2, …, m; i=1, 2, …, n),
between the analysed objects and an element which is an anti-pattern (lower development pole for the parameters above working as a stimulant) – determined by the relation

$$z_{0j} = \min_{i}\{z_{ij}\} \quad (7)$$

The constructed synthetic variable was named the synthetic index of value migration (*SIOVM*). Its values fall within the range ⟨0÷1⟩. The construction is based on the concept of the taxonomic measure of development introduced for the first time by [Hellwig 1968].
Linear ordering of industries in relation to the synthetic variable is non-growing. Lower values of $SIOVM_i$ correspond to a lower level of value migration. Remarks on the ways of creating synthetic variable can be found in the following studies: Hellwig [1968], Gatnar, Walesiak [2004], Grabinski, Wydmus, Zeliaś [1989], Witkowska [2010], Jaworska, Kożuch [2012], Łuniewska, Tarczyński [2006], Łuniewska [2008], Malina [2004], Młodak [2006], Nowak [1990], Ostrowska [2007], Panek [2009], Pociecha, Podolec, Sokołowski, Zając [1988], Walesiak [1996], [2006], and Zeliaś [2000]. The problem of the normalization of independent variables is addressed in the works of: Kukuła [2000], [2012], and Pawełek [2008].
The grounds for synthetic variables (abbreviation *SIOVM*) with regard to the rationale for choosing: diagnostic variables, applied system of weights, methods of normalization and aggregation can be found in Siudak's work (2013a).High estimates of the discriminatory property of the synthetic index of value migration (*SIOVM*) using the measure analysis (*G*) were provided in [Siudak, 2013a, p. 154-168]. The description of the measure (G) can be

found in the following studies [Pociecha Podolec, Sokołowski, Zając 1988], [Nowak 1990].

Figure 1. Algorithm of the classification of the analyzed objects in relation to the three stages of value migration



Me - median

Source: based on D. Siudak [2013a, s. 162]

As diagnostic variables contain outliers, there cannot be applied standard procedures for analysis of the considered set using available methods of cluster analysis. The application of median approach in the presented algorithm of classification makes the classification robust.

RANKING AND CLASSIFICATION OF INDUSTRIES IN TERMS OF THE DEVELOPMENT OF VALUE MIGRATION

Table 3 presents the ranking and division of the analyzed industries in terms of the development of value migration.
The number of industries belonging to the stability stage is 9. In 7 industries value migrated to 3 other industries – respectively in non-growing order of *SIOVM* – DEVELOPERS; POWER INDUSTRY; MEDIA.

Table 3. Ranking and division of industries into three stages of value migration

Threshold value (u)		Median d_i			
0.2395		0.5658			
Industry		$SIOVM = d_i$	d'_i	$d'_i \rightarrow u$	Migration stage
1	Developers	0,9434	0,3775	Larger	Inflow stage
2	Power industry	0,9241	0,3583	Larger	
3	Media	0,8326	0,2668	Larger	
4	Financial industry	0,6988	0,1330	Smaller	Stability stage
5	Retail	0,6268	0,0610	Smaller	
6	Chemical industry	0,6268	0,0609	Smaller	
7	Metal industry	0,6062	0,0404	Smaller	
8	Wholesale	0,6040	0,0382	Smaller	
9	Food industry	0,5665	0,0007	Smaller	
10	Hotels and restaurants	0,5658	0,0000	Smaller	
11	Electromechanical industry	0,5377	0,0281	Smaller	
12	Building industry	0,5145	0,0513	Smaller	
13	Telecommunications	0,3263	0,2395	Equal	Outflow stage
14	Services	0,1892	0,3767	Larger	
15	Construction materials	0,1772	0,3886	Larger	
16	Light industry	0,1626	0,4032	Larger	
17	Finance-other	0,0727	0,4931	Larger	
18	Wood and paper industry	0,0693	0,4966	Larger	
19	Computer science	0,0195	0,5463	Larger	

source: own calculations

What follows is a verification of the hypothesis for the equality of means of the synthetic variable $(\overline{SIOVM})$ among the three classes of industrial value migration using one-way analysis of variance (one-way ANOVA). Before performing the analysis of variance, this method's assumption of equality of variances in groups is tested. Lavene's test for equality of variances provides the following result: $F(2, 16)=1{,}516$; $p=0.249$, which implies that the variances in the individual groups are equal at the level of significance $\alpha=0.05$. This conclusion is confirmed by the test statistics: (1) Hartley's *F-max*=3.435, (2) Cochran's *C*=0.615 and (3) Bartlett's

Chi-square=2.546; *p*=0.280 (the variances in the three sets are equal at the level of significance α=0.05).
The formal representation of the hypotheses for the equality and inequality of the values of the means for the synthetic variable is as follows:
H_0: $\mu_1=\mu_2=\mu_3$
H_1: $\exists j_1 j_2 : \mu_{j_1} \neq \mu_{j_2}$
Table 4 presents the statistics of the *F*-test.

Table 4. Statistics of the *F*-test

Specification	Sum square (*SS*)	*df*	Mean square (*MS*)	*F*	*p*-value
Between groups	1,431	2	0,715	121,38	0,000
Within groups	0,094	16	0,006		
Total	1,525	18			

Source: own calculations

The *F*-test statistics is *F*(2; 16)=121.38 (a value which is much higher than one) and is statistically significant at the level of significance α=0.01. As a result we the hypothesis H_1 is supported, which unambiguously points at a statistically significant difference in the values of the means of the synthetic variable $(\overline{SIOVM})$ between at least two groups.
With multiple comparisons using post-hoc HSD tests proposed by Turkey (for different *N* in groups) and Scheffe, we determine between which classes there are statistically significant differences in the values of the synthetic variable which cause the support of the hypothesis H_1. Table 5 presents approximate *p*-levels for Turkey's and Scheffe's HSD tests.

Table 5. Approximate p-levels for post-hoc tests

Test	Phases of value migration	Inflow stage	Stability stage	Outflow stage
HSD Turkey	Inflow stage		0,0006	0,0002
	Stability stage	0,0006		0,0002
	Outflow stage	0,0002	0,0002	
Scheffe	Phases of value migration	Inflow stage	Stability stage	Outflow stage
	Inflow stage		0,0001	0,0000
	Stability stage	0,0001		0,0000
	Outflow stage	0,0000	0,0000	

Source: own calculations

Both tests show statistically significant differences in the values of means for all comparisons between the individual groups of the industrial value migration, at the level of significance α=0.001.
The mean values of the synthetic variable in the distinct stages of the industrial value migration are as follows: (1) inflow stage: 0.9000; (2) stability stage: 0.5941 and (3) outflow stage: 0.1452. Obviously the largest difference between the mean

values of the synthetic variable is in two extreme classes (inflow stage-outflow stage), which results from the non-growing linear ordering of industries in relation to *SIOVM*.

The proper taxonomic division should have a high diversity of objects between various groups and a low diversity within the individual classes [D. Witkowska, 2002, p. 90]. For the evaluation of the results of the classification we use between groups dissimilarity (high values denote a high degree of dissimilarity of objects between groups) and within group dissimilarity (low values denote a low degree of dissimilarity and simultaneously low diversity of objects within the individual classes), using respectively [Witkowska, 2002, p. 91; Nowak, 1990, p. 190]:

1. Average between groups distance

$$D_{pq} = \frac{1}{N_p N_q} \sum_{O_i \in A_p} \sum_{O_j \in A_q} d(O_i, O_j) \tag{8}$$

where: D_{pq} – average between group distance,
A_p – concentration of i objects O_i (i=1, 2, …, Np), ,
A_q – concentration of j objects O_j (j=1, 2, …, Nq), ,
N_p – number of objects in group A_p,
N_q – number of objects in group A_q,
$d(O_i, O_j)$ – distance between i element of group A_p and j object of group A_q.

2. Average within group distance

$$D_{pp} = \frac{1}{N_p (N_q - 1)} \sum_{O_i \in A_p} \sum_{O_j \in A_p} d(O_i, O_j) \tag{9}$$

where: D_{pp} – average within group distance,
A_p – concentration of O_i, O_j (i, j=1, 2, …, Np),
N_p – number of objects in group A_p,
$d(O_i, O_j)$ – distance between individual elements of group A_p.

Table 6 shows measures of the evaluation of the classification based on mean between groups distance and average within group distance.

Table 6. Average between groups distance and average within group distance

Phases of value migration	Inflow stage	Stability stage	Outflow stage
Inflow stage	0,4964	0,7677	1,4198
Stability stage	0,7677	0,2986	1,0691
Outflow stage	1,4198	1,0691	0,1848

Source: own calculations

We observe lower values of the average within group distance as compared to the values of average between groups distance. Objects are more similar to each other within the individual groups (stages of value migration) and simultaneously more

diversified between the stages in question. It proves that the division of the industries in question into the three stages of value migration is correct.
It should be emphasized that the diversity of industries between the extreme groups, i.e. inflow and outflow of value is higher than in the two other pairs – (1) inflow stage-stability stage and (2) stability stage-outflow stage. It proves that the division is valid.

SUMMARY

The current study has proven the validity of the introduced division of the analysed industries in terms of the three stages of value migration using measures to evaluate the classification and the test of the differences in the values of the means of the synthetic variable in the individual groups.
Importantly, it should also be emphasized that there are more industries at the outflow stage than those at the inflow stage. Three industries captured the value flowing out of seven others, which indicates a concentration of an industrial allocation of capital.

REFERENCES

Ceduła Giełdy Warszawskiej, (2007), Oficjalny Biuletyn, Nr 249/2007 (3710).

Gatnar E., Walesiak M. (2004), Metody statystycznej analizy wielowymiarowej w badaniach marketingowych, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.

Grabiński T., Wydymus S., Zeliaś A. (1989), Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych, PWN, Warszawa.

Hellwig Z. (1968), Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę kwalifikowanych kadr, Przegląd Statystyczny, nr 4.

Jaworska M., Kożuch A. (2012), Ocena przydatności wybranych metod WAP w analizie samodzielności finansowej gmin, Metody ilościowe w badaniach ekonomicznych. T. XIII/1, s. 131-137, Warszawa.

Kukuła K. (2000), Metoda unitaryzacji zerowanej, PWN, Warszawa.

Kukuła K. (2006), Propozycja budowy rankingu obiektów z wykorzystaniem cech ilościowych i jakościowych, Metody ilościowe w badaniach ekonomicznych, T. XIII/1, s. 5-16, Warszawa.

Łuniewska M., Tarczyński W. (2006), Metody wielowymiarowej analizy porównawczej na rynku kapitałowym, PWN, Warszawa.

Łuniewska M. (2008), Ekonometria finansowa, Analiza rynku kapitałowego, PWN, Warszawa.

Malina A. (2004), Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw, Wydawnictwo Akademii Ekonomicznej
w Krakowie, Kraków.

Młodak A. (1990), Analiza taksonomiczna w statystyce regionalnej, Difin, Warszawa.

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

Ostrowska E. (2007), Rynek kapitałowy, PWE, Warszawa.

Panek T. (2009), Statystyczne metody wielowymiarowej analizy porównawczej, Szkoła Główna Handlowa w Warszawie, Warszawa.

Pawełek B. (2008), Metody normalizacji zmiennych w badaniach porównawczych zjawisk ekonomicznych, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.

Phillips P. (2012), Time to Revisit Value Migration. Business Corner, Strategies & Analysis. Rodman Publishing, July.

Pociecha W., Podolec B., Sokołowski A., Zając K. (1988), Metody taksonomiczne w badaniach społeczno-gospodarczych, PWN, Warszawa.

Siudak D. (2013a), Pomiar procesów migracji wartości przedsiębiorstw na polskim rynku kapitałowym, CH Beck, Warszawa.

Siudak D. (2013b), Klasyfikacja przedsiębiorstw ze względu na trzy fazy migracji wartości z wykorzystaniem metod porządkowania liniowego, Przegląd statystyczny, 60/2, s. 251-268.

Siudak M. (2001), Zarządzanie wartością przedsiębiorstwa, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa.

Slywotzky A.J. (1996), Value Migration. How to Think Several Moves Ahead of the Competition. Harvard Business School Press, Boston, Massachusetts.

Stewart G.B. (1991), The Quest for Value. The EVA Management Guide, Harper Business, New York.

Szczepankowski P. (2007), Wycena i zarządzanie wartością przedsiębiorstwa, PWN, Warszawa.

Walesiak M. (1996), Metody analizy danych marketingowych, PWN, Warszawa.

Walesiak M. (2006), Uogólniona miara odległości w statystycznej analizie wielowymiarowej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.

Witkowska D. (2002), Sztuczne sieci neuronowe i metody statystyczne. Wybrane zagadnienia finansowe. CH Beck, Warszawa.

Witkowska D. (2010), Zastosowanie syntetycznych mierników taksonomicznych do pomiaru efektywności chińskich banków, Metody ilościowe w badaniach ekonomicznych, T. XI/2, s. 281-292, Warszawa.

Zeliaś A. (red.), Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 242 – 251

DISTRIBUTIONS OF THE NUMBER OF CLAIMS BY AGE GROUPS OF INSURED IN CIVIL LIABILITY MOTOR INSURANCE PORTFOLIO

Anna Szymańska
Department of Statistical Methods, University of Lodz
e-mail: szymanska@uni.lodz.pl

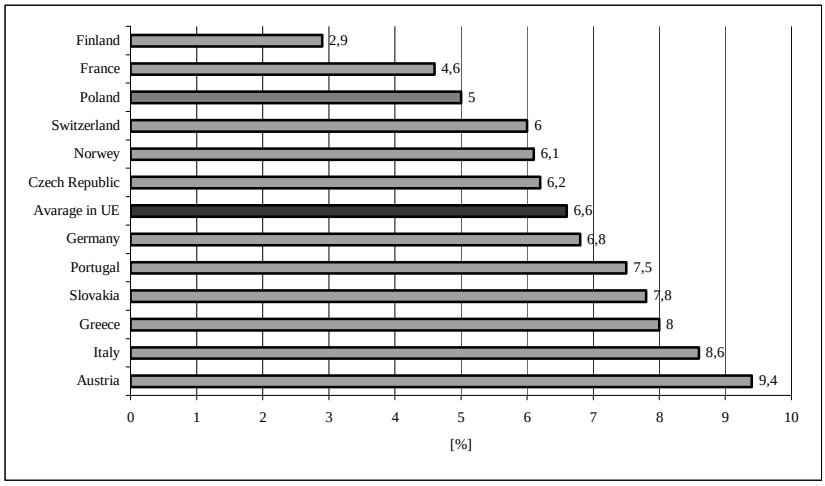
Abstract: In the process of calculation of the insurance premium in civil liability motor insurance the knowledge of distribution of the number and value of paid claims is required. The paper presents the methods of assessing the degree of the fit of theoretical distributions of the number of claims to empirical distributions in civil liability motor insurance in the example of data from one of the insurance companies in Poland. The distributions of the number of claims in separate age groups of drivers were also analyzed and their compatibility to the theoretical distribution of the number of claims in the portfolio was assessed.

Keywords: distribution of the number of claims, civil liability motor insurance of vehicle owners, driver's age

INTRODUCTION

Civil liability motor insurance of vehicle owners are most often concluded insurance in our country. Polish motor insurance market, in comparison to other European countries, has one of the lowest frequency of claims (cf. Figure 1).

Figure 1. Frequency of the damage in chosen European countries in 2007



Source: Europejski Rynek Ubezpieczeń Komunikacyjnych, PIU, Warszawa 2010

At the moment of establishing the insurance premium insurance company does not know the future costs of compensation, but it can estimate them on the basis of historical data. In civil liability motor insurance ratemaking is a two-step process [Antonio et al. 2012]. In the first stage - called *a piori* - the base premium based on known risk factors is determined. Then in the base contribution discounts and increases, mainly resulting from the mileage of claims are included in the previous insurance period (bonus-malus system) and of such factors like e.g. age of insured, the period of holding the driving license, or the age of the car. This stage is called *a posteriori* ratemaking, and its' result is an assigned premium

Both tariffs, *a priori*, and *a posteriori* require from the actuary the determination of distributions of theoretical random variables describing the number and value of paid claims. In the actuarial literature, the tests which are usually used to evaluate the relevance of the theoretical distribution to empirical data, are: goodness-of-fit test χ^2 and test statistics based on λ – Kołmogorow [Panjer et al. 1992, Domański 1990]. However, in the case of the distribution of the number of claims in car automobile insurance the number of classes is often not larger than four, which means that the number of degrees of freedom of the chi - squared test is too small. Moreover, most policies in the insurance portfolios are concentrated in the number zero class, which results in the distortion of the distribution. Portfolios are usually large, resulting in the chi-squared test generally rejecting the null hypothesis even though empirical data match theoretical distribution closely. In such cases, measures assessing the degree of the fit of the theoretical distribution to empirical data may be found in statistical literature, such as the standard deviation of the differences in relative frequencies, the index of structures similarity, index of distribution similarity, ratio of the maximum

difference of relative frequencies, ratio of the maximum difference of cummulative distribution functions [Kordos 1973].

The study analyzed the distribution of the number of claims in civil liability motor insurance portfolio of passenger cars of individuals of one of the insurance companies operating on the Polish market in 2006. In the audited insurance company base premium is determined based on two factors: region of registration of the vehicle and engine capacity. One of the factors of a posteriori tariff is the age of the driver, the insured persons under the age of 25 years are charged by increase of the basic premium of 100% to 200%. Insured from 25 to 28 years of age have increase in the amount of 30% to 70% of the basic premium. Height of the increase depends on many factors, such as whether the insurance contract is concluded with the company for the first time, or is a continuation of previous insurance, if the insured has another vehicle insured in this company and whether he continues insurance without claims.

The aim of the study is to assess the legitimacy of used in studied insurance company premium increases by virtue of age of the insured based on an analysis of distributions of the number of claims in particular age groups of drivers. In the paper 5000 attempt drawn from the portfolio of the civil liability motor insurance of one of insurance companies was disposed.

It should be highlighted that in Poland access to individual data is very difficult and insurers are reluctant to provide them. Therefore, in the study the name of the insurance company was not given.

THE CHOICE OF THE DISTRIBUTION OF THE NUMBER OF CLAIMS IN THE CIVIL LIABILITY MOTOR INSURANCE

The choice of the distribution of the number of claims in civil motor liability insurance depends on the relationship between the sample expected value and variance [Heilmann 1988]. Three distributions are considered: binomial, Poisson and negative binomial, which belong to the class $(a,b,0)$ [Klugman et al. 2004].

Definition: Let p_k be the probability function of a discrete random variable. It is a member of the $(a,b,0)$ class of distributions, provided that there exists constants a and b such that:

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k}. \tag{1}$$

Lemma: The only members of the class of discrete distributions on the non-negative integers satisfying (1) are the Poisson, binomial, negative binomial and geometric distributions.

Let the random variable X represent the number of claims from individual policy or a policy portfolio. According to the paper [Panjer et al. 1992] the pre-selection of the theoretical distribution of the number of claims can be based on the calculated moments of the sample and the frequency coefficients.

Let $X_1, X_2, ..., X_n$ be an i.i.d. random sample. In case of aggregated data, where we know only the number of policies for the number of claims, simple sample moments usually are:

$$M_r = \frac{1}{n}\sum_{l=1}^{\infty} k^r N_k, \quad r = 1,2,... \tag{2}$$

where N_k is the number X_i, for which $X_i = k$, (k=0,1,2,.....), $n = \sum_{k=0}^{\infty} N_k$.

The first three central moments of the sample are: $\overline{X} = M_1$; $S^2 = M_2 - M_1^2$; $K = M_3 - 3M_2M_1 + 2M_1^3$.

For class (a, b,0) and since we might expect $\frac{N_k}{n}$ to be close p_k, it follows that:

$$T(k) = (a+b) + ak, \quad k = 0,1,2,... \tag{3}$$

When the function given by equation (3) is linear, whose slope coefficient:

- is zero and $\overline{X} = S^2$; then to describe the distribution of the number of claims the Poisson distribution is suggested;
- is negative and $\overline{X} > S^2$; then the binomial distribution can be assumed;
- is positive and $\overline{X} < S^2$; then the negative binomial distribution should be chosen.

When the function described by equation (3) grows faster than linearly, the skewness of the distribution should be taken into account.

Let us denote: $W = 3S^2 - 2\overline{X} + 2\frac{(S^2 - \overline{X})^2}{\overline{X}}$. If the equation: $K = W$ holds, the negative binomial distribution should model the number of claims well. If inequality $K < W$ holds, the generalized Poisson Pascal distribution, or its special case the Poisson-inverse normal distribution can be used to describe the distribution of the number of claims [Tremblay 1992]. If the inequality $K > W$ holds, the Neyman type A, Polya - Aeppli , Poisson - Pascal or negative binomial distributions are suitable for modeling the distribution of the number of claims.

STATISTICAL MEASURES OF FIT OF THE EMPIRICAL AND THEORETICAL DISTRIBUTIONS

Deviation of the differences in relative frequencies is a measure given by:

$$S_r = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(\gamma_i - \hat{\gamma}_i)^2}, \tag{4}$$

where: k - the number of classes γ_i - empirical frequencies $\hat{\gamma}_i$ - theoretical frequencies. The measure is equal to zero in case of full compliance of the empirical and theoretical distribution. Practice shows that the value $S_r \leq 0{,}005$ is an evidence of high compliance of schedules, if $0{,}005 \leq S_r < 0{,}01$ the compatibility of tested distributions is satisfactory and $S_r \geq 0{,}01$ shows significant deviations between the studied distributions.

The index of *structures similarity* is given by:

$$w_p = \sum_{i=1}^{k} \min(\gamma_i, \hat{\gamma}_i). \tag{5}$$

The index value is in the range [0,1]. The closer the value is to the unity, the more similar the structures of the studied distributions are.

Index of distribution similarity is determined by the equation:

$$W_p = 1 - \frac{1}{2} \sum_{i=1}^{k} |\gamma_i - \hat{\gamma}_i| . \tag{6}$$

Distribution similarity index is equal to 100% for fully compatible distribution. The distributions show high compatibility when $W_p \geq 0{,}97$. If $W_p < 0{,}95$ distributions show significant differences.

Ratio of the maximum difference of relative frequencies is given by the formula:

$$r_{\max} = \max_i |\gamma_i - \hat{\gamma}_i| . \tag{7}$$

This ratio is equal to zero for distributions fully compatible. If $r_{\max} < 0{,}02$, it is believed that the distributions are quite compatible.

Ratio of the maximum difference of cummulative distribution functions is given by the equation:

$$D_{\max} = \max_i \left| F_i - \hat{F}_i \right| , \tag{8}$$

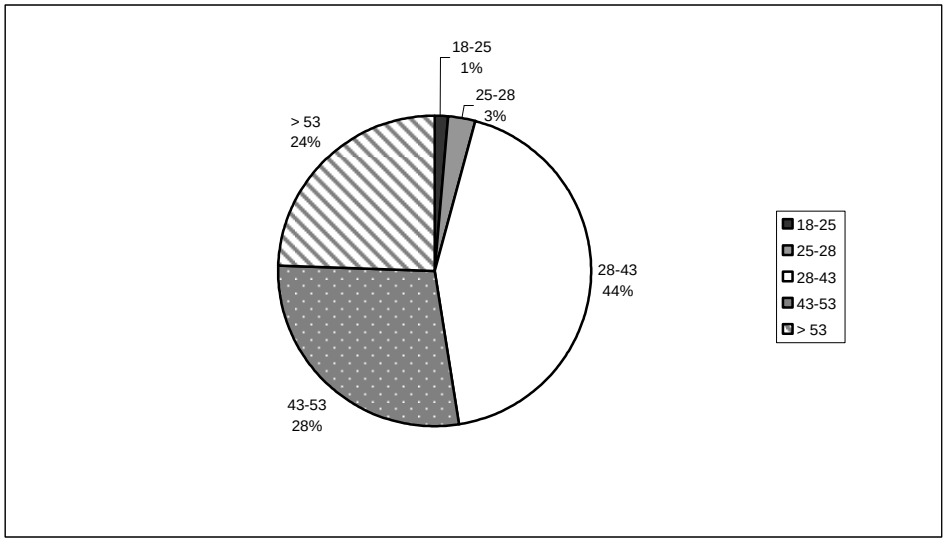
where: $F_i = \sum_{j=1}^{i} \gamma_j$ - value of the empircial cummulative distribution function, $\hat{F}_i = \sum_{j=1}^{i} \hat{\gamma}_i$ - value of the theoretical cummulative distribution function. This ratio is equal to zero for fully consistent distributions.

EMPIRICAL EXAMPLES

In this part of the study the distribution of the number of claims paid in civil liability motor insurance portfolio and in separate age groups of drivers of analyzed insurance company in 2006 was investigated.

Figure 2 shows the structure of according to the age of the insured of the civil liability motor insurance portfolio of the analyzed insurance company.

Figure 2. Portfolio structure by age of the insured in 2006



Source: own calculations

On the basis of empirical data about the number of paid claims the numerical characteristics of distributions in particular age groups of the insured and in the portfolio were calculated (see Table 1).

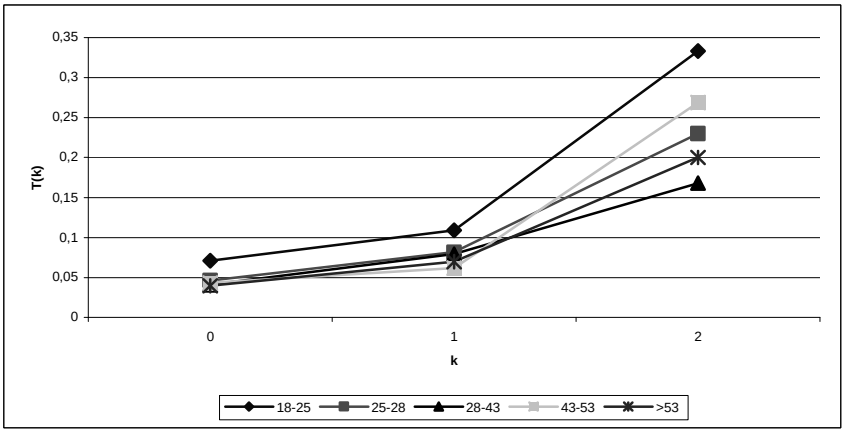
Table 1. Numerical characteristics of the empirical distributions of the number of claims in the age groups of the insured in civil liability motor insurance portfolio insurance company in 2006

Numerical characteristics of distributions	portfolio	Age of insured [years]				
		18-25	25-28	28-43	43-53	>53
$\overline{X}$	0.0441	0.0743	0.0477	0.0425	0.0442	0.0411
S^2	0.0457	0.0783	0.0490	0.0443	0.0455	0.0426
a	0.0308	0.0383	0.0356	0.0386	0.0189	0.0298
K	0.0492	0.0875	0.0514	0.0482	0.0485	0.0459
W	0.0490	0.0869	0.0518	0.0481	0.0481	0.0457

Source: own calculations

In the next stage the compatibility of the empirical distribution in the portfolio and in the particular age groups of the insured with selected theoretical distributions was examined (results are shown in Table 2). In the selection of theoretical distributions the linearity of the function frequency and the relationships between the parameters of distributions from the sample were taken into account.. The frequency functions for particular age groups of the insured are not linear (see Figure 3), which suggests consideration of the skewness of distributions. For each of the considered distributions, in addition to the age group of 25-28 years, the following relations hold: $a > 0, \overline{X} < S^2$ and $K > W$ (see Table 1). In further analyzes the following theoretical distributions were considered: Poisson (*Poi*), negative-binomial (*NB*), the Poisson-inverse normal (*PGI*) and Neyman A (*NA*). Generalized Poisson-Pascal distribution in this case could not be considered due to the assumptions about the parameters of this distribution not fulfilled by the empirical distribution. The parameters of distributions for each age group and for the portfolio estimated by the maximum likelihood method in the case of the Poisson distributions, by the method of moments in the case of the negative binomial distributions and by means of recursive formulas for the Poisson-inverse normal and Neyman A distributions.

Figure 3. The frequency functions in the age groups of insured



Source: own calculations

Table 2. Measures of the degree of distributions' fitting

portfolio				
measure	distribution			
	Poi	*NB*	*PIG*	*NA*
S_r	0.00240807	0.00170703	0.00175441	0.01632648
w_p	0.99597445	0.99716675	0.99658889	0.95944651
W_p	0.99998550	0.99999272	0.99999231	0.99933362
r_{max}	0.00001620	0.00000786	0.00000791	0.00132328
D_{max}	0.00355339	0.00280417	0.00281276	0.04055349
age group:18-25				
measure	distribution			
	Poi	*NB*	*PIG*	*NA*
S_r	0.00164310	0.00026806	0.00098381	0.02734804
w_p	0.99692915	0.99945118	0.99790153	0.93510215
W_p	0.99999325	0.99999982	0.99999758	0.99813021
r_{max}	0.00000942	0.00000017	0.00000325	0.00372750
D_{max}	0.00170156	0.00026007	0.00200691	0.06474332
age group: 25-28				
measure	distribution			
	Poi	*NB*	*PIG*	*NA*
S_r	0.00076233	0.00011774	0.00059518	0.01879855
w_p	0.99859679	0.99976459	0.99876536	0.95617589
W_p	0.99999855	0.99999997	0.99999911	0.99911654
r_{max}	0.00000192	0.00000003	0.00000137	0.00176378
D_{max}	0.00070171	0.00011864	0.00103962	0.04382411
age group: 28-43				
measure	distribution			
	Poi	*NB*	*PIG*	*NA*
S_r	0.00086505	0.00002989	0.00044491	0.01676594
w_p	0.99839956	0.99993788	0.99924540	0.96093832
W_p	0.99999813	1.00000000	0.99999951	0.99929726
r_{max}	0.00000256	0.00000000	0.00000056	0.00140318
D_{max}	0.00083739	0.00002819	0.00075343	0.03903725
age group: 43-53				
measure	distribution			
	Poi	*NB*	*PIG*	*NA*
S_r	0.00053232	0.00012409	0.00037277	0.01766546
w_p	0.99900607	0.99975023	0.99916067	0.95915272
W_p	0.99999929	0.99999996	0.99999965	0.99921983
r_{max}	0.00000099	0.00000004	0.00000045	0.00155876
D_{max}	0.00054717	0.00012324	0.00078425	0.04078068

Table 2. cont.

age group: >53				
measure	distribution			
	Poi	*NB*	*PIG*	*NA*
S_r	0.00068258	0.00006351	0.00038041	0.01638094
w_p	0.99873325	0.99987112	0.99928145	0.96200831
W_p	0.99999884	0.99999999	0.99999964	0.99932916
r_{max}	0.00000160	0.00000001	0.00000042	0.00133998
D_{max}	0.00067190	0.00006244	0.00069646	0.03795317

Source: own calculations

Analyzing the values of the measures of fit from Table 2 it follows that in the portfolio, as well as in each considered age group of insured the theoretical distribution, best fit to the empirical data on the number of claims is the negative binomial distribution. The question is: how compatible are the distributions of the number of claims in each age groups of the insured with the distribution of the number of claims in the whole portfolio? In the next stage of the analysis the compatibility of empirical distributions in particular age groups of insured with the determined for the portfolio negative binomial distribution[1] with parameters $\alpha = 1.1227; \beta = 27.5032$ was rated (see Table 3).

Table 3. Measures of the degree of fitting of the compatibility of empirical distributions of the number of claims in the age groups with negative binomial distribution of the portfolio

measure	Age of insured [years]				
	18-25	25-28	28-43	43-53	>53
S_r	0.01667476	0.00216927	0.00102821	0.00030703	0.00170703
w_p	0.97252388	0.99643296	0.99834551	0.99940429	0.99716675
W_p	0.99930488	0.99998824	0.99999736	0.99999976	0.99999272
r_{max}	0.00075481	0.00001229	0.00000265	0.00000029	0.00000786
D_{max}	0.02747384	0.00350534	0.00162692	0.00029194	0.00280417

Source: own calculations

Analyzing the results presented in Table 3, we find that in the case of insured above 25 years old the values of measurements of fitting of distributions show a high compatibility of distributions of the number of claims in these groups with distribution of the number of claims in the portfolio. The highest compatibility was obtained in group of insured aged 43 to 53 years and above 53 years. Fit of distributions of is unsatisfactory in the group of insured who are under the age of

[1]negative binomial distribution: $P(X=k) = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)k!}\left(\frac{\beta}{1+\beta}\right)^{\alpha}\left(\frac{1}{1+\beta}\right)^{k}$

25 years. Perhaps it is the result of a small number of policies in this age group (50 policies). Analysis should be performed on larger sample.

CONCLUSIONS

In assessing the consistency of distributions, in most cases, due to the nature of the data on the number of claims in motor liability insurance, the chi-square and λ-Kolmogorowa test cannot be used. Measures proposed in the paper offer a possibility to assess the goodnes-of-fit of empirical and theoretical distributions. Distribution of the number of claims in all groups of insured civil liability motor insurance portfolio of the studied insurance company is negative binomial.

In the analyzed civil liability motor insurance portfolio clearly differed in terms of the number of claims was a group of under 25 years old. Average number of compensations paid in 2006 of a single policy in this group was 0.0743, 0.0441 in the portfolio The average value of compensation paid in 2006 in a group of the drivers aged to 25 years was equal to about 8 thousand zlotys, in the portfolio of approximately 5.5 thousand zlotys. Despite the fact that the insured up to the age of 25 constituted only 1% of the portfolio, the insurance premium for this group of the insured should be estimated separately. Distribution of the number of claims in a group of drivers from 25 to 28 years old does not confirm the need to use such large increases of premiums to them. Treating the age of the insured as variable in *a priori* tariffication, could reduce the base premiums for the insured who are over the age of 25 years, while increasing the competitiveness of the insurer in the market.

REFERENCES

Antonio K., Valdez E. (2012) Statistical concepts of a priopri and a posteriori classification in insurance, AStA Adv Stat Anal 96, str. 187-224.

Domański Cz. (1990) Testy statystyczne, PWE, Warszawa.

Heilmann W.R. (1988) Fundamentals of Risk Theory, Verlag Versiecherungswirtschaft, Karlsruhe.

Klugman S. A., Panjer H.H., Willmot G.E. (2004) Loss Models. From Data to Decisions, J. Wiley &Sons, New York.

Kordos J. (1973) Metody analizy i prognozowania rozkładów płac i dochodów ludności, PWE, Warszawa.

Panjer H.H., Willmot G.E. (1992) Insurance risk models, Society of Actuaries, Schaumburg.

Tremblay L. (1992) Using the Poisson Inverse Gaussian in Bonus-Malus Systems, ASTIN Bulletin 22(1), str. 97-106.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 252 – 260

THE APPLICATION OF DISCRIMINANT ANALYSIS IN FORECASTING OF INVESTORS' REACTION TO MACROECONOMIC NEWS ANNOUNCEMENTS

Tomasz Wójtowicz[1]
Department of Applications of Mathematics in Economics
AGH University of Science and Technology in Cracow
e-mail: twojtow@agh.edu.pl

Abstract: Macroeconomic news announcements, particularly concerning the U.S. economy, have a significant impact on stock markets. Recent studies show that stock prices react significantly as soon as macroeconomic news is announced. However, the strength of the reaction and its duration depends on the market and on the news announced. In this paper, we study the applicability of discriminant analysis in the prediction of direction of changes of the main indices of stock exchanges in Warsaw and Vienna after release of the Employment Report by the U.S. Bureau of Labor Statistics.

Keywords: macroeconomic news announcements, intraday data, discriminant analysis, nonfarm payrolls

INTRODUCTION

Every day, stock markets are deluged by a variety of news. Among them, announcements of macroeconomic news describing the state of economies seem to be most important because they impact whole stock markets. European stock markets are mainly influenced by news announcements from European economies and from the U.S. economy. However, only U.S. macroeconomic news announcements are released during trading hours of European stock markets. News from European economies are mainly released before opening of stock markets or after their closure. This fact and the place of the U.S. in the global economic

[1] Financial support for this paper from the National Science Centre of Poland (Research Grant DEC-2012/05/B/HS4/00810) is gratefully acknowledged.

system is the main reason of the importance of U.S. macroeconomic news announcements for investors on European stock markets. It is also a source of strong and immediate impact of U.S. macroeconomic data on stock prices on European stock markets.

The impact of U.S. macroeconomic news announcements on European stock markets was first studied (and confirmed) on developed European markets. Nikkinen and Sahlström [Nikkinen and Sahlström 2004] study the impact of monthly announcements of CPI, PPI and Unemployment Rate on German and Finnish stock markets. On the basis of daily data from January 1996 to December 1999 they show that both markets react only on announcements of U.S. data while they remain unaffected by domestic macroeconomic news announcements. The strongest reaction on the both markets is observed when UR or PPI is announced.

Nikkinen et al. [Nikkinen et al. 2006] extend the above results and show differences in reaction to U.S. macroeconomic data announcements between emerging and developed markets in various parts of the world. Among others, Nikkinen et al. conclude that European developed markets strongly react to news about the American economy, particularly to announcements about inflation and employment situation. On the other hand, the reaction of European emerging markets, including Poland is insignificant.

Quite different result is presented by Gurgul et al. [Gurgul et al. 2012] who examine the effect of U.S. macroeconomic data announcements solely on the Warsaw Stock Exchange. They show that information about CPI and Industrial Production significantly affects daily returns of the WIG20 (the main index of the WSE), but there is no visible reaction to Unemployment Rate announcements. The insignificant impact of UR announcements on the WSE is probably due to different interpretation of unemployment rate during expansion and contraction periods (see for example [Andersen et al. 2007]).

The abovementioned paper deal with daily data, however, application of intraday data allows to more detailed description of stock market reaction to U.S. macroeconomic data announcements. For example, Andersen et al. [Andersen et al. 2007] show significant impact of unexpected news on high-frequency stock, bond and exchange rate data in U.S., Germany and UK. Significant reaction of intraday returns on European stock markets is also confirmed by Harju and Hussain [Harju and Hussain 2011] who examine five-minute returns of the CAC40, DAX, SMI and FTSE 100. One of the conclusion of Harju and Hussain is that the impact of U.S. macroeconomic data is immediate. Significant changes in returns are observed in the first five minutes after news announcements. The strongest impact is implied by Unemployment Rate and Durable Goods Orders announcements. Similar results for five-minute returns are presented by Będowska-Sójka [Będowska-Sójka 2009] who examines reaction of the CAC40, DAX and WIG20. Also Gurgul et al. [Gurgul et al. 2013] show strong and immediate reaction of the WIG20 in the first five minutes after U.S. macroeconomic data

announcements. However, they indicate that the strongest reaction is observed after Nonfarm Payrolls announcements.

Analysis of more frequent data, i.e. 1-minute returns of the DAX index, allows Dimpfl [Dimpfl 2011] to prove that investors on the FSE react to U.S. data in the first minute after a news release and the significant reaction is observed in the first 10 minutes. Similar conclusion for the Polish stock market is presented by Gurgul and Wójtowicz [Gurgul and Wójtowicz 2014].

All the abovementioned papers prove significant impact of U.S. macroeconomic news announcements on European stock markets. In fact, investors react as soon as news is released. This significant reaction is observed after announcements of variety of macroeconomic indicators describing the U.S. economy. However, Andersen et al. [Andersen et al. 2007] underlines the importance of Nonfarm Payrolls announcements. This is also confirmed by the results for the Warsaw Stock Exchange [Suliga and Wójtowicz 2013, Gurgul et al. 2013, Gurgul and Wójtowicz 2014].

It is natural to ask if stock prices reaction to U.S. macroeconomic data announcements (its direction or strength) can be predicted and what variables should be applied to ensure correctness of such prediction. In this paper we examine whether investors' reaction on the Warsaw Stock Exchange and on the Vienna Stock Exchange can be forecasted on the basis of announced news about the U.S. labor market. More precisely, we apply discriminant analysis to predict the direction of changes in stock market indices implied by announcements of the Employment Report. This study will show whether such prediction is feasible and what is its effectiveness. The analysis of both stock markets, the WSE and the VSE, also allows us to compare results for emerging market and small developed market.

The structure of this paper is as follows. Next section describes the data under study. Empirical results are presented and discussed in the third section. Short summary concludes the paper.

DATA

Nonfarm Payrolls (NFP) is one of the macroeconomic indicators included in the Employment Report published monthly by the U.S. Bureau of Labor Statistics. The other important indicators in the report are: Unemployment Rate (UR), Average Hourly Earnings (AHE) and Average Workweek (AW). The Employment Report is usually released on first Friday of the month at 8:30 EST (Eastern Standard Time) i.e. at 14:30 CET (Central European Time). However, due to differences in introduction of the Daylight Saving Time in the U.S. and Europe some of the announcements (in November) are released at 13:30 CET.

The Report describes the U.S. labor market in the month prior to release date. Thus, its importance comes from the fact that it is released at the beginning of the month and its publication precedes other macroeconomic data

announcements. It is usually first official publication in the month that describe such important aspect of economic situation in U.S. Thus, values of other macroeconomic indices (e.g. about inflation of production) can be partially forecasted on the basis of information contained in the Employment Report.

The Employment Report contains four main macroeconomic indicators. Each of them describes different aspect of the U.S. labor market and its changes can be differently interpreted by investors depending on the situation of the U.S. and global economy. However, as indicated by [Suliga and Wójtowicz 2013] on the example of the Warsaw Stock Exchange, investors' reaction is in line with unexpected news about NFP. The other indicators in the Report (AW, AHE and UR) play less important role in determining stock prices changes.

Unexpected news is the difference between announced value of an indicator and its expected value approximated by consensus forecast published by news agencies few days before the announcement. Released value of NFP greater than consensus is good news and in general it implies positive reaction of investors and increase in stock prices. On the other hand, NFP smaller than expected is bad news and is followed by negative returns.

The analysis in this paper is based on the Employment Report announcements released between January 2009 and December 2013. The whole sample contains 56 announcements released on trading days on the WSE and 57 announcements released on trading days on the VSE. We also consider consensus published by Bloomberg a few days before an announcement. Bloomberg publications contain consensus value of each indicator together with consensus range (minimum and maximum). In the whole dataset for the WSE good news was released 27 times and bad news was also released 27 times. In two cases announced value of NFP was identical with consensus. In the case of the VSE there are 28 good and 27 bad news announcements.

In order to study the predictability of investors' reaction to NFP announcements on the WSE and the VSE we apply intraday data. The ATX, the main index of the VSE, is quoted every second, but the WIG20 is quoted every 15 seconds. Hence, to make results for both markets comparable, the analysis in the paper is based on 15-sec log-returns of the ATX and WIG20. This also allows us to take into account the fact that stock markets react immediately to U.S. macroeconomic news announcements.

The Employment Report is usually announced at 14:30 CET and thus the first return when new information can be observed is computed at 14:30:15 on the basis of index prices at 14:30:00 and 14:30:15. To shortly characterize the impact of unexpected news about NFP on the WIG20 Figure 1 presents means of cumulative returns of the WIG20. To simplify the description we restrict our attention only to news announced at 14:30. Cumulative returns R_t are calculated every 15 seconds between 14:30 to 16:00 according to the formula:

$$R_t = 100(lnP_t - lnP_0),$$

where P_0 is the value of the WIG20 at 14:30 and P_t is the value of the WIG20 at time t. These cumulative returns describe the behavior of the WIG20 relative to the moment of news release. Means presented in Figure 1 are computed in two clusters: "good news" when announced the value of NFP was greater than consensus and "bad news" when the value of NFP was smaller than expected.

Figure 1. Means of cumulative returns of the WIG20 implied by good and bad news about NFP in the period January 2009 - December 2013.

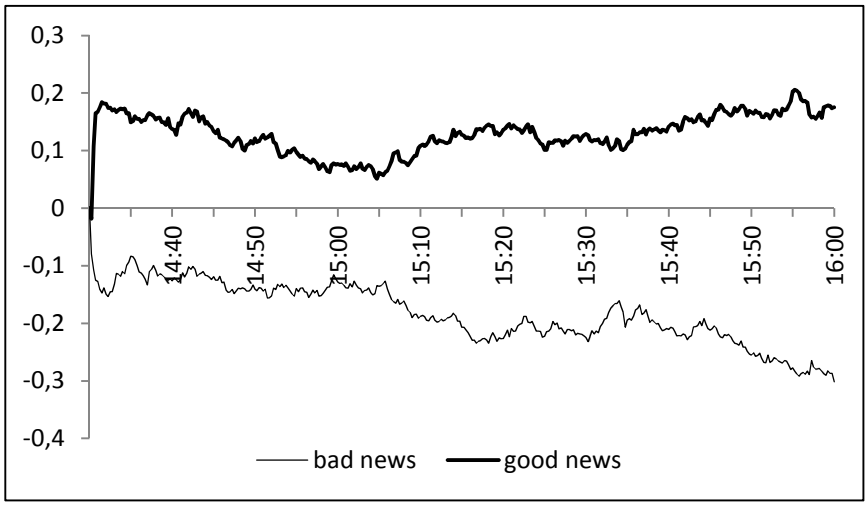


Figure 1 confirms the immediate reaction of the WIG20 to NFP announcements. The change in means of cumulative returns is observed as soon as news is released. Means of cumulative returns implied by good news quickly increases to reach 0.18% at 14:32. Then the impact of good news seems to disappear. After about 40 minutes cumulative returns stabilize. Similarly, very fast and strong reaction of the WIG20 is also observed after bad news announcements.

EMPIRICAL RESULTS

In order to describe more precisely the reaction of the WIG20 and ATX to NFP announcements and to make it more readable we restrict our attention to cumulative returns computed in the following horizons after news release: 15 seconds, 30 seconds, 1 minute, 5 minutes, 10 minutes, 20 minutes. When an announcement is released at 14:30 these cumulative returns are computed at: 14:30:15, 14:30:30, 14:31, 14:35, 14:40 and at 14:50. Additionally, we consider cumulative returns from the moment of news release to the end of a trading session. Reaction to news announcement is positive when cumulative return at given time horizon is greater than zero. Otherwise, the reaction is seen as negative.

There is no need to consider separate class of neutral reaction, because only one cumulative return under study is equal to 0.

NFP is the most informative indicator included in the Employment Report. Hence, forecast investors' reaction after the Report release only on the basis of value and consensus of NFP. Each announcement is described by four variables. First of them is the actual value of NFP published by the Bureau of Labor Statistics. The other three are values published by Bloomberg before news release: consensus value, consensus minimum and consensus maximum. The difference between the last two values describes the accuracy of the estimation of NFP made by market experts. It is also a measure of a discrepancy between investors' expectations about the future value of NFP.

In order to evaluate predictability of investors' reaction we apply discriminant analysis to construct appropriate classification functions and then validate them. We describe the analysis procedure in detail in the case of the WSE. For each time horizon under study the procedure is as follows. From the whole sample of 56 4-dimensional vectors containing described above data associated with the announcements we randomly generate the subsample of length 46. This is a training set. In the case of the VSE, the training set contains 47 elements. Remaining 10 data will be applied later. On the training set we perform linear discriminant analysis and construct linear classification functions. Independent variables are: value of NFP, consensus, consensus minimum and consensus maximum. The grouping factor is a reaction of investors (positive or negative) described by the sign of cumulative returns computed on given time horizon. Then, on the basis of 10 remaining data that were not used to estimate classification functions we compute classification matrix. It contains numbers of correctly and incorrectly classified data from testing set and it describes the predictive validity of the classification functions. Because we have very limited dataset we repeat the whole above procedure 200 times. It means that for each time horizon we perform 200 discriminant analyzes and verify them on the basis of 2000 randomly chosen data. This allows us to compute posterior probabilities of correct classification to each group. They describe what is the probability that announcement classified as implicating positive (negative) investors' reaction really implies positive (negative) cumulative returns at given time horizon.

Table 1 presents the results of discriminant analysis performed accordingly to described above procedure. For each market index under study and each time horizon we report two probabilities that forecast of negative (positive) investors' reaction is correct. As a comparison we report also percentage of correct naive forecast that good (bad) news implies positive (negative) cumulative returns. Comparison of corresponding probabilities indicates that in the case of the WSE the applications of discriminant analysis to forecast the WIG20 changes impied by release of the Employment Report gives, in general, more adequate results than naive forecasts based only on the comparison of announced NFP value with its

consensus. In the case of the ATX, similar results hold true only when negative changes are forecasted.

Application of discriminant analysis is the most efficient when changes in the WIG20 in the first 15 seconds after news release are forecasted. The estimated posterior probability of a correct forecast is equal to 0.92 and 0.898 for negative and positive investors' reaction, respectively. The probability of correct prediction decreases when its horizon increases. In general, prediction about closing prices is more certain when the naive method is applied. However, prediction of changes in market indices under study even in a 20-minute period after news announcements is correct in about 75% of cases.

Table 1. Posterior probabilities of correct classification when cumulative returns are calculated on the basis of index prices at the time of announcement.

time horizon	discriminant analysis				naive prediction			
	WIG20		ATX		WIG20		ATX	
	negative reaction	positive reaction	negative reaction	positive reaction	negative reaction	positive reaction	negative reaction	positive reaction
15 sec	0.920	0.898	0.813	0.759	0.815	0.741	0.704	0.750
30 sec	0.875	0.845	0.868	0.825	0.778	0.778	0.741	0.821
1 min.	0.837	0.787	0.892	0.777	0.741	0.815	0.778	0.821
5 min.	0.847	0.785	0.722	0.743	0.667	0.704	0.630	0.750
10 min.	0.813	0.827	0.709	0.708	0.593	0.815	0.667	0.750
20 min.	0.747	0.690	0.754	0.742	0.667	0.667	0.667	0.750
close	0.642	0.598	0.606	0.617	0.704	0.593	0.741	0.679

Practical applications of the results from Table 1 may be questioned because prediction is made at exactly the same moment as U.S. data are published and all the information may not yet be available. To resolve these doubts we also perform classification 15 seconds after news announcements. In this case cumulative returns are computed relative to index prices at 14:30:15 (at 13:30:15 when news is released at 13:30). Results of this analysis are presented in Table 2 together with naive classification. When analyzing the posterior probabilities in Table 2 we must take into account that stock markets reacts immediately to U.S. macroeconomic data announcements. It is very well visible in Figure 1. The average of WIG20 returns in the first 15 seconds after good news announcements is about 0.11%. In the same period after bad news announcements the WIG20 decreases about 0.08%. These are the large portions of the total changes implied by the announcements under study. Hence, after 15 seconds from an announcement changes in stock market indices are much slower and cumulative returns stabilize.

In general, the probabilities of correct predictions in Table 2 are smaller than the corresponding probabilities in Table 1. However, there is still over 70% chance that the forecast of the WIG20 behavior made in 10-minute horizon will come true.

It is particularly important when we take into account that significant impact of U.S. macroeconomic news announcements is observed only in few minutes after news release. Similarly to results in Table 1, the quality of predictions decreases when time horizon increases. However, in the case of the WIG20, application of discriminant analysis in short horizon gives better results than naive prediction. When the behavior of the ATX is predicted, discriminant analysis outperforms naive method only when negative reaction is forecasted. When positive changes of the ATX are predicted it is still better to simply compare NFP value with its consensus.

It is worth noting here that the posterior probabilities of correct classification presented in Tables 1 and 2 for the WIG20 and the ATX are quite close. Predictions of cumulative returns computed on the basis of index prices at the moment of the announcements are more certain in the case of the WIG20. On the other hand, better results for the ATX are observed when the analysis starts 15 seconds later. This difference is mainly due to slower and weaker reaction of investors on the Vienna Stock Exchange.

Table 2. Posterior probabilities of correct classification when cumulative returns are calculated on the basis of index prices 15 seconds after news release.

time horizon	discriminant analysis				naive prediction			
	WIG20		ATX		WIG20		ATX	
	negative reaction	positive reaction	negative reaction	positive reaction	negative reaction	positive reaction	negative reaction	positive reaction
1 min.	0.758	0.708	0.795	0.789	0.667	0.704	0.667	0.786
5 min.	0.645	0.677	0.673	0.710	0.593	0.556	0.593	0.786
10 min.	0.699	0.744	0.677	0.722	0.593	0.741	0.593	0.750
20 min.	0.643	0.582	0.700	0.657	0.630	0.593	0.667	0.679
close	0.514	0.473	0.591	0.581	0.630	0.593	0.704	0.679

In order to verify that NFP is the best explanatory variable among macroeconomic indicators published in the Employment Report we repeat the above analysis also for consensus and announced values of AW, AHE and UR. None of them, however, give better results than NFP. In majority, posterior probabilities of correct forecast are even smaller than naive prediction.

CONCLUSIONS

This paper analyses applicability of discriminant analysis to predict the behavior of investors after announcements of the Employment Report. The Report is one of the earliest publications of the month that describe the U.S. economy. Hence investors around the world pay special attention on it. As indicated by empirical studies announcements of the Report implies immediate and very strong

reaction of stock markets. This reaction is mainly in line with unexpected news about Nonfarm Payrolls contained in the Report.

On the basis of data from January 2009 to December 2013 we show that reaction of investors on the Warsaw Stock Exchange and on the Vienna Stock Exchange can be successfully predicted on the basis of consensus published by news agencies and on the basis of released value of Nonfarm Payrolls. The best results are obtained when prediction is performed in the moment of news release. Significant reaction to U.S. macroeconomic news announcements is observed in first few minutes after news release. Hence, we study the performance of forecast on very short horizon. From the analysis presented in the paper it follows that the probability of a correct forecast of investors' behavior up to 20 minutes after the Employment Report announcements is greater than 0.7 on each of the market under study.

REFERENCES

Andersen T., Bollerslev T., Diebold F., Vega C. (2007) Real-time price discovery in global stock, bond and foreign exchange markets, Journal of International Economics, vol. 73, pp. 251-277.

Będowska-Sójka B. (2010) Intraday CAC40, DAX and WIG20 returns when the American macro news is announced, Bank i Kredyt, vol. 41(2), pp. 7-20.

Dimpfl T. (2011) The impact of US news to the German stock market - an event study analysis, The Quarterly Review of Economics and Finance, vol. 51, pp. 389–398.

Gurgul H., Suliga M., Wójtowicz T. (2012) Responses of the Warsaw Stock Exchange to the U.S. Macroeconomic Data Announcement, Managerial Economics, vol. 12, pp. 41-60.

Gurgul H., Suliga M., Wójtowicz T. (2013) The reaction of intraday WIG returns to the U.S. macroeconomic news announcements, Quantitative Methods in Economics, vol. XIV, pp. 150-159.

Gurgul, H., Wójtowicz T. (2014) The impact of US macroeconomic news on the Polish stock market. The importance of company size to information flow, Central European Journal of Operations Research, [DOI 10.1007/s10100-014-0343-x].

Harju K, Hussain S (2011) Intraday seasonalities and macroeconomic news announcements, European Financial Management, vol. 17, pp. 367–390.

Nikkinen J., Sahlström P. (2004) Scheduled Domestic and US Macroeconomic News and Stock Valuation in Europe, Journal of Multinational Financial Management, vol. 14, pp. 201-245.

Nikkinen J., Omran M., Sahlström M., Äijö A. (2006) Global stock market reactions to scheduled U.S. macroeconomic news announcements, Global Finance Journal, vol. 17(1), pp. 92-104.

Suliga M., Wójtowicz T. (2013) The reaction of the WSE to U.S. employment news announcements, Managerial Economics, vol. 14, pp. 39-60.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 261 –272

APPLICATION OF MULTIVARIATE DISCRIMINANT ANALYSIS FOR PREDICTION OF BANKRUPTCY OF SELECTED CONSTRUCTION AND DEVELOPMENT COMPANIES

Monika Zielińska-Sitkiewicz
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
e-mail: monika_zielinska_sitkiewicz@sggw.pl

Abstract: Functioning of the company in the conditions of the free market competition depends on its flexible reactions to changes and the response speed to perturbations in the unstable economy. Entities, which are not able to keep up with the current changes, enter the path of crisis in the company, which last stage may be the bankruptcy. The paper presents an attempt to use and evaluate five Polish models of the multivariate discriminant analysis in forecasting the threat of bankruptcy. The analysis was conducted for the years 2008 – 2013. For the study there were selected 10 construction and real estate development companies, listed on the main market of the Warsaw Stock Exchange, which profit and loss account is made in the calculation model and for which in the years 2012-2014 there were initiated the bankruptcy proceedings.

Keywords: Polish real estate market, construction company, financial ratios, discriminant analysis

INTRODUCTION

The variability of the economic environment and the increase of the competition on the market makes the assessment of the economic condition the key element of the management process in the company. In addition, the reliable information about the companies “building” the Polish residential market is of fundamental importance for the security of relations between the participants of the real estate market and the mutual trust of all its participants.

However, maintaining a stable financial situation of construction and real estate development companies may sometimes be more difficult due to the nature of their activities, which feature a long process of the construction investment.
Recent data indicate that in 2013 213 construction companies went bankrupt. This means the decrease of 2,3% compared to the previous year. However, this is still almost four times more bankruptcies than in 2008. (cf. Table 1)
The collapse of the construction market from 2012 moved to 2013. Many companies, which did not manage to obtain the sufficient number of orders and get financing for their implementation ceased their activities. A lot of companies of the sector are still facing the liquidity problems. The additional problem is still the small amount of new investments, what inhibits the growth of companies, enhances the competitive struggle and pressure on margins. Experts predict that 2014 may bring some recovery in the industry, but so far, in the ongoing year, construction companies are still responsible for every fourth bankruptcy in the Polish economy.

Table 1. Figures concerning the bankruptcy in construction

Year	Construction companies	Share in the total number of bankruptcies	Companies serving the real estate market
2013	213	24,1 %	16
2012	218	24,9 %	37
2011	143	19,8 %	28
2010	98	15,0 %	12
2009	82	11,9 %	14
2008	59	14,3 %	6

Source: Coface report on bankruptcies of companies in Poland in 2013

Bankruptcy of companies may be considered both in the economic and legal aspect.
From the economic point of view, bankruptcy of the company means that it is not able to independently continue the activity without outside help. The company may be in a critical condition but still this is not revealed in financial data.

In legal terms, bankruptcy is determined in court. It is a procedure introduced in order to satisfy claims, in case of insolvency of the debtor and addressed to his whole property. Bankruptcy proceedings are conducted under systemic or liquidation bankruptcy. Systemic bankruptcy is to restructure the company and to conclude an arrangement with creditors. The consequence of the liquidation bankruptcy is the sale of the assets of the bankrupt company and the satisfaction of creditors from thus obtained assets.

PURPOSE AND DESCRIPTION OF THE STUDY

The purpose of the article is the attempt to use and evaluate some Polish models based on the discriminant analysis in the field of forecasting the threat of bankruptcy of 10 selected construction and real estate development companies,

listed on the main market of the Warsaw Stock Exchange, which profit and loss account is made by the calculation model. The analysis was conducted for the years of 2008 – 2013.

For the study there were selected companies: ABM SOLID S.A., ALTERCO S.A., BUDOPOL-WROCŁAW S.A., DSS S.A., ENERGOMONTAŻ-POŁUDNIE S.A., GANT S.A., PBG S.A., HYDROBUDOWA POLSKA S.A., INTAKUS S.A., POLIMEX MOSTOSTAL S.A. and PBG S.A. towards which in the years 2012-2014 there were initiated the bankruptcy proceedings. (cf. Table 2)

Table 2. Information concerning the bankruptcy proceedings in the studied companies

Company	Bankruptcy proceeding
ABM SOLID	from 6.09.2012 – systematic bankruptcy from, 13.09.2012 – liquidation bankruptcy from 3.04.2013 – systematic bankruptcy again
ALTERCO	from 27.11.2012 – systematic bankruptcy from 22.02.2013 – cancellation of bankruptcy proceedings May 2014 – submission of the creditor's petition for liquidation bankruptcy
BUDOPOL-WROCŁAW	from 6.09.2012 – systematic bankruptcy from 30.07.2014 – liquidation bankruptcy
DSS	from 17.04.2012 – liquidation bankruptcy from 29.06.2012 – systematic bankruptcy
ENERGOMONTAŻ-POŁUDNIE	from 26.01.2013 – systematic bankruptcy from 28.08.2013 – liquidation bankruptcy
GANT	from 2.01.2014 – systematic bankruptcy from 7.07.2014 – liquidation bankruptcy
HYDROBUDOWA POLSKA	from 22.06.2012 – systematic bankruptcy from 1.10.2013 – liquidation bankruptcy
INTAKUS	from 8.05.2012 – systematic bankruptcy
POLIMEXMS	from 10.2012 creditors submit petitions for liquidation bankruptcy
PBG	from 13.06.2012 – systematic bankruptcy

Source: Stock Exchange reports

The study used five discriminant models with the greatest ability to predict bankruptcy. The presented methods have been developed for the Polish market and selected based on the ranking of Z-score models created by P. Antonowicz.

The first place in the ranking was taken by the model of $Z_{7\ INE\ PAN}$, which best predicts the bankruptcy with the average prognosis efficiency of 94,82% and the error of 5,18%. The second place was occupied by the model of $Z_{6\ INE\ PAN}$ with a slightly lower average of the prognosis efficiency of 94,20% and the error of 5,80%. Both models were created in the Institute of Economics PAN under the direction of E. Mączyńska.

Financial indicators used for the construction of both functions have the form of:

X_1 = operational result/value of assets,
X_2 = value of equity/value of assets,
X_3 = (net result + depreciation)/total liabilities,
X_4 = current assets/short-term liabilities,
X_5 = sales revenue/value of assets.

The discriminant function of the $Z_{7\ \text{INE PAN}}$ model used the group of first four variables and it is presented by the equation:

$$Z_{7\,\text{INE PAN}} = -1{,}498 + 9{,}498 \cdot X_1 + 3{,}566 \cdot X_2 + 2{,}903 \cdot X_3 + 0{,}452 \cdot X_4 \quad (1)$$

The form of the $Z_{6\ \text{INE PAN}}$ model used all five indicators:

$$Z_{6\,\text{INE PAN}} = -2{,}478 + 9{,}478 \cdot X_1 + 3{,}613 \cdot X_2 + 3{,}246 \cdot X_3 + 0{,}455 \cdot X_4 + 0{,}802 \cdot X_5 \quad (2)$$

Entities with values for both discriminant functions above zero ($Z_{[7\text{ and }6\text{ INE PAN}]} > 0$) are determined as those not at risk of bankruptcy. While companies for which the results take the values not greater than zero ($Z_{[7\text{ and }6\text{ INE PAN}]} \leq 0$), are the companies at risk of bankruptcy in the perspective of 1 year. [Mączyńska 2006]

The Poznan model developed by three authors: M. Hamrol, B. Czajka and M. Piechocki took the third place in the ranking of P. Antonowicz, with an average prognosis efficiency of 93,78% and the error of 6,22%. The model used four financial indicators:

X_1 = net financial result/total assets,
X_2 = (current assets – inventories)/short-term liabilities,
X_3 = fixed capital/ total assets,
X_4 = sales financial result/sales incomes.

When interpreting the discriminant function:

$$Z_{\text{HCP}} = -2{,}368 + 3{,}562 \cdot X_1 + 1{,}588 \cdot X_2 + 4{,}288 \cdot X_3 + 6{,}719 \cdot X_4 \quad (3)$$

we should be based on the following principles: $Z_{HCP} \leq 0$ is a company at risk of bankruptcy in the perspective of 1 year, and when $Z_{HCP} > 0$ the company is not at risk of bankruptcy.

The fourth and fifth places in the ranking of P. Antonowicz were occupied by two models of B. Prusak: Z_{BP1} and Z_{BP2}. The first model predicts the bankruptcy for a year ahead, with an average prognosis efficiency of 92,52% and its error of 7,48%. The second function allows you to extrapolate bankruptcy two years in advance, with an average prognosis efficiency of 91,82% with the error of 8,19%.

Z_{BP1} model uses four financial indicators:

X_1 = results from the operational activity/average value of the balance sum,
X_2 = operating expenses (without other operating expenses)/short-term liabilities (without special funds and financial liabilities),
X_3 = current assets/short-term liabilities,
X_4 = result from the operating activity/net incomes from sales,

and is expressed by the formula:

$$Z_{\text{BP1}} = -1{,}5685 + 6{,}5245 \cdot X_1 + 0{,}1480 \cdot X_2 + 0{,}4061 \cdot X_3 + 2{,}1754 \cdot X_4 \quad (4)$$

Making the correct interpretation of results of the function Z_{BP1} is possible based on the certain boundary values:

$Z_{BP1} > 0{,}65$ company unthreatened by bankruptcy,
$Z_{BP1} < (-0{,}13)$ company threatened by bankruptcy in the perspective of 1 year,
$Z_{BP1} \in \langle -0{,}13; 0{,}65 \rangle$ area of uncertainty, so-called “gray area”,
Cut-off point = (-0,13)

The Z_{BP2} function used three financial indicators:

X_1 = (net results + depreciation)/total liabilities,
X_2 = operating expenses/short-term liabilities,
X_3 = sales profit/balance sum.

Linear discriminatory model estimated on their basis adopted the following form:

$$Z_{BP2} = -1{,}8713 + 1{,}4383 \cdot X_1 + 0{,}1878 \cdot X_2 + 5{,}0229 \cdot X_3 \quad (5)$$

Function interpretations are performed based on the following criteria:
$Z_{BP2} > 0{,}2$ company not threatened by bankruptcy
$Z_{BP2} < (-0{,}7)$ company threatened by bankruptcy in the perspective of 2 years
$Z_{BP2} \in \langle -0{,}7;0{,}2 \rangle$ area of uncertainty, that is the "gray area"
Cut-off point = (-0,295). [Antonowicz 2007 p. 61-62]

Based on data from the financial statements of the studied companies there were determined values of discriminant functions for the years of 2008-2013. Then, in accordance with the rules of interpreting the values of individual models, there was performed the assessment of the threat of bankruptcy.

RESEARCH RESULTS

In tables numbered from 3 to 12 there are presented results of the discriminant analysis for all studied companies. There were adopted the following labelling according to the appropriate criteria determined for particular studied models:

- a company not threatened by bankruptcy
- a company in a situation of uncertainty (grey area)
- a company at risk of bankruptcy

The results of the discriminant analysis for the ABM SOLID S.A. company were presented in table 3. After a period of moderate gains in the years of 2008-2010, in 2011 the company started to generate losses of PLN 47 million, and in 2012 of PLN 185 million. Since 2012 against the company there has been conducted the bankruptcy proceeding.

Table 3. Results of early warning models for the ABM SOLID company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-5,2624	-12,3315	-2,1561	0,4269	1,0091	1,4180
$Z_{6\ INE\ PAN}$	-5,6893	-12,7279	-1,9318	0,5043	1,2724	1,6780
Z_{HCP}	-5,8043	-8,3755	-1,0285	1,4160	1,6308	2,9453
Z_{BP1}	-1,3648	-6,2366	-1,4542	-0,1400	-0,0211	0,4608
Z_{BP2}	-1,6151	-2,6230	-1,6218	-0,7921	-0,4845	-0,3315

Source : own calculations

Z_{BP1} and Z_{BP2} models already since 2010 have shown the risk of bankruptcy of the company, and since 2008 the uncertainty regarding the bankruptcy prognosis. Other functions alarmed the bankruptcy since 2011. In the years of

2008-2010 both PAN models and the Poznan model indicated the lack of risk of bankruptcy in the perspective of one year.

Table 4 presented the results of the discriminant analysis for the ALTERCO S.A. company. This company since 2011 has had serious problems with the financial liquidity, as evidenced by the petitions for bankruptcy submitted by creditors in the years of 2011 and 2012. The paradoxical situation in case of this company is the fact of redemption for bankruptcy in February 2013 due to the lack of resources in the company to conduct the bankruptcy proceedings.

Only the Z_{BP2} function already since 2008 has signalled troubles of this company. Rother models detected the threat of bankruptcy in 2012, with a single warning alarm in 2009 of the Z_{BP1} function.

Table 4. Results of early warning models for the ALTERCO company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-12,5324	-28,1509	3,2807	4,0611	2,6761	0,7448
$Z_{6\ INE\ PAN}$	-13,6010	-29,3803	2,5675	3,4879	1,8123	0,0879
Z_{HCP}	-10,2503	-15,1740	9,8822	7,6251	10,8209	3,3866
Z_{BP1}	-9,5842	-18,0824	1,7851	0,7899	-29,6340	-[1]
Z_{BP2}	-1,9412	-3,8294	-1,2558	-0,5927	-1,5413	-1,1630

Source : own calculations

The company BUDOPOL-WROCŁAW (dominant shareholder is the development company GANT) in 2012 reported a dramatic increase in financial expenses, caused by losses on financial assets, what was probably connected with the reckless financial policy of the company. In September 2012 the court issued a decision declaring the bankruptcy of the company with the possibility of an arrangement. Table 5 illustrates the results of the discriminant analysis for the BUDOPOL-WROCŁAW company. The most vulnerable were the ZBP1 and ZBP2 models, which already in, respectively, 2009 and 2008 have signalled the deterioration of the company's situation, which in the years of 2008-2011 reached good financial results. All functions have alerted the thread of bankruptcy in 2012.

Table 5. Results of early warning models for the BUDOPOL-WROCŁAW company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-5,2054	-4,1476	2,2514	2,0400	2,6573	2,9321
$Z_{6\ INE\ PAN}$	-6,0225	-4,5021	2,3929	2,2345	2,5056	3,3915
Z_{HCP}	0,5692	-0,8568	3,7873	2,9147	5,8204	5,1536
Z_{BP1}	-5,6527	-2,4355	0,4827	-0,0399	0,1708	0,9987
Z_{BP2}	-1,6777	-2,2258	-0,6646	-0,5447	-0,7385	-0,0593

Source : own calculations

[1] The financial statements of the company have been prepared according to IFRS since 2009, what means that the earliest data relate to 2008. The Z_{BP1} model refers to the average balance sum of the current and previous year. The result is the lack of capacity to perform function calculations for 2008, due to referring to data from 2007.

The company of Lower Silesian Rock Raw Materials (Dolnośląskie Surowce Skalne) (DSS) in the years of 2008 - 2013 was experiencing losses of a few to even several hundred (in 2011) millions of zlotys. The consequence of poor financial results was the company's bankruptcy announced by the court in 2012. (cf. Table 2)

Almost all analysed discriminant functions paid attention to the risk of the DSS company's bankruptcy already since 2008. Only the Poznan model Z_{HCP} in the years of 2009-2010 estimated the company positively. (cf. Table 6)

Table 6. Results of early warning models for the DSS company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-14,4024	-6,2103	-10,7606	-0,2150	-0,5554	-0,3671
$Z_{6\ INE\ PAN}$	-15,1875	-7,0201	-11,0610	-1,0164	-1,2899	-1,2407
Z_{HCP}	-14,3441	-6,4193	-7,3586	0,5936	4,1502	4,7984
Z_{BP1}	-5,4356	-1,3858	-6,5406	-1,3620	-1,5417	-[2]
Z_{BP2}	-1,9502	-1,5122	-3,0377	-1,6517	-0,9343	-1,4563

Source: own calculations

The ENERGOMONTAŻ-POŁUDNIE company since 2009 has generated several million negative financial results, so that in 2012 to register the highest loss of PLN 345 million. Crisis in the whole construction industry resulted in the company's trouble, towards which in January 2013 there was initiated the bankruptcy proceeding. (cf. Table 2)

Indices $Z_{6\ INE\ PAN}$ and Z_{BP1} indicated the threat of bankruptcy of the ENERGOMONTAŻ-POŁUDNIE company since 2009. The Z_{BP2} model has signalled bankruptcy since 2008, the $Z_{7\ INE\ PAN}$ function since 2011, and the Z_{HCP} index has maintained optimism the longest alarming the threat only since 2012. (cf. Table 7)

Table 7. Results of early warning models for the ENERGOMONTAŻ-POŁUDNIE company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-22,5409	-46,4777	-0,7148	0,1856	0,3268	1,4937
$Z_{6\ INE\ PAN}$	-23,3026	-45,9320	-1,0072	-0,1115	-0,1733	1,1900
Z_{HCP}	-24,9915	-27,9479	0,4686	1,7352	2,8325	2,6186
Z_{BP1}	-6,1621	-13,2109	-1,1211	-0,7814	-0,6905	0,2889
Z_{BP2}	-2,6538	-6,7029	-1,4299	-1,2798	-1,2265	-0,8358

Source: own calculations

[2] Compare footnote 1

GANT is a large real estate development company, building, among others, by own forces using the acquired construction companies, e.g., BUDOPOL-WROCŁAW.

In the years of 2010-2012 the company recorded an impressive increase in sales revenues (by 158 % in 2012 in relation to 2010). However, at the same time, cost of sales increased by 194 % and the company updated the value of inventories, financial assets and investment properties, what resulted in the enormous loss in the amount of PLN 450 million in 2012. The year 2013 was closed by the developer with the loss of PLN 71 million, and at the beginning of 2014 declared bankruptcy.

Table 8 shows the results of the discriminant analysis for the GANT company. First three models $Z_{7\ INE\ PAN}$, $Z_{6\ INE\ PAN}$ and Z_{HCP} indicated the bankruptcy threat since 2012. The Z_{BP1} function already in the years of 2010-2011 paid attention to the deterioration of the developer's situation, and the Z_{BP2} index signalled bankruptcy in the perspective of two years already since 2008. (cf. Table 8)

Table 8. Results of early warning models for the GANT company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-0,4459	-4,5500	1,2210	1,3963	2,3408	2,4447
$Z_{6\ INE\ PAN}$	-1,2981	-5,3876	0,4252	0,5405	1,6639	1,7163
Z_{HCP}	-0,8761	-1,8726	1,3616	1,8282	3,4493	4,6081
Z_{BP1}	-0,5380	-3,3542	0,3355	0,2888	1,1762	1,1533
Z_{BP2}	-1,9007	-2,5806	-1,6312	-1,6894	-1,0950	-1,1316

Source: own calculations

The HB POLSKA company in the years of 2008-2011 achieved the net incomes from sales of PLN 1,5 billion and the profit in the range from PLN 45 to 113 million. In 2012 the company reduced the incomes only to PLN 61 million and recorded a gigantic loss of PLN 1 190 million. This resulted in the transition of the company in 2012 into bankruptcy.

Table 9. Results of early warning models for the HB POLSKA company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-[3]	-28,0780	0,3434	0,5189	1,0554	0,4066
$Z_{6\ INE\ PAN}$	-	-29,2816	0,2551	0,4549	1,0446	0,2827
Z_{HCP}	-	-98,8002	1,0695	1,2662	1,7199	1,3777
Z_{BP1}	-	-46,8762	-0,4019	-0,3712	0,0124	0,1331
Z_{BP2}	-	-10,0113	-1,3132	-1,1858	-0,8598	-0,9888

Source: own calculations

[3] 28.03.2013 r. WSE withdraw from trading shares of the company HYDROBUDOWA POLAND SA and the company has been exempted from the requirement to publish financial statements.

The $Z_{7\ INE\ PAN}$, $Z_{6\ INE\ PAN}$ and Z_{HCP} functions signalled the bankruptcy threat only in the bankruptcy year of the company. The Z_{BP1} function since 2010 paid attention to the bankruptcy threat in the perspective of one year, and the Z_{BP2} model indicated this threat in the perspective of two years already in 2008. (cf. Table 9)

The INTAKUS company has clearly reduced its incomes in the years of 2010 – 2013 from the level of PLN 64 million to PLN 12 million. From 2011 to 2013 the company has constantly generated losses of several million zlotys, resulting in the declaration of bankruptcy in 2012.

The Z_{HCP} model as the only one in the years of 2008-2013 did not indicate the company's problems. The functions of $Z_{7\ INE\ PAN}$, $Z_{6\ INE\ PAN}$ and Z_{BP1} signalled the bankruptcy threat in the perspective of one year since 2011, and the Z_{BP2} index predicted this risk in the perspective of two years already since 2008. (cf. Table 10)

Table 10. Results of early warning models for the INTAKUS company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-1,6259	-2,3484	-2,5836	1,0894	0,8696	1,2004
$Z_{6\ INE\ PAN}$	-2,5435	-3,1925	-3,3417	0,4543	0,1143	0,6217
Z_{HCP}	2,2978	0,0614	0,1972	4,8850	4,3951	3,7916
Z_{BP1}	-2,7001	-2,4348	-2,4890	0,5469	0,3066	-[4]
Z_{BP2}	-1,6038	-1,9065	-1,9645	-0,9239	-1,1991	-0,8836

Source: own calculations

The POLIMEX MOSTOSTAL company in the years of 2008-2011 achieved net incomes from sales of PLN 4-5 billion and the profit of PLN 102 to 175 million. In 2012 the company still achieved incomes at the level of PLN 4 billion but recorded a gigantic loss of PLN 1244 million. In 2013 the company has reduced its incomes to PLN 2 billion, but it has also reduced the loss to PLN 261 million. Despite this, its creditors since October 2012 has been systematically submitting motions to declare the liquidation bankruptcy of the company.

The $Z_{7\ INE\ PAN}$, $Z_{6\ INE\ PAN}$ and Z_{HCP} models paid attention to the bankruptcy threat of the company since 2012. The Z_{BP1} function since 2011 has signalled the bankruptcy risk, and the Z_{BP2} model alarmed troubles in the perspective of two years already since 2008. (cf. Table 11)

Table 11. Results of early warning models for the POLIMEX MOSTOSTAL company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-1,4695	-4,5819	0,6848	1,1717	1,3308	1,1811
$Z_{6\ INE\ PAN}$	-1,8972	-4,8078	0,6042	1,0913	1,3320	1,2546
Z_{HCP}	-0,6488	-1,6005	1,4416	2,5217	2,7604	2,3041
Z_{BP1}	-1,5263	-3,2845	-0,4728	-0,1062	0,0323	0,0964
Z_{BP2}	-1,7393	-2,7367	-1,0520	-0,7533	-0,5701	-0,6373

Source: own calculations

[4] Compare footnote 1

The specificity of the PBG company is that the significant part of its assets are the already listed companies of HB POLSKA and ENERGOMONTAŻ-POŁUDNIE. This company in the years of 2008 – 2013 achieved incomes from PLN 1,5 to about 4 billion, and in the years of 2008 -2011 the profits at the level of PLN 200 million. However, the crisis in the construction industry, own troubles and of the subsidiaries made the company in 2012 record a huge loss in the amount of PLN 3 690 million. Hence, in the same year, the decision of court to declare the company in the systematic bankruptcy.

Table 12. Results of early warning models for the PBG company

MODEL	2013	2012	2011	2010	2009	2008
$Z_{7\ INE\ PAN}$	-0,5814	-14,6167	0,7147	1,3614	1,6770	1,6890
$Z_{6\ INE\ PAN}$	-1,1533	-15,4425	0,2251	0,8926	1,2715	1,3631
Z_{HCP}	-1,4141	-8,6531	2,0054	3,4712	3,5799	3,4784
Z_{BP1}	0,2246	-9,6399	-0,4446	0,0083	0,2175	0,5717
Z_{BP2}	-1,6694	-4,1537	-1,3368	-1,1541	-0,9849	-0,8116

Source: own calculations

The values placed in table 12 indicate that indices $Z_{7\ INE\ PAN}$, $Z_{6\ INE\ PAN}$ and Z_{HCP} detected the bankruptcy risk of the PBG company since 2012, that is from the year of declaring the company's bankruptcy. The Z_{BP1} function paid attention to the bankruptcy threat since 2011, and the Z_{BP2} model alarmed the threat in the perspective of two years already since 2008.

SUMMARY

Analysis of the financial threat of the studied companies indicates that the use of particular discrimination models does not guarantee the clear assessment of their economic condition.

The characteristic feature for construction companies is the recognition of the incomes and costs of construction services based on the International Accounting Standard 11 – Construction Contracts. Overall, this provision states that the result of the contract for the construction service is estimated based on incomes and costs connected with the construction contract estimated as incomes and costs respective for the advancement state of the implementation of the contract at the reporting day. The completion state of the contract is determined based on the proportion of the costs incurred for works performed to date in relation to the estimated total contract costs. The incurred costs only consider those contract costs, which reflect the state of the performed works.

However, when the entity recognises that the implementation of the given contract will be connected with incurring losses (even in the distant future), then it is obliged to the single increase of costs of the current period.

This is connected with a number of effects, among others:

- the contracts include the costs of the failure to accomplish or the untimely fulfilment of the contract, therefore the current period will be charged with future contractual penalties;
- a single deterioration of the financial result may result in the reaction of the funders of the company, in particular banks and lessors. Long-term loans may become short-term, with higher interest rate. However, while the companies often come to an agreement with banks, the lessors are ruthless. The result of this are the drastic increases of the financial costs (interests) and other operating costs (creating provisions for liabilities towards lessors and impairment losses for the value of fixed assets);
- company's environment, and especially entities, which ordered services, may terminate contracts.

In the light of the above, it should be noted that the indicators included in the models only indirectly refer to the other operating and financial costs. From the construction of many indicators it also results that the effect of influence of these costs is mitigated by the fact that these items are included indirectly in the numerator and denominator of the indicator.

Three of the analysed models: $Z_{7\ INE\ PAN}$, $Z_{6\ INE\ PAN}$ and Z_{HCP} indicated the companies as threatened with bankruptcy too late, because in the year in which this bankruptcy was announced. These functions refer in the indicators used to the broadly understood assets. However, in companies there is often a situation that only at the time of announcing the bankruptcy, a large part of assets (e.g. receivables and inventories) is subject to verification as to their real value. Suddenly, from day to day, there are made powerful write-downs updating assets, while on the other hand the financial result of the company deteriorates. The financial statement of the company made in good faith may not reflect the essential threats resulting from, e.g., the collapse of the payments in the chain of service recipients.

Other two models of B. Prusak: Z_{BP1} and Z_{BP2} were more sensitive to threats of bankruptcy and in some cases of the studied companies in great advance alarmed the potential troubles of companies. These functions in most part refer to the values of the profit and loss account. Their sensitivity, and hence the success in the assessment of the studied companies, may result from the use of the indicator, referring directly to the incurred costs by the company, which informs about the cycle of implementation of commitments in relation to operating costs. Moreover, the Z_{BP1} model takes into account the changes of the balance sum over time.

Each of the analysed discriminant functions is based on another set

of indicators and it better or worse analyses the state of the finances of the construction and real estate development companies. In assessing the discriminatory model one should focus on the dynamics of financial results of the given company from previous years. Such an analysis may indicate the long-term

factors determining the company's activity, which in the future may cause financial problems.

REFERENCES

Antonowicz P., Metody oceny i prognozowania kondycji ekonomiczno-finansowej przedsiębiorstwa, ODDK, Gdańsk 2007.

Jajuga K., Walesiak M. (1998) Klasyfikacja i analiza danych. Teoria i zastosowania, AE, Wrocław

Kisielińska J., Waszkowski A. (2010) Polskie modele do prognozowania bankructwa przedsiębiorstw i ich weryfikacja, EiOGŻ nr 82, Wydawnictwo SGGW, Warszawa

Lichota W. (2009) Metody wczesnego ostrzegania o zmianach sytuacji finansowej przedsiębiorstw, Wiadomości Statystyczne nr 10, Warszawa

Hamrol M., Chodakowski J. (2008) Prognozowanie zagrożenia finansowego przedsiębiorstwa. Wartość predykcyjna polskich modeli analizy dyskryminacyjnej, Badania Operacyjne i Decyzje nr 3, Wrocław

Mączyńska E. Zawadzki M. (2006) Dyskryminacyjne modele predykcji bankructwa przedsiębiorstw, Ekonomista nr 2, Warszawa

Prusak B. (2005) Nowoczesne metody prognozowania zagrożenia finansowego przedsiębiorstwa, Difin, Warszawa

Prusak B. (2004) Jak rozpoznać potencjalnego bankruta?, Wydawnictwo Politechniki Gdańskiej, Prace Naukowe Katedry Ekonomii i Zarządzania Przedsiębiorstwem, Tom 3, Gdańsk

Raport Coface nt. upadłości firm w Polsce w 2013 r., 2 stycznia 2014, http://www.coface.pl/Aktualnosci-i-Media/Biuro-prasowe

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 273 – 283

METHODS OF DETERMINING THE PREFERENCE FOR PURPOSES OF THE CONSTRUCTION OF THE COMPUTERISED DECISION SUPPORT SYSTEM

Aneta Becker
Department of Economics, West Pomeranian University of Technology
e-mail: aneta.becker@zut.edu.pl
Jarosław Becker
The Department of Technology,
The Jacob of Paradyż University of Applied Sciences in Gorzów Wielkopolski
e-mail: jbecker@pwsz.pl

Abstract: The article presents the functionality of preference modelling for purposes of the multi-methodical, multi-criteria decision analysis implemented in the computerised decision support system. Research results can be considered through the prism of preference: authoritatively determined by the decision maker or democratically by the group of beneficiaries and from the perspective of scientific views of the council of experts. Preferences in the system can be determined using several different methods.

Keywords: analysis and preference determination methods, decision support system

INTRODUCTION

The term of preference is a concept found in economics (microeconomics) and is associated with the theory of consumer's choice. Consumer's preferences reflect his taste and depend on approval, satisfaction, habits. The buyer submits such combinations of goods, which maximise its utility, that is satisfaction derived from their consumption. The concept of preference is very often confused or equated with the concept of relation. Preference means the man's attitudes and is the feature of the real world, the feature of people. This term can be formally defined as a kind of relation [Ostasiewicz 2003]. In formal terms, preferences are understood as the pre-order (reverse and transitive binary relation) or the linear order (return, transitive and consistent binary relation) determined in the area of profiles

of goods and services (basket of goods). The preference relation enables the assignment of the individual preferences scale to the consumer, on which the profiles of products can be valued and the choices can be optimised [Bąk 2004].

In the article, the term of preference is considered in the context of quantifying utility, which cannot be directly measured. Preferences specified in terms of the choice options allow the defining of the utility function, which in turn allows you to connect every option with the specified number characteristic [Bąk 2004]. The utility theories belong to the area of interest of microeconomics, while the preference testing methods are the research tool of micro-econometrics [Pełka, Rybicka 2012]. Preferences are the expression of the principles of the decision-maker's rationality. Their incidence means that in the set of all possible pairs made of the decision variants there was distinguished a subset, which constitutes the relation. While the utility is a concept, which allows to assign decisions with some contractual values, constituting their merged assessment due to preferences of the decision-maker. Assigning utility to decisions allows to bring the decision problem to the automatic choice of decisions of the highest value (utility) [Shapiro 1993].

The aim of the article is to present functionality of the computerised decision support system within the preference modelling for purposes of the multi-methodical, multi-criteria decision analysis. The research procedure in the system involves the selection (optimization of MLP – multi-criteria linear programming), ranking (AHP – Analytic Hierarchy Process) and grouping (Electre Tri) of decision variants (called the objects of analysis). Objects can be examined from the point of view of preference of the given person or group (coalition), that represents the specific side in the decision-making proceedings: decision-makers, beneficiaries, experts. Wherein it is assumed that preferences can be determined using several different methods.

ANALYSIS OF PREFERENCES AND METHODS OF THEIR MEASUREMENT

The preference analysis is the research approach, which involves the qualification of objects in the particular scale, what results in the hierarchy of objects' importance. The analysis provides the adequate measurement and objectification, and its aim is the multi-criterion evaluation focused primarily on comparative studies and on the selection of one from the set of alternative solutions.

The basic methods used within the preference analysis include the method of:

- *ranking*, which consists of determining the importance of a specific object in the given set, due to preferences and is used for their arrangement for purposes of the comparative and diagnostic studies;
- *scoring* consisting of the evaluation of objects using points (in any scale of real numbers); studies with this method include the evaluation:

- comparative, which means the qualification of importance of particular objects (systems, parameters, assessment criteria, etc.) based on relations between them;
- testing (diagnostic) which is the test of a degree of respecting the determined requirements by the given system;
- variation of direction towards the rational choice (optimal) solution [Stabryła 2002].

One should also mention the methods of the identification and diagnostic nature, which include, among others: surveys, interviews and checklists. These preferences of the preference analysis can be used in connection with other research methods.

The preference measurement is made on the basis of determined declarations expressed on the respective measurement scales or they are revealed through observations of the real market choices. One of the proposals of division of the preference measurement methods is their classification, resulting from the data theories, based on two criteria:

- *nature of relations between them* – data can be similar (closeness) or dominant (preference) in nature,
- *number of comparison of the object type* – comparisons are made in the area of one set or two sets.

As a result of different combinations of the presented division, one can obtain the following types of data:

- a single stimulus,
- preferential choice,
- comparison of stimuli,
- similarities between stimuli [Sagan 2009].

In practical studies, especially the marketing ones, the preference analysis used the historical observations and data describing intensions of consumers. We can distinguish the methods of preference measurement:

- disclosed, these are the analysis methods of historical data, which reflect the real market decisions of consumers – the source of data are the information about the past market choices of consumers from the direct or indirect polls,
- expressed regarding the suspected market behaviours of consumers – data reflect the intentions of consumers during measurement and are collected using the direct or indirect polls; there are used the methods representing the approach: composite (assessment methods of levels and attributes), decomposed (traditional methods of the conjoint analysis, methods based on choices) or mixed (hybrid methods of the conjoint analysis, the adaptive method of the conjoint analysis) [Bąk 2004].

The selection of a method is influenced by the researcher's decisions concerning the aim, subject and scope of the study, costs and technical capabilities.

For many years, the relation of the individual preference is the subject of interest of social sciences. Especially the issues for a better understanding of the de-

cision-making problem. That is, the decision support, in which an important role is played by the analyst supporting the decision-maker in formulating decision variants, constructing criteria for variant assessment and the form selection of the assessment aggregation performed towards individual criteria. There is emphasised the multi-criteria nature of the decision process, and decisions usually lead to meeting the whole set of the decision-maker's needs [Nowak 2004].

In methods of multi-criteria analysis, information reflecting preferences of particular participants found in the decision support process are provided prior to the start of the calculation procedure. This is done by an analyst, who is a bystander, responsible for the whole decision-making process and communication with the decision-maker or the decision-maker himself, who can be an individual or a collective body (a group of people), and sometimes create a group of interests. Such a structure has an impact on the way of determining the preference. In case the decision-maker participates directly (or via the analyst) in the decision support process, he determines his preferences directly. While the group approach requires conducting of, for example, the direct questionnaire during the specially organised meetings [Thiel 2009]. According to the theory of making group decisions, such a choice takes place when it is performed by more than one person. While in the group thinking, striving to maintain the integrity of the group is more important than the facts. A good method of the group choice should have the following properties: limitless field, rationality of the group preference, Pareto optimality, independence of the irrelevant alternatives, no dictator [Sosnowska 1999].

The decision-maker making decisions may have to deal with a large number of decision-making variants, assessed by many criteria. The intuitive approach in the choice situation may lead to false conclusions. Therefore, it is important to use the appropriate procedure, which will enable the avoidance of decision errors. Among the decision support methods one can distinguish the multi-attribute technique (multi-criteria) of the decision assessment (MADA – *multi-attribute decision analysis*, MCDA – *multi-criteria decision analysis*). In literature there are many methods, which could be used for solving the decision-making problems. However, the mere choice of the appropriate procedure is a multi-criteria issue, because there should be considered many aspects, which include both the nature of the decision-making issue under consideration, possibilities of the tool used, its perception, flexibility and ease of use. There are methods of proven, universal character [Dytczak et al. 2010]. Such procedures include the method of multi-criteria decision-making AHP (*Analytic Hierarchy Process*), which is used for ranking the decision-making variants and indirectly to support their choice. It was developed by the American mathematician Prof. T. L. Saaty, whose works on the algorithm construction were started in the 70s [Saaty 1977, 1980]. The approach proposed by Saaty combines the elements of mathematics and psychology. It is used to solve decision-making problems, especially in situations, when criteria have the qualitative nature, and assessments are subjective and result from knowledge and experience o the analyst.

From the theoretical side, the problem of merging individual preferences is also contemplated on the basis of the utility theory. Inside it are formed some conditions, which should be met by the utility function describing the preferences of the whole group. Depending on the arrangement of these conditions, one can obtain a specific form of utility for the group. This function allows to obtain the problem solution and allows to formulate the analytical way of finding the final decision, just like the selection of the voting method allows to achieve the final decision based on the results of the vote [Shapiro 1993].

The approach towards assessment aggregation of the considered attributes, provided by experts, depends mainly in the way of allocating their assessments. We can distinguish two grasps of the work organisation of the team of experts:

- cooperation in assessing relations between the attributes of the group in question; assessments provided by individual experts are averaged, usually by the geometric mean;
- approach based on the autonomy of the individual experts in making of the assessments; in this grasp the rankings are aggregated, obtained from the separately operating experts [Dytczak et al. 2010].

Among the methods used to determine the preferences there can be observed, among others, the following suggestions:

- methods using the optimisation tools, based on the mathematical programming technique (linear programming),
- procedure for combining the opinions of the group of experts into focus, using the cosinus distance measure between the preference vectors, obtained by individual experts,
- aggregation of individual preference structures, (AIPS),
- the use of the Bayesian estimation procedure, which becomes particularly useful in case of issues with a great number of experts,
- using the theory of games by applying the criteria of minimal regret in searching for the agreement between experts [Dytczak et al. 2010].

MODELLING PREFERENCES IN THE COMPUTERISED DECISION SUPPORT SYSTEM

The topic of the discussion is the functionality of preference modelling in the computerised decision support system – DSS (version DSS 2.0 – authors Budziński R., Becker J., 2008-2014). It is a hybrid solution, which with the help of engineering techniques of the computerised data processing combines and provides the algorithms of different decision support methods in a simple and usable form. Integration of methods in the computer system is not accidental. The multi-criteria decision analysis covers the issues of choice, arrangement and grouping of objects (decision variants) from the point of view of the determined set of criteria and pref-

erences and with the possibility of taking into account the set of restrictive conditions.

The information and decision process was divided into:

- *decision optimisation* based on the multi-criteria linear programming (MLP) with the utility function and including the choice of the strongest preferred objects, which on one hand is considered from the point of view of the disposer of resources, and on the other, supports the beneficiaries competing for these resources (e.g. tasks of allocating the EU funds or reverse auctions),
- *analysis*, which uses two approaches connected with the achievements of schools: American (AHP) and European (ELECTRE); this is about the preference analysis, rankings and grouping of objects (e.g. formulating the clients' profiles for purposes of the marketing analysis),
- *identification* in terms of quantitative methods of the econometric analysis and methods based on linguistic data (RST – rough set theory).

Integration of methods in the decision support system comes down to using their functionality on a common set of input data determined within the considered decision problem. The foundation of the method integration in the computer system is the acceptance of the coherent and flexible structure based on the notation of the MLP information method [Becker 2010]. It allows to define any model information structure (the so-called mathematical model template) for the decision task. In the template construction into account are taken the decision-maker's requirements determining the substantive scope of the object analysis (decision variants: W_1, W_2, …, W_n). They are expressed by: decision variables, restrictive conditions, one- or two-level structure of the assessment and preference criteria.

Based on the mathematical model template, into the system there is introduced the set of homogenous objects. Every decision variant is a record (row) in the table of the relational database and at the same time an autonomous, partial mathematical model, otherwise formalised form of the linear programming task, which has the solution (is not a contrary system). Technical and economic parameters in the partial model of each variant can be expressed in the form of numerical values and linguistic assessments determined by one expert of their team.

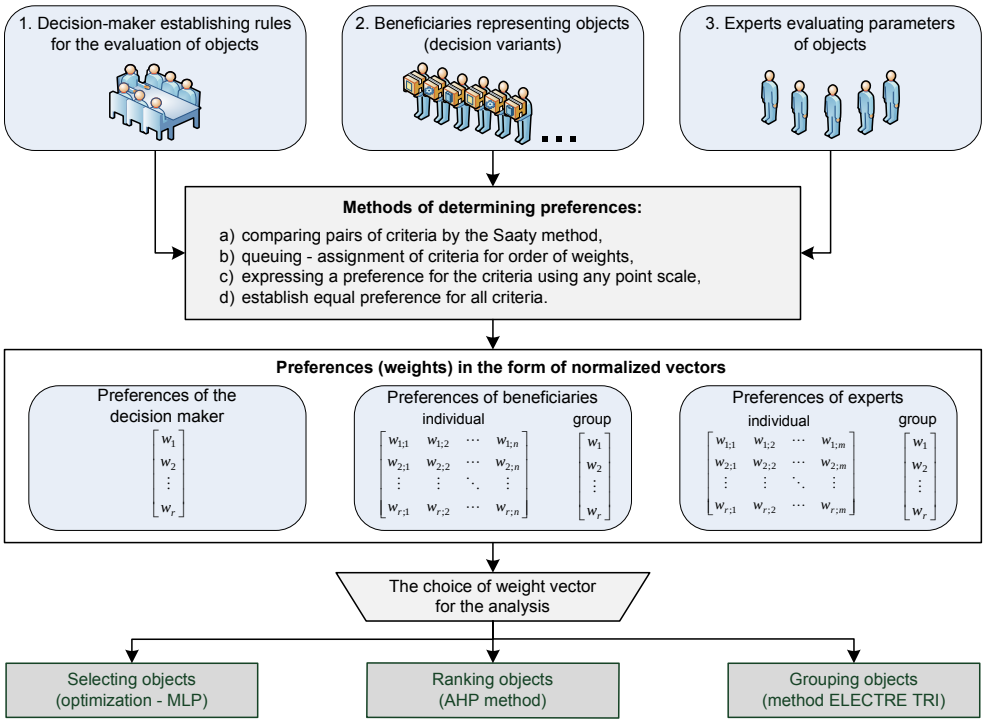
An essential part of the input data, common to the methods used in the system: MLP, AHP and ELECTRE TRI, includes:

- *value vectors of major criteria* $\mathbf{d}_t = \{d_{1;t}, d_{2;t}, \ldots, d_{r;t}\}$, which are declared with the partial models for the objects W_t ($t = 1, 2, \ldots, n$),
- *preference vector* (weight) $\mathbf{w} = \{w_1, w_2, \ldots, w_r\}$ mandatory for all analysed objects W_t, where $\sum_{k=1}^{r} w_k = 1$.

In the construction of the system there is predicted the possibility to generate results of the object analysis from the point of view of preferences of different participants, representing one of three parties in the decision process (fig. 1), we can distinguish here:

1) *decision-maker* fulfilling the superior function in the whole process, in most of the decision problems it appears in the role of the distributor of resources, for which the applicant beneficiaries apply for (then, the request is the object of analysis),
2) *objects of the analysis*, which are subordinated to the rules of the decision game and compete to obtain the highest score (subordinate function); the compared objects must constitute a category, which may include: people, organisations, proposals, service or product offers, plan variants, etc.,
3) *experts* (or respondents), who meet the measurement function (advisory, providing opinion), for an evaluation of parameters of each object.

Figure 1. Preferences in the decision support system



Source: own study

In the context of many users, an important issue is to provide them with ergonomic tool to express own preferences, which at the output generates weight factors in the form of a normalised eigenvector. In the DSS system, this function is performed by the special program (fig. 2), which enables the determination of preferences for main criteria and for non-obligatory level of sub-criteria.

The main and yet the substantively advanced function of determining preferences in the system is fulfilled by the popular Saaty method (used in the AHP method), which supports the articulation of decision-maker's preferences and vali-

dation of consistency of the expressed judgements, and their usage in the assessment aggregation process. The essence of this method is the presentation of results of comparison of assessment criteria in the form of square matrices, and calculations are based on the vector processing and non-linear algebra in terms of:

- *one user*, where by the pair comparison the relations between them are determined linguistically using the Saaty scale [Saaty 1980],
- *group of users*, who make the assessment as if they were the single users, and determination of the group preference takes place using the geometric mean.

The procedure of determining the normalised vector with the Saaty method takes place in four stages [Trzaskalik 2006]:

- summing up the assessment α_{ij} in each column of the comparison matrix $\mathbf{A} = [\propto_{ij}]_{i,j=1,\ldots,r}$ according to the provision:

$$\sigma_j = \sum_{i=1}^{r} \alpha_{ij} , \tag{1}$$

  where: α_{ij} − a number from the Saaty assessment scale, which is the result of pair comparison of r criteria (group notes are the values of the geometric mean from individual assessments),

- construction of the normalised matrix $\mathbf{B} = [\beta_{ij}]_{ij=1,\ldots,r}$ − where:

$$\beta_{ij} = \frac{\alpha_{ij}}{\sigma_j}; \tag{2}$$

- calculation of the approximate scale vector $\mathbf{w}$ according to the formula:

$$w_i = \frac{1}{r} \sum_{j=1}^{r} \beta_{ij} ; \tag{3}$$

- determination of the approximate eigenvalue of the matrix **A**:

$$\lambda_{\max} = \frac{1}{r} \sum_{i=1}^{r} \frac{(\mathbf{Aw})_i}{w_i} , \tag{4}$$

  where $(\mathbf{Aw})_i$ means the i-th element of the vector formed as a result of multiplying the matrix **A** by vector **w**.

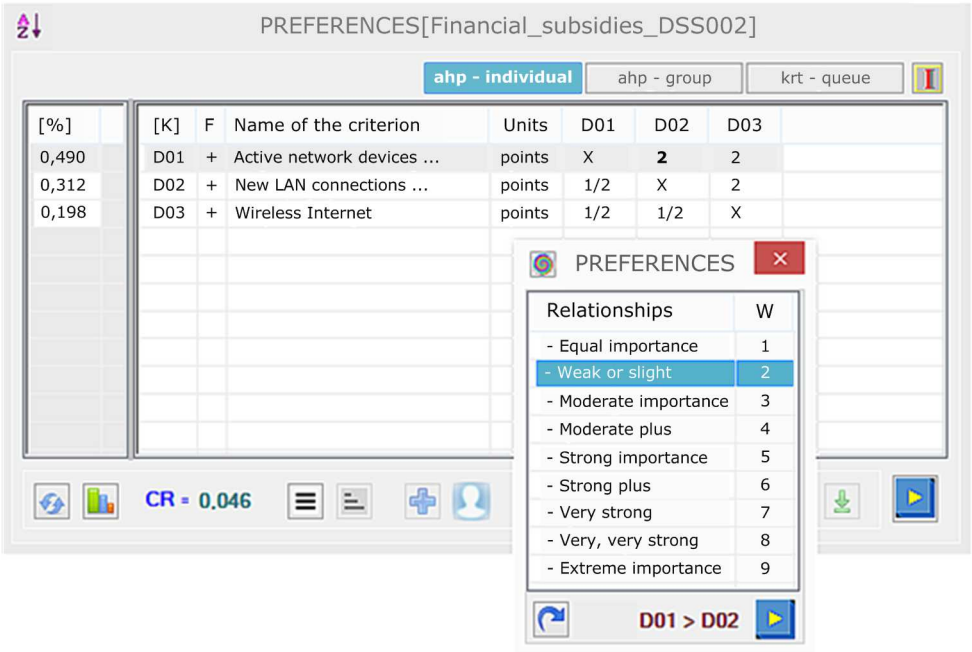
During the introduction of estimations, the special algorithm of the DSS system checks to what extent of the decision-maker's assessments written in the matrix $\mathbf{A} = [\propto_{ij}]_{i,j=1,\ldots,r}$ are consistent. For this purpose, the compliance factor is calculated

$$C = \frac{\lambda_{\max} - r}{\varphi(r-1)} , \tag{5}$$

where: λ_{max} − the largest eigenvalue, r − comparison matrix size, φ – number read from the table of compliance rates for the r size [Saaty 1990]. If the ratio $C \leq 0{,}1$

the occurrence of the assessment conformity is considered. Otherwise, the pair comparisons should be performed once again.

Figure 2. Preference determination functions using the Saaty method in the DSS 2.0 system



Source: own study based on the DSS 2.0. system

A simpler solution, compared to the Saaty method, is the use of the so-called view recorder, thanks to which each of the assessing participants presents his judgment in the form of the aim order (queue of criteria) perceived by him. To increase the scope of perceiving the goal line, there was introduced the straight and curved parameterisation. Then, there is a possibility to map the situation, in which, e.g., higher goals are harder to achieve than the lower ones (or vice versa). The relevant mathematical effect was achieved by introducing conversion factors for the line quantification: proportional $w_l = l/\sum_{l=1}^{r} l$, increasing $w_l = l^2/\sum_{l=1}^{r} l^2$ and decreasing $w_l = (1/l)/\sum_{l=1}^{r} 1/l$, calculated for l = 1, 2,…, r (where r is the number of criteria). For the selected conversion factor, there is created the initial assignment of D_k criteria to the values of generated weights w_l on the principle of $k = l$ = 1, 2,…, r. Each shift (change of order) D_k in the order changes the assignment of the w_l value. In case of line groups, for each k-th criterion there is calculated the arithmetic mean of the assigned values.

The supplementation of the option of determining the partial preferences in the system includes two use cases, in which all weights are introduced in the direct way. The first case is an arbitrary determination of positive weight values of par-

ticular criteria using any scoring scale. The provided values are automatically normalised into a vector, which sum of elements is equal to unity. The second case concerns the situation, in which preferences should be identical, e.g., for three criteria ($r = 3$) there will be the sizes of $w_k = 1/r$, i.e. 0,33 for each $k = 1, 2,\ldots, r$. It should be added that the most popular are exactly these two cases of expressing preferences for criteria.

SUMMARY

Depending on the specifics of the decision problem, type of objects assessed in the DSS system and laws applicable to his process, as well as guidelines, procedures and rules, the decision-maker's task is to determine clear rules of conduct of this analysis. The most important of them can be determined by answering the following questions. What or who is the object of analysis? What assessment criteria of objects should be taken into account? Whose preferences for criteria should be included in the decision analysis: decision-maker's, expert's, group of experts or beneficiaries' representing the objects?

If the decision analysis has the character of cognitive research, the look at the results of optimisation, ranking and grouping of objects from the perspective of every side of the decision proceedings becomes interesting. Preferences may express the view:

a) authoritative – of a decision-maker being in the role of, e.g., manager or the board,
b) democratic – a group of beneficiaries, who are assessed (e.g. employees, students) or represent the analysis objects (e.g. proposals, offers, companies and others),
c) scientific – expert or group of experts (in case of a team of experts the resultant preference may also have the interdisciplinary character).

In the decision-making process with a group of experts, what is interesting is the study of relations between their individual beliefs (preferences) and assessments of objects (criteria values) and the analysis of the influence of the beliefs on the results of ranking and grouping.

REFERENCES

Bąk A. (2004) Mikroekonometryczne metody badania preferencji, Zeszyty Naukowe Uniwersytetu Szczecińskiego nr 394, Prace Katedry Ekonometrii i Statystyki nr 15, s. 22.

Becker J. (2010) Integracja metod w informatycznym systemie wspomagania decyzji DSS (rozwiązania autorskie), Kongres Badań Operacyjnych i Systemowych – 2010, Seria: Studia i Materiały Polskiego Stowarzyszenia Zarządzania Wiedzą nr 32, Wydawnictwo: BEL Studio Sp. z o. o., Bydgoszcz.

Dytczak M., Ginda G., Kwiesielewicz M. (2010) Wybrane narzędzia grupowego wspomagania decyzji w zarządzaniu miastem, Materiały z konferencji KZZ, Zakopane (http://www.ptzp.org.pl/files/konferencje/kzz/artyk_pdf_2010/44_Dytczak_M.pdf)

Nowak M. (2004) Metody Electre w deterministycznych i stochastycznych problemach decyzyjnych, „Decyzje" (półrocznik), XII-2004, s. 35.

Ostasiewicz W., red. (2003) Pomiar statystyczny, Wydawnictwo Akademii Ekonomicznej, Wrocław, s. 31.

Pełka M., Rybicka A. (2012) Pomiar i analiza preferencji wyrażonych z wykorzystaniem pakietu conjoint programu R, Przegląd Statystyczny, R.LIX – Zeszyt 3, s. 303.

Saaty T. L. (1977) A scaling method for priorities in hierarchical structures, Journal of Mathe-matical Psychology 15(3).

Saaty T. L. (1980), The analytic hierarchy process: Planning, priority setting, resource allocation, McGraw-Hill International Book Co., New York.

Sagan A. (2009) Analiza preferencji konsumentów z wykorzystaniem programu Statistica – Analiza Conjoint i skalowanie wielowymiarowe, StatSoft Polska, s. 40, (www.statsoft.pl/czytelnia.html).

Sosnowska H., red. (1999) Grupowe podejmowanie decyzji. Elementy teorii, przykłady zastosowań, Wydawnictwo Naukowe SCHOLAR, Warszawa, s.: 63, 64, 74-76.

Stabryła A. (2002) Zarządzanie strategiczne w teorii i praktyce firmy, PWN, Warszawa-Kraków, s. 92-94.

Szapiro T. (1993) Co decyduje o decyzji? PWN, Warszawa, s. 47, 49.

Thiel T. (2009) Analiza wielokryterialna – jej miejsce i rola w fazie przedinwestycyjnej, materiały z Polskiego Kongresu Drogowego – Region Wielkopolski (17.12.2009), Politechnika Poznańska, Poznań (http://www.pkd.org.pl/pliki/Analiza_wielokryt_Thiel.pdf)

Trzaskalik T., red. (2006) Metody wielokryterialne na polskim rynku finansowym, PWE, Warszawa, s. 69.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 284 – 295

TRANSFORMATIONS OF LINGUISTIC AND NUMERICAL DATA IN THE COMPUTERISED DECISION SUPPORT SYSTEM

Jarosław Becker
The Department of Technology
The Jacob of Paradyż University of Applied Sciences in Gorzów Wielkopolski
e-mail: jbecker@pwsz.pl
Ryszard Budziński
Faculty of Economics and Management, University of Szczecin
e-mail: ryszard.budzinski@wneiz.pl

Abstract: The article discussed the issue of the bi-directional data transformation (from the linguistic form into numerical, and vice versa) in the computerised decision support system (DSS). The system uses multi-methodical research approach, which is to provide a simple and useful form of functionality of different, complementary decision support methods. Data transformations are based on the use of fuzzy set logic and specially developed for thus purpose scoring and linguistic scales of the ordinal nature.

Keywords: linguistic and numerical data transformations, fuzzy quantifiers, computerised decision support system.

INTRODUCTION

In business practice very common are situations, in which information expressed orally are necessary to make a decision. This type of expression form is cognitive in nature and closer to the human perception of reality. By nature, we perceive and describe objects and phenomena in the imprecise and blurred way. Only the need to make precise calculations, for example, of the engineering nature, forces the use of right tools and measurement methods and expressing some properties using precise numerical values.

Literature [Bouyssou, Roy 1993], [Greco et al. 2001, 2002], [Słowiński 2007] contains a variety of procedures and methods of multiple criteria decision making (MCDM). According to Greco et al. (2001) they can be divided into methods based on the functional model (American school) and relational model (European school). The vast majority of these methods depends on the input data ex-

pressed numerically. The remaining group, constituting the complement in this context, are the research methods created on the basis of statistics, artificial intelligence and psychology, in which the numerical parameters characterising the research subject are not specified (phenomenon, object). They are called the non-parametric methods, often there are no assumptions in them as to the completeness or precision of data. This group, for example, includes the symbolic methods of data classification [Gatnar 1998] and most of the methods based on the theory of rough sets, applied to the analysis of data consistency, their grouping and induction of decision-making rules [Pawlak 1982].

Integration of many complementary methods of decision-making in the information system requires, first of all, the development of such a model of data organization which will be more adjusted to the theory of decision-making. Secondly, the integration requires arming of the decision-making analysis process on its each step with computer algorithms of transformation of various data forms in such a way that in the context of the problem there is used one common set of input data (numeric, linguistic or mixed).

The article focused on the issue of bi-directional transformation of linguistic and numerical data, which was used in the computerised decision support system – DSS (version DSS 2.0 – author Budziński R., Becker J., 2008-2014). This system distinguishes the multi-methodical research method, consisting of sharing the simple and useful form of algorithms of different, complementary decision support methods. It covers the issues of selection, ordering and grouping of objects (decision variants) from the point of view of the determined set of criteria and preferences and the possibility to take in to account the set of restrictive conditions. Apart from this, it enables the econometric valuation of objects and induction of decision rules. Data transformations were based on the use of the fuzzy set logic and specially developed for this purpose profiles of ordinal and linguistic scales. The broader context for this data conversion in the computer system is the integration of knowledge sources – measurement data, expert opinions, unified structures of mathematical models and collections of methods – at an important rime of the information and decision process, that is the decision game, which purpose is the selection of the best solutions from the available ones.

FUNCTIONAL SCOPE OF THE COMPUTERISED DECISION SUPPORT SYSTEM

The functional scope of supporting the decisions was determined as the solving of decisive tasks connected with multi-criteria selection, grouping (sorting) and organising (ranking) of any decision variants, understood as objects of the analysis representing the given category of events or things. These objects must have a uniform information structure. The additional functionality of the system is the analysis and the evaluation *ex post* of the obtained results of the decision-making pro-

cess. It should be noted that the studies carried out in the system can have the formal nature (official), taking on the form of the legally sanctioned procedure (e.g. public tender, where the offers are evaluated) or less official, cognitive, where the decision maker is repeatedly supported through simulations (e.g. evaluation of employees, products, services, variants of planning, etc.). The fact that the theory of decisions creates methodological foundations for the analysis and generating best solutions is not about the utility of the information system in practice. In fact, the needs of management translate into the essential factors that should be taken into account in the design of system supporting decision-making, namely:

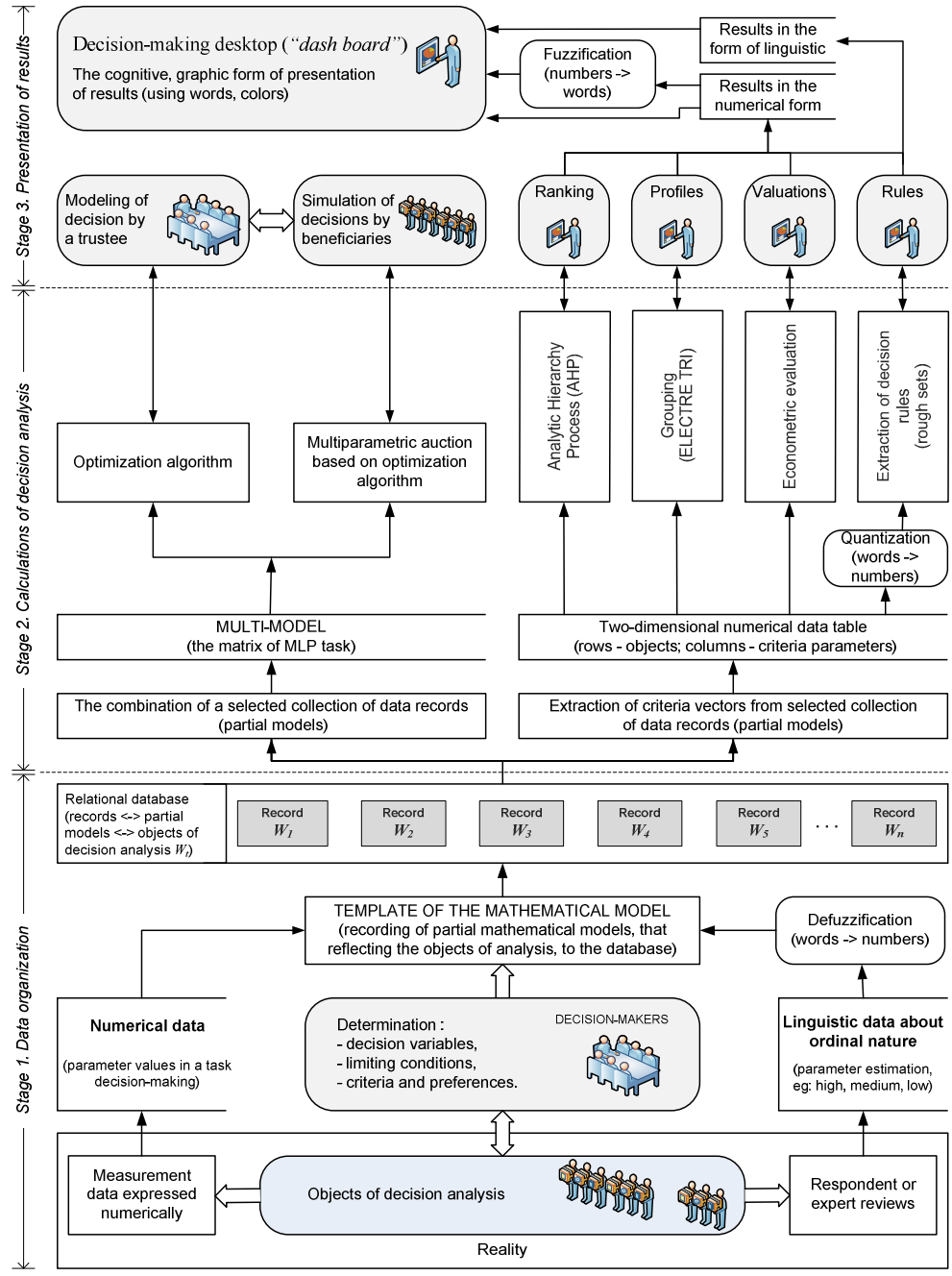
- multi-stage nature of the decision-making process,
- multi-criteria nature, in which the structure of criteria is simple (criteria vector) or complex (hierarchical or network dependencies),
- number of decision-makers and experts,
- scale of the decision problem (few or mass problems),
- flexibility of decision variants (customising the parameter values),
- linguistics of data (statements of experts or respondents).

The complexity of the description of the decisive situation causes that it is difficult to emerge the method that would be universal, to which we could attribute the possibility to obtain the best solution of many different decision-making problems.

The discussed system of supporting decision-making is a hybrid solution, which using the engineering techniques of the computer processing of data connects and shares in a simple useful form algorithms of various supplementary and implementing the paradigm of the methods supporting the decisions. The research procedure included in it is performed in three stages, it includes: (1) organization of data, (2) calculations of the decision analysis and (3) presentation of results (Figure 1). The intention of the proposed scheme of thought comes from the understanding of the support of decisions as a process, in which based on the fact base (data) we analyse and conclude, and then we make decisions. This takes into account the knowledge of users and most of all of experts, who analyse facts, express their opinions using the ordinal scale of linguistic assessments and use the mapping methods proposed in the system.

Organizing data (Figure 1) as the base of integration of methods there was accepted the coherent and flexible information structure of the system, which was subordinated to the construction of MLP models (Multi-criteria Linear Programming). It allows you to define the template for the decision-making task (standard mathematical model, Figure 1). This construction takes into account the requirements of the decision maker, which relate to the potentially analysed set of objects and they are expressed through: decision variables, limiting conditions, one- or two-level structure of criteria of assessment and the corresponding preferences [Becker 2008].

Figure 1. General architecture of the computerized decision support system (DSS 2.0)



Source: own work

According to the template to the system there are introduced data of objects (decision variants: W_1, W_2, …, W_n). Technical and economic parameters of each variant can be expressed in the form of numerical values and linguistic assessments (fuzzy values) from the ordinal scale defined by experts or respondents. For the optimization calculations all linguistic forms of data must get transformed into numerical values. The basis for the conversion of verbal expressions into numerical (defuzzification) and vice versa (fuzzification) is the methodology of the construction of linguistic quantifiers based on the theory of fuzzy sets. After the introduction and confirmation of data, each variant becomes the record (writing) in the relational database and at the same time is the autonomous, partial mathematical model. The object takes the form of the formalised task of the linear programming, which after obtaining the positive optimization result (where it is not the contrary system) is saved in the database with the admission status to the stage of decision analysis calculations.

The second stage (Figure 1) includes the issues of combining data records – partial mathematical models identical to objects of the decision-making analysis – to the form of a multi-model (MLP task matrix) for the needs of the multi-criteria optimization and transformation to the simple, tabular structure of data required on other inputs of the multi-methodical analysis. Integration of methods in the system of supporting decisions consists of the use of their functionality on a common set of data (objects) within a coherent, logical and comprehensive information-decisive process consisting of:

A. *optimization of decisions* – considered from the point of view of interests of the trustee's resources and from the perspective of beneficiaries competing for the resources,
B. *multi-criteria analysis*, in which there were used the approaches: connected with the achievements of the American school (AHP method [Saaty 1980]), European (ELECTRE TRI [Roy 1991]) and Polish school (Rough Set Theory – [Pawlak 1982]),
C. *identification* in terms of quantitative methods of the econometric analysis.

The third stage (Figure 1) includes the presentation of detailed results for each method separately and together, in the form of the decision-making desktop *("dash board")*, within which the applied methods (points B and C) function on the basis of a consultation of experts diagnosing the state of the tested objects. The desktop integrates the results of methods supporting decisions in the utility aspect. It is an interactive system enabling the multi-dimensional (multi-methodical) diagnostics of the selected object W_t (or a new one W_{n+1}) against the results of the whole set (W_1, W_2, …, W_n). It has the cognitive, graphic form of presentation of results of the applied methods. It is a kind of machine graphics, which consolidates the graphic visualization with cognitive processes taking place in the man's mind at the moment of making the decision. The structure of the desktop is based on the premise that knowledge about the object (its rating) expressed by shape and colour is absorbed faster than information in the form of numbers and text.

THE USE OF LINGUISTIC QUANTIFIERS FOR THE CONSTRUCTION OF ORDINANCE SCALE PROFILES

In the computer decision support system there are distinguished three areas of the use of linguistic quantifiers. The first one involves the transformation of linguistic data (e.g. assessment of experts) to the numerical form (defuzzification) required in the MLP mathematical model. It may relate to all values or only some parameters (defined in the template of the decision task, Figure 1), which characterise the W_t objects considered in the decision task. In the second area there is the quantification of data (also referred to as discretisation, Figure1), which is carried out for the purpose of induction of decision rules. The basic quantification stage, connected with the division of the scope of attribute values into separate sections, can be implemented in an automated manner using the selected scoring and linguistic scale or arbitrarily determined by the user. The third area of the use of linguistic quantifiers is the decision display (Figure 1). Data transformations are used in order to unify, consolidate and cognitively visualise the analysis results, obtained with different methods.

The ideal, to which we aspire, are the lossless transformations of numerical data to the linguistic forms, and vice versa. Tools, which relatively well allow for this type of conversions, are provided by the fuzzy logic. It is based on the term of *fuzzy sets*, meaning those, which do not have strictly defined boundaries. In 1965, Zadeh provided an idea and the first concept of a theory, enabling the fuzzy description of real systems. The fuzzy set is an object including the elements of some area of considerations, wherein each of these elements can fully belong to the fuzzy set, do not belong to it at all or belong to it to some degree [Łachwa 2001]. The fuzzy set A in the space (area of considerations) $X = \{x\}$, what can be written as $A \subset X$, is a set of pairs

$$A = \{(f_A(x), x)\}, \quad \forall x \in X, \tag{1}$$

where f_A: $X \rightarrow [0, 1]$ is the membership function, which assigns each element of the X space with the grade of membership to the given fuzzy set: from membership $f_A(x) = 0$ through partial membership $0 < f_A(x) < 1$ to complete membership $f_A(x) = 1$ [Kacprzyk 1986].

In fuzzy sets, the transition from membership to non-membership is gradual, and not abrupt, as in the conventional set. The concept of a fuzzy set is used for the formal recognition and quantitative expression of blurry, imprecise, ambiguous terms. They are commonly used for the qualitative assessment of physical quantities, conditions of objects and systems, and their comparison [Piegat 1999].

The concept of the scoring and linguistic scale profile ($scale^{(\tau)}$) in the computerised decision support system means the user-determined configuration of the adjustable elements of the ordinal scale, i.e.:

- *number of degrees* $\tau = 2, 3, \ldots, 11$ – the system distinguishes 9 variants of the span of $scale^{(\tau)}$ (they were given Latin names: ‘duo’, ‘tria’, ‘quatuor’,

'quinque', etc., in which the next degrees were given the absolute, non-negative ordinance values $\alpha = 0, 1, 2, \ldots, \tau\text{-}1$, always starting from zero,

- *linguistic values (names of degrees)* $a^{(\alpha)}$ – for example, for $scale^{(\tau=3)}$ these may include: $a^{(\alpha=0)}$ = *'low'*, $a^{(\alpha=1)}$ = *'average'*, $a^{(\alpha=2)}$ = *'high'*,
- *type of characteristics of the linguistic quantifier* – this is a non-linear dependency ($y = ax^2 + bx$) or linear ($y = x$) applied to generate, for any span of $scale^{(\tau)}$, triangular or pentagonal membership functions for individual linguistic values $a^{(\alpha)}$. [Becker 2014].

Linguistic quantifier consists of the membership functions, which number corresponds to the number of degrees τ on the given scale. These functions are created based on linear or non-linear function transformations. For each scale there can be determined many different linguistic quantifiers. In the computer system there are prepared five basic variants of the linguistic quantifier with: a) proportional, b1) strongly growing, b2) moderately growing, c1) strongly decreasing and c2) moderately decreasing distances between linguistic values $a^{(\alpha)}$. Disproportionate versions (b1, b2, c1, c2) can consist of triangles or pentagons, what in total gives nine proposals. For advanced users there is predicted the possibility of adjusting the shape of the characteristics of the linguistic quantifier according to the relationship

$$f_\eta(x) = y = (\eta\text{-}1)x^2 + (2\text{-}\eta)x, \tag{2}$$

in which the η parameter adopts the values from the range of $\langle 0; 2\rangle$, and $x = \alpha/(\tau\text{-}1)$. If $\eta = 1$, the characteristics (2) is linear, distanced obtained on its basis between $a^{(\alpha)}$ are identical, and membership functions for each $a^{(\alpha)}$ have the form of equilateral triangles (except for extreme values $a^{(0)}$ and $a^{(\tau-1)}$, for which the half-figures are always taken into account). In other cases (when $\eta \neq 1$) the inscription (2) determines non-linear relationships, and membership functions obtain the selected shape, of a triangle or pentagon. When $\eta \in \langle 0; 1)$, the system generates quantifiers of decreasing distances between successive $a^{(\alpha)}$ ($\alpha = 0, 1, 2, \ldots, \tau\text{-}1$), while for $\eta \in (0; 2\rangle$ proportions of these distances move in the opposite direction.

In order to simplify the notation of the function determining the degree of variable membership $x' \in \langle 0; 1\rangle$ to linguistic values $a^{(\alpha)}$, determined for particular degrees $\alpha = 0, 1, 2, \ldots, \tau\text{-}1$, the functional relationship has been transformed (2) to the following form

$$f_\eta(\alpha) = \frac{(\eta-1)\alpha^2}{(\tau-1)^2} + \frac{(2-\eta)\alpha}{\tau-1}. \tag{3}$$

Triangular membership functions are constructed for any span of $scale^{(\tau)}$ – assuming that the variable $x' \in \langle 0; 1\rangle$ and represents the numerical value subject to conversion into the linguistic form – can be expressed in the form of the following entries (Figure 2, an example for the five-point scale):

- for the first linguistic value $a^{(\alpha)}$ ($\alpha = 0$)

$$f_{a^{(0)}}(x') = \begin{cases} \frac{f_\eta(1)-x'}{f_\eta(1)} & for\ f_\eta(1) \geq x' \geq 0 \\ 0 & for\ other\ x' \end{cases}, \tag{4}$$

- when $\tau > 2$, then for every $a^{(\alpha)}$ satisfying the condition $0 < \alpha < \tau\text{-}1$ particular membership functions can be generalised to the form of

$$f_{a^{(\alpha)}}(x') = \begin{cases} \frac{x'-f_\eta(\alpha-1)}{f_\eta(\alpha)-f_\eta(\alpha-1)} & for\ f_\eta(\alpha-1) \leq x' < f_\eta(\alpha) \\ \frac{f_\eta(\alpha+1)-x'}{f_\eta(\alpha+1)-f_\eta(\alpha)} & for\ f_\eta(\alpha) \leq x' \leq f_\eta(\alpha+1) \\ 0 & for\ other\ x' \end{cases}, \tag{5}$$

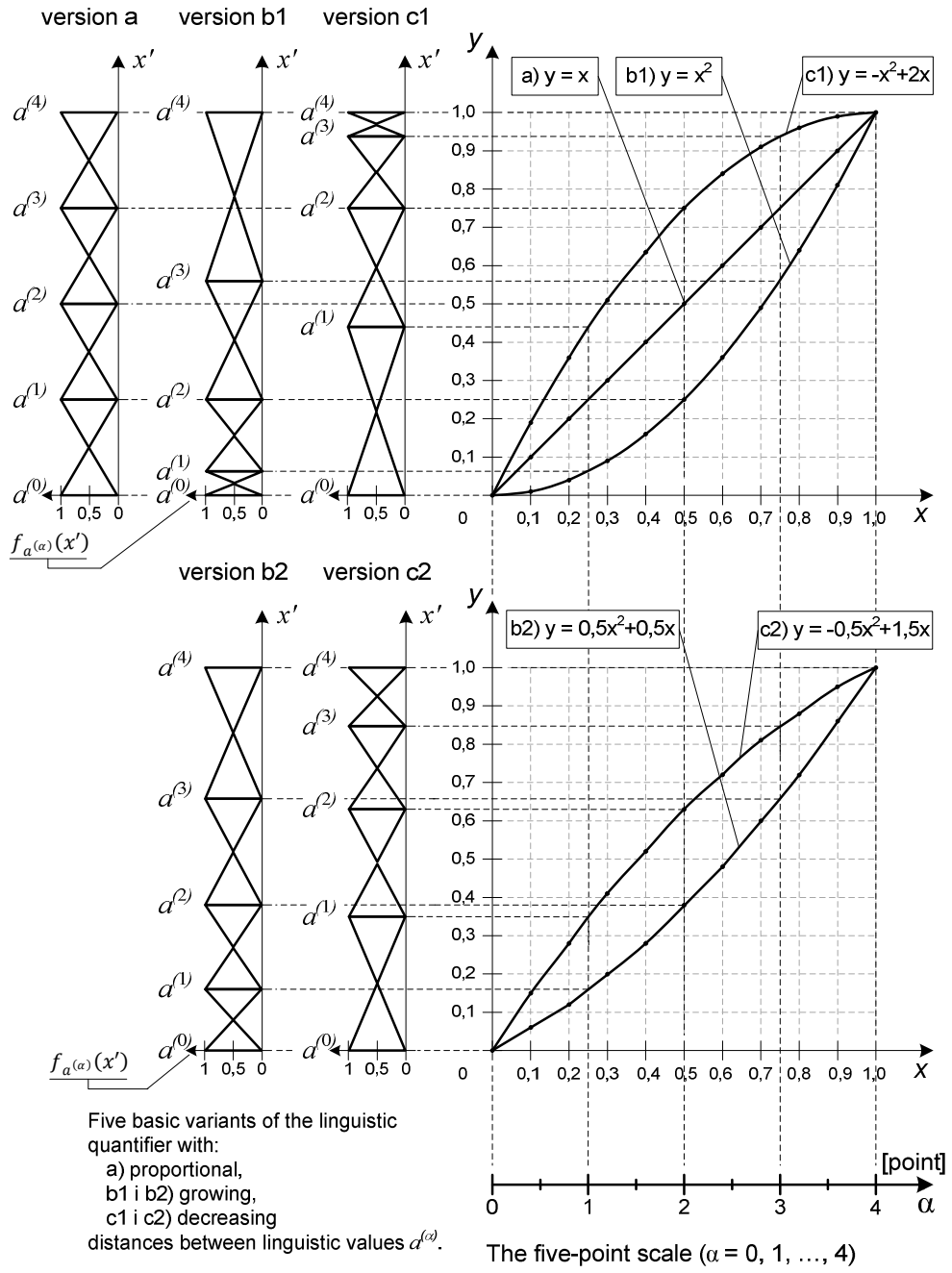
- for the last linguistic value $a^{(\alpha)}$ ($\alpha = \tau\text{-}1$)

$$f_{a^{(\tau-1)}}(x') = \begin{cases} \frac{x'-f_\eta(\tau-2)}{1-f_\eta(\tau-2)} & for\ f_\eta(\tau-2) \leq x' \leq 1 \\ 0 & for\ other\ x' \end{cases}. \tag{6}$$

In a similar manner are constructed the linguistic quantifiers equipped with the membership functions shaped as a pentagon. Pentagonal functions in relation to triangular ones are more approximate to the shape of the non-linear characteristics (2). This is due to the fact that apart from the values $x'_\alpha = f_\eta(\alpha)$ calculated for each $\alpha = 0, 1, 2, \ldots, \tau\text{-}1$, for which $f_{a^{(\alpha)}}(x'_\alpha) = 1$ and reaches the extreme, with the same rule there were also determined the intersecting points $x'_{\alpha+} = f_\eta(\alpha + 0{,}5)$ i $x'_{\alpha-} = f_\eta(\alpha - 0{,}5)$ of the adjacent membership functions, where $f_{a^{(\alpha)}}(x'_{\alpha-}) = 0{,}5$ i $f_{a^{(\alpha)}}(x'_{\alpha-}) = 0{,}5$ [Becker 2014].

According to Łachwa (2001) the issue of assigning the parameters describing the specific objects with the right membership degrees to linguistic expressions of the ordinal nature, it should be stated that doing this in a good way is difficult. This procedure is usually of the subjective nature and depends on the situational context. Clarifying this issue, membership degrees which are individual and depend on the circumstances indicate a kind of trend, which reflects on the set of studied objects from the given area of considerations some arrangement, created by association with the set of specific features. To determine the membership degrees there is used, for example, the questionnaire method common in statistics. The membership value is calculated as the relation of the number of affirmative answers to the number of all answers provided by responders. Another, popular method is determination of membership degrees by the expert. However, the expert often determines only the general shape of the membership function, and the accurate parameter values are selected experimentally.

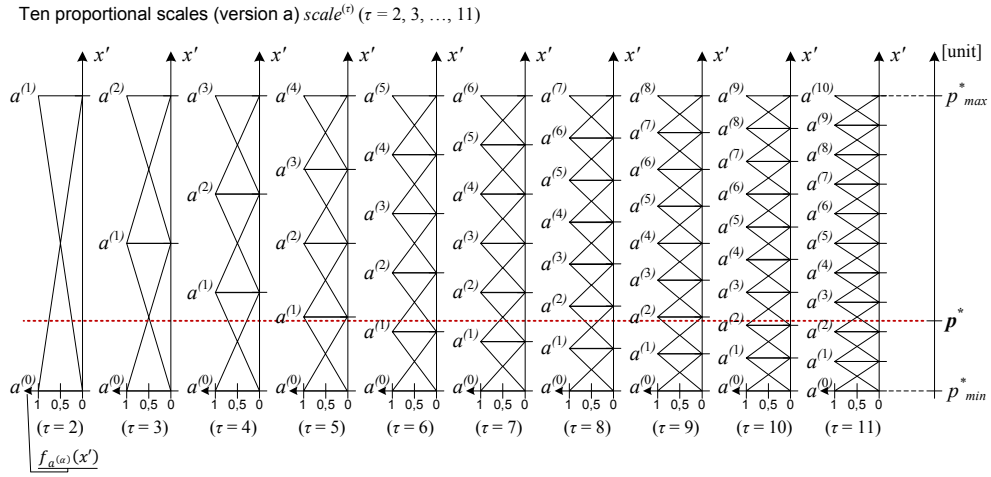
Figure 2. Constructing linguistic quantifiers based on triangular membership functions (example for the five-point scale)



Source: [Becker 2014]

In the decision support system there was proposed a tool for creating individual profiles of scoring and linguistic scales, which act as ready-to-use models of the linguistic quantifiers with the selected membership functions. These may include the equilateral triangles (type a – proportional scales, Figure 3), irregular triangles or pentagons (variants of the b and c type) of varying proportions on the ordinal scale, what is understood as different distances between degrees and focal points (the point of intersection of two functions), additionally they can have the growing or decreasing trend. A multitude of parameters configuring the profile of scoring and linguistic scale allows to define in the given decision task (z) the individual, required for each p^* parameter in the template of the mathematical model, linguistic quantifier.

Figure 3. The use of the linguistic quantifiers in the DSS system (projection of values of the p* parameter on the axis of the proportional scales)



Source: [Becker 2014]

The choice of the quantifier characteristics can be the result of the expert's suggestion, discussion of a group of several people (e.g. decision-makers and experts) or the survey, which van be performed in the phase of obtaining data (e.g. in the form of additional questions in the offer proposals submitted by beneficiaries representing the objects).

After determining the character of the membership function for the quantifier and for the given scale profile, there should be determined the number of its degrees (τ). In the system there are two kinds of allocation of this size to the given p^* parameter. The group variant, in which τ is determined for all experts assessing p^* and the individual one, where τ is selected by each expert giving opinion on p^* according to his preferred (intuitive, best perceived) structure of the assessment value system. In the computer system there were made available 9 scale models

(*scale*$^{\tau}$), which have from $\tau = 2$ to $\tau = 11$ degrees (Figure 3). It should be noted that the greater number of degrees on the scale may influence the extension of the process of parameter assessment, but instead there is obtained the higher precision of transformation of linguistic notes onto the numerical values, and in a large group of assessing people this may also result in their greater diversity.

The use of linguistic quantifiers in the decision support system is related to the bi-directional conversion of numerical and linguistic data. Defuzzification means the conversion of signals from the qualitative field to the quantitative one. In the system it relates to the parameter assessment (p^*) expressed with imprecise measures $a^{(\alpha)}$ within the determined profile of the scoring and linguistic scale. For the given $a^{(\alpha)}$, according to (3), the value $x'_\alpha = f_\eta(\alpha)$, is calculated, for which $f_{a^{(\alpha)}}(x'_\alpha) = 1$. Then, the value $x'_\alpha \in \langle 0; 1\rangle$ is proportionally converted into the acceptable parameter scope $p^* \in \langle p^*_{min}; p^*_{max}\rangle$ (Figure 3). The conversion process taking place in the opposite direction, where quantitative data (precise) are converted to qualitative is called fuzzification or dissolving. The numerical value of the parameter $p^* \in \langle p^*_{min}; p^*_{max}\rangle$ is transformed proportionally to $x' \in \langle 0; 1\rangle$, then for each scale process (with the selected model *scale*$^{\tau}$ and quantifier characteristics – type *a*, *b* or *c*) there are calculated the values of membership functions for every $a^{(\alpha)}$ ($\alpha = 0, 1, 2, \ldots, \tau\text{-}1$) according to the entries (4; 5 and 6) for triangular functions (or similarly for the pentagon-shaped functions – more in the paper [Becker 2014]). The highest value $f_{a^{(\alpha)}}(x')$ from the calculated ones determines the linguistic category $a^{(\alpha)}$. It should be noted that the applied data transformation based on the theory of fuzzy sets – in which based on the linear characteristics and different non-linear ones the scopes of membership functions are determined (equal, increasing or decreasing) – is adequate to the process of determining the quantization intervals (discretisation) of the attribute values in the induction studies of decision rules. The generated data scopes can be clarified by the system user.

SUMMARY

The article presented the issues of the linguistic and numerical transformation of data used within the uses of the computer decision support system (DSS 2.0). They most often concern the problems solved with experts, who express their opinions using imprecise terms, for example, assessment: of employees or recruits, grant or loan applications, tenders including the specialised services or devices, etc. The place of experts may be occupied by respondents, e.g., the representative group of students assessing individual departments, directions of teaching.

Linguistic quantifiers, included in the form of fuzzy and ordinal scale profiles, on one hand, are used in order to bring data to the specific form and provide them to the input of the appropriate decision support method. On the other hand, numerical method results are converted to the linguistic form and integrated with others on the decision desktop. The aim of this operation is the synthesis of results

obtained with different methods based on the cognitive presentation and interpretation (using words and spectrum of colours).

What is interesting is the use of scale profiles of a different number of degrees describing the conditional attributes and the decision attribute to the search of such quantization, due to which there will be generated the most valuable rules. Generalising the description of attributes (reducing the number of categories), we admittedly influence the structure and consistency of the data set, but we are moving towards the deep knowledge, expecting the rules of a more general content and greater coverage. This rule can be reversed and the shallower knowledge may be sought, that is more precisely formulated rules in the description of reality.

REFERENCES

Becker J. (2008), Architektura informatycznego systemu generowania wielokryterialnych rozwiązań decyzyjnych (cz. 1). Koncepcja budowy modelu WPL oparta na niestandardowych zadaniach decyzyjnych, Seria IBS PAN: Badania Systemowe, Tom 64, Wyd. Instytut Badań Systemowych PAN & Polskie Towarzystwo Badań Operacyjnych i Systemowych, Warszawa.

Becker J. (2014), Integracja źródeł wiedzy w informatycznym systemie wspomagania decyzji (podstawy metodyczne i inżynierskie), maszynopis monografii naukowej po uzyskaniu pozytywnych recenzji, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie.

Bouyssou D., Roy B. (1993), Aide multicritere a la decision: Methodes et cas, „Economica", Paris.

Gatnar E. (1998), Symboliczne metody klasyfikacji danych, PWN, Warszawa.

Greco S., Matarazzo B., Słowiński R. (2001), Rough sets theory for multicriteria decision analysis, European Journal of Operational Research (129).

Greco S., Matarazzo B., Słowiński R. (2002), Axiomatization of utility, outranking and decision-rule preference models for multiple-criteria classification problems under partial inconsistency with the dominance principle, „Control and Cybernetics", vol.: 31/4, p: 1005-1035.

Kacprzyk J. (1986), Zbiory rozmyte w analizie systemowej, PWN, Warszawa, s. 43.

Łachwa A. (2001), Rozmyty świat zbiorów, liczb, relacji, faktów, reguł i decyzji, Akademicka Oficyna Wydawnicza „Elit", Warszawa.

Pawlak Z. (1982), Rough sets, Int. J. Computer and Information Sci., 11, s. 341-356.

Piegat A. (1999), Modelowanie i sterowanie rozmyte, Akademicka Oficyna Wydawnicza EXIT, Warszawa, s. 23 i 28.

Roy B. (1991), The outranking approach and the foundations of Electre methods, Theory and decision, no. 31.

Saaty T. L. (1980), The analytic hierarchy process: Planning, priority setting, resource allocation. McGraw-Hill International Book Co., New York.

Słowiński R. (2007), Podejście regresji porządkowej do wielokryterialnego porządkowania wariantów decyzyjnych, in: Kulczycki P., Hryniewicz O., Kacprzyk J. (eds.), Techniki informacyjne w badaniach systemowych, Wyd. Naukowo-Techniczne, Wyd. Naukowo-Techniczne, Warszawa.

Zadeh L. (1965), Fuzzy sets, "Information and Control" vol. 8, s. 338-353.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 296 – 306

CONSTRUCTION OF THE ENVIRONMENTAL PERFORMANCE INDEX USING DEA

Ewa Chodakowska
Faculty of Management
Bialystok University of Technology
e-mail: e.chodakowska@pb.edu.pl

Abstract: There is wide range of environmental performance indicators. The more sophisticated they are, the more arbitrary viewpoint they take. The EPI weights are established based on experts' judgments form Yale University and Columbia University team. In the article was analysed the changes in the EPI when the weights had been replaced with the weights determined by Data Envelopment Analysis. DEA determines weights that maximize the performance of each unit with some assumptions. Not for every European country this approach turned out to be the better option.

Keywords: EPI (environmental performance index), DEA (data envelopment analysis)

INTRODUCTION

Environmental Performance Index

The sustainable development and environmental performance are the subjects of unwavering popularity. The popularity has given rise to a considerable wealth of research in this area. Increasing consciousness about environmental problems was the origin of introduction of measurements like the Environmental Performance Index (EPI). "The Environmental Performance Index (EPI) ranks how well countries perform on high-priority environmental issues in two broad policy areas: protection of human health from environmental harm and protection of ecosystems" [Hsu at al. 2014]. The EPI can be used by environmental advocates, business leaders, politicians to improve management decisions and enable more sustainable choice. Although, it is used more often only as a public relations and marketing tool.

The EPI indicators are constructed in several steps in detail described on its project website: http://epi.yale.edu/. Generally speaking, first, the raw data values are transformed by dividing by population, GDP or some other denominator in order to make the data comparable across countries. Second, a logarithmic transformation is performed on most of the variables. Third, the transformed and logged data are converted into indicators using a proximity-to-target methodology. The proximity-to-target methodology measures each country's performance on any given indicator based on its position within a range established by the lowest performing country (equivalent to 0 on a 0-100 scale) and the target (equivalent to 100). Then explicit weights are assigned to the indicators, policy categories, and objectives in order to create the aggregate EPI score [Emerson at al. 2012; Hsu at al. 2014]. The weights and indicators are presented in Table 1.

Table 1. Statistical Weightings Used for the 2014 Environmental Performance Index (EPI)

EPI	Objective	Issue Category	Indicator
Environmental Performance Index (EPI)	Environmental Health (40%)	Health Impacts (33%)	Child Mortality (100%)
		Air Quality (33%)	Household Air Quality (33%)
			Air Pollution - Average Exposure to PM2.5 (33%)
			Air Pollution - PM2.5 Exceedance (33%)
		Water and Sanitation (33%)	Access to Drinking Water (50%)
			Access to Sanitation (50%)
	Ecosystem Vitality (60%)	Water Resources (25%)	Wastewater Treatment (100%)
		Agriculture (5%)	Agricultural Subsidies (50%)
			Pesticide Regulation (50%)
		Forests (10%)	Change in Forest Cover (100%)
		Fisheries (10%)	Coastal Shelf Fishing Pressure (50%)
			Fish Stocks (50%)
		Biodiversity and Habitat (25%)	Terrestrial Protected Areas (National Biome Weights) (25%)
			Terrestrial Protected Areas (Global Biome Weights) (25%)
			Marine Protected Areas (25%)
			Critical Habitat Protection (25%)
		Climate and Energy (25%)	Trend in Carbon Intensity (weighting varies according to GDP)
			Change of Trend in Carbon Intensity (weighting varies according to GDP)
			Trend in CO2 Emissions per KWH (33%)
			Access to Electricity (N/A)

Source: based on [Hsu at al. 2013]

The weights are determined based on expert judgments on the suitability of the data or the quality of the underlying data through an iterative process. The

EPI developers are aware that "the selection of weights is not a completely objective process and that disagreements are inevitable based on political preferences and even the performance of individual countries on different facets of environmental performance" [Hsu at al. 2014]. And there may be legitimate differences of opinion regarding the relative importance of selected indicators [Emerson at al. 2012].

Data Envelopment Analysis

Data Envelopment Analysis (DEA), developed by [Charnes at al. 1978], is a well-established method for evaluating the relative efficiency of a set of comparable entities — decision making units (DMUs). Due to the fact that the method allows to evaluate systems with multiple inputs and outputs, DEA has been widely investigated and applied in various areas. Since DEA does not necessarily require the use of financial data and can take into account uncontrolled inputs (such as environmental circumstances) is well suited especially for the evaluation non-profit organizations [Chodakowska at al. 2010; Nazarko 2010]. DEA has also gain the popularity in environmental performance measurements [Callens at al. 1999; Meng at al. 2013; Zhou at al. 2008; Zhou at al. 2007]

To measure the EPI by DEA often are used the concept of environmental DEA technology described inter alia in [Meng at al. 2013]. In the technology all outputs are classified into desirable (e.g. GDP) and undesirable outputs (e.g. CO_2). It is assumed that outputs are weakly disposable which implies that the proportional reduction in desirable and undesirable outputs is possible, whereas it may not be feasible to reduce undesirable outputs solely. Desirable and undesirable outputs are null-joint. In other words, the assumptions mean that undesirable outputs must be produced in order to produce desirable outputs and the only way to remove all the undesirable outputs is to cease the production process [Meng at al. 2013]. The DEA Radial Environmental Index models for measuring the environmental performance of (DMUo) can be written as [Meng at al. 2013]:

$$REI = (X_o, Y_o, Q_o) = \min \theta \tag{1}$$
$$\sum_{i=1}^{I} \lambda_i x_{im} \le x_{om}, \quad m = 1,\ldots,M$$
$$\sum_{i=1}^{I} \lambda_i y_{in} \ge y_{on}, \quad n = 1,\ldots,N$$
$$\lambda_i \ge 0, \quad i = 1,\ldots,I$$

where:
$X_i = (x_{i1}, x_{i2}, x_{i3}, \ldots, x_{iM})$ – input vector,
$Y_i = (y_{i1}, y_{i2}, y_{i3}, \ldots, y_{iN})$ – desirable output vector,
$Q_i = (q_{i1}, q_{i2}, q_{i3}, \ldots, q_{iJ})$ – undesirable output vector,
λ_i – intensity levels at which the production activities are conducted by the DMUs,
I – number of DMUs.

If specific DMU has a larger REI, it has better environmental performance [Zhou at al. 2008].

The EPI created by researchers from Yale University and Columbia University contains only desirable outputs. Despite the sometimes confusing names such as child mortality, higher index value indicates a better situation in the country in terms of the environment. Undesirable outputs are included indirectly in some indices. It means that to apply environmental DEA technology it is necessary to use raw data. Taking into account the proximity-to-target methodology used for the construction of these indicators it would be difficult to compare the results of DEA and the EPI, because in fact they would use different data.

In the article DEA was applied to choose the weights of the indicators used in the construction of the EPI so as to maximize the position of each country in the ranking of environmental performance. By confronting purely mathematical approach with substantive approach involving experts the sensitivity of the EPI to the assumptions was tested.

For this purpose the following primal mulitplier CCR DEA model was used [Ramanathan 2003]:

$$\begin{gathered} \max \sum_{n=1}^{N} \upsilon_{no} y_{no} \\ \sum_{n=1}^{N} u_{mo} y_{mo} = 1 \\ \sum_{n=1}^{N} \upsilon_{ni} y_{ni} - \sum_{n=1}^{N} u_{mi} x_{mi} \le 0, \qquad i = 1,\ldots,I \\ u_{mi}, \upsilon_{ni} > 0, \quad n = 1,\ldots,N, \quad m = 1,\ldots,M \end{gathered} \tag{2}$$

In this linear programming problem the weights (u_{mi}, υ_{ni}) are chosen to maximize the weighted sum of outputs to the condition that the sum of the weighted inputs is equal to 1, and that the efficiencies of other DMUs (calculated using the same set of weights) is restricted to values between 0 and 1.

EPI FOR EUROPEAN COUNTRIES — A CASE STUDY

Units, Variables and Weights

The EPI in 2014 was calculated for 178 countries, 42 of them are located in Europe. Assuming a constant, identical level of inputs for each European country, weights for outputs were adjusted to maximize the assessment of environmental performance. Selected weights for other DMUs – European countries – (calculated using the same set of weights) do not exceed the range 0 and 1.

Without going into the construction of aggregated indicators, at the beginning weights were chosen for two variables: Environmental Health (EH) and Ecosystem Vitality (EV). In the 2014 EPI they have fixed weights: 40% and 60%. Weights for 42 European countries adjusted using DEA methodology have 2 patterns. European countries and theirs weights are presented in Table 2.

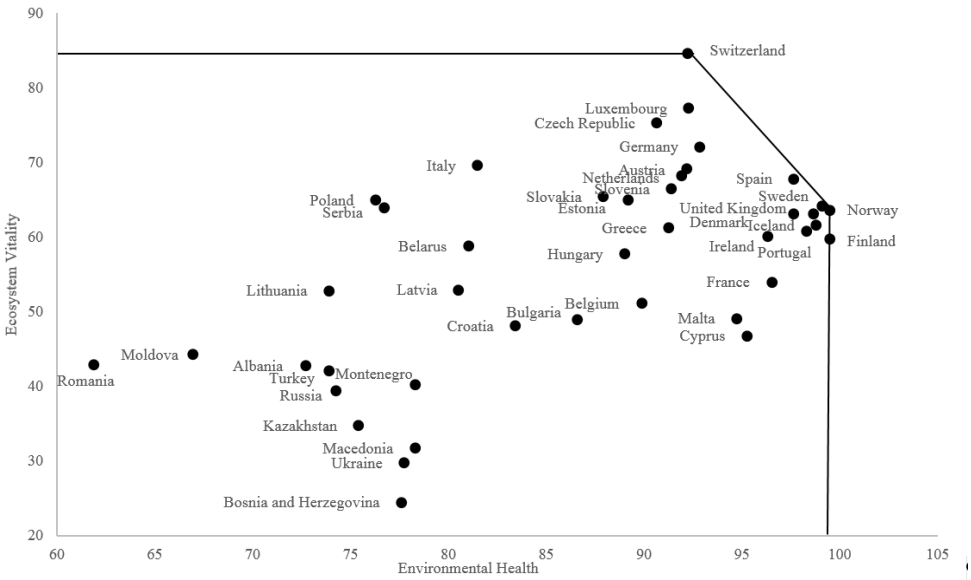
Table 2. Weights Calculated Using DEA Methodology

Country	Environmental Health (EH)	Ecosystem Vitality (EV)
Albania Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Finland, France, Iceland, Ireland, Kazakhstan, Macedonia, Malta, Montenegro, Norway, Portugal, Russia, Turkey, Ukraine, United Kingdom	0,0100563	0,0000001
Austria, Belarus, Czech Republic, Denmark, Estonia, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Luxembourg, Moldova, Netherlands, Poland, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland	0,0082263	0,0028536

Source: own calculations

The EPI for the first group of states should base only on the variable EH, while for the second group of states should take into account 0,008 EH and 0,003 EV. This is due to the fact that the aggregated EH indices are higher for all countries. It is worth noting that the weights determined by DEA method does not add up to 1 and are chosen to maximize the weighted sum of outputs to the condition the efficiencies of other DMUs calculated using the same set of weights is between 0 and 1. In Figure 1 is shown the EH and EV with the frontier imposed by the best DMUs.

Figure 1. Environmental Health and Ecosystem Vitality



Source: 2014 Environmental Performance Index (2014 EPI)

Weights for variables exploited in prior level of aggregation of the EPI were also determined. Due to the lack of data there were used eight out of nine variables that make up the index EH and EV with the following weights: Health Impacts (HI) – 33%, Air Quality (AQ) – 33%, Water and Sanitation (W&S) – 33%, and Water Resources (WR) – 25%, Agriculture (A) – 5%, Forests (F) – 10% , Biodiversity and Habitat (B&H) – 25%, Climate and Energy (C&E) – 25%. In this case, each country received its own unique set of weights highlighting its strengths. The only exceptions are Greece and Italy (Table 3).

Table 3. Weights Calculated Using DEA Methodology

Unit name	EH - HI	EH - AQ	EH - W&S	EV - WR	EV - A	EV - F	EV- B&H	EV - C&E
Albania	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,01162
Austria	0,00001	0,00001	0,00976	0,00001	0,00028	0,00001	0,00001	0,00003
Belarus	0,00001	0,00001	0,00001	0,00355	0,00943	0,00060	0,00001	0,00001
Belgium	0,00001	0,00001	0,00978	0,00001	0,00024	0,00001	0,00001	0,00003
Bosnia and Herz.	0,00021	0,00001	0,00001	0,00001	0,00001	0,00974	0,00005	0,00001
Bulgaria	0,00001	0,00001	0,00645	0,00001	0,00001	0,00293	0,00201	0,00001
Croatia	0,00933	0,00001	0,00001	0,00001	0,00017	0,00046	0,00032	0,00001
Cyprus	0,00993	0,00001	0,00002	0,00001	0,00002	0,00001	0,00001	0,00001
Czech Republic	0,00315	0,00001	0,00001	0,00001	0,00161	0,00001	0,00472	0,00167
Denmark	0,00128	0,00001	0,00001	0,00270	0,00839	0,00001	0,00001	0,00094
Estonia	0,00001	0,00005	0,00001	0,00001	0,00001	0,00001	0,00992	0,00001
Finland	0,00001	0,00001	0,00740	0,00013	0,00334	0,00001	0,00001	0,00037
France	0,00001	0,00001	0,00994	0,00001	0,00002	0,00001	0,00001	0,00001
Germany	0,00001	0,00001	0,00001	0,00583	0,00066	0,00001	0,00398	0,00001
Greece	0,00994	0,00001	0,00001	0,00001	0,00002	0,00001	0,00001	0,00001
Hungary	0,00001	0,00001	0,00005	0,00001	0,00001	0,00991	0,00002	0,00001
Iceland	0,00001	0,00383	0,00001	0,00001	0,00001	0,00035	0,00230	0,00584
Ireland	0,00001	0,00001	0,00001	0,00001	0,00001	0,00995	0,00001	0,00001
Italy	0,00994	0,00001	0,00001	0,00001	0,00002	0,00001	0,00001	0,00001
Kazakhstan	0,00001	0,01012	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001
Latvia	0,00001	0,00935	0,00001	0,00001	0,00065	0,00001	0,00053	0,00001
Lithuania	0,00534	0,00167	0,00001	0,00001	0,00329	0,00001	0,00117	0,00001
Luxembourg	0,00001	0,00032	0,00001	0,00580	0,00062	0,00001	0,00379	0,00001
Macedonia	0,00335	0,00001	0,00001	0,00001	0,00429	0,00001	0,00001	0,00479
Malta	0,00001	0,00238	0,00706	0,00001	0,00001	0,00124	0,00001	0,00001
Moldova	0,00001	0,00001	0,00001	0,00001	0,00002	0,00994	0,00001	0,00001
Montenegro	0,00011	0,00001	0,00001	0,00001	0,00007	0,00981	0,00001	0,00001
Netherlands	0,00001	0,00001	0,00001	0,01007	0,00001	0,00001	0,00001	0,00001
Norway	0,00001	0,00003	0,00992	0,00001	0,00001	0,00001	0,00001	0,00002
Poland	0,00697	0,00001	0,00001	0,00003	0,00063	0,00159	0,00187	0,00001
Portugal	0,00001	0,00131	0,00016	0,00001	0,00001	0,00001	0,00143	0,00944
Romania	0,00586	0,00001	0,00001	0,00001	0,00269	0,00034	0,00051	0,00251
Russia	0,00001	0,01011	0,00001	0,00001	0,00001	0,00001	0,00002	0,00001
Serbia	0,00001	0,00001	0,00001	0,00001	0,00001	0,00995	0,00001	0,00001

Unit name	EH - HI	EH - AQ	EH - W&S	EV - WR	EV - A	EV - F	EV- B&H	EV - C&E
Slovakia	0,00001	0,00001	0,00417	0,00001	0,00315	0,00001	0,00140	0,00364
Slovenia	0,00013	0,00001	0,00001	0,00001	0,00006	0,00005	0,00978	0,00001
Spain	0,00001	0,00001	0,00001	0,00804	0,00170	0,00001	0,00001	0,00170
Sweden	0,00001	0,00001	0,00679	0,00001	0,00374	0,00001	0,00001	0,00094
Switzerland	0,00001	0,00001	0,00001	0,00968	0,00001	0,00001	0,00001	0,00073
Turkey	0,00001	0,01010	0,00001	0,00001	0,00003	0,00001	0,00002	0,00001
Ukraine	0,00001	0,00589	0,00001	0,00001	0,00621	0,00005	0,00001	0,00001
United Kingdom	0,00001	0,00001	0,00001	0,00975	0,00001	0,00094	0,00001	0,00001

Source: own calculations

Due to the design of the EPI with multiple weighted indexes, DEA for choosing the weights can be used at any stage. There are plenty of combinations of weights obtained by solving the linear program tasks and weights determined by experts. Moreover, DEA can be used to raw, untransformed data. It is worth mentioning that DEA models can also take into account additional constraints and experts' knowledge of weight.

Rankings of Environmental Performance

The rankings of countries according to the experts' EPI and the one obtained using DEA were compared. Countries' environmental performance indices and rankings are presented in Table 4.

Table 4. EPI and DEA results

Unit name	EPI Score	EPI Rank	EPI Europe Rank	DEA 1 Efficiency	DEA 1 Rank	DEA2 Efficiency	DEA 2 Rank
Albania	54,73	67	35	73,04%	40	100,00%	1
Austria	78,32	8	6	95,53%	16	99,96%	24
Belarus	67,69	32	25	83,42%	28	99,20%	34
Belgium	66,61	36	27	90,33%	23	99,91%	28
Bosnia and Herz.	45,79	107	42	78,01%	35	99,85%	30
Bulgaria	64,01	41	30	87,06%	25	100,00%	1
Croatia	62,23	45	31	83,86%	27	95,84%	38
Cyprus	66,23	38	28	95,77%	15	99,90%	29
Czech Republic	81,47	5	3	96,06%	14	100,00%	1
Denmark	76,92	13	11	98,31%	9	100,00%	1
Estonia	74,66	20	17	91,87%	21	100,00%	1
Finland	75,72	18	15	100,00%	1	100,00%	1
France	71,05	27	21	97,02%	11	100,00%	1
Germany	80,47	6	4	96,98%	12	100,00%	1
Greece	73,28	23	20	92,53%	20	99,93%	26
Hungary	70,28	28	22	89,69%	24	100,00%	1
Iceland	76,5	14	12	99,33%	6	100,00%	1

Unit name	EPI Score	EPI Rank	EPI Europe Rank	DEA 1 Efficiency	DEA 1 Rank	DEA2 Efficiency	DEA 2 Rank
Ireland	74,67	19	16	96,81%	13	100,00%	1
Italy	74,36	22	19	86,88%	26	99,95%	25
Kazakhstan	51,07	84	38	75,82%	37	97,35%	36
Latvia	64,05	40	29	81,38%	30	99,59%	33
Lithuania	61,26	49	32	75,85%	36	96,55%	37
Luxembourg	83,29	2	2	97,96%	10	100,00%	1
Macedonia	50,41	89	40	78,70%	33	100,00%	1
Malta	67,42	34	26	95,23%	17	99,92%	27
Moldova	53,36	74	37	67,67%	41	99,85%	30
Montenegro	55,52	62	33	78,75%	32	100,00%	1
Netherlands	77,75	11	9	95,07%	18	100,00%	1
Norway	78,04	10	8	100,00%	1	100,00%	1
Poland	69,53	30	23	81,29%	31	99,85%	30
Portugal	75,8	17	14	98,79%	8	100,00%	1
Romania	50,52	86	39	63,14%	42	90,12%	40
Russia	53,45	73	36	74,64%	38	95,78%	39
Serbia	69,13	31	24	81,39%	29	100,00%	1
Slovakia	74,45	21	18	90,99%	22	99,19%	35
Slovenia	76,43	15	13	94,13%	19	100,00%	1
Spain	79,79	7	5	99,69%	5	100,00%	1
Sweden	78,09	9	7	99,77%	4	100,00%	1
Switzerland	87,67	1	1	100,00%	1	100,00%	1
Turkey	54,91	66	34	74,27%	39	85,43%	42
Ukraine	49,01	95	41	78,15%	34	88,85%	41
United Kingdom	77,35	12	10	99,17%	7	100,00%	1

Source: 2014 Environmental Performance Index and own calculations

In Table 4 are shown the EPI values (EPI scores) and places in a ranking based on the EPI of all 178 classified countries (EPI Rank) and of 42 European countries (EPI Europe Rank). There are also presented the indicators calculated using DEA for determining the weights to the variables of the last phase of the EPI procedure (DEA 1) and for the weights to the indicators of the previous step (DEA 2).

The inclusion of DEA in the final stage of the EPI to determine the weights of EH and EV slightly changed the countries' positions. The three countries occupy the first place: Finland, Switzerland and Norway. Switzerland is classified in the first place regardless the way of choosing the weights. Norway to the first position among European countries moved from 8th place, Finland from the 15th. General, use of DEA turned out to be a favourable alternative to 21 countries. In particular, for countries with large, more than 30, discrepancy between the values of EH and EV.

However, the use of the linear programming model (2) for 8 variables forming EH and EV radically changed the countries' ranking. As could be expected, DEA taking into account the strengths of each country, is able to find for each at least one distinguishing element. And in this way 31 of 42 countries received 100% environmental performance score and occupy the 1st place.

Unfortunately Poland does not belong to this group. In the case of Poland, experts' recommendations for weights are optimal. The weights for Poland, chosen so as to maximize its index under assumption that other countries' efficiencies do not exceed 1 at Polish set of weights, do not improve Poland position in the rankings.

The results of the compatibility between rankings based on the Pearson and Spearman correlation coefficient is given in Table 5.

Table 5. Correlation Coefficient

	EPI Score	EPI Rank	EPI Europe Rank	DEA 1 Effic.	DEA 1 Rank	DEA2 Effic.	DEA 2 Rank	
EPI Score		-0,986	-0,972	0,883	-0,866	0,520	-0,587	Pearson
EPI Rank	-1,000		0,944	-0,863	0,834	-0,525	0,565	
EPI Europe Rank	-1,000	1,000		-0,865	0,868	-0,472	0,621	
DEA 1 Efficiency	0,871	-0,871	-0,871		-0,964	0,541	-0,621	
DEA 1 Rank	-0,870	0,870	0,870	-1,000		-0,490	0,643	
DEA2 Efficiency	0,593	-0,593	-0,593	0,606	-0,606		-0,593	
DEA 2 Rank	-0,642	0,642	0,642	-0,664	0,664	-0,988		
	Spearman							

Source: own calculations

Strong relationship was noted between the results of DEA 1 (efficiency and rank) and the ranking of the EPI (score, rank). Other correlations, although statistically significant with $p < 0,05$, are weak.

SUMMARY

There is wide range of environmental performance indicators, from simple indexes to more sophisticated ones, in which there is a possibility to take arbitrary viewpoints [Tyteca 1996]. The aggregated measurements of environmental performance, which is often in the form of an environmental performance index, can provide condensed information for analysts and decision makers dealing with energy and environmental related issues [Zhou at al. 2007]. Whatever weightings and aggregation methodologies are eventually chosen, there will always be individuals who disagree with the final decision [Hsu at al. 2013]. DEA has gained great popularity because it provides synthetic indicator that take into account the

strengths of each analysed entity and do not require the specification of any a priori weight on the variables.

In this article DEA was used to adjust weights for variables that construct the 2014 EPI created by scientists of the Yale University and Columbia University. Depending on which steps of constructing the EPI DEA was incorporated, weights estimating via solving linear programing tasks provide better assessments in case of 21 or 31 countries. Other countries should rather rely on the experts' recommendation.

The article can be regarded as part in the discussion about the choice of weights to achieve a predefined order of a ranking. DEA applied to two variables and 42 objects improves assessments for 50% of the units. If more variables are considered, the easier is to find the strengths of a larger number of units.

It is worth noting that there are also the possibility of including the DEA method for choosing the weights at other stages of the EPI. In particular, it may be interesting to implement DEA to the raw data or at least not converted into indicators using a proximity-to-target methodology.

REFERENCES

Callens I., Tyteca D. (1999) Towards indicators of sustainable development for firms. A productive efficiency perspective, Ecological Economics, vol. 28, issue 1, pp. 41–53.

Charnes A., Cooper W.W., Rhodes E. (1978) Measuring the Efficiency of Decision Making Units, European Journal of Operational Research, vol. 2, issue 6, pp. 429–444.

Chodakowska E., Komuda M. (2010) Zgodność rankingów jednostek sektora edukacji uzyskanych za pomocą metody obwiedni danych, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, Taksonomia 107, pp. 492–500.

Emerson J.W., Hsu A., Levy M.A., Sherbinin de A., Mara V., Esty D.C., Jaiteh M. (2012) Environmental Performance Index and Pilot Trend Environmental Performance Index. New Haven: Yale Center for Environmental Law and Policy. Available: http://www.epi.yale.edu/.

Hsu A., Emerson J., Levy M., Sherbinin de A., Johnson L., Malik O., Schwartz J., Jaiteh M. (2014) The 2014 Environmental Performance Index. New Haven, CT: Yale Center for Environmental Law & Policy. Available: www.epi.yale.edu.

Hsu A., Johnson L.A., Lloyd A. (2013) Measuring Progress: A Practical Guide From the Developers of the Environmental Performance Index (EPI). New Haven:Yale Center for Environmental Law & Policy.

Meng F.Y., Fan L.W., Zhou P., Zhou D.Q. (2013) Measuring environmental performance in China's industrial sectors with non-radial DEA, Mathematical and Computer Modelling, vol. 58, issue 5-6, pp 1047–1056.

Nazarko J. (2010) DEA method in efficiency assessment of public higher education institutions, Performance Management and Measurement with Data Envelopment Analysis, American University of Beirut, Lebanon, pp. 177–182.

Ramanathan R. (2003) An introduction to Data Envelopment Analysis: a tool for performance measurement, Sage Publications, New Delhi.

Tyteca D. (1996) On the Measurement of the Environmental Performance of Firms — A Literature Review and a Productive Efficiency Perspective, Journal of Environmental Management, vol. 46, issue 3, pp. 281–308.

Zhou P., Ang B.W., Poh K.L. (2008) Measuring environmental performance under different environmental DEA technologies, Energy Economics, vol. 30, issue 1, pp. 1–14.

Zhou P., Poh K.L., Ang B. W. (2007) A non-radial DEA approach to measuring environmental performance, European Journal of Operational Research, vol. 178, issue 1, pp. 1–9.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 307 – 316

APPLICATION OF ARTIFICIAL NEURAL NETWORK SUPPORTING THE PROCESS OF PORTFOLIO MANAGEMENT IN TERMS OF TIME INVESTMENT ON THE WARSAW STOCK EXCHANGE

Marcin Halicki
Department of Regional Policy and Food Economy, University of Rzeszow
e-mail: mhalicki@ur.edu.pl
Tadeusz Kwater
Faculty of Mathematics and Natural Sciences, University of Rzeszow
e-mail: tkwater@univ.rzeszow.pl

Abstract: The paper presents the use of artificial neural networks as a tool expert, which supports decision-making for the quarterly period investing on the stock exchange. The authors also proposed a set of 12 features of the economy and the stock market, which has a universal character so that the approach presented in the publication of this configuration data can be useful for any chosen market. Tests were carried out on the basis of actual data from WSE (GPW in Warsaw) and the Polish economy.

Keywords: artificial intelligence, stock exchange, features, investment

INTRODUCTION

The period of globalization with the gradual removal of barriers to the cash flows should contribute generally to the increase in the volume of financial transactions in the world. The practice has become so until 2007, since the collapse of Lehman Brothers started a new period for the capital markets in which we can distinguish two trends. Firstly, the value of financial transactions in the world is about 30% lower than in the strong development of the capital markets, it is in the years between 2006-2007 [Report Life after Lehman 2013, p.6]. Secondly, it has negatively impacted on the equity markets of countries belonging to the "emerging

markets". One such a country is Poland[1]. Practice shows that due to the globalization, the analysis should take into account not only indigenous capital markets but also foreign markets, due to their interdependencies [Murphy 2004, p. 236]. Therefore, financial institutions desiring to raise cash customers for their efficient investment in shares, must effectively choose a stock exchange with an international portfolio diversification.

In this light, the problems which are still under discussion is the method of selecting the securities exchanges and the duration of the investment on the selected exchange. Due to the current state of stock exchanges of the countries belonging to the emerging markets, it is the second focus of the presented problems. Artificial neural networks can be used as a tool for supporting the decision on stock market investments for a specific period of consideration. Therefore, the further part of the publication is dedicated to discuss it. The purpose of this article is to propose a support tool to the decision to invest in the selected stock exchange on the emerging markets on the example of the Stock Exchange in Warsaw. Artificial neural networks are many times used in economic research, but for a specific purpose, which is a choosing the best investment period in shares listed on the stock market, applying this tool has not been met by the authors yet.

It should be added that in the article we used a set of empirical data and artificial neural networks as a modeling tool. Experiments were performed using quarterly data because too short investment period with a given currency may be too less effective due to transaction costs.

GENERAL PRESENTATION OF THE STOCK EXCHANGE IN WARSAW (ON 06.30.2014)

Warsaw Stock Exchange in SA (abbreviated WSE) is a relatively young creation, because it was founded on 12 April 1991 and from 16 April 1991 it began to take its trading sessions. It should also be recalled that the first session of the shares recorded only five companies [gpw.pl]. The main task is to organize characterized exchange trading in financial instruments in order to ensure that the concentration in one place and time, offers buyers and sellers an adequate determination of the course and fulfillment. Currently, Polish stock market faces the challenge as the value of transactions made on it and its capitalization classify it as a small one, especially compared with the stock exchanges of countries with mature economies [Halicki 2013, pp.164-165]. Data characterizing the WSE, which are relevant from the perspective of the analysis are presented in the Table 1.

[1] Despite strong economic growth, Poland is still perceived as a country of emerging market. This is confirmed by the fact that renowned MSCI Emerging Markets Index includes more than 800 shares from 23 countries of this type, including the shares of the Poland.

Table 1. The basic data relating to the Warsaw Stock Exchange (end of June 2014)

Features of WSE	**Values of Features of WSE**
Market capitalization	EUR 148,283.44 mln
Number of listed companies	897
Number of listed bonds	481
Number of listed ETFs	3
Value of Equity Trading of the stocks in period between January - June 2014	EUR 25,743.8 mln

Source: own studies based on FESE - Federation of European Securities Exchanges, (http://www.fese.eu/en/)

ARTIFICIAL NEURAL NETWORKS IN THE CONTEXT OF DETERMINING THE DECISION TO INVEST IN THE SELECTED STOCK EXCHANGE (AN EXAMPLE OF WSE)

Artificial neural networks can be considered as a tool to assist decision-term investments on the selected stock exchange and the aim of this article is to propose a tool for determining the investment decisions for the quarterly period on the example of the Stock Exchange in Warsaw from the perspective of maximizing the return on investment in shares. The neural network is formed by a large number of elements [Tadeusiewicz 1993, p.13], called neurons, in order to process information, which can also be described as a binary element [Arbib 2003, p. 7]. Neurons are connected together in a network with specific weights that are modified in the course of the learning process. It can be divided into 3 types. The first one is defined as a "supervised" learning with a teacher [Ghosh-Dastidar, Adeli 2009, pp.1419-1431]. This method is available as a pair consisting of a learning input vector and the desired response vector. It is assumed then that we know in advance how the network will behave and is known for correct answers to some vectors of the input space. The second was called learning without the supervision [Acciani and other 2003, pp. 427-436] (unsupervised learning), that is, without a teacher, for which there is no information available (either from the teacher or critic) describing the correctness of answers provided by the network. The network still has inputs and outputs, but there is no feedback from the environment. Learning with a critic or reinforcement [Noel, Pandian 2014, pp. 444-451] (reinforcement learning) is the third kind, and does not assume the existence of a teacher, and the only critic who does not give a specific answer, but evaluates its correctness.

The use of artificial neural networks in the context of determining the decision to invest in a given period on the selected stock exchange can be seen as a method based on computer pattern recognition (called. Pattern recognition) based on neural networks. In the case of the analysis period to invest on the stock exchange, the objects are those periods. This solution also requires the assignment

of objects to specific characteristics. Ultimately, during the learning process, artificial neural networks should recognize the different situations of economic and stock market, which will provide the basis for a decision to assign them to the appropriate class. However, the application of the proposed scheme of conduct, it is necessary to determine the specific characteristics of these periods. The basic difficulty in the application of this tool lies in the fact that it is not easy to determine the team universal features that are an essential factor in the process of its use. This is evidenced by the fact that in the literature on portfolio management there can not be found a universal set of properties, that is possible to be used by each managing portfolios, giving under analysis the holding period for the selected stock exchange in the world. With this in mind, and limiting the considerations to shares, it is worth suggesting the 12 universal traits for the described periods (along with the features of countries where there is a stock exchange), which can be a learning data for artificial neural networks. A set of these data should allow to generate a 3-class decision. The first would include the period during which you should not invest in stocks (based on suggestions by the expert, the pattern in this case is "-1"), and the second - in which you should invest (the pattern is "1"). While the third would include a decision on the period in which you can continue to invest, but rather not to increase the portfolio (the pattern is "0"). This division is made on the basis of the suggestions of an expert. It is easy to show that the division of classes is due to the separate periods in order to evaluate their attractiveness from the perspective of obtaining a high return on investment in equities. At the same time we should take into account the fact that the placement of cash for short periods of time in the shares of the exchange could become unprofitable because of transaction costs and low liquidity of certain shares. The article shows only the best quarter of investing in stock market, except for the aspect of the choice of financial instrument, therefore the research does not concern individual shares. Key assumptions concerning their empirical study are as follows:

- A single period of the investment on the Warsaw Stock Exchange is 3 months and tested quarterly periods from 02 Jan. 2001 to 30 June 2014 (the number was 54). It was assumed, therefore, that the attractiveness of the period is examined from the perspective of 3 months. On the basis of the characteristics of the individual periods, their attractiveness was rated(in the form of expert's suggestions) and there was assigned the value of the standard, amounting to "-1", "0" and "1". This will allow the study of artificial neural networks to evaluate future 3-month periods based on the expected value characteristics. This means that after 3 months, you can decide to continue investing or enlarging the portfolio or sell stocks.
- The research process used "supervised" learning where after learning network based on actual data, an experiment was conducted for 20 hypothetical periods predictable and the possible values of the features, but that may be ambiguously evaluated by an expert.

- Experiments were performed using a variety of network configurations, wherein: the number of entries was always 12 and a multiple of the learning experience was constantly changed, the number of neurons in the hidden layer was between 4 and 20. Finally a network architecture comprised of the hidden layer, which consisted of the 6 neurons and the output layer (thus it was two layer feed forward neural network), wherein the transfer function in the hidden layer was sigmoid function (TANSIG), and in the output layer – linear function (PURELINE), multiple learning was 50.
- The network was trained by Back Propagation Method, according to the algorithm of L-M (Levenberg-Marquardt), and the goal of learning was to obtain the smallest value of the sum of squares of the difference between the output of the network and the value of the pattern constructed by an expert.

The precise nature of the proposed universal set of 12 features is presented in tabular form (the 6 traits relates to the stock exchanges, and the state, and other 6 features is the change in the value of the former). It should be noted that as the value of a change becomes greater, it becomes the better - except for a change in the risk index because its growth is evaluated negatively.

Table 2. Set of features quarters of investing in WSE (related to the stock segment)

Name of the features	**The indicator used to calculate the value of feature**	**Essence in relation to the returns on investment in shares**
Polish economic growth (in relation quarter to quarter) and its change in % compared to the previous quarter.	Growth of GDP	Real processes in the economy interact (sometimes with delay) with stock's price
The quarterly growth rate of the main index WIG 20 and the change in % compared to the previous quarter.	WIG20 index	The index will enable a synthetic presentation of the situation on the stock market.
Risk WIG20 quarterly and its change in % compared to the previous quarter.	The standard deviation of the rate of return index based on the last five quarterly periods.	The higher is the standard deviation, the greater is the risk.
The growth rate of the stock exchange capitalization in% - relation of quarter from the previous quarter and the change in% compared to the previous quarter .	The total value of stock market capitalization of all companies is the capitalization of the stock exchange.	The growth rate of market capitalization reflects directly the increase in prices of all shares listed on the stock exchange.
The rate of increase in the value of trading in shares-quarter relation to the same quarter of the previous year (a database	Turnover in terms of value is calculated as the product of the course and the	When the marketing of the instrument is higher than the other, we can say that the liquidity of the instrument is

Name of the features	The indicator used to calculate the value of feature	Essence in relation to the returns on investment in shares
with only the cumulative value) and its change in% compared to the previous quarter.	number of sold and purchased instruments (counted twice).	higher. As the trade in shares grows, we can expect a growing interest in the stock market, which affects the share's price.
The increase in the number of listed companies and its change compared to the previous quarter.	The increase in the number of new companies listed on the Stock Exchange during the quarter.	A growing number of companies shows a positive trend in the stock market because it reflects the expectations of issuers.

Source: own study

The presented set of features aimed to reflect generally the situation on the stock market in the period of 3 months. All the features were calculated on the basis of reputable databases OECD[2], FESE[3] and Stooq[4]. The expert's suggestion became a standard, which was used for artificial neural networks to learn. Almost all the features are expressed in "%", only the last 2 take the values of natural numbers, positive or negative. Sample time along with the actual data is presented in Table 3.

Table 3. Examples of the actual data values of 12 features in a second quarter of 2014. (that is on 06.30.2014)

Name of the features	Value of the feature	Change of feature value
Polish economic growth	1.06%	0.0%
The rate of return of the main index WIG 20	-2.18%	-4.74%
The risk of WIG20	4.48%	-1.46%
The growth rate of market capitalization of the Stock Exchange	-2.26%	-4.29%
The rate of increase in the value of trading in shares	-3.99%	-11.30%
The increase in the number of listed companies	5	8
Expert suggestion	Z (pattern value)	-1

Source: own study

[2] OECD (http://stats.oecd.org/index.aspx?queryid=350) [Accessed 29 August 2014]
[3] FESE, (http://www.fese.eu/en/)[Accessed 29 August 2014]
[4] Stooq, (http://stooq.pl/)[Accessed 29 August 2014]

USING ARTIFICIAL NEURAL NETWORKS TO DETERMINE THE QUARTERLY INVESTMENT DECISION ON THE SELECTED STOCK EXCHANGE (THE EXAMPLE OF WSE)

Implementation of empirical research required to carry out a multi-stage cycle simulation running in the MATLAB software environment. On the basis, the network has been taught using actual data. The aim was to obtain simulation results consistent with the suggestions of an expert to be able to undertake investment decisions in the future. It should be noted that in the study, there were tested variants of a neural network architecture with hidden layer neurons and receptors, where the input amount was always 12, and one neuron in the output layer. In addition, the number of epochs was set in the range (50-2000). The network generated results in the form of numbers from about -1 to 1. As it is known, the initialization of the neural network implemented by a random selection of the initial weights and biases generated different starting points in the learning process, hence each learning process was associated with not identical end results. However, in the initial learning phase, the neural network did not establish the weights properly because its answers were different from the expert's indications. Therefore, there was proposed an innovative approach of multiple learning of artificial neural network in the way that the result of each learning (that is, the distribution of weights and biases - input threshold) was the beginning of the next learning. The number of these repetitions (iterations) were matched experimentally and the smallest number was 50, after which the distribution of weights and biases provide the correct response network. Most interesting elements of the experiments in the learning process are presented in the form of table (table number 4).

To improve the quality of the results, the outcome of the experiments is presented in a quantitative form. The values are calculated based on the arithmetic mean of the three results for the learning network. In Table 4, the "C" is the sum of the squares of the difference between the output of an artificial neural network and the value of the pattern (the recommendation of an expert). The lower is the value of "C", the result generated by the network is improved, and the ratio "D" is the number of indications (data network) which differ from the experts' not less than 0.99. The lower is the ratio "D", the better is the result of the network and „n" is the number of the examined period.

In addition, in experiments there were also used modifications of startup data, using multipliers for some input. There was used:

a – multiplier, which is multiplied by the value of two features (each of the studied period) that is growth and change in the value of the growth.

b – multiplier, which is multiplied by the value of two features (each of the studied period) that is the increase in the number of listed companies and the change in the value of the growth.

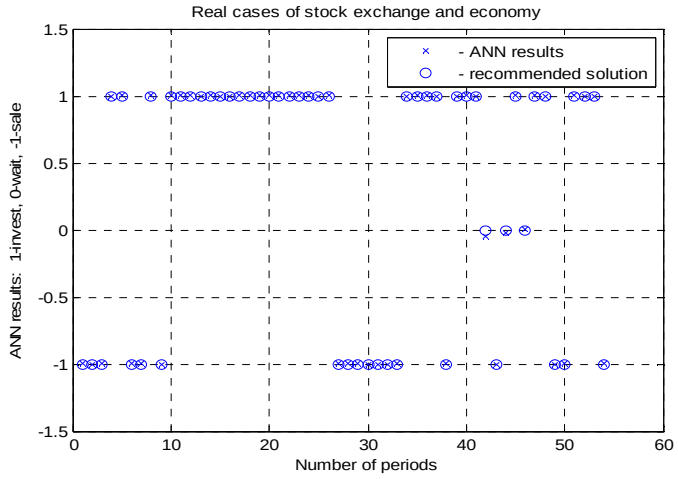
Table 4. Selected information about the experiments of the learning process.

No.	Description of the experiment	Type of data	The averaged result	Comment
1	Standard learning networks without multipliers	The real (n=54)	C=17.36 D=5	This approach does not guarantee satisfactory results generated by the network.
		The hypothetical (n=20)	C=11.59 D=4.33	
2	Multiple learning network without the use of multipliers	The real (n=54)	C=0.44 D=0	The indicated method proved to be moderately satisfactory, however, did not yet expert network tool.
		The hypothetical (n=20)	C=4.53 D=2	
3	Multiple learning network of multipliers for parameter data (a=10, b=3)	The real (n=54)	C=0.00 D=0	This method proved to be the best. In relation to the actual data, the network recognizes perfectly expert suggestions, reaching 100% efficiency (Fig. 1). For the hypothetical data results are also good.
		The hypothetical (n=20)	C=3.16 D=2.33	

Source: own study

In the summary of empirical study, a standard network training without multipliers does not always result in satisfactory end results, but multiple network learning multipliers proved to be the best method (which is presented in Figure 1, in which the OX axis is the number of the examined period, while the axis OY - Values the network results in a marked "O" and the suggestions of the expert - in the form of "X").

Figure 1. Answer of network with data multipliers based on the real values



Source: own study on the basis of the result of MATLAB

Figure 1 presents the results of simulation Option 3 in Table 2, which visually can be considered satisfactory (in all cases the position of the "O" and "X" is almost identical).

CONCLUSIONS

The aim of the publication is to present the research of artificial neural networks as an expert tool, supporting the decision to invest cash in the shares of the selected stock exchange on the basis of the knowledge obtained from the history of the quarterly observation periods. To solve the research problem, which is to determine the investment decision, it was first offered a universal set of 12 features of the economy and the stock exchanges on which the artificial neural network (implemented in MATLAB) could be subjected to the learning process. In the process of learning, the artificial neural network was configured in a variety of ways.The best results were obtained for the two layer feed forward neural network when:

- the network consists of one hidden layer composing of 6 neurons, and the output layer, where the transfer function in the hidden layer was "TANSIG", and in the output layer - "PURELINE",
- the network was trained by Back Propagation method, according to the algorithm of L-M (Levenberg-Marquardt).

Such realized net, unfortunately not always generates results similar to expert's forecasts. Therefore, further modifications were proposed in both data and learning process, and in particular the use of:

- multipliers that increase the value of the first 2 and last 2 features (in the case of WSE there were used numbers 10 and 3),
- multiple learning neural network (it equaled 50).

It should be emphasized that the effective maneuver enhancing the quality of the results was the introduction of multipliers "a" and "b" and the use of multiple learning. Finally, there were obtained satisfactory research results, whose main characteristic was that the artificial neural network obtains results consistent with expert's suggestions based on real data. On this basis, it is reasonable to conclude that the artificial neural network can be regarded as an expert system, supporting decisions for the quarterly investing period on the stock exchange. Furthermore, it appears that the proposed set of 12 features has such a universal character that the presented way to configure the network with this set of data can be useful in any chosen market. Therefore, the presented approach in this article and supplemented in other prognostic tools for the value of features in future 3 month periods can be a comprehensive expert system for portfolio management process of shares listed on various stock exchanges.

REFERENCES

Acciani G., Chiarantoni E., Fornarelli G., Vergura S. (2003) A feature extraction unsupervised neural network for an environmental data set, Neural Networks, Vol. 16, Issue 3-4, pp. 427–436.

Arbib M.A. (ed.) (2003) The Handbook of Brain Theory and Neural Networks, Massachusetts Institute of Technology, London, p. 7.

FESE, (http://www.fese.eu/en/) [Accessed 29 August 2014]

Ghosh-Dastidar S., Adeli H. (2009) A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection, Neural Networks, Vol. 22, Issue 10, pp. 1419–1431.

gpw.pl, (http://www.gpw.pl/root) [Accessed 29 August 2014]

Halicki M. (2013) Capital Market in Poland and Germany – the comparative law analysis, Scientific Papers of the Ministry of Education and Science of Ukraine, No. 4/69, Kiev, pp. 164-165.

Murphy J.J. (2004) Intermarket analysis: profiting from global market relationship, John Wiley&Sons, Inc., p. 236.

Noel M. M., Pandian B. J. (2014) Control of a nonlinear liquid level system using a new artificial neural network based reinforcement learning approach, Applied Soft Computing, Vol. 23, pp. 444-451.

OECD, (http://stats.oecd.org/index.aspx?queryid=350) [Accessed 29 August 2014]

Raport (2013) Life after Lehman, Five years on, Allen & Overy LLP 2013, p. 6.

Stooq, (stooq.pl) [Accessed 29 August 2014]

Tadeusiewicz R. (1993) Sieci neuronowe, A. O. W., Warszawa, p. 13.

WEALTH-X, (http://www.wealthx.com/wealthxubswealthreport/) [Accessed 29 August 2014]

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 317 – 329

A MULTI-PRODUCT VERSION OF THE DEA+ METHOD [1]

Artur Prędki
Department of Econometrics and Operational Research
Cracow University of Economics
e-mail: predkia@uek.krakow.pl

Abstract: The paper presents the DEA+ method as a tool for estimating the production function and the measure of technical efficiency in data points. A multi-product case is considered. Presentation of the underlying semiparametric frontier model is followed by demonstration of the very algorithm of DEA+ and a discussion of its validity. Finally, the method is illustrated with an empirical example with selected model distributions for each random variable constituting the composed error.

Keywords: DEA+, semiparametric frontier model, production function, technical efficiency

INTRODUCTION

DEA+ is a two-stage procedure of point estimation of the production function (transformation) and the measure of technical efficiency of a production unit within the semiparametric frontier model. It was first presented by Gstach [Gstach 1998, 1999] but did not gain popularity. However, chronologically, it is the first method in which DEA (*Data Envelopment Analysis*) is connected with the composed error term. Construction of the model and the method is based on SFA (*Stochastic Frontier Analysis*) – see, e.g., [Kumbhakar, Lovell 2000]. It is thus a way of linking DEA with the methods of production process analysis based on parametric models. Additionally, it can be considered a predecessor of now commonly used StoNED (*Stochastic Non-smooth Envelopment of Data*) – see [Kuosmanen, Kortelainen 2012] or a paper in Polish [Prędki 2012].

[1] The study conducted with financial support from the Faculty of Management, Cracow University of Economics.

The paper presents a multiproduct version of the method briefly described in the source paper [Gstach 1999]. Originality of this paper can be seen in, firstly, organizing and describing the assumptions of a corresponding semiparametric model, which make the considerations that follow clearer. Certain assumptions are not explicitly mentioned by Gstach, but their introduction results from the context and sparse hints, while others are slightly changed in comparison to their original form. Secondly, critical comments included in this paper can provide an explanation to the reasons behind the lack of popularity of this method. The empirical example in which the method is used is also original, as it assumes a different distribution of one of the components of composed error term than in the source paper.

STATISTICAL MODEL WITH A DISCUSSION OF ASSUMPTIONS

Let us begin with defining the semiparametric model mentioned above with a set of assumptions. The idea was, to a great extent, borrowed from the theory of parametric frontier models, introduced in the late 1970s – see [Aigner et al. 1977] and [Meeusen, Van den Broeck 1977].

Assumption 1. Economic units produce *s* sorts of outputs out of *m* sorts of inputs and use the same technology, represented by T – a compact and convex production possibility set, satisfying the inefficiency condition.

Assumption 2. The quantity of inputs and outputs is given for *n* production units by sample $\mathcal{X}_n = ((\mathbf{x}_j, \mathbf{y}_j) \in T, j = 1, \ldots, n)$. Vector $\mathbf{x}_j$ is characterized by the density function h_{Xj}:

$$\forall \mathbf{x} \in (\mathbf{0}, \underline{\mathbf{x}}): h_{Xj}(\mathbf{x}) > 0, \tag{1}$$

where

$\mathbf{x}_j = [x_{1j}, \ldots, x_{mj}]$ – the vector of the quantity of inputs of j^{th} production unit,

$\mathbf{y}_j = [y_{1j}, \ldots, y_{sj}]$ – the vector of the quantity of outputs of j^{th} production unit.

The vector notation of belongingness of vector $\mathbf{x}$ in formula (1) should be understood "by coordinates". Most probably, the identity of distribution of $\mathbf{x}_j$, $j = 1, \ldots, n$ is assumed here as well. The support of all densities is the same, although the density is indexed by j.

The description of generating vector $\mathbf{y}_j$ is performed separately. Using set T, we first define the so called *Farrell output measure of technical efficiency* for feasible production plan $(\mathbf{x}_o, \mathbf{y}_o)$:

$$\theta_P(\mathbf{x}_o, \mathbf{y}_o) = \max\{\theta \in R: (\mathbf{x}_o, \theta\mathbf{y}_o) \in T\}. \tag{2}$$

What is important here is the fact that its value is greater than or equal to unity and equal to unity for objects technically efficient. Because the form of T is unknown, we also do not know the value of the measure in point $(\mathbf{x}_o, \mathbf{y}_o)$. So, it will be estimated in the model.

Next, we define the set:

$$F = \{(\mathbf{x}, \mathbf{y}) \in T: \theta_P(\mathbf{x}, \mathbf{y}) = 1\}, \tag{3}$$

called *the production possibility frontier*. The Farrell technical efficiency measure here plays the role of a transformation function in its implicit form[2]. Introducing F, leads to formulation of the next model assumption.

Assumption 3.

$$\mathbf{y}_j = \mathbf{y}_{jF} \cdot e^{v_j - u_j} = \mathbf{y}_{jF} \cdot e^{w_j}, \text{ for } (\mathbf{x}_j, \mathbf{y}_{jF}) \in F, \quad (4)$$

where $w_j = v_j - u_j$ is so called the composed error term.

To obtain a full description of DGP for $\mathbf{y}_j$, it is necessary to provide a way in which quantities on the right side of equation (4) are generated. Let us start with the assumption regarding component v_j

Assumption 4a. Noise components v_j, j = 1, …, *n*, have independent and identical parametric distributions set by density function f_V which depends on parameters (θ_V, v_{max}). Additionally:

$$E(v) = 0 \text{ oraz } f_V(v) = 0, \text{ dla } v > v_{max}. \quad (5)$$

Bounding the support of noise by parameter v_{max} was Gstach's idea. Its justification will be provided in the next section. In his earlier paper [Gstach 1998], he presented a slightly different version of this postulate.

Assumption 4b. Noise components v_j, j = 1, …, *n*, have independent and identical parametric distributions set by symmetric density function f_V with the support $(-v_{max}, v_{max})$ dependent on parameters (θ_V, v_{max}).

It should be noticed that a more general version of 4a can be derived from Assumption 4b. This change was probably caused by two reasons. Firstly, in order to prove properties of estimators obtained by using the DEA+ method, it is enough to bound the noise up. Secondly, a lower bound is problematic while introducing the joint density function of the composed error term, which is not mentioned by the author. The details can be found in the paper [Prędki 2014]. The author of this paper supports version 4b as the more practical one. There are numerous well-known and widely-used distributions which satisfy Assumption 4b, yet it is difficult to construct a useful distribution satisfying more general Assumption 4a and, at the same time, not satisfying Assumption 4b. Besides, bounding random noise arbitrarily only up contradicts the idea of disturbances which are to be modelled by this component. That is why in practical applications it is not possible to avoid introducing lower bound of the support of noise and difficulties involved in it.

Both source papers mentioned above offer the same assumption regarding component u_j.

Assumption 5. Components u_j connected with modelling inefficiency have independent and identical distributions set by density function f_U with support R_+ dependent on parameters θ_U.

[2] It means that we have the equation linking inputs and maximal outputs, but it is usually not possible to use it for deriving an analytical formula for the maximal quantity of outputs depending on the quantity of inputs.

Theoretically, θ_V, θ_U can be the vectors of parameters, most often, however, they are just single parameters.

In the DEA+ method, the maximum likelihood method (ML) is used for estimating parameters, and to do so, the likelihood function for composed errors is introduced. This poses a question regarding independence of these components as well as independence of components u_j and v_j. In order to estimate the production frontier, it is necessary to know whether factor $\mathbf{y}_{jF}$ is not dependent[3] on a multiplicative version of the composed error e^{w_j} – see factors in equation (4). The source papers do not address this problem, and that is why it is necessary to introduce an additional assumption.

Assumption 6. Components u_j, v_j are not dependent on each other nor on vector $\mathbf{y}_{jF}$.

The last model assumption is connected with a way of generating $\mathbf{y}_{jF}$ values.

Assumption 7. $\mathbf{y}_{jF}$ value is generated from conditional distribution $Y_F | X_j$ set by density $g_{Y_F|X_j}$ with the support Int(Isoq($\mathbf{x}_j$)).

In the paper [Gstach 1999], the support of density was called *the interior of an input-isoquant* and was not precisely defined. Using the context, it can be deduced that it probably refers to the topological interior of set:

$$\text{Isoq}(\mathbf{x}_j) = \{\mathbf{y}_{jF}: (\mathbf{x}_j, \mathbf{y}_{jF}) \in F\}, \tag{6}$$

where the quantity of inputs $\mathbf{x}_j$ is fixed.

However, the author of this paper has certain doubts regarding correctness of this definition. It is true that there might be numerous combinations of outputs, for which $\theta_P(\mathbf{x}_j, \mathbf{y}_{jF}) = 1$, yet this equation suggests that this set can be a measure-zero set in the space of outputs, so its topological interior might be an empty set as well.

We take the logarithms of both sides of equation (4), itemise it for components according to sorts of output , and then we obtain:

$$\forall j = 1, \ldots, n \; \forall r = 1, \ldots, s: w_j = \ln(y_{rj}) - \ln(y_{rjF}). \tag{7}$$

Let us pay attention to the fact that the composed error term is not dependent on the sort of output. In [Gstach 1998, p. 163] it is called the equi-proportional impact of error terms on particular outputs. So, it is a stochastic equivalent of radial property of the Farrell technical efficiency measure, which means a proportional change of all outputs, thus not dependent on the sort of output.

DEA+ METHOD

Let us start with introducing certain auxiliary terms. The set:

$$\tilde{F} = \{(\mathbf{x}, \mathbf{y}) \in T: (\mathbf{x}, \mathbf{y}\, e^{-v_{max}}) \in F\}. \tag{8}$$

[3] In parametric models it is usually assumed that components u_j, v_j are not dependent on $\mathbf{x}_j$, which is caused by the analytical form of the production frontier dependent explicitly on the elements of vector $\mathbf{x}_j$.

is called a production pseudo-frontier.
The error term:

$$\tilde{w}_j = v_j - v_{max} - u_j = w_j - v_{max} \leq 0, \tag{9}$$

is called a pseudo-efficiency of j^{th} object.
Let us notice that :

$$\mathbf{y}_j = \mathbf{y}_{jF} \cdot e^{w_j} = \mathbf{y}_{jF} \cdot e^{w_j - v_{max} + v_{max}} = \tilde{\mathbf{y}}_{jF}\, e^{\tilde{w}_j}, \tag{10}$$

where $(\mathbf{x}_j,\ \tilde{\mathbf{y}}_{jF}) \in \tilde{F}$.

The sign of pseudo-efficiency and the sequence of equalities (10) follows directly from Assumptions 3-5. This means that an observed quantity of outputs $\mathbf{y}_j$ can be looked at from two perspectives. In fact, we can observe optimal quantities of outputs connected with the production frontier disturbed by two types of shocks (exogenous and inefficiency). On the other hand, it can be assumed that it is the optimal quantity of outputs connected with the pseudo-frontier, disturbed by a shock called pseudo-efficiency. This second concept will be used during Stage I of the DEA+ method.

As it directly follows from definition (8), the shape of a pseudo-frontier is identical with the shape of a real frontier, only shifted by the quantity of $e^{v_{max}}$. Similarly, the same happens to the distribution of deviations $\tilde{w}_j$, j = 1, …, *n*. From Assumption 4 it follows that pseudo-efficiencies form i.i.d. sequence, and their distribution is the distribution of the composed error term w_j shifted downwards by the quantity of v_{max}.

As far as the DEA+ procedure itself is concerned, during Stage I we calculate the estimate of the technical efficiency measure for all objects in the sample:

$$\hat{\theta}_P(\mathbf{x}_j, \mathbf{y}_j) = \max\{\theta \in R: (\mathbf{x}_j, \theta\mathbf{y}_j) \in \hat{T}\}, \tag{11}$$

where

$$\hat{T} = \{(\mathbf{x}, \mathbf{y}) \in R_{0+}^{m+s}: \exists \lambda_j \geq 0: \sum_{j=1}^{n} \lambda_j = 1, \mathbf{x} \geq \sum_{j=1}^{n} \lambda_j \mathbf{x}_j, \mathbf{y} \leq \sum_{j=1}^{n} \lambda_j \mathbf{y}_j\}. \tag{12}$$

A deterministic version of the DEA is used to estimate the value of the technical efficiency measure, precisely, the envelopment form of the BCC model – see [Banker et al. 1984]. Here a set $\hat{T}$ is an approximation of an unknown production possibility set T.

The estimate of the multidimensional equivalent of a production pseudo-frontier in point $\mathbf{x}_j$ is then:

$$\hat{\tilde{\mathbf{y}}}_{jF} = \hat{\theta}_P(\mathbf{x}_j, \mathbf{y}_j) \cdot \mathbf{y}_j. \tag{13}$$

Consequently, the estimate of pseudo-efficiency is expressed by:

$$\forall r = 1, \ldots, s: \hat{\tilde{w}}_j = \ln(y_{rj}) - \ln(\hat{\tilde{\mathbf{y}}}_{jF}) = -\ln[\hat{\theta}_P(\mathbf{x}_j, \mathbf{y}_j)]. \tag{14}$$

Equation (14) indicates that the estimates of pseudo-efficiency are also not dependent on the sort of output. Besides, they are closely connected with the estimate of the Farrell technical efficiency measure obtained by applying the DEA method. The minus in equation (14) indicates that the quantity $\hat{\tilde{w}}_j$ should actually be called an estimate of *pseudo-inefficiency*, according to the convention adopted in parametric models. Let us stress once more that radiality of the Farrell technical efficiency measure is desirable here, because it corresponds with Assumption 3 regarding equi-proportionality of the composed error term.

Stage I of the DEA+ method is justified by the following theorem:

Theorem 1. Under Assumptions 1-7, the asymptotic distribution of the estimator $\hat{\tilde{w}}_j$ is identical to deviation $\tilde{w}_j$ in the interior of set F.

Proof: According to the author of the method, the assumptions of Theorems 5 and 6 from [Banker 1993], which implicate the above theorem, are satisfied.
However, using set intF evokes the same doubts as using set Int[Isoq($\mathbf{x}_j$)] previously.

During Stage II the estimates of parameters $\theta = (\theta_U, \theta_V, v_{max})$ are calculated by the ML method, on the basis of the estimates of pseudo-efficiency $\hat{\tilde{w}}_j$ obtained during Stage I. This means that:

$$(\hat{\theta}_U, \hat{\theta}_V, \hat{v}_{max}) = \operatorname{argmax}_\theta \ln\left[\prod_{j\in J} f_{\tilde{W}}\left(\hat{\tilde{w}}_j|\theta\right)\right], \tag{15}$$

where

$$f_{\tilde{w}}\left(\hat{\tilde{w}}\right) = \int_{\hat{\tilde{w}}}^{0} f_V\left(v + v_{max}\right) f_U\left(v - \hat{\tilde{w}}\right) dv \text{ oraz } J = \{j:\ \hat{\tilde{w}}_j < 0\}. \tag{16}$$

Introducing set J means that in the estimation we do not take into account pseudo-efficient objects ($\hat{\tilde{w}}_j = 0$). For such objects it may happen that $\mathbf{x_j} \notin$ intX, and then consistency of Stage I of the DEA+ method is not guaranteed. Secondly, the next stage of DEA+ is infeasible due to degeneration of limits of integration in formula (16) for joint density function. Gstach claims that, asymptotically, a fraction of objects outside set J is neglected, that is:

$$\lim_{n\to\infty} \frac{\left[\sum_{j=1}^{n} I_{j\notin J}(j)\right]}{n} = 0, \tag{17}$$

takes place, where I(·) – is the indicator function. So, in an asymptotic sense, it will not matter whether the procedure is performed on all observations or only on the pseudo-inefficient ones from set J.

Once we have the estimates of all parameters characterizing distributions of both components of the composed error, we can, firstly, estimate the actual production frontier in data points:

$$\hat{\mathbf{y}}_{jF} = \hat{\theta}_P(\mathbf{x}_j, \mathbf{y}_j)\cdot\mathbf{y}_j\cdot e^{-\hat{v}_{max}} . \quad (18)$$

Secondly, using the same methods as the ones used in the SFA approach, we can also obtain the estimate of the efficiency measure of j^{th} object - see, e.g., [Kumbhakar, Lovell 2000, p. 78]. The author of the method only mentions such a possibility in the paper [Gstach 1999, p. 102], not performing it[4]. It should be remembered that, if we obtain the estimates of variances of both components of the composed error, it is not necessary to apply the method of moments nor the pseudo-likelihood method. The first one, although simple, is not free from drawbacks and limitations - see [Kumbhakar, Lovell 2000, p. 92]. The second one, however, requires laborious computations – see, e.g., [Kuosmanen, Kortelainen 2012, p. 18].

According to Gstach, consistency of the estimator used during Stage I is conditioned by introducing the bounded support of noise. Parameter v_{max} assures one-sidedness of the error term $\tilde{w}_j$, which is consistent with the nature of the DEA estimator, which is also one-sided. The author of this paper would like to highlight another role of parameter v_{max}. It is a necessary component of the correction of the initial production pseudo-frontier. It is thus a similar procedure as in case of the COLS method or the MOLS, where the initial estimator of the production frontier is also corrected, by the largest residual or by $E(u_j)$ characteristic, respectively.

In his source paper [Gstach 1998, pp. 165-167], he attempts to prove consistency of the whole DEA+ procedure, although only its single-product version. Particularly, on the basis of Theorem 1 (consistency of Stage I) and the theorem from paper [Bierens 1994], Gstach proves the theorem regarding consistency of the estimator of the production frontier in data points, obtained as a result of using the DEA+. In paper [Gstach 1999, p. 102], the author of the method claims that this result can be transferred to a multi-product case. Let us write down a corresponding theorem.

Theorem 2. Under Assumptions 1-7, the DEA+ method provides a consistent estimator $(\mathbf{x}_j, \hat{\mathbf{y}}_{jF})$ of point $(\mathbf{x}_j, \mathbf{y}_{jF})$ in the interior of set F.

It should be noted, however, that consistency of the DEA+ procedure is questionable, because of the reason mentioned by its author – see [Gstach 1998, p. 165]. In the ML procedure, in the notation of the likelihood function (15), we obtain the product of densities of random variables $\hat{\tilde{w}}_j$, even though they do not have to be independent. Further elaborations on this issue can be found in paper [Prędki 2014].

[4] Only the characteristic of $E(U|\hat{\theta}_U)$, called average inefficiency, is computed, while the main objective of the paper is a comparison of the SFA and the DEA+ methods.

EMPIRICAL STUDY

The study is based on the data from 2000 gathered by the paper's authors [Osiewalski, Osiewalska 2006], describing 240 county, urban and municipal public libraries in Poland[5]. These libraries are non-profit institutions, which can be treated as production (service) units using certain production factors to manufacture specific products (library services). Suggestions included in the source paper were used to select a set of inputs and outputs.

The following factors are taken as inputs:
x_1 – the number of job positions,
x_2 – the number of books,
x_3 - the number of magazine titles,
x_4 – the usable area of a library,
x_5 – the number of seats in reading rooms

Outputs of libraries include the following quantities, that is the ones which could generate profits if a library were a private firm – following the suggestions from paper [DeBoer 1992].
y_1 – the number of library members registered in a library,
y_2 – the number of books borrowed,
y_3 – the number of visits in reading rooms and reading corners.

Due to the numerousness of data, they are presented as selected empirical characteristics – see Table 1.

Table 1. Selected empirical statistics of data

	median	mean	deviation	min.value	max.value
x1	3,49	8,39	13,27	0,88	93
x2	35779	66886,81	88461,84	345	525441
x3	22	49,64	77,82	1	559
x4	211	524,84	949,04	40	10545
x5	30	53,23	61,94	2	441
y1	1364	4395,34	8135,12	263	74003
y2	29797	85932,70	162064,67	698	1643662
y3	2836	10567,74	20841,96	47	232300

Source: own elaboration

The first stage of the DEA+ method was performed by computing the technical efficiency measures $\hat{\theta}_p(\mathbf{x}_j, \mathbf{y}_j)$, j = 1,…,*n* using formula (11) and formula

[5] Due to the multidimensionality of the model and asymptotic properties, numerous multiproduct data were selected. The author of the paper would like to thank here Prof. Jacek Osiewalski and Dr. Anna Osiewalska for granting him access to their data.

(12). Then their logarithms were taken and pseudo-efficiencies $\hat{\tilde{w}}_j$, j = 1,…,*n* were obtained, which allowed us to move to Stage II. Following the source paper, [Gstach 1998], half-normal distribution $N^+(0, \sigma_u^2)$ was assumed for component u modelling inefficiency. Normal distribution $N(0, \sigma_v^2)$ was assumed for noise *v*, truncated to interval $(-v_{max}, v_{max})$, unlike in the source paper, where truncated symmetric beta distribution was used.

As it was mentioned above, the lower bound of the support of noise causes certain problems while introducing joint density of the random variable $\tilde{W}$. That is why it was necessary to modify the formula suggested by Gstach to the following form:

$$f_{\tilde{w}}(\tilde{w}) = \begin{cases} \int_{\tilde{w}}^{0} f_V(v + v_{max}) f_U(v - \tilde{w}) dv, & \tilde{w} > -2v_{max} \\ \int_{-2v_{max}}^{0} f_V(v + v_{max}) f_U(v - \tilde{w}) dv, & \tilde{w} < -2v_{max} \end{cases}. \quad (19)$$

After a number of arduous but simple transformations, the following form of density was obtained:

$$f_{\tilde{w}}(\tilde{w}) = \frac{\sqrt{\frac{2}{\pi(\sigma_u^2 + \sigma_v^2)}} \exp\left\{\frac{-1}{2}\left[\frac{(v_{max} + \tilde{w})^2}{\sigma_u^2 + \sigma_v^2}\right]\right\}\left[\Phi\left(\frac{-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right]}{2\Phi\left(\frac{v_{max}}{\sigma_v}\right) - 1}, \quad (20)$$

where $\Phi(\cdot)$ is cumulative distribution function of standard normal, and:

$$\mu = \frac{-v_{max}\sigma_u^2 + \tilde{w}\sigma_v^2}{\sigma_u^2 + \sigma_v^2}, \sigma = \frac{\sigma_u \sigma_v}{\sqrt{\sigma_u^2 + \sigma_v^2}}, a = \begin{cases} \tilde{w}, & \tilde{w} > -2v_{max} \\ -2v_{max}, & \tilde{w} < -2v_{max} \end{cases}. \quad (21)$$

Next, reparametrization was performed:

$$\tilde{\lambda} = \frac{\sigma_u}{\sigma_v}, \ \tilde{\sigma} = \sqrt{\sigma_u^2 + \sigma_v^2}, \quad (22)$$

and, using the residuals $\hat{\tilde{w}}_j$ of pseudo-inefficient units, the formula for the log-likelihood function was derived:

$$\ln\left[\prod_{j \in J} f_{\tilde{W}}(\hat{\tilde{w}}_j | \theta)\right] = \quad (23)$$

$$\sum_{j \in J}\left\{\ln\sqrt{\frac{2}{\pi}} - \ln\tilde{\sigma} - \frac{1}{2}\left(\frac{v_{max} + \hat{\tilde{w}}_j}{\tilde{\sigma}}\right)^2 + \ln\left[\Phi\left(\frac{v_{max}\tilde{\lambda}^2 - \hat{\tilde{w}}_j}{\tilde{\sigma}\tilde{\lambda}}\right) - \Phi(b_j)\right] - \ln\left[2\Phi\left(\frac{v_{max}\sqrt{\tilde{\lambda}^2 + 1}}{\tilde{\sigma}}\right) - 1\right]\right\},$$

where

$$b_j = \begin{cases} \frac{\left(\hat{\tilde{w}}_j + v_{max}\right)\tilde{\lambda}}{\tilde{\sigma}}, & \hat{\tilde{w}}_j > -2v_{max} \\ \frac{-\tilde{\lambda}^2 v_{max} - 2v_{max} - \hat{\tilde{w}}_j}{\tilde{\lambda}\tilde{\sigma}}, & \hat{\tilde{w}}_j < -2v_{max} \end{cases} . \quad (24)$$

In order to obtain convergence of the ML procedure, the following restrictions on parameters in the log-likelihood function were imposed:

1. The arguments of cumulative distribution functions of standard normal from formula (23) were bounded to interval [-5;5].
2. The arguments of natural logarithms from formula (23) and parameters $\tilde{\sigma}$, $\tilde{\lambda}$, v_{max} are not lower than 10^{-4}.
3. Additionally, more substantial restrictions[6] were assumed:

$$2\Phi\left(\frac{v_{max}\sqrt{\tilde{\lambda}^2+1}}{\tilde{\sigma}}\right) - 1 \geq 0{,}955,\ v_{max} \geq 2\sigma_v. \quad (25)$$

After numerous tests, the following starting points were selected:

$$[\tilde{\sigma}_o, \tilde{\lambda}_o, v_{max,o}] = [0{,}83;\ 1{,}65889674;\ 0{,}85956661], \quad (26)$$

whose choice was not purely accidental. The starting value $\tilde{\lambda}_o$ results from assuming equal dispersion of two distributions at the start: the normal distribution of nontruncated noise v and half normal distribution of component u. The starting value $v_{max,o}$ is linked with residuals by:

$$\min_{j \in J} \hat{\tilde{w}}_j = -2v_{max,o}, \quad (27)$$

and this relation is strongly connected with the author's correction of joint density of the random variable $\tilde{W}$ given by formula (19). The starting value $\tilde{\sigma}_o$ is close to the local maximum of the log-likelihood function, with starting values of other parameters obtained earlier.

As a result of ML procedure in selected starting points, convergence was obtained reaching the values:

$$[\hat{\tilde{\sigma}}, \hat{\tilde{\lambda}}, \hat{v}_{max}] \approx [0{,}37889;\ 1{,}46353;\ 0{,}42850], \quad (28)$$

which were assumed as the final estimates of unknown parameters of corresponding distributions. Only the restriction $2\Phi\left(\frac{v_{max}\sqrt{\tilde{\lambda}^2+1}}{\tilde{\sigma}}\right) - 1 \geq 0{,}955$ turned out to be valid.

Using $\hat{v}_{max}$, the production frontier in data points was estimated from formula (18). Next, the estimates of parameters were used to compute the value of

[6] Unfortunately, they turned out necessary to obtain convergency of the procedure. It should be noticed that truncating the support of the noise by v_{max} occurs deeply in tails of corresponding distribution, and does not seem particularly limiting.

the efficiency measure for particular objects following the scheme described in the paper [Kumbhakar, Lovell 2000, pp. 78-82]. This measure is given by[7]:

$$TE_j = \exp(-\hat{E}(u_j|\tilde{w}_j)),\ j = 1,\ldots,n. \tag{29}$$

To use the above formula, first of all, conditional distribution $U|\tilde{w}$ is needed. Using known forms of the densities of random variables U and V, after simple transformations, we obtain:

$$U|\tilde{w} \sim D(\tilde{w})\cdot N^{+}(\mu-\tilde{w},\sigma^2), \tag{30}$$

where

$$D(\tilde{w}) = \frac{1-\Phi\left(-\frac{\mu-\tilde{w}}{\sigma}\right)}{\Phi\left(-\frac{\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)}. \tag{31}$$

To compute the expected value $E(u|\tilde{w})$, formulae from the paper [Kumbhakar, Lovell 2000, p. 78] were used, and a missing $D(\tilde{w})$ was added:

$$E(u|\tilde{w}) = D(\tilde{w})[\mu_* + \sigma\frac{\varphi(-\mu_*/\sigma)}{\Phi(\mu_*/\sigma)}], \tag{32}$$

where $\mu_* = \mu - \tilde{w}$ and $\varphi(\cdot)$ is the density function of standardized normal distribution. Formula (32) was realized for particular observations on residuals $\hat{\tilde{w}}_j$ and on the estimates of parameters obtained from the DEA+ method, which yielded the estimate of the appropriate expected value, and, consequently, the value of the efficiency measure TE_j.

The results obtained are presented in Table 2, and, again, selected empirical characteristics are provided. Pseudo-efficient objects, for which the DEA+ method degenerates, (as it was mentioned above) were omitted .

Table 2. Selected empirical statistics of the maximal products and the technical efficiency measure

	median	mean	deviation	min.value	max.value
y_{1F}	2072,01	3846,242	5255,20266	293,6984	33269,87
y_{2F}	38936,2	73289,39	91931,9138	1227,143	504389,4
y_{3F}	3724,3	8739,437	11864,0055	68,62053	71196,99
TE	0,82382	0,781215	0,13075852	0,085188	0,896236

Source: own elaboration

Lines 2-4 in Table 2 refer to estimated maximal values of respective outputs – formula (18). Line 5 refers to the value of the technical efficiency measure calculated using formula (29). Due to a methodological nature of the paper and

[7] It is one of the ways of measuring efficiency presented in the source paper.

limited space available, the author did not include a broader economic interpretation of the results.

CONCLUSIONS

The results of the author's methodological research (see also [Prędki 2014]) indicate two major reasons behind only scarce popularity of DEA+. Firstly, ordinary, deterministic DEA also yields technical efficiency measure as well as the value of production frontier at the data points. Hence, according to Ockham's razor, there appears to be no point in turning to a more complex approach to obtain (however different) estimates of the above quantities. Presumably, what Gstach intended was to render DEA a method of estimation of these quantities within a semiparametric statistical model comparable to a more common approach based on stochastic frontier analysis. However, it still appears to lack validity, for such a model does not allow one to obtain dispersion measures of the new estimator of efficiency measure. What is more, it is not possible to make statistical inference about either the model assumptions or the production process.

Secondly, as pointed by the author of the current paper, the efficiency measure estimator resulting from the DEA+ method has not been proven to be consistent so far. Moreover, the algorithm itself is rather cumbersome, entailing numerous methodological and numerical obstacles. In addition, the statistical model itself, as formulated by its original author, seems to be lacking in its underlying assumptions – one of the relevant assumptions is clearly missing, whereas the formulation of the other is disputable. Therefore, the author of the current article would not recommend employing DEA+ in practice. Simultaneously, it should be noted that such apparently negative conclusions are of genuine scientific merit, and the methodological objectives of the study (presented in the Introduction), constituting major contribution of the article, have been successfully achieved.

REFERENCES

Aigner D., Lovell C. A. K., Schmidt P. (1977) Formulation and estimation of stochastic frontier models, Journal of Econometrics, Vol. 6, pp. 21–37.

Banker R.D. (1993) Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation, Management Science, Vol. 39 (10), pp. 1265-1273.

Banker R., Charnes A., Cooper W. 1984 Some models for estimating technical and scale inefficiencies in DEA, Management Science, Vol. 30 (9), pp. 1078-1091.

Bierens H. J. (1994) Topics in advanced econometrics, Cambridge University Press, Cambridge.

DeBoer L. (1992) Economies of scale and input substitution in public libraries, Journal of Urban Economics, Vol. 32, pp. 257-268.

Gstach D. (1998) Another approach to Data Envelopment Analysis in noisy environments: DEA+, Journal of Productivity Analysis, Vol. 9, pp. 161-176.

Gstach D. (1999) Technical efficiency in noisy multi-output settings, CEJOR, Vol. 7, pp. 93-110.

Kumbhakar S. C., Lovell C. A. K. (2000) Stochastic frontier analysis, Cambridge University Press, Cambridge.

Kuosmanen T., Kortelainen M. (2012) Stochastic non-smooth envelopment of data: semi-parametric frontier estimation subject to shape constraints, Journal of Productivity Analysis, Vol. 38, pp. 11-28.

Meeusen W., Van den Broeck J. (1977) Efficiency estimation from Cobb-Douglas production functions with composed error, International Economic Review, Vol. 18 (2), pp. 435-444.

Osiewalski J., Osiewalska A. (2006) Stochastic cost frontier function for Polish public libraries, in "Space-time modelling and forecasting of economic phenomena" (ed. Zeliaś), pp. 179-193, Cracow Academy of Economics Publishing House, Cracow (in Polish).

Prędki A. (2012) Selected estimation methods in the semiparametric frontier model, Statistical Review, Vol. 59 (3), pp. 215-232 (in Polish).

Prędki A. (2014) Single-product version of the DEA+ method, Statistical Review, Vol. 61 (2), pp. 115-129 (in Polish).

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 330 – 338

THE APPLICATION OF RELATIVE TAXONOMY METHODS TO THE STUDY OF TECHNICAL INFRASTRUCTURE DEVELOPMENT IN RURAL AREAS ACROSS THE PROVINCES OF POLAND

Jarosław Lira, Romana Głowicka-Wołoszyn
Department of Finance and Accouting, University of Life Sciences in Poznan
e-mails: jlira@up.poznan.pl; roma@up.poznan.pl
Andrzej Wołoszyn
Department of Informatics, University of Physical Education in Poznan
e-mail: woloszyn@awf.poznan.pl

Abstract: The article applies dynamic multi-dimensional methods of relative taxonomy in an attempt to evaluate disparities in development of technical infrastructure in rural areas between the provinces of Poland. The results show that although the indices that describe the level of infrastructure development have been rising between 2004 and 2012, regional inequalities have remained high. Moreover, some of the provinces least developed in 2004 have not taken advantage of the opportunities provided by the EU accession to develop technical infrastructure in rural areas and boost their attractiveness.

Keywords: relative taxonomy, technical infrastructure, rural areas

INTRODUCTION

Quality of life and economic growth in rural areas depend heavily on the level of development of their technical infrastructure. The access to its components: water supply, sewerage, gas, electricity, transportation, and communication systems, determines the investment attractiveness of these areas [Chudy 2011]. Due to high capital intensity, infrastructure development requires substantial investments and significant State participation in their financing. Polish accession to the European Union provided access to EU funds and opportunity to quicken the pace of rural infrastructure development. To be sure, the indices of this

development were climbing steadily between 2004 and 2012. The pace, however, has not been uniform across the provinces.

Study of complex phenomena that are typically described by several diagnostic features often employs multi-dimensional comparative analysis to reduce the space of the features to one-dimensional synthetic index. Temporalized taxonomic methods, which are described i.a. in Grabiński [1984] or Zeliaś [2000], permit not only classification of objects in a given time period, but also analysis of evolution of the synthetic index. Hydzik [2011] suggests to study the progression simultaneously in two synthetic indices by what may be called *object development matrix*. A different method to study changes in synthetic indices between objects is proposed by Wydymus [2013]. His method consists of constructing the indices based on relativized diagnostic features.

The article aims to evaluate the scale of disparities in rural technical infrastructure development among the Polish provinces between 2004 and 2012. The dynamic approach to multidimensional methods of relative taxonomy not only allowed to compare provincial levels of this development but also to examine the process of levelling regional disparities following the accession of Poland to the European Union.

MATERIALS AND METHODS

For the evaluation of rural technical infrastructure development among the provinces of Poland in 2004-2012, five diagnostic features were selected[1] (all stimulants):

1) length of public extra-urban communal roads of improved hard surface in km per 100 km^2 of province's rural areas (road network density),
2) users of water supply network as percentage of total rural population[2],
3) users of sewerage network as percentage of total rural population,
4) users of gas network as percentage of total rural population,
5) users serviced by sewage treatment facilities as percentage of total rural population.

The method used for studying disparities in infrastructure development was devised by Wydymus [2013] and consists of constructing relative synthetic indices. The values of individual features for each object (province) and each time period (year) were relativized according to the formula:

$$d_{(b/c)jt} = x_{bjt} / x_{cjt}$$

where: $b \neq c$, $b=1,\ldots, n$, $c=1,\ldots, n$

[1] the selection process followed a thorough appraisal of their merits but also statistical analysis of diagonal elements of the inverse correlation matrix of the features, which helped to avoid excessive correlation in the diagnostic set [Lira, Wysocki 2004]

[2] taken as the number of actual inhabitants as of December, 31 of any given year

x_{ijt} – denoted the observation in the *i-th* object (*i=1, …, n*) of the *j-th* feature (*j=1, …, m*) in time period *t* (*t=1, …, k*).

Thus transformed infrastructure indices of the *c-th* object relative to other objects for feature *j* and time period *t* could be presented in the following form:

$$\mathbf{D}_{jt} = \begin{bmatrix} 1 & d_{(2/1)jt} & \cdots & d_{(n/1)jt} \\ d_{(1/2)jt} & 1 & \cdots & d_{(n/2)jt} \\ \vdots & \vdots & \ddots & \vdots \\ d_{(1/n)jt} & d_{(2/n)jt} & \cdots & 1 \end{bmatrix}$$

In order to classify the objects with respect to all diagnostic features simultaneously the subsequent matrices were calculated:

$$\mathbf{D}^{*}_{jt} = \mathbf{A} \cdot \mathbf{D}_{jt}$$

where the matrix A was defined as:

$$\mathbf{A} = \begin{bmatrix} 0 & \cdots & \frac{1}{(m-1)} \\ \vdots & \ddots & \vdots \\ \frac{1}{(m-1)} & \cdots & 0 \end{bmatrix}$$

The diagonal elements of $\mathbf{D}^{*}_{jt}$ formed matrices $\mathbf{W}_t$ (for each time period):

$$\mathbf{W}_t = \begin{bmatrix} w_{11t} & w_{12t} & \cdots & w_{1mt} \\ w_{21t} & w_{22t} & \cdots & w_{2mt} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1t} & w_{2nt} & \cdots & w_{nmt} \end{bmatrix}$$

The higher the value of w_{ijt} index, the greater was the advantage of the *i-th* object over remaining ones in the *j-th* feature and in the *t-th* period.

Next, the $\mathbf{W}_t$ matrices were used to compute the S_{it} matrix of relative synthetic indices of development for given objects and time periods:

$$S_{it} = \frac{1}{m} \sum_{j} \frac{1}{w_{ijt}}$$

The values of S_{it} smaller than 1 signified relative advantage of the *i-th* object over others in period *t*.

Research material was obtained from the Local Data Bank published by the Central Statistical Office in Warsaw. Calculations were performed using the R program; the script of the method's algorithm is available in the Appendix.

RESEARCH RESULTS

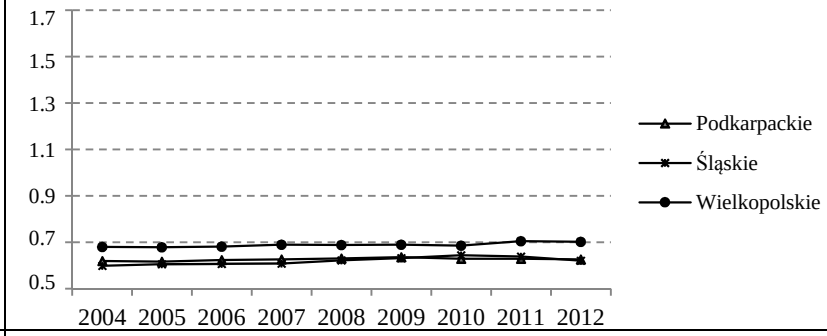
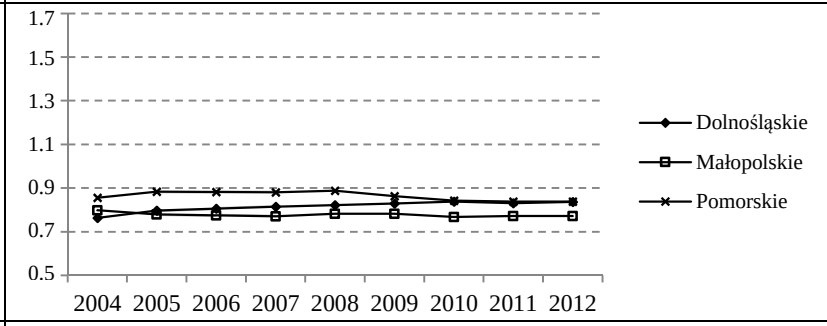
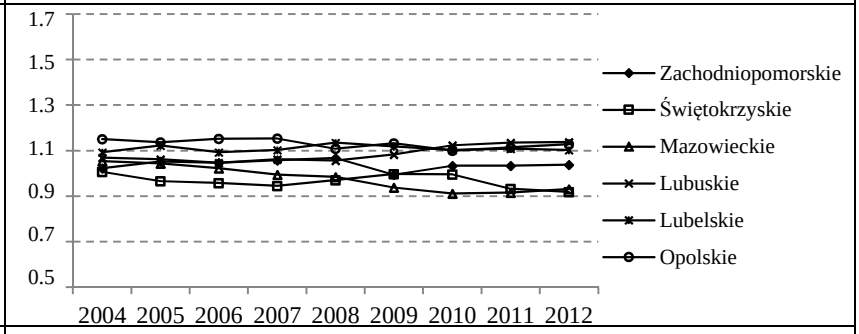
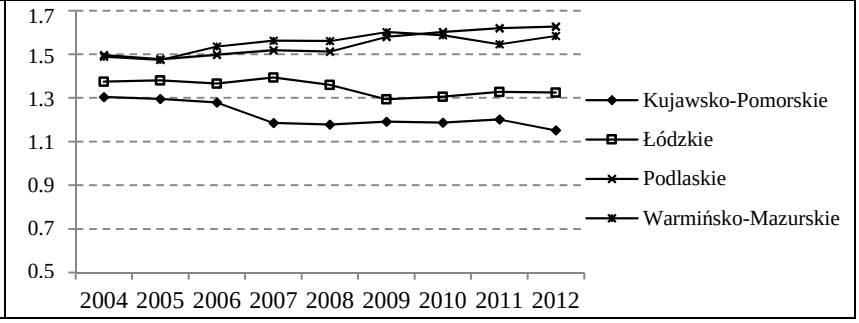
The values of rural technical infrastructure development relative synthetic index S_{it} calculated for 2004 period were used for linear ordering of the provinces: from the highest values of the index to the lowest. Next, the differences between adjacent provinces were computed and used to classify all the provinces into four typological classes. The decision to split a class was made when the differences came out relatively high. Class I of high relative development level comprised the provinces of Śląskie, Podkarpackie and Wielkopolskie, class II of medium high relative level: Dolnośląskie, Małopolskie and Pomorskie, class III of medium low relative level: Lubelskie, Lubuskie, Mazowieckie, Opolskie, Świętokrzyskie and Zachodniopomorskie, and class IV of low relative level: Kujawsko-Pomorskie, Łódzkie, Podlaskie and Warmińsko-Mazurskie.

During the first stage of the analysis the S_{it} values of each province were evaluated for the whole 2004-2012 period. Table 1 depicts them graphically in four 2004 typological classes. All four plots of Table 1 preserved the same scale of the vertical axis in order to facilitate comparison of S_{it} values.

In the whole 2004-2012 period the most noticeable improvement in relative estimates of technical infrastructure development was observed in four provinces: Kujawsko-Pomorskie, Mazowieckie, Świętokrzyskie and Łódzkie, which in 2004 belonged to the medium low (III) or low (IV) relative classes of development.

The maximum drop in the S_{it} index, from 1.30 in 2004 to 1.15 in 2012, which indicated marked improvement in development, occurred in Kujawsko-Pomorskie (class IV). Nonetheless, even with such a striking progress the relative estimate of rural infrastructure in the province remained low. Kujawsko-Pomorskie was followed by Mazowieckie (class III) with the second highest drop in the index, from 1.05 to 0.93. Less pronounced falls in the index, or in other words smaller progress in relative estimates of development, were noted in Świętokrzyskie (class III) and Łódzkie (class IV) and amounted to 0.09 and 0.05, respectively. Four provinces: Dolnośląskie (class II), Lubuskie (class III), Warmińsko-Mazurskie and Podlaskie (class IV) showed worsened estimates of relative development. In Dolnośląskie, despite the increase in the index from 0.76 to 0.84, the relative estimate of infrastructure was still better than in the third and fourth relative classes. Similarly, Lubuskie, where the index reached 1.14 in 2012, was still better off than the provinces of the fourth class. The most worrying changes were observed in Warmińsko-Mazurskie and Podlaskie. Not only had they been considered the provinces least developed in terms of rural infrastructure in 2004, their relative position further deteriorated in 2012. The remaining eight provinces (including the three that in 2004 counted among the first class) showed stable relative estimates: the changes of S_{it} index did not exceed 0.03 either way.

Table 1. Values of relative synthetic index S_{it} for all Polish provinces in 2004-2012

Relative class of development	Values of S_{it} index
I	
II	
III	
IV	

Source: own calculation based on Local Data Bank, Central Statistical Office, Warsaw

Analysis that was conducted in the second stage involved the relative classes created in 2004. Table 2 shows the basic characteristics of analyzed diagnostic features and relative synthetic indices in 2004 and 2012.

The first relative class, which comprised the provinces of Śląskie, Podkarpackie and Wielkopolskie, covered 18.4% of rural areas and close to 25% of population actually living in the rural areas between 2004 and 2012. This class stood out for the highest values of gas and sewerage network, and sewage treatment diagnostic features during the whole 2004-2012 period. The rate of technical infrastructure development in this class was better only than in the slowest second class (except in water supply network) and slower than in the third and fourth classes (except for sewerage network in class IV).

The second relative class with the provinces of Dolnośląskie, Małopolskie and Pomorskie covered 16.6% of rural areas and roughly 22% of rural population. This class was marked by increasing advantage in access to the sewerage network relative to classes III (by 1.1 p.p.) and IV (by 4.1 p.p.) and in access to sewage treatment facilities (by 2.6 p.p. and 4.3 p.p., respectively). Moreover, it maintained its advantage over the two lower classes in road network density (by 13 and 17 km/100 km^2, respectively) and in access to the gas network (by 16.7 p.p. and 27.2 p.p., respectively). A peculiarity of this class lay in its poor access to water supply network, poorer than in any of the remaining classes. As mentioned before, the rate of development of this class was lower than in class I (except in water supply network), class III, and class IV (except in water and sewerage systems).

The largest third class of six provinces covered 38.5% of rural areas and 34.2% of rural population. It had advantage over the fourth class in road network density and in access to gas network, and leveled out the fourth class' advantage in access to sewerage network and sewage treatment plants. Moreover, it had the fastest rate of development of all classes (except for road network density in class IV).

The fourth class with the provinces of Kujawsko-Pomorskie, Łódzkie, Podlaskie and Warmińsko-Mazurskie covered 26.5% of rural areas and 18.7% of rural population. It was the least developed of all four classes, but showed the fastest rate of development in road network density and only slightly slower than the third class' in gas network and sewer treatment access.

The individual diagnostic features in the classes that were characterized by notably faster annual growth rate in 2004-2012 showed positive values of relative annual growth rate (Table 2).

Figure 1 illustrates the values of relative synthetic index S_{it} calculated for the four relative classes. There was little of the levelling out in 2012 of the differences existing in 2004. Slight deterioration was observed in the second class of medium high development (S_{it} rose by 0.04) and equally slight improvement in the third class of medium low development (S_{it} dropped by 0.06).

Table 2. Relative classification of provinces and inter-class disparity in rural technical infrastructure development in 2004 and 2012

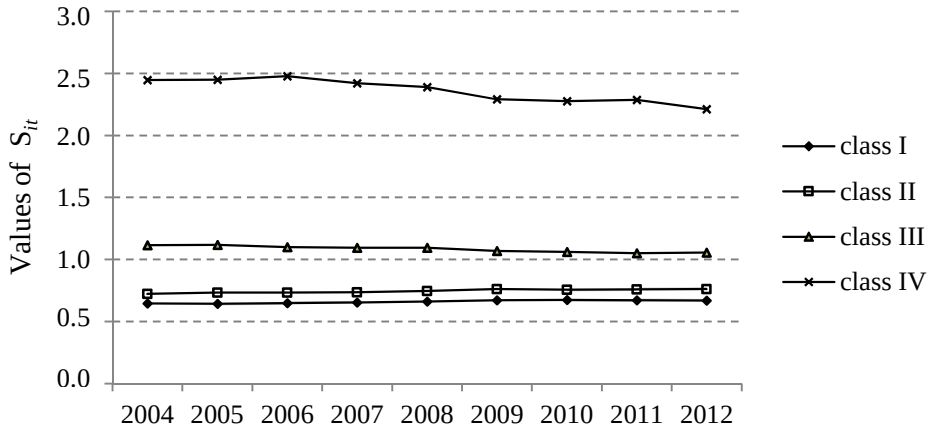
Technical infrastructure	Diagnostic features		Class of rural areas				Poland
			I-high	II-medium-high	III-medium-low	IV- low	
road network density [km/100 km^2]	mean	2004	30.60	31.82	18.17	14.12	21.66
		2012	38.54	38.41	25.45	21.75	29.04
	Average Annual Growth Rate[3] (%)		2.49	1.51	3.28	4.28	2.83
	relative indices w_{ijt}	2004	1.60	1.68	0.82	0.56	1.46
		2012	1.43	1.42	0.83	0.66	1.47
	relative Average Annual Growth Rate[3] (%)		-0.73	-1.87	0.12	1.60	0.23
users of water supply network [%]	mean	2004	75.44	64.77	69.80	76.17	71.32
		2012	78.75	70.53	75.80	80.32	76.20
	Average Annual Growth Rate (%)		0.43	0.91	0.96	0.61	0.73
	relative indices w_{ijt}	2004	1.08	0.88	0.97	1.09	0.98
		2012	1.05	0.90	0.99	1.07	0.99
	relative Average Annual Growth Rate (%)		-0.40	0.23	0.30	-0.17	0.03
users of sewerage network [%]	mean	2004	21.56	19.59	13.39	15.94	17.26
		2012	36.59	32.26	24.99	24.51	29.45
	Average Annual Growth Rate (%)		5.52	5.27	6.65	4.65	5.64
	relative indices w_{ijt}	2004	1.35	1.20	0.71	0.91	1.12
		2012	1.36	1.16	0.83	0.80	1.08
	relative Average Annual Growth Rate (%)		-0.06	-0.35	1.40	-1.10	-0.52
users of gas network [%]	mean	2004	31.77	27.52	9.93	2.45	17.81
		2012	36.55	31.00	14.32	3.79	21.69
	Average Annual Growth Rate (%)		1.14	0.98	3.93	3.90	1.81
	relative indices w_{ijt}	2004	5.77	4.96	1.58	0.14	3.53
		2012	4.46	3.73	1.54	0.16	2.76
	relative Average Annual Growth Rate (%)		-2.26	-2.42	0.52	1.09	-2.06
users serviced by sewage treatment facilities [%]	mean	2004	22.49	20.77	15.17	16.29	18.43
		2012	41.24	36.23	28.07	27.42	33.09
	Average Annual Growth Rate (%)		6.45	6.11	6.86	6.60	6.53
	relative indices w_{ijt}	2004	1.32	1.19	0.78	0.86	1.10
		2012	1.37	1.16	0.83	0.80	1.08
	relative Average Annual Growth Rate (%)		-0.09	-0.50	0.37	0.18	-0.22

Source: as in Table 1.

[3] Average Annual Growth Rate was computed from all elements of the time series [Lira, Wysocki 2004]

The only clear progress was noted in the fourth class, where Sit dropped by 0.23. However, one should notice that the class itself was far from homogenous in terms of rural infrastructure development between 2004 and 2012. Łódzkie and Kujawsko-Pomorskie improved noticeably but the other two provinces showed further deterioration from an already low relative estimates of development.

Figure 1. Values of relative synthetic index S_{it} for relative classes in 2004-2012



Source: as in Table 1.

CONCLUSIONS

The use of dynamic approach to multi-dimensional methods of relative taxonomy facilitated the analysis of the process of smoothing out the differences in development of the rural technical infrastructure between the provinces. These methods proved to be a useful tool in the analysis of changes in the development of infrastructure between individual provinces relative to all others.

The existing disparities in rural technical infrastructure development between the classes of provinces based on 2004 relative synthetic index values were observed to even out slightly between 2004 and 2012. The advantage of the class of high relative development level remained unchanged, but improvement was noticed of the low relative class, spearheaded by two provinces, Kujawsko-Pomorskie and Łódzkie.

APPENDIX

```
# database is expected in the normal (or molten) form with columns
#  + named PERIOD, FEATURE, OBJECT, VAL
# set up the data matrices x
lPERIOD = length(levels(z$PERIOD))
lFEATURE = length(levels(z$FEATURE))
```

```
lOBJECT = length(levels(z$OBJECT))
x = array(z$VAL, dim=c(lOBJECT, lFEATURE, lPERIOD),
 + dimnames = list(levels(z$OBJECT), levels(z$FEATURE),
 + levels(z$PERIOD))
# calculate matrices of relative indices D
D = array(, dim=c(lOBJECT, lOBJECT, lFEATURE, lPERIOD),
 + dimnames = list(levels(z$OBJECT), levels(z$OBJECT),
 + levels(z$FEATURE), levels(z$PERIOD)))
for (t in 1:lPERIOD)
  for (j in 1:lFEATURE)
    D[,,j, t] = (1/x[,j, t]) %*% t(x[,j, t])
# set up the averaging matrix A
A = array(1 / (lOBJECT-1), dim=c(lOBJECT, lOBJECT)); diag(A) = 0
# calculate matrices W of relative indices
W = array(, dim=c(lOBJECT, lFEATURE, lPERIOD),
 + dimnames = list(levels(z$OBJECT), levels(z$FEATURE),
 + levels(z$PERIOD))
for (t in 1:lPERIOD)
  for (j in 1:lFEATURE)
# vectors of W are formed by the diagonal elements of scaled
# D matrices
    W[,j, t] = diag(A %*% D[,,j, t])
# calculate the S matrix: reciprocals of elements of W
# averaged over the diagnostic features
S = apply(1/W, c(1,3), mean)
```

REFERENCES

Local Data Bank, www.stat.gov.pl

Chudy W. (2011) Rozwój infrastruktury obszarów wiejskich, Infrastruktura i Ekologia Terenów Wiejskich, Nr 10/2011, PAN, Kraków, s. 97-106.

Grabiński T. (1984) Wielowymiarowa analiza porównawcza w badaniach dynamiki zjawisk ekonomicznych, Zeszyty Naukowe, Seria specjalna: Monografie, Nr 61, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

Hydzik P. (2012) Zastosowanie metod taksonomicznych do oceny poziomu rozwoju społeczno-ekonomicznego powiatów województwa podkarpackiego, Zeszyty Naukowe Politechniki Rzeszowskiej Nr 286, Ekonomia i Nauki Humanistyczne z. 19 (2/2012), s. 17-32.

Lira J., Wysocki F. (2004) Zastosowanie pozycyjnego miernika rozwoju do pomiaru poziomu zagospodarowania infrastrukturalnego powiatów, Wiadomości Statystyczne, 9, GUS PTS, Warszawa, s. 39-49.

Wydymus S. (2013) Rozwój gospodarczy a poziom wynagrodzeń w krajach unii europejskiej – analiza taksonomiczna, Zeszyty Naukowe Uniwersytetu Szczecińskiego Nr 756, Finanse, Rynki Finansowe, Ubezpieczenia Nr 57, s. 631-645.

Zeliaś A. (2000) Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 339 – 348

TECHNICAL EFFICIENCY OF POLISH COMPANIES OPERATING IN THE COURIERS AND MESSENGERS SECTOR - THE APPLICATION OF DATA ENVELOPMENT ANALYSIS METHOD

Justyna Kozłowska
Department of Business, Informatics and Logistics
Bialystok University of Technology
e-mail: j.kozlowska@pb.edu.pl

Abstract: The main goal of this paper is an attempt to assess condition of polish couriers and messengers sector in terms of technical efficiency. The sector in Poland, in comparison to the Western Europe countries, is still under development. The research was conducted using Data Envelopment Analysis method, which is successfully applied worldwide both in public and private sector organizations. DEA allows to identify the best performers in analyzed group and build a ranking in terms of efficient resources utilization. Malmquist index was also calculated to assess changes in relative productivity over time and to detect which factors ("frontier shift" or "catch-up") affects them.

Keywords: couriers and messengers services, technical efficiency, DEA method, Malmqiust index

INTRODUCTION

First courier companies appeared in the polish market around the 1980s. Although the sector is relatively young, it plays a significant role in the national economy. Polish couriers and messengers market still dynamically develops, in comparison to the Western Europe countries, where the market has already stabilized. The use of courier services per capita is there a few times higher than in Poland, therefore the rapid growth will presumably continue in the near future. The main player in this market is public company – *Poczta Polska* (2,1 billion services provided in 2011), nevertheless the package segment is dominated by private operators (3,7 billion services provided in 2011) [Marcysiak at al. 2013] and the analysis presented hereof focuses on this group specifically. In the paper the

author made an attempt to assess the performance of companies operating in the mentioned above sector in terms of technical efficiency. Not only does the results of such analysis allow to identify entities which are the best in efficient utilization of its resources, and thereby indicate benchmarks for other players in the market, but also it can be the basis for further analysis of inefficient units and the first step to improve such entity's performance.

Efficiency, including technical efficiency, evaluation of any entity involves the study of its productivity [Kosieradzka A. 2000]. Productivity is define in literature generally as a single input-output ratio. However, more common problem pose analyses of multiple-input and multiple-output cases. The term "efficiency" is understood as a production without any wastefulness. Thus an economy (organization, company, unit etc.) is efficient when there is no way to increase a production of one commodity without a drop in production of another, what means it attains the production-possibility frontier or, in other words, achieves the highest level of production possibilities given the resources [Samuelson at al. 1995]. In this meaning, the term "technical efficiency" is also applied.

METHODOLOGY BACKGROUND

DEA method

Data Envelopment Analysis (DEA) is worldwide successfully applied method for evaluating the performance of a set of peer entities, in literature called "decision making units" (DMU), that transform multiple inputs into multiple outputs. DMU can operate in public (see studies of [Ruggiero J. 1996] or [Nazarko at al. 2008]) and private sectors (e.g. [Parlińska M. at al. 2008] and [Rangan N. 1988]), nonetheless a homogeneity of the units is a critical postulation. DEA was first introduce in 1978 by Charnes, Cooper and Rhodes as a Linear Programing based method and, since then, it has become a very popular means for productivity[1] assessment in academic research and in practice as a decision support tool for management level employees.

The basic DEA models enable to measure the efficiency of a DMU relative to similar DMUs in order to estimate a 'best practice' frontier [Cooper W. at al. 2004]. Primary DEA model is named CCR after its authors (Charnes, Cooper, Rhodes), assumes constant returns to scale, and is defined as follows (output-oriented model):

$$\max \; z = \sum_{j=1}^{J} v_{j,m} y_{j,m} \qquad (1)$$

subject to $\sum_{i=1}^{I} u_{i,m} x_{i,m} = 1$

[1]From this point, terms: productivity, efficiency and technical efficiency are alternatively used in the paper, however the author always means the efficiency in accordance to the DEA framework.

$$\sum_{j=1}^{J} v_{j,m} y_{j,m} - \sum_{i=1}^{I} u_{i,m}\, x_{i,m} \leq 0$$

$$u_{i,m}\,, v_{j,m} \geq 0$$

where:
z – weighted sum of outputs calculated for DMU_m (m=1,..,N);
$y_{j,m}$ – denotes DMU_m outputs (j=1,..,J);
$x_{i,m}$ – represents inputs (i=1,…,I) of DMU_m;
$u_{i,m}\,, v_{j,m}$ – weights that are unknown variables.

In input-oriented model the weighted sum of inputs of each DMU is minimalized and the level of its' inputs remains unchanged. In practice, a dual problem of (1) is solved. Primary model had several limitations (e.g. the assumption about returns to scale), thus many modification has been developed since its introduction. However, it is still regularly used in analyses of technical efficiency.

The standard DEA models allows to build a ranking of inefficient units, setting the best performers at the same level (of 100% efficiency). In 1993 Andersen and Petersen developed a new approach, the super-efficiency DEA (SE-DEA) model, which allows to rank DMU on the efficiency frontier [Andersen P. at al. 1993]. This innovation was a subject of many paper afterwards (e.g. [Thrall M.R. 1996], [Yao C. 2003] or [Zhu J. 2001]), and applied to real data researches (see [Helta M. 2009] or [Chen Y. 2004]). The basic idea of the new approach is to exclude the efficient DMU from the comparison set, so it cannot be compared against itself (which is the reason of receiving the score of 100% or 1), thus it can obtain score better or less than one in the input-oriented or output-oriented super-efficiency evaluation, respectively. The difference in the score is usually interpreted as an amount by which the DMU can exceed the performance if it peers.

It is highly desirable to differentiate units with the best results, as it can encourage them to improve theirs output and be better than others. Hence, this extension of DEA method was applied in the research presented hereof.

Malmquist index

Malmquist index represents a change in a unit's productivity over a period of time. In the non-parametric framework (as DEA method) it is defined as the product of two terms – a "frontier shift" (likewise technology change or innovation) term and a "catch-up" (also relative efficiency change or recovery) term. The first term represents changes in efficiency resulting from "industry" gains - i.e. the degree to which the surroundings has changed. The second term shows the degree to which the individual unit's efficiency has improved relative to the frontier - i.e. how far it as from the frontier at time t versus how far it was from the frontier at t+1. [Cooper W. et al. 2004]. Accordingly, Malmquist index (MI) is calculated with the following formula:

$$MI = \left[\frac{\delta^t((x_m,y_m)^{t+1})}{\delta^t((x_m,y_m)^t)} \times \frac{\delta^{t+1}((x_m,y_m)^{t+1})}{\delta^t((x_m,y_m)^t)}\right]^{1/2} \quad (2)$$

where:

$(x_m, y_m)^t$ denotes the vector of inputs (x_i, i=1,…,I) and output (y_j, j=1,…,J) of the DMU_m (m=1,..,N) in the period t,

$\delta^{t+1}((x_m, y_m)^t$ expresses the efficiency score of DMU_m measured by the frontier technology in the period $t+1$.

Malmquist index value of 1 signifies no change in total factor productivity, while values greater than 1 or less than 1 indicates the progress or regress, respectively.

It is essential for the productivity analysis to determine the index value as it allows to recognise the degree of efficiency changes, their direction and major factor ("catch-up" or "frontier shift") that affects them.

THE ANALYSIS AND ITS RESULTS

DEA model selection

Preliminary analysis with the CCR-DEA model indicated that the low score of few units' productivity were due to variable returns to scale, thus in the main research the BCC-DEA model was used. As the output in the service sector is rather difficult to control, the input-oriented model seemed to be more appropriate for this study. To build a ranking of all DMUs, the super-efficiency extension was applied.

Data and variables selection

Data for the analysis was collected from EMIS (Emerging Markets Information Service) database. The time period as well as the number of companies under examination were determined by the availability of the data at the first step. EMIS database gathers mainly financial data and reports so that the evaluation of technical efficiency covers an economic activity of the companies. Then Pearson correlation coefficient between variables were analyzed and variables with strongly linear dependence were chosen. As the author's intention was to obtain the results that can be used both for building general opinion about the sector and for practical reasons from entrepreneurs point of view, therefore following variables were finally chosen for further analysis: total assets and number of employees as inputs and total operating revenue as the output. The analysis was carried out for 23 largest companies from the sector. As the EMIS data are relatively new, the analysis covers the period of 3 years only: 2010, 2011 and 2012.

Table 1 presents values of Pearson correlation coefficient for selected variables.

Table 1. Pearson correlation coefficient values

	Operating revenue		
	2010	2011	2012
Number of employees	0,932236	0,935515	0,927888
Total assets	0,9016044	0,795428	0,754033

Source: own calculations

The league table

The table below presents the result of analysis conducted with Frontier Analyst Application software. It is organized in descending order, so the best performers in terms of technical efficiency are at the top of the table. *Polska Grupa Pocztowa S.A.* takes consistently the first place in every year of analyzed period. *UPS Scs (Polska) Sp z o.o.* takes the second position regularly. The third score received the *UPS Polska Sp. z o.o.*, but in the first two years of analyzed period only. In 2012 another unit appeared in the leading group, namely *World Courier (Poland) Sp. z o.o.*, which hitherto indicated inefficiency. Similarly, the *Madagra Sp. z o.o.* improved its output and joined the "best performers" group.

Table 2. Rankings built with the usage of super-efficiency DEA model

2010		2011		2012	
Unit name	Score	Unit name	Score	Unit name	Score
Polska Grupa Pocztowa S.A.	**413,90%**	**Polska Grupa Pocztowa S.A.**	**268,90%**	**Polska Grupa Pocztowa S.A.**	**321,20%**
Ups Scs (Polska) Sp. z o.o.	**128,50%**	**Ups Scs (Polska) Sp. z o.o.**	**170,60%**	**Ups Scs (Polska) Sp. z o.o.**	**205,60%**
UPS Polska Sp. z o.o.	**104,70%**	**UPS Polska Sp. z o.o.**	**120,30%**	**World Courier (Poland) Sp. z o.o**	**104,10%**
Transline Polska Sp. z o.o.	98,90%	Transline Polska Sp. z o.o.	94,70%	**Madagra Sp. z o.o.**	**103,20%**
Viva Xpress Logistics Polska Sp.	95,10%	Viva Xpress Logistics Polska Sp.	86,30%	UPS Polska Sp. z o.o.	83,90%
World Courier (Poland) Sp. z o.o	93,00%	Federal Express Poland Sp. z o.o	84,60%	Transline Polska Sp. z o.o.	82,10%

2010		2011		2012	
Unit name	Score	Unit name	Score	Unit name	Score
Federal Express Poland Sp. z o.o	88,20%	Madagra Sp. z o.o.	80,40%	Viva Xpress Logistics Polska Sp.	77,90%
Go! Express & Logistics Wrocław	68,90%	Money Save Polska Sp. z o.o.	72,00%	Federal Express Poland Sp. z o.o	75,20%
Fedex Express Polska Sp. z o.o.	67,40%	Go! Express & Logistics Wrocław	69,30%	TNT Express Worldwide Poland Sp.	65,60%
Patron Service Sp. z o.o.	64,90%	World Courier (Poland) Sp. z o.o	68,90%	Patron Service Sp. z o.o.	58,00%
Dpd Polska Sp. z o.o.	63,10%	Patron Service Sp. z o.o.	63,50%	Fedex Express Polska Sp. z o.o.	53,60%
Madagra Sp. z o.o.	62,70%	DHL Express (Poland) Sp. z o.o.	60,70%	DHL Express (Poland) Sp. z o.o.	52,20%
Riders Express Sp. z o.o.	55,70%	TNT Express Worldwide Poland Sp.	60,60%	Matmarex Sp J sp. j.	51,20%
DHL Express (Poland) Sp. z o.o.	55,30%	Riders Express Sp. z o.o.	59,80%	Money Save Polska Sp. z o.o.	46,30%
Money Save Polska Sp. z o.o.	47,60%	Fedex Express Polska Sp. z o.o.	57,50%	Riders Express Sp. z o.o.	44,30%
Matmarex Sp J sp. j.	47,10%	Matmarex Sp J sp. j.	56,60%	Go! Express & Logistics Wrocław	39,60%
Lande Sp. z o.o.	46,00%	Integer.pl S.A.	51,90%	Dpd Polska Sp. z o.o.	37,90%
TNT Express Worldwide Poland Sp.	45,00%	Dpd Polska Sp. z o.o.	44,70%	Integer.pl S.A.	37,70%
Siódemka S.A.	43,60%	Forposta S.A.	39,10%	Inpost Sp. z o.o.	32,80%
Trade System Sp. z o.o.	42,40%	Inpost Sp. z o.o.	36,30%	Lande Sp. z o.o.	32,70%
Inpost Sp. z o.o.	41,10%	Lande Sp. z o.o.	35,00%	Forposta S.A.	18,20%

2010		2011		2012	
Unit name	Score	Unit name	Score	Unit name	Score
Integer.pl S.A.	30,40%	Trade System Sp. z o.o.	24,10%	Siódemka S.A.	18,00%
Forposta S.A.	29,20%	Siódemka S.A.	12,60%	Trade System Sp. z o.o.	16,60%

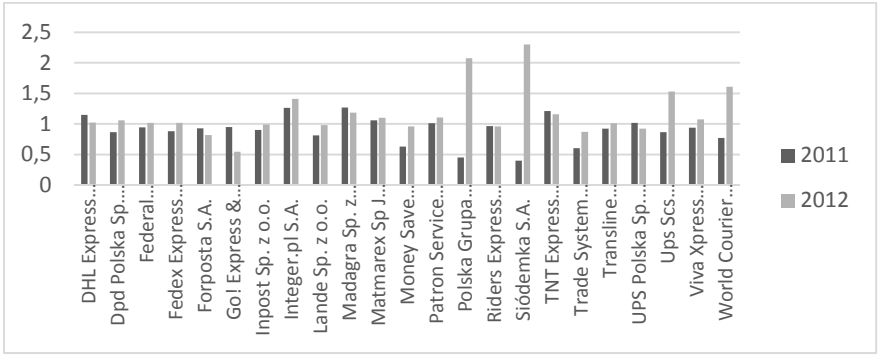
Source: own study based on *Frontier Analyst Application* software calculation

At the bottom of the table (the last few position) the same unit names appear every year, so one can conclude that those companies should reconsider the way they transform their inputs into effects. The minimum value of productivity scores is slightly above 10% in 2011 and 2012, whereas in 2010 it was close to 30%, which is rather alarming and worth to study in details. However, the average technical efficiency in analyzed period of time is around 60%, calculated using standard DEA model, and over 75% determined by supper-efficiency model, thus in general companies seem to manage their resources rather reasonably.

Changes of the efficiency in the 2010 - 2012

Figure 1 presents a magnitude of changes in productivity for each DMU in analyzed period of time. Most of the units (e.g. *DHL Express (Polska) Sp. z o.o*. or *Federal Express Poland Sp. z o.o*.) indicates sustainable changes in its productivity scores (around 1 e. i. without significant degree of change in both direction – up and down). Though, there are few entities (namely *Siódemka S.A.*, *Polska Grupa Pocztowa S.A.*, *World Courier Poland Sp. z o.o*.) which distinguish itself from the group with a very high value of Malmquist index in 2012 (MI >1,5).

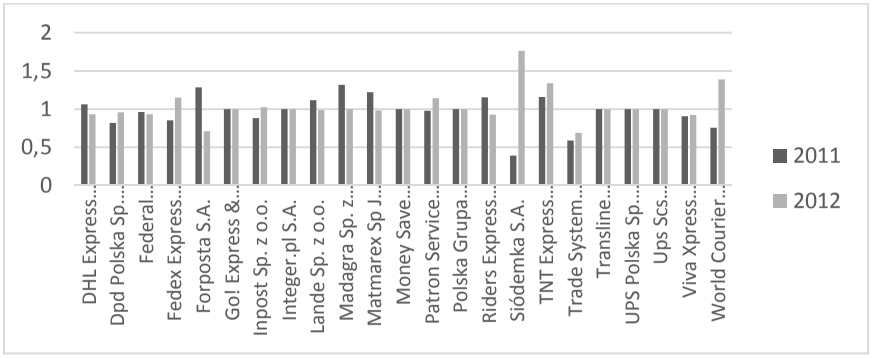
Figure 1. Malmquist index values for each DMU



Source: own calculations

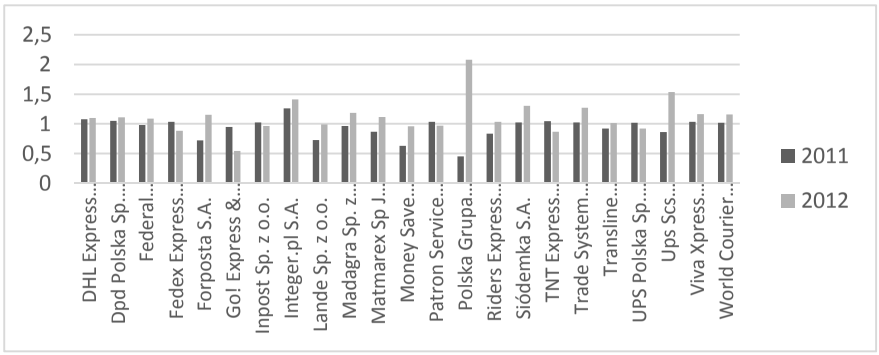
The study of the figures 2 and 3 reveals that these changes are affected mainly by the second term of the Malmquist index, e.i. a “frontier shift”. Moreover, the MI value of units in question was relatively low in the previous year, so presumably some inside issues impacted these units performance as well.

Figure 2. "Catch-up" for each unit – i.e. the relative efficiency change component



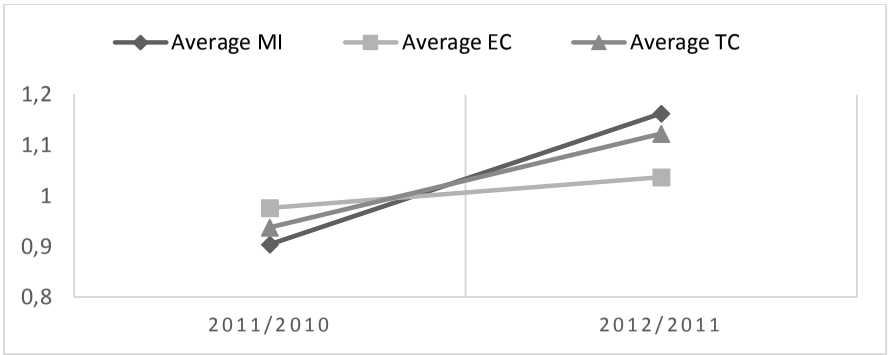
Source: own calculations

Figure 3. "Frontier shift" for each unit – i.e. technology change component (TC)



Source: own calculations

Figure 4. The average change of Malmquist Index and its components over analyzed period of time



Source: own calculations

Figures 2 and 3 show values of decomposed indices of the index for each unit. Every unit's Malmquist index can be greater than 1 even if one of its component decreases over time as long as the other component balances the difference. The average change of Malmquist index over the years 2010-2012 is 1,034, what means 3,4% overall increase of productivity in the group under evaluation, thus companies indicate an improvement of input-output ratio.

Figure 4 presents the average value of Malmquist index (MI) in 2011 (in reference to 2010) and 2012 (in reference to 2011) and its components at the time. Even though in 2011 there had been decrease of MI observed (its value was 0,904), the next year brought significant increase of the index (1,16 – i.e. average the index was 16% higher in comparison to 2011) therefore in general the input-output ratio in the sector has changed in positive way. The figure also shows the average change of decomposed indices and it suggests (as it was already mentioned above) that technology change term (TC) had greater impact on productivity changes in the sector in analyzed period of time than EC component.

SUMMARY

Couriers and messengers sector in Poland is relatively young, nevertheless its role in national economy is significant. The main player in this market is public company – Poczta Polska, however the package segment is dominated by private operators. Results of technical efficiency analysis conducted on 23 polish companies operating in the sector are presented hereof. The author intended to build a ranking of analyzed units in terms of productivity, therefore the DEA (Data Envelopment Analysis) method with super-efficiency extension was applied. The research allowed to identify the best performers in the analyzed period of time – the first two positions took consecutively *Polska Grupa Pocztowa S.A.* and *UPS Scs (Polska) Sp z o.o*. The third score received *UPS Polska Sp. z o.o.,* in the 2010 and 2011. In the last year (i.e. 2012) *World Courier (Poland) Sp. z o.o*. took the third position. The worst performers were: in 2010 - *Forposta S.A.*, in 2011 – *Siódemka S.A.*, and in 2012 - *Trade System Sp. z o.o*. The mean of scores in all analyzed period of time were over 60% and it suggests that couriers' companies rather reasonably utilize their resources. Malmquist index, which was also calculated for years 2010-2012, showed a 3,4% average increase in productivity of the group under evaluation and the main factor of these changes occurred to be the "frontier shift" component (technology change). In general the sector seems to indicate an improvement of the input-output ratio. The results of this study may be the basis for further analysis of the sector and of the specific DMU as well. The next step of the research can be also a recognition of best performers' strengths and worst performers' weaknesses, to indicate possible ways to enhance scores of efficiency in future.

REFFERENCES

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

Andersen P., Petersen N.C. (1993), A procedure for ranking efficient units in data envelopment analysis, Management Science, Vol. 39, pp. 1261-1264.

Chen Y., Motiwalla L., Khan M. R. (2004), Using super-efficiency DEA to evaluate financial performance of E-business initiative in the retail industry, International Journal of Information Technology & Decision Making, Vol. 3, no. 2, pp. 337−351.

Cooper W.W., Seiford L.M., Zhu J. (2004), Handbook on data envelopment analysis, Kluwer Academic Publishers, Boston.

Helta M. (2009), Zastosowanie metody DEA do opracowania rankingu efektywności spółek Agencji Nieruchomości Rolnych w 2006 roku, Roczniki Nauk Rolniczych, seria G, Vol. 96-3, pp. 107-111.

Kosieradzka A.(2000), Ocena i analiza produktywności w przedsiębiorstwie [w:] Strategie wzrostu produktywności firmy, Science conference materials (red. A. Stabryła), Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków, pp. 284–293.

Marcysiak A., Pieniak-Lendzion K., Lendzion M. (2013) Usługi kurierskie na rynku usług logistycznych w Polsce, Zeszyty Naukowe Uniwersytetu Przyrodniczo-Humanistycznego w Siedlcach, Seria: Administracja i Zarządzanie (23)2013, pp. 29-38.

Nazarko J., Komuda M., Kuźmicz K., Szubzda E., Urban J. (2008), Metoda DEA w badaniu efektywności instytucji sektora publicznego na przykładzie szkół wyższych, Badania operacyjne i decyzje 2008, Vol. 4, pp. 89-105.

Parlińska M., Bezat A. (2008), Efficiency of the polish wholesale markets . Validation based on the data envelopment analysis, Stowarzyszenie ekonomistów rolnictwa i agrobiznesu, Roczniki Naukowe tom X, Vol. 5, pp. 122-124.

Rangan N., Grabowski R., Y. Aly H., Pasturka C. (1988),The technical efficiency of US banks, Economy Letters, vol 28, Issue 2, pp. 169-185.

Ruggiero J. (1996), On the measurement of technical efficiency in the public sector, European Journal of Operational Research, Elsevier, Vol. 90, p.553–565.

Samuelson P., Nordhaus W. (1995): Ekonomia. T. 1. PWN, Warszawa, p. 185.

Thrall R.M. (1996), Duality, classification and slack in data envelopment analysis, Annals of Operational research, vol. 66, pp.109-138.

Yao C. (2003), Measuring super-efficiency in DEA in presence of infeasibility, European Journal of Operational research, vol. 161-2, pp.429–444.

Zhu J. (2001),Super-efficiency and DEA sensitivity analysis, European Journal of Operational Research, vol. 129, pp.443-455.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 349 – 358

THE INS AND OUTS OF UNEMPLOYMENT IN POLISH VOIVODESHIPS

Stanisław Jaworski
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
Department of the Prevention of Environmental Hazards and Allergology
Medical University of Warsaw
e-mail: stanislaw_jaworski@sggw

Abstract: This paper measures and compares the job-finding and separation transition rates of Polish voivodeships from 2005:Q1 to 2014:Q2. It uses readily accessible data of Local Data Bank in Poland: registered unemployment person by duration of unemployment and registered unemployment rate. The method of measurement stems from influential paper by Shimer on "Reassesing the Ins and Outs of Unemployment". The main sources of variability between the rates are investigated in terms of functional principal components.

Keywords: unemployment rate, Poisson process, smoothing, functional principal component data analysis

INTRODUCTION

Unemployment rate in Poland varied substantially across time but not across voivodeships over the last ten years. This phenomenon may occur as a result of variation in the rate at which workers flow the unemployment pool, variation in the rate at which workers exit the unemployment pool, or a combination of the two. It is largely these flows that drive the overall unemployment rate which is the main indicator of the health of the labour market. These flows generally follow a pattern in a typical business cycle. As the economy enters a downturn, separations start rising and job-finding rates start falling. These movements cause the overall unemployment rate to rise. The diagnosis of the labour markets should take into account the rates of reallocation of labour force. Reduced rates of worker turnover are usually associated with the more regulated labour markets. The reduced

turnover may contribute to higher unemployment. Lower turnover could reflect the fact that searching for jobs and workers has gotten costlier or harder, resulting in poorer worker-job matches, and therefore lower productivity. Low rates would indicate that firms and workers are settling for matches that are substantially suboptimal. What is optimal, of course, depends on the quality of the match and on the costs of searching.

To learn more about these flows in Polish voivodeships, I track the job-separation and job-finding rates. I show that the split job-finding to job-separation rate is closely related to the unofficial distinction to Poland A and B. Although the distinction is oversimplified, it is widely acknowledged and discussed in Poland. It refers to the historical, political, and cultural distinction between west and east areas of Poland, with the west (Poland A) being significantly more economically developed. Historically it can be traced to the period of the partitions of Poland: the Prussian partition (Poland A), compared to the Austrian and Russian partitions (Poland B). Poland B encloses for example Podlaskie, Mazowieckie, Lubuskie, Łódzkie, Małopolskie and Podkarpackie voivodeships. I show that these voivodeships display the lowest rates of reallocation of labour force. In contrast I show that the west voivodeships have the highest rate of worker turnover.

METHODOLOGY

I relate to the method developed by Shimer (2012). The model relates to continuous time environment in which data are available at discrete dates. For every $t \in [0, T]$ the interval $[t, t+1)$ will be referred to as 'period t'. It is assumed that all unemployed workers find a job according to Poisson process with arrival rate $f_t = -\log(1 - F_t)$ and all employed workers lose their job according to Poisson process with arrival rate $s_t = -\log(1 - S_t)$, where $F_t \in [0,1]$ is the job finding probability and $S_t \in [0,1]$ is the separation probability during period t. Two critical assumptions in the model should be underlined: (i) movements in and out of the labour force are ignored, (ii) arrival rates s_t and f_t are constant during period t. For evidence that the assumptions are good approximation of reality see Shimer (2012) and Elsby et al. (2013).

Let $\tau \in [0,1]$ and let $e_{t+\tau}$ denote the number of employed workers at time $t+\tau$, $u_{t+\tau}$ the number of unemployed workers at time $t+\tau$, and $u_t^s(\tau)$ workers who are employed at some time $t' \in [t, t+\tau]$. For $t \in \{0,1,\dots\}$ and $\tau \in [0,1)$, unemployment and short unemployment evolve according to

$$\begin{aligned} \frac{d}{d\tau} u_{t+\tau} &= e_{t+\tau} - u_{t+\tau} f_t \\ \frac{d}{d\tau} u_t^s(\tau) &= e_{t+\tau} s_t - u_t^s(\tau) f_t \end{aligned} \tag{1}$$

Note that $u_t^s(0) = 0$ and, under assumption (i), $e_{t+\tau} + u_{t+\tau} = l_t$ holds for every $\tau \in [0,1)$, where l_t denotes labour force in period t. Thus equations (1) have the following solutions in discrete time t:

$$\begin{aligned} \exp\{-f_t\} &= \frac{u_{t+1}-u_t^s(1)}{u_t} \\ u_{t+1} &= \frac{(1-\exp\{-f_t-s_t\})s_t}{f_t+s_t} l_t + \exp\{-f_t-s_t\} u_t \end{aligned} \tag{2}$$

We can estimate f_t and s_t for $t \in \{0,1,\dots\}$ as we know u_t and $u_t^s(1)$ from available data. The second equation in (2) implies that if $u_{t+1} = u_t$ then the unemployment rate $\tilde{u}_t := \frac{u_t}{l_t} = \frac{s_t}{s_t+f_t}$ $\left(\frac{s_t}{s_t+f_t}\right.$ will be denoted as $\tilde{u}_t^\star$ and be called the flow steady-state unemployment rate$\left.\right)$. The equation can be reformulated into the form

$$\tilde{u}_{t+1} = \frac{l_t}{l_{t+1}}(\lambda_t \tilde{u}_t^\star + (1-\lambda_t)\,\tilde{u}_t)\text{ , where } \lambda_t = 1-\exp(-f_t-s_t) \tag{3}$$

In practice $\frac{l_t}{l_{t+1}} \simeq 1$ and $\lambda_t \simeq 0$ so $\tilde{u}_{t+1} \simeq \tilde{u}_t^\star$. In particular, log differentiation of $\tilde{u}_{t+1} \simeq \tilde{u}_t^\star$ yields

$$\frac{d\log \tilde{u}_{t+1}}{d\log(t)} \simeq (1-\tilde{u}_t)\left[\frac{d\log(s_t)}{d\log(t)} - \frac{d\log(f_t)}{d\log(t)}\right]. \tag{4}$$

The approximation provides a decomposition in which the contributions of the inflow and outflow rates are separable and may be compared on an equal footing with respect to their impact on the unemployment rate. To obtain a transparent view of the relative contributions of the inflow and outflow rates it is advisable to plot charts (see Elsby et al. (2013)):

$$\begin{aligned} &\{(t,\log(s_t)) : t \in \{0,1,\dots\}\}, \\ &\{(t,\log(f_t)) : t \in \{0,1,\dots\}\}. \end{aligned} \tag{5}$$

In this paper I plot the charts for each voivodeship. Next I smooth the obtained series and apply functional component analysis[1] to the smoothed data. The goal is to recover the dominant modes of variation of the smoothed series. For example I reveal that the main source of variability between the inflow rates and between the outflow rates does not change in time and relates to Poland A and B.

Functional principal component analysis for functions $y_i(t),\ i = 1,\dots,N$, is designed to find weight functions $\xi_i,\dots,\xi_M$, where each function $\xi_m, m \in \{1,\dots,M\}$, maximizes

[1] For a theory of functional principal analysis the book by Ramsay and Silverman (2005) is recommended

$$\frac{1}{N}\sum_i \left(\int_0^T \xi_m(s) y_i(s) ds\right)^2$$

subject to

$$\int_0^T \xi_k^2(s)\, ds = 1 \text{ and } \int_0^T \xi_m(s)\xi_k(s) ds = 0, \quad (\forall\, k < m).$$

The vector $f_m = (f_{1m}, \dots, f_{Nm})$ where $f_{im} = \int_0^T \xi_m(s) y_i(s) ds$ is called the m -th principal component. The percentage of variability of the first m components is expressed as

$$V_m = \frac{\sum_{i=1}^N \sum_{l=1}^m f_{jl}^2}{\sum_{i=1}^N \sum_{l=1}^M f_{jl}^2}. \tag{6}$$

The method found to be helpful in interpreting the components is to examine plots of the overall mean function $\hat{\mu}(t) = \frac{1}{N}\sum_i y_i(t)$ and the functions obtained by adding and subtracting a suitable component functions:

$$\hat{\mu} \pm C_k \xi_k, \ k \le m, \tag{7}$$

where $C_k^2 = \frac{1}{N}\sum_i f_{ik}^2$. The plots of the components along with the principal components scores plots

$$\{(f_{ik_1}, f_{ik_2}) : i = 1, \dots, N\}, \tag{8}$$

$k_1 < k_2 \le m$, give good insight into the differences between objects $y_i, i = 1, \dots, N$.

In this paper y_i represents smoothed $\log(f_t)$ or $\log(s_t)$ for the i-th voivodeship, $N = 16$ is the number of voivodeships and the interval $[0, T]$ is the range time of observations ($t \in [0, T]$, the interval corresponds to time from 2005:Q1 to 2014:Q2).

ANALYSIS

I use readily accessible data of Local Data Bank in Poland: quarterly registered unemployment by duration of unemployment, quarterly registered unemployment rate and gross domestic product GDP. The data covers the period since 2005 to the second quarter of 2014. Figure 1 shows the time series for the job finding rate f_t, separation rate s_t constructed according to equations (2) and unemployment rate, all of in logarithmic scale. Both rates exhibits seasonal variability of length 4. At each first quarter the rates are relatively high and at each fourth quarter are relatively low. The rates of worker turnover are thus the highest in each second quarter and the lowest at each fourth quarter. It corresponds to the gross domestic product, that is the highest turnover corresponds to the lowest GDP. The high turnover precedes the growth of GDP in next three quarters.

The curves in Figure 1 are almost parallel to each other. It means that we can expect the differences between the voivodeships hold over time. I seasonally adjust the time series, $\log(f_t)$ and $\log(s_t)$, by smoothing and apply functional principal analysis.

The first two components of $\log(f_t)$ explain 95% of variability ($V_1 = 0.90, V_2 = 0.05$; calculated according to the formula (6)). The first and the second component are plotted in Figure 3 and in Figure 4 accordingly. We can note from Figure 3 that the job-finding rate vary between the majority of voivodeships in the same pattern over the whole observational period. The difference between voivodeships is constant since 2008. The order of voivodeships with respect to the level job-finding rate holds from 2005. However, there are some of 5% of voivodeships which changed their ranking in 2009 (for example Świętokrzyskie and Podlaskie). The change is represented by the second functional component (see Figure 4}): the curves $\hat{\mu} \pm C_2\xi_2$ crossed in 2009. It can be noted from functional principal scores plot (Figure 5}) that the curve $\hat{\mu} + C_2\xi_2$ represents Podlaskie voivodeship and the curve $\hat{\mu} - C_2\xi_2$ Świetokrzyskie voivodeship.

The first two components of $\log(s_t)$ explain 99% of variability ($V_1 = 0.98, V_2 = 0.01$). The first and the second component are plotted in Figure 6 and in Figure 7 accordingly. The first component imply the long-term upward and downward movement of job-separation rates go simultaneously across voivodeships. The 1% of variability (see Figure 4 and compare with Figure 8) relates to the period before the end of 2007 when the job-separation rate of Podkarpackie became lower than the job-separation rate of Dolnośląskie voivodeship.

The majority (over 95%) of diversity of both flows are included in first components. Figure 9 depicts these components. The range of the component for $\log(f_t)$\$ is half the range of the component for $\log(s_t)$, what can be compared with Figure 1. Thus we can expect that the separation rate causes the gross of diversity in unemployment rate across voivodeships. Plotting the components in the same plot gives us possibility to find voivodeships with low and high dynamics of reallocation of labour force. The first group of the low dynamics consists of Mazowieckie, Lubelskie, Podlaskie, Małopolskie and Łódzkie voivodeships. The group of the high dynamics consists of Lubuskie, Zachodniopomorskie, Opolskie and Dolnośląskkie voivodeships.

SUMMARY

This paper measures and compares the job-finding and separation transition rates of Polish voivodeships from 2005:Q1 to 2014:Q2. It uses readily accessible data of Local Data Bank in Poland: registered unemployment person by duration of unemployment and registered unemployment rate. The following conclusions are drawn:

1. The job-finding rate vary between the majority of voivodeships in the same pattern over the whole observational period. The difference between voivodeships is constant since 2008. The order of voivodeships with respect to the level job-finding rate holds from 2005. However, there are some of 5 \% of voivodeships which changed their ranking in 2009 (for example Świętokrzyskie and Podlaskie).
2. The long-term upward and downward movement of job-separation rates go in majority simultaneously across voivodeships. Only the job-separation rate of Podkarpackie became significantly lower then the job-separation rate of Dolnośląskie voivodeship. The separation rate causes the gross of diversity in unemployment rate across voivodeships.
3. The split job-finding to job-separation rate is closely related to the unofficial distinction to Poland A and B. We can separate two groups with respect to the dynamics of reallocation of labour force. The first group of the low dynamics consists of Mazowieckie, Lubelskie, Podlaskie, Małopolskie and Łódzkie voivodeships (Poland A). The group of the high dynamics consists of Lubuskie, Zachodniopomorskie, Opolskie and Dolnośląskkie voivodeships (Poland B).

The difference between voivodeships with respect to unemployment rate is strongly influenced by the inflow to unemployment. However, the outflow of unemployment has to be taken into account. What is most important nothing heralds changes in the flows across Poland A and B. Note that Poland A characterise of low job-finding rate and low separation rate. It is typical to so called sclerotic labour markets, where the unemployment rise more in a downturn and stay persistently high. Thus we should expect that unemployment differentiate Poland A from Poland B in a downturn. But it is not the case of Polish voivodeships. The remark is of paramount importance. The research of Polish labour market, relating to differences with respect to sex, age, social groups and similar, should take into account the type of part of Poland (A and B) as strongly influencing factor.

REFERENCES

Besse P.; Ramsay J.O. (1986) Principal components analysis of sampled functions. Psychometrika, 51, 285-311.

Elsby M.W., Hobijn B., .Sahin A. (2013) Unemployment Dynamics in the OECD, Review of Economics and Statistics (2013), MIT Press, vol. 95(2), pages 530-548.

Ramsay J.O., Silverman B. W. (2005), Functional Data Analysis. Second Edition, Springer. NY.

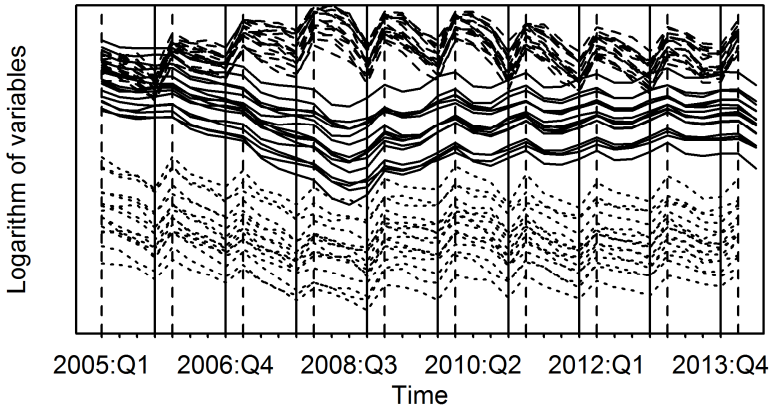
Shimer R. (2012) Reassesing the Ins and Outs of Unemployment Review of Economic Dynamics, Elsevier for the Society for Economic Dynamics, vol. 15(2), 127-148.

FIGURES

In some Figures the following abbreviations are used:

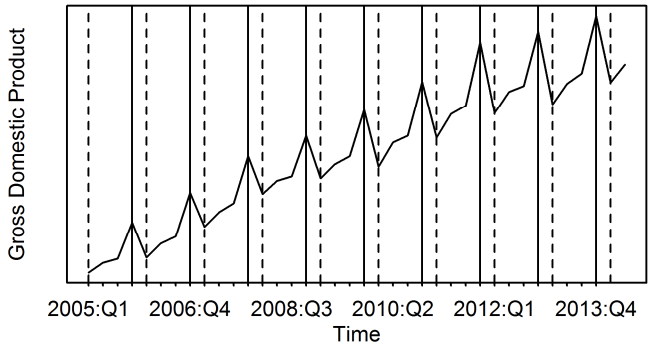
DOLNOŚLĄSKIE *DOLN*	KUJAWSKO-POMORSKIE *KUJA*
LUBELSKIE *LUBE*	LUBUSKIE *LUBU*
ŁÓDZKIE *ŁÓDZ*	MAŁOPOLSKIE *MAŁO*
MAZOWIECKIE *MAZO*	OPOLSKIE *OPOL*
PODKARPACKIE *PODK*	PODLASKIE *PODL*
POMORSKIE *POMO*	*ŚLĄSKIE* *ŚLĄS*
WIELKOPOLSKIE *WIEL*	ZACHODNIOPOMORSKIE *ZACH*
ŚWIĘTOKRZYSKIE *ŚWIĘ*	WARMIŃSKO-MAZURSKIE *WARM*

Figure 1. Dashed lines: job-finding rates in voivodeships; dotted lines -- separation rates; solid lines -- unemployment rates in voivodeships. Solid vertical lines -- fourth quarters. Dashed vertical lines -- first quarters.



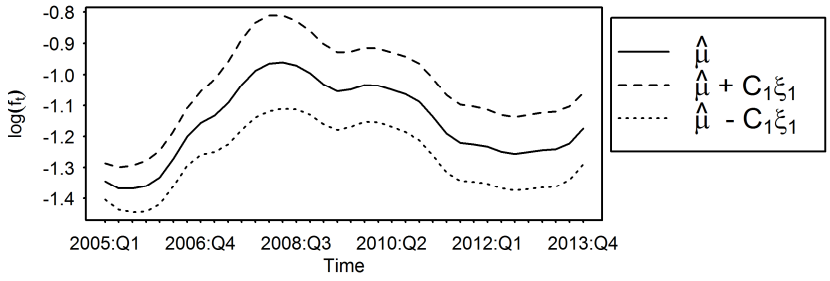
Source: own preparation

Figure 2. Gross Domestic Product. Vertical solid lines -- fourth quarters. Vertical dashed lines -- first quarters.



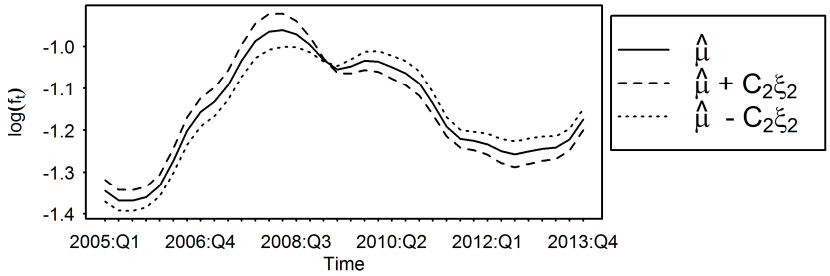
Source: own preparation

Figure 3. First component function of $\log(f_t)$.



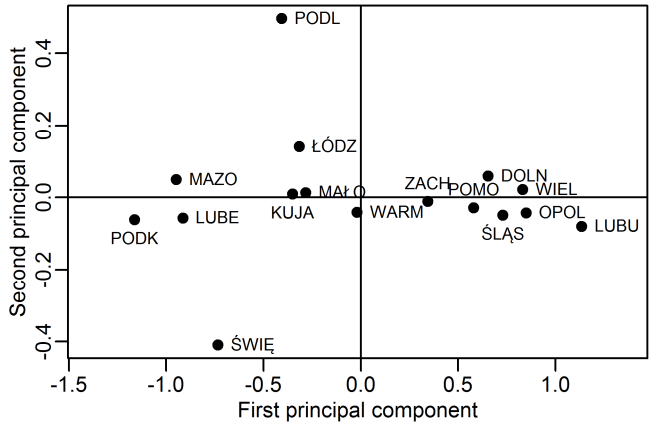
Source: own preparation

Figure 4. Second component function of $\log(f_t)$.



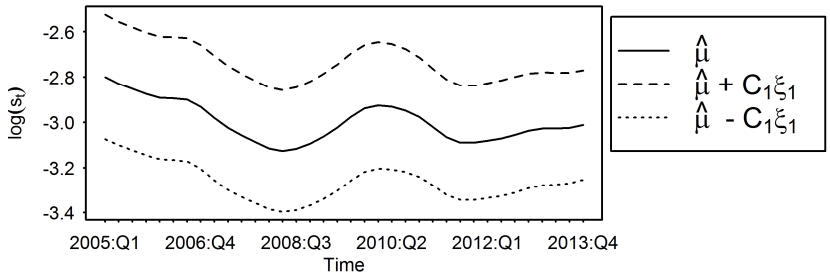
Source: own preparation

Figure 5. Principal components plot of $\log(f_t)$.



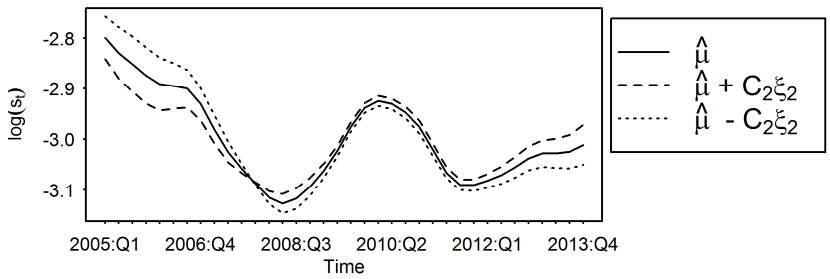
Source: own preparation

Figure 6. First component function of $\log(s_t)$.



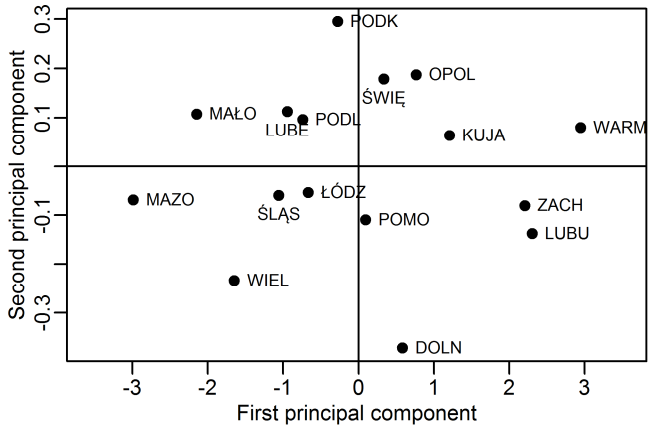
Source: own preparation

Figure 7. Second component function of $\log(s_t)$.



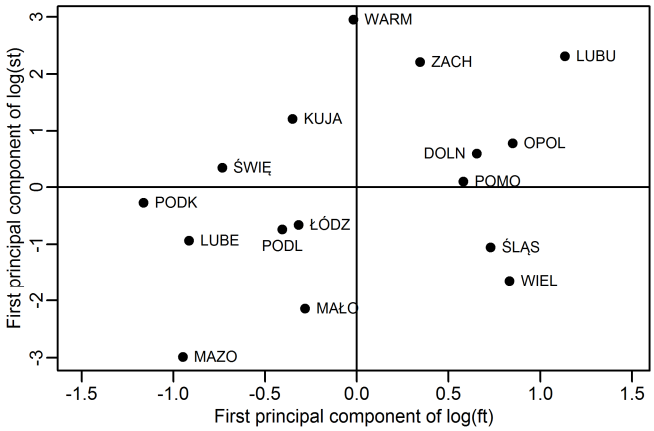
Source: own preparation

Figure 8. Principal components plot of $\log(s_t)$.



Source: own preparation

Figure 9. Plot of first principal components.



Source: own preparation

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 359 – 368

THE APPLICATION OF SCORING METHOD IN ESTIMATING RISK IMPLEMENTATION OF ERP SYSTEM

Rafik Nafkha
Department of Informatics
Warsaw University of Life Sciences – SGGW
e-mail: rafik_nafkha@sggw.pl

Abstract: Good implementation of an ERP (Enterprise Resource Planning) system differs from unsuccessful one before everything the ability of both foreseeing and reducing the amount and the size of numerous traps which appear at every stage of the life cycle project implementation. In this article, based on the results of questionnaire sent to 50 companies with different employment size, events affecting the failures of the ERP system implementation were identified and on their base a level of risk as well as additional costs related to preventive actions (reducing the probability or effects of the problem occurrence) were investigated. To evaluate the risk values of chosen ERP system implementation tasks, a spot risk assessment method was used.

Keywords: ERP implementation system, risk value, risk assessment methods, risk register, scoring risk method

INTRODUCTION

Information systems suppliers in particular ERP (Enterprise Resource Planning) systems avoid clearly in their presentations risk analysis of the system mainly for two reasons: the first is the lack of or limited knowledge regarding the risks in individual sectors of the economy, the other for reasons of sales and marketing. Risk has always aroused panic among both customers and retailers offering the system. Disclosure of threats by the supplier in the first steps of the sale may be subject to conflicts of interest. One general principle that is in force is that the risk in the first stages of the project is a forbidden word. Unfortunately, in the next stages of the system implementation, it becomes the not needed word and for its analysis is too late, it remains only mitigate the impact of rising incurred costs. This article presents, a sample list of risks for typical ERP systems

implementation and the risk assessment calculation method, that can be useful running own risk implementation analysis, especially for medium and small enterprises (MSP).

A special case of projects are the ERP implementation projects, which are subject to adjustment previously produced software to the specific conditions of the company in order to achieve certain benefits. Risks in these projects arise on each stage of the implementation of the ERP system [Khvalev E.A, 2009]. According to report (Business Software Report, Management Institute of Warsaw, 2001) and analyzing the implementation management system suppliers methodology [Nafkha 2014], the implementation of a ready system is usually implemented in five phases:

- preparation of the organization for change – work out a project organization and the rules for its implementation,
- determination the business concept - elaborate a list of business processes that will be implemented by the system,
- implementation - development of a prototype solution ,
- preparation for work in the target environment - installation, launch (test integration of the prototype, user training, data transfer , preparation of the working environment) and transmission system operation,
- start and supervise the work of the system in the real environment.

In the following article, based on the results of surveys sent to 50 companies with different structures of employment, we identified events affecting the information management system implementation failures that occur at every level in the life cycle implementation project. To identify the implementation project key risk factors, we asked both customers and experts in the field of ERP systems implementation to indicate repeated and common in their opinion, implementation failures factors. Participants in the study indicated more than 42 different problems occurring during the implementation of the ERP system. In this study only 25 of them have been identified as having a negative impact on the time, budget of the project and product conformity with the project objectives. Table 1 shows the risk critical factors ranked by the number of reported problem.

Table 1.Types and quantities of identified problems

Id	Critical risk factors	Number of reported problem
1	Lack of Top Management commitment and support	20
2	Poor project management team	19
3	Lack of Departmental cooperation	19
4	Unclear goals and objectives	18
5	Incorrect project management	18
6	Ineffective communications	17
7	Improper management of expectations	17

Id	Critical risk factors	Number of reported problem
8	Incompetent project leader	16
9	Lack of vendor or supplier support	16
10	Improper change management, risk and scope of the project	15
11	lack of knowledge of their own business processes	15
12	incorrect system selection	12
13	Analysis and data conversion	12
14	Limitation in resources	12
15	Insufficient training of end-users	10
16	Lack of new business processes familiarity	10
17	Non-acceptance of organizational structure change and business processes	10
18	Poor integration of the infrastructure systems	9
19	Poor conflict management	9
20	Using tools supplier	8
21	Ineffective project cost and time management	6
22	Lack of metrics for evaluating project efficiency and benefits	6
23	Lack of competence of ERP's consultants	5
24	Data losses	2
25	Insufficient testing phase	2

Source: own study based on survey result

PROJECT METHODOLOGIES

Risk assessment according to scoring method for risk assessment consist in the identification of risky tasks that can lead to implementation failure and give them a quantitative measure of the risk level according to an adopted scale. Tasks classified as risky are grouped into specified categories. The grouping should be made by the risk manager (executing the obligations described in the Policy and Risk Management Plan). Survey participants filling the questionnaire do not need to be familiar with risk management, it is sufficient that they present significant implementation threats in their opinion, and then the grouping and the formalization of the risk list is made by an expert in this area. In this article, the following risk categorization has been provided:

1. Organizational - subcategories include (top management, business processes, strategy, employment policy, company culture, process planning, finance, staff).

2. Project - subcategories include (project management methods, quality and implementation team, business development, project integration).
3. Technical and technological - subcategories include (system functionality, support, critical IT infrastructure,).
4. External - subcategories include (legislation, the economic situation, exchange rate, competition, lobbying …).

The most common risk categories and their weight listed on a scale of 1 to 5 are shown in Table 2

Table 2. Risk categories

Category	Weight	Sign
Technical and technological	3	T
Organizational	4	O
External	5	E
Project	4	P

Source: own study and assumptions

To carry out a comparative analysis, each problem has been prescribed a certain value on a scale of 1 (least important) to 5 (the biggest problem). The final value of each problem is the sum of all values fulfilled by various participants in the interview. Since the determination of the probability is done intuitively, the intuitive probability scheme is defined as follows:

Table 3. Scoring risk value

Range	Risk value (scoring)	Probability	Designation
1 - 4	1	0,1	very low
5 - 8	2	0,3	low
9 - 12	3	0,5	medium
13 - 16	4	0,7	high
17 - 20	5	0,9	very high

Source: own study an assumptions

Please note that there is no verifiable method that precisely determine the risk probability. Therefore, it was decided only to determine the range within which the probability is located. In table 2 the allocation for each risk category weight values was entirely subjectively. In one publication [Frączkowski, 2003] weight constitute a general validity assessment of evaluated risk regarding the entire project. According to this way of thinking weights can be implemented for revaluation (normalization) the risks in each category and then the whole project. To calculate risk values for each category before and after preventive actions (in order to minimize the risk value), in his book, Frączkowski uses the following formulas:

Non-standardized risk value (before preventative action):

$$R_x = \frac{\sum_v R_v}{n} \tag{1}$$

Standardized risk value (after preventative action):

$$R_{zn\_x} = \frac{R_x W_x}{W_{avr}} \tag{2}$$

Non-standardized risk total value:

$$R_{total} = \frac{\sum_x R_x}{k} \tag{3}$$

Standardized risk total value:

$$R_{zm\_total} = \frac{\sum_x R_{zm\_x}}{k} \tag{4}$$

where:
n - number of tasks belonging to given category,
R_v – task risk value belonging to given category,
K - number of categories,
W_x - category risk weight,
R_{avr} - average weight value calculated from the formula

$$R_{avr} = \frac{\sum_x W_x}{k}$$

After summing up indicated scores and assigning ratings to each risk factor, it is necessary to evaluate its effects in order to apply any simplest strategies for its elimination or restriction by adding an appropriate cost estimation (to handle emerging problems) to estimated schedule "margin of safety".

APPLICATION OF SCORING RISK ASSESSMENT METHOD ON THE EXAMPLE OF SAP SPRINT SYSTEM IMPLEMENTATION

The scope and cost of the proposed example is specified using SAP Business All-in-One the Configurator (http://www.sap.com/solution/sme/software/erp/all-in-one/buy/rds.html), enabling the calculation of the predicted and the estimated SAP Business All-in-One rapid deployment solution price including hardware, software and system implementation (without software maintenance cost).

SAP Business All-in-One is a complex, integrated ERP solution, prepared by SAP partners for medium-sized companies. Implementation scope for a typical enterprise SMEs (Small and medium-sized enterprises) adopted in this example includes the following areas:

- activities related to logistics process in terms of sales, distribution and invoicing including , offer to the customer, customer contract, customer order, sales, refunds and claims adjustment.

- Activities related to process of ensuring supply including: warehouse management, purchase offer, a supply contract, batch management, stock transfer, inventory and purchase settlement.
- Financial Accounting and Management which includes: general ledger, accounts receivable and suppliers, liquidity management, accounting and reporting of fixed assets for finance.

The following assumptions and cost estimation are adopted: total number of employees 100, number of users 20, licenses cost 214 200 PLN, services 300 000 PLN, total solution cost 584 200 PLN. After working out an implementation timetable, which established the duration and the resources assigned to the project tasks, the next step focus on tasks identification that are risky during their implementation and then assign each of them to adopted in table 3, range of risk allocation. According to ASAP (Accelerated SAP methodology for implementation) [Nafkha, 2014], the road map is composed of five well-known consecutive phases and includes within the following tasks:

- Project Preparation **(P)** – some important milestones that need to be accomplished for this stage include: identifying clear project objectives, architect an efficient decision-making process, creating an environment suitable for change and re-engineering, building a qualified and capable project team.
- Business Blueprint **(B)** - to help extract pertinent information about the company that is necessary for implementation. Information like future business processes and business requirements are essential in this stage.
- Realization **(R)** - With the completion of the business in phase 2, "functional" experts are now ready to begin configuring SAP. The Realization phase is broken into two parts: Baseline configuration based on the information provided in the blueprint document and Fine-tuning the system to meet all business and process requirements necessary to fit company special needs. Fine tuning usually deals with the exceptions that are not covered in baseline configuration
- Final Preparation **(F)** - this stage also concentrates on the fine tuning of the configuration before Go-live and more importantly, the migration of data from old system or systems to SAP. Workload testing (including peak volume, daily load, and other forms of stress testing), and integration or functional testing are conducted to ensure the accuracy of data and the stability of SAP system
- Go-Live and support (L) - The Go-live milestone is itself is easy to achieve.

Examples of risky tasks for the adopted implementation are shown in Table 4.

Table 4 Types and risk values of selected tasks in different project implementation phases

Id	Tasks	Cat	Critical risk factor	Risk Value	Project phase				
					P	B	R	F	L
1	Strategic Analysis	O	Unclear goals and objectives, lack of Top Management commitment and support.	5	X	X	X	X	X
2	Preliminary project plan	O	Unclear goals and objectives, lack of Top Management commitment and support.	5	X				
3	Pre-implementation analysis including modeling	P	Incompetent project leader, lack of Departmental cooperation, lack of knowledge of their own business processes.	4		X			
4	License purchase	Z	Delay of license delivery.	1			X		
5	Needed shopping and infrastructure preparation	Z	Lack of co-operation with supplier, delay of devices delivery.	3			X		
6	Installation and Technical Configuration	T	Poor integration of the infrastructure systems.	3			X		
7	Installation and functional configuration – Logistics	T	Lack of dedicated resources, module is not in time.	3			X	X	
8	Installation and functional configuration - Materials Management	T	Lack of dedicated resources, module is not in time.	3			X	X	
9	Installation and functional configuration- Financial Accounting and Management	T	Lack of dedicated resources, module is not in time.	3			X	X	
10	Business processes modifying according to accepted company needs.	P	Non-acceptance of organizational structure change and business processes, lack of new business processes familiarity.	3		X	X	X	
11	Administrators training	T	Delay in training, lack of competent trainers.	1				X	
12	Training users with regard to purchased modules	T	Insufficient training of end-users.	3				X	
13	Data migration	T	Lack of prepared appropriate forms, lack of supplier tools for data conversion , loss of data.	2				X	
14	Data input	T	System not ready yet, lack of data prepared.	2				X	
15	System performance testing	T	System not ready yet, ineffective project time management, lack of metrics for evaluating project efficiency and benefits.	1				X	
16	system testing in terms of system functionality including interfaces	T	System not ready yet, lack of all functionality system testing.	1				X	
17	Technical support during system startup	T	Lack of supplier support , lack of competence of ERP's project team fine-tunes.	4					X

Source: own study and calculation

The risk analysis purpose is to determine the quantitative value and identified risks impact on the project implementation. Data are collected in a risk register updated with score risk value and measurable financial and non-financial consequences reducing identified risks. Table 5 shows risks and costs of introducing preventive actions of individual project tasks.

Table 5. Risk and costs of introducing preventive actions

Id	Preventive actions	Cost	Risk value	Effects after introducing preventive actions
1	• All stakeholders identification • Kick-off.	- 5 000	3	Defined the purpose and scope of the project. Top Management commitment and support.
2	Internal training, coaching	1 000	3	Clear goal and Project purpose. Top Management commitment and support.
3	• High power decision for PM • External consultants support • Process modeling training	- 3 000 2 000	3	A competent project manager. Knowledge of business process modeling techniques.
4	Collecting supplier references	-	1	License in time.
5	• Collecting supplier references • Early orders and transfers in time.	- -	3	Devices in time.
6	• Additional technologies testing before implementation work.	-	3	Familiarity with new technologies and easier solution selection.
7	• Determine the necessary time dedicated for project implementation • Provide separate room.	- -	3	A dedicated project team. Module provided in time.
8	• Determine the necessary time dedicated for project implementation • Provide separate room	- -	3	A dedicated project team. Module provided in time.
9	• Determine the necessary time dedicated for project implementation • Provide separate room	- -	3	A dedicated project team. Module provided in time.
10	• Collecting supplier references • Client management support and co-operation with supplier .	- -	3	Business processes adopted to company needed.
11	References and trainers certificates		1	Training on time and meets the user’s needs.

Id	Preventive actions	Cost	Risk value	Effects after introducing preventive actions
12	• References and trainers certificates • Additional targeted training	- 1 000	3	Training on time and meets the user's needs.
13	• Collecting supplier references • Request data migration methodology	- -	2	Data migration in time.
14	Determine the necessary time dedicated to preparation and data input	-	2	Data input in time.
15	Making simulation tests	-	1	Right performance testing in time.
16	Functionality testing in particular phases of the project	-	1	Functionality testing in time.
17	• Collecting supplier references • Request additional consultancy or support	- 2 000	3	Contract technical support.

Source: own study and calculation

A summary of measured key risk indicators shows table 6. The proposed preventive actions reduce the non-standardized risk - in case of acceptance a risk reduction cost at level 6000 PLN. (which represent only 2% of the whole implementation budget). In this case the risk probability is reduced to one level less from 3.22 to 2.97. The standardized risk value after preventive actions (formula 4) is reduced from 3.19 to 2.94. In case of acceptance risk reduction cost 14 000 PLN which increases the implementation budget by 4.67%, then the risk probability is reduced from 3.22 to 2.72 and for standardized risk value after preventive actions from 3.19 to 2.69.

Tabel 6. Estimated risk indicators for the implemented case

Nr	Risk indicators	Cost = 6 000 PLN	Cost = 14 000 PLN
1	Non-standardized project risk	3,22	3,22
2	Standardized project risk value	3,19	3,19
3	Risk value after preventive actions	2,97	2,72
4	Standardized risk value after preventive actions	2,94	2,69

Source: own study and calculation

According to risk management theory [Korczowski, 2009], the risk owner decides how to deal with risk. If the threats reduction cost does not exceed 5% total implementation budget, which represent in most cases an acceptable risk level, the project is realized without any corrections. In the adopted example, additional cost of 14 000 PLN. reduces the ERP system implementation failure probability by 2

levels and increases the system implementation budget by approximately 4.67% of total implementation budget. In another case, decision of selecting an appropriate scenario to deal with the risk value belongs to the steering committee.

SUMMARY

Successful management systems implementation is dependent on many factors related both to the type of activities carried out by company, and the way of managing project in particular the selection and use of risk management methods. There are no verifiable methods that will accurately determine system implementation failure or success probability, but we can determine the interval in which the success/failure probability of each task is located. Risk assessment according to scoring method for risk assessment consist in the identification of risky tasks that can lead to implementation failure. For such risky tasks, a set of key factors that have directly influence on project failure are selected. Next for each problem a range of risk value was adopted. Since the determination of the implementation risk probability takes place intuitively, based on the experiences and statements of specialists risk values are determined on the basis of indicated key factors number related to given implementation tasks execution. Subsequently, and after assigning each task to one of set categories (organizational, project, technical or technological, external), risk indicators before and after preventive actions were calculated.

REFERENCES

Beynon-Davies, P. (1999), Human Error and Information Systems Failure: The Case of The London Ambulance Service Computer-aided Despatch System Project, Interacting with Computers, vol. 11, no. 6, pp. 699-720.

Frączkowski K, (2003), Zarządzanie projektem informatycznym, Projekty w środowisku wirtualnym czynniki sukcesu i niepowodzeń projektów, Politechnika Wrocławska.

Khvalev E.A, (2010) Key characteristics in ERP implementation Identification and Analysis by Project Phases: Conceptual model for analysis - Proceedings of the 4th Conference on Theory and Practice of Modern Science, Moscow, Russia.

Korczowski A, (2009) Zarządzanie ryzykiem w projektach informatycznych. Teoria i praktyka, Hellion.

Lech, P. (2003). Zintegrowane systemy Zarządzania ERP/ERP II. Wykorzystanie w biznesie, wdrażania. Difin, Warszawa.

Lyytinen, K. (1988), Expectation Failure Concept and Systems Analysts' View of Information System Failures: Results of an Exploratory Study. Information and Management 14(1): 45-56.

Nafkha, R (2014) Analiza wybranych metod zarządzania projektem informatycznym we wsparciu procesów biznesowych i organizacyjnych firmy, IX Scientific Conference, Internet in the information socjety.

Sorupka D, Kuchta D, Górski M,(2012) Zarządzanie ryzykiem w projekcie, WSOWL, Wrocław

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 369–381

DEFAULT PREDICTION FOR SME USING DISCRIMINANT AND SURVIVAL MODELS, EVIDENCE FROM POLISH MARKET

Aneta Ptak-Chmielewska
Institute of Statistics and Demography, Warsaw School of Economics
e-mail: aptak@sgh.waw.pl
Anna Matuszyk
Institute of Finance, Warsaw School of Economics
e-mail: anna.matuszyk@sgh.waw.pl

Abstract: The aim of this paper was to compare the new technique (survival analysis) used in the credit risk models with the traditional one (discriminant analysis), analyse the strengths and weaknesses of both methods and their usage in practice. This study attempts to use macroeconomic data to build models and examine its impact to the prediction. For this purpose, a number of models was built on the basis of the sample of 1547 enterprises including 494 defaults. The time range covered by sample was 2002-2012.

Keywords: survival analysis, discriminant analysis, macro variables, rating model

INTRODUCTION

Credit risk is the most important type of risk to which banks are exposed. This is due to their role as financial intermediaries. Also different factors have an impact on the credit risk level in banks, including the credit conditions and creditworthiness methods. The main factor limiting the risk of the credit portfolio is good economic and financial situation of the clients, in this case - company. Therefore, a special attention to the proper assessment of customers' creditworthiness should be paid and then the subsequent monitoring of their financial situation should be carried out.

The wide range and increasing availability of the credit in modern societies have led to the inordinate indebtedness of many borrowers (Allen and Rose, 2006). Since the problem of insolvency is getting bigger and bigger, the interest in the effective management of customer debt repayment is also growing.

The recent developments in the financial literature demonstrate the various questions posed in front of the financial analysts who build models based on survival analysis. Over the past 25 years, this method was implemented and successfully used in many fields of finance. In 1989, Altman suggested measuring the expected mortality of bonds and, consequently, loss rates in a similar way to that which actuaries use to evaluate human mortality. In 2000, Altman and Suggitt (2000) applied this analysis to assess the risk of corporate loans. In 1998, Lando estimated the bond default time, using a proportional hazard model for survival analysis, and applied macroeconomic variables as predictors (Lando 1998). The same approach was applied to modelling credit risk in the valuation of bonds and other financial instruments by Pierides (1997).

However it is recognized that the idea of the application of survival analysis in the credit risk models (credit scoring) was first used by Narain in 1992, and further developed by Carling et al. (1998), Stepanova and Thomas (2002), Allen and Rose (2006), Malik and Thomas (2006). In all these papers the parametric or semiparametric regression techniques for modelling the time to default (duration models) were used.

An interesting research was done by Nunes at al. (2014)[1] where authors using probit regressions and on the basis of two research samples: 1589 family-owned SMEs and 485 non family-owned SMEs checked whether there are significant differences between family-owned SMEs and non-family-owned SMEs for determinants of survival. The results obtained show the existence of significant differences between these two types of companies for the determinants of survival. In the context of family-owned SMEs, authors think that: size, age and R&D expenditure are neither positive nor restrictive determinants of survival; cash flow and labour productivity are positive determinants of survival; and, debt, interest paid and risk are restrictive determinants of survival. In case of the second group of SMEs, size, age, cash flow, debt and R&D expenditure are positive determinants of survival, with interest paid, risk and labour productivity being neither positive nor restrictive determinants of survival.

In the paper written by Glennon and Nigro (2005)[2], it was presented that the default behaviour of the analysed loans is time sensitive. The likelihood of default is pretty high at the beginning, then peaks in the second year, and declines thereafter. Authors used a discrete-time hazard model and from the received results showed that the likelihood of default is conditional on customer, lender, loan characteristics, and changes in economy.

[1] P. M. Nunes, Z. Serrasqueiro, J.V. da Silva: "Family-owned and non family-owned SMEs: empirical evidence of survival determinants ",Economics and Business Letters, 3(1), pp.68-76, 2014

[2] D.C.Glennon, P.Nigro: „Measuring the Default Risk of Small Business Loans: A Survival Analysis Approach", Journal of Money, Credit, and Banking, Volume 37, Number 5, October 2005, pp. 923-947.

Giovannetti et al. (2013) wanted to check the thesis that the firms' survival is often seen as crucial for economic growth and competitiveness[3]. In their research they considered business demography of Italian firms, using an original database. They considered consider the size effect, technology, trade, FDIs and innovation on companies' survival probability. The results obtained suggest that size and technological level positively affect the likelihood of firms' survival. It was also interesting that the internationalized firms showed higher failure risk: on average competition is stronger in international markets, forcing firms to be more efficient. However, large internationalized firms were more likely to 'survive'.

Different approach was proposed by Moon and Sohn (2011)[4]. According to authors, the scorecards are often filled-in based on the evaluator's total perception rather than the individuals' scores of which the scorecards are built. Authors proposed a survival model that considers the time to default as well as the total perception scoring phenomenon. Their approach can be used during the decision-making process in various areas of technology, (for example in R&D), alliances, transfers, and loans.

Papers regarding such types of models in Poland have much shorter history. The implementation of the western models to the market of enterprises which function in the transition economy such as Poland failed. It appeared that those models are not successful in conditions of the political and economic changes. Unsatisfactory effects of using foreign models in Polish conditions contributed to developing research into domestic models. The most popular became the models based on discriminatory analysis as it is the case abroad.

We propose a model for companies' prediction based on the survival analysis. This model will be compared with the model using the traditional method, i.e. discriminant analysis.

The paper is structured as follows. Section 2 describes the approaches used to estimate the probability of default. In Section 3 the data set was described. Results of the models are presented in Section 4. Section 5 contains conclusions.

DISCRIMINANT ANALYSIS

Discriminant analysis is used to determine which variables can be used to identify two or more groups from the analyzed data set. It allows identifying these

[3] G. Giovannetti, G. Ricchiuti, M. Velucchi: Size, innovation and internationalization: a survival analysis of Italian firms", Applied Economics,Volume 43, Issue 12, 2011, pp. 1511-1520.

[4] T.H.Moon, S.Y.Sohn: "Survival analysis for technology credit scoring adjusting total perception", Journal of the Operational Research Society (2011) 62, pp.1159–1168.

variables which allows classification of different groups with higher accuracy than the random ones.

The purpose of discriminant analysis is a correct classification of observations into two subspaces defined as groups. Discriminant function is defined as maximization of the distance between subpopulations (groups). In discriminant analysis the classification of units as defaults or non-defaults is based on minimum two explanatory variables. Simultaneously the analysis is carried out taking into account all selected ratios. It is crucial to find out dependencies between variables enabling a correct distinction of entities. In discriminant analysis the dependent variable is qualitative (binary). The classification of entities is based on linear discriminant function. Synthetic ratio arose as a result of applying the model (value of the function) makes it possible to classify the entity. However the discriminant analysis are limited to certain extent. It is possible to apply it when the analysed ratios are normally distributed. It is also necessary to meet the assumption of their independence and completeness. The lack of fulfillment of assumptions influence negatively the classification capacities of the model. Checking whether the assumptions about ratios were fulfilled can be verified by applying relevant tests and statistical procedures.

Discriminant functions, which are used to build multivariate warning bankruptcy models, take different forms – linear, square, etc.

Linear discriminant function takes the following form (Ptak-Chmielewska, Pęczkowski 2009):

$$Z = a_0 + a_1X_1 + a_2X_2 + \ldots + a_nX_n, \tag{1}$$

where:
Z – target (dependent variable),
a_0 – constant,
a_1 – a_n – parameter estimates (weights),
$X_1, X_2, \ldots, X_n$ – explanatory variables (financial ratios).

The presented discriminant function is also known as a Fisher discriminant function. Parameters a_i, are called discriminatory factors (weights). After determining the discriminant function the next step is to define the limit, allowing for classification of the individual being at a financial risk or not. Usually a mean value of the discriminant function is being determined for each group and then the cut-off value between the means. If the z value for the current company is lower than $Z_{cut\text{-}off}$ then the company is classified as being at bankruptcy risk otherwise is classified as a good company. The model's efficiency is assessed calculating the type I and type II error. The first one determines the percent of companies classified as being at bankruptcy risk, while the second category represents the percentage of companies classified as good ones.

Altman, who presented his model in 1968, is considered to be the precursor of multi-dimensional model. This model is a combination of ratio analysis and statistical method - multivariate discriminant analysis. The author analyzed 22 factors on a group of 66 companies (33 bankrupt and 33 non- bankrupt ones). In the subsequent stages of the analysis he discarded the weaker predictors and the final model included 5 indicators. In 1977, Altman and his team conducted further studies related to the companies' bankruptcy prediction. He analyzed 58 bankrupt and 58 healthy companies. The result is a model consisting of 7 variables without specified weights, and hence a discriminant function was not designated. Prediction of bankruptcy a year before the event reaches 90% and 5 years before the event amounts to 70%.

The next version of Altman's model was developed in 1983. E. I. Altman has made changes in the weights assigned to the variables of the first model. The value of the misclassification error was 6%. The next improvement of the Altman model concerned reducing the impact of economic and industry specificity to the Z index value. Altman models have been developed for companies operating in the U.S. market. Its application for companies operating in other conditions does not give very high discriminatory power.

The Z-Score Model selected for this paper application was based on the following revised model:

$$Z = 0.717 \cdot X_1 + 0.847 \cdot X_2 + 3.107 \cdot X_3 + 0.420 \cdot X_4 + 0.998 \cdot X_5$$

where:

X_1 = Working capital/Total assets

X_2 = Retained Earnings/Total assets

X_3 = Earnings before interest and taxes/Total assets

X_4 = Book value of equity/Book value of total debt

X_5 = Sales/Total assets

Z = Total Index

In the estimation data, the Z-Score Model classified correctly 90.9% of bankrupt firms and 97.0% of the non-bankrupt firms. In all, the data included 66 firms (33 + 33). Thus, the classification accuracy was only slightly less impressive than for the original model. Altman did not test the model on a secondary sample due to lack of a private firm data base. In concluding remarks, Altman (1983) regards the general applicability of his Z-Score Model as debatable. The model did not exclude very large and very small firms, the observation period was quite long, and the analysis included only manufacturing companies. Ideally, development of a bankruptcy predicting model should be done based on homogenous group of bankrupt companies and data as near to the present as possible. Altman advised the analysts interested in practical utilization of the Z-Score Model to be careful.

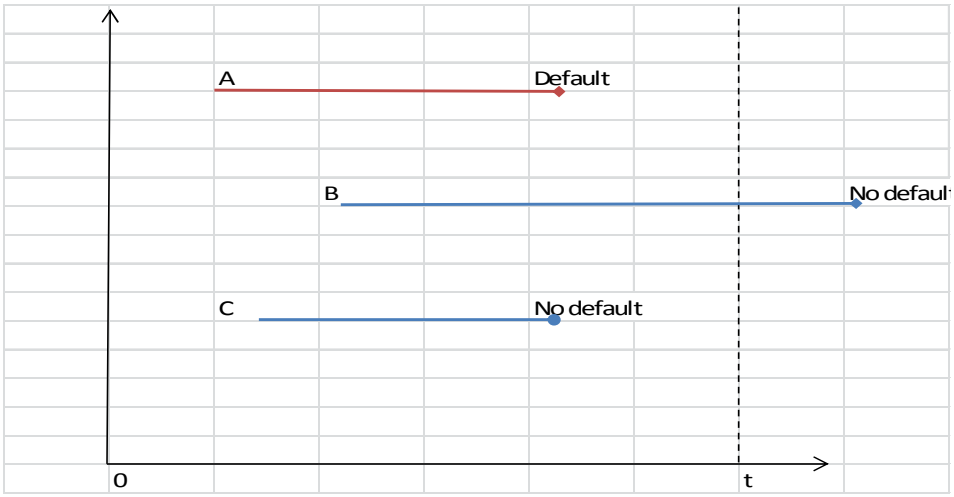
SURVIVAL ANALYSIS

Survival analysis is a collection of statistical procedures for data analysis, for which the analyzed variable is the time of the event. So it is a study in time, counted from the time when a case "came into" observation until the occurrence of the event.

The idea of using the survival analysis to assess credit risk, and more particularly to model PD (Probability of Default), is shown on Figure 1. It shows three cases that may occur in practice during the lifetime of a company.

The first customer (A) defaulted before the end of the credit. In this case, the lifetime of the customer (time to default), is observable during the analysed period. Customers: (B) and (C) present two different situations. In both of them it is not possible to observe the time of default, so the status of them is censored. In case of customer (B) it is only the time from the start of the loan to the end of the study, while in case of customer (C) presents a situation where the end of the loan occurred before default (i.e. early repayment).

Figure 1. Idea of the survival analysis



Source: own elaboration

Survival and hazard function

Two important functions in survival analysis are the survival function and the hazard function. The first function is a continuous function representing the probability that the 'failure time' T of an individual (company in this case) is greater than time t.

$$S(t|x) = P(T > t|X = x) = \int_t^{\infty} f(u|x)du, \tag{2}$$

The hazard function $h(t)$ represents the point in time default 'intensity' at time t conditional upon survival up to time t.

$$h(t|x) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t | T \ge t, X = x)}{\Delta t} = \frac{f(t|x)}{S(t|x)}, \quad (3)$$

There are many different models used in survival analysis. Models are differentiated according to assumptions about functional form of hazard rate and its variability in time. In practice the most frequently used model is proportional hazards Cox regression model. For this reason this model was presented in more details in this paper.
For Cox regression model the hazard function is given by formula:

$$h(t \mid x_1,...,x_k) = h_0(t)\exp(\alpha_1 x_1 + ... + \alpha_k x_k), \quad (4)$$

where: $h_0(t)$ - means base hazard, parametrically non-specified function of time and $x_1,x_2,\ldots x_k$ - means explanatory variables (including time dependent variables).

Cox proposed also the special type of estimation method called pseudo-likelihood (Cox, 1972). This method divides the likelihood function for proportional hazards model into two parts: first including only information about parameters and second, including information about parameters, and hazard function. Division into two components is justified because first depends only on sequence of events occurrence, does not depend on exact time of occurrence, and the second is 0 and is omitted.

Main advantage of Cox model (and other semi-parametric models) is assessment of many variables (including time dependent variables) influence on the process without necessity of base hazard $h_0(t)$ specification. The main disadvantage of Cox model is hazard proportionality assumption. This assumption impose that for each pair of individuals in any time the hazard rate is fixed. The relative hazard (ranking) for individuals is stable in time.

Despite this limitation of Cox model, it is especially attractive for researchers in case of (Blossfeld and Rohwer, 2002):

- unknown shape of hazard in time;
- no theoretical bases for parameterization;
- no possibility of functional shape of hazard specification;
- main interest is focused on explanatory variables influence on hazard.

Above mentioned advantages in using Cox regression model make this model useful in risk of enterprises' liquidation modeling. The only disadvantage of this model is proportionality assumption which implies fixed proportion of hazard for individuals during the observation time period. This problem may be solved by including additional time dependent variables in the model (like interaction between variable and the time). For checking the proportionality assumption the

easy way is to include the interaction with time, the significance of this parameters confirm that the proportionality assumption is violated. In this case model is named non-proportional hazards Cox regression model. Results of Cox model estimation are parameters describing the influence of explanatory variables on the probability of event occurrence and on the base hazard (the same for all individuals, dependent only on time).

DATA DESCRIPTION

The available data comes from one source and covers a period from 2002 to 2010 (2004-2012 for defaults history), so the whole economic cycle was covered and therefore, the condition required for parameter estimation of PD is fulfilled. There are 1053 good and 494 bad companies in the sample. There are 2910 FS[5] in total. The sample was limited to companies with turnover between 2-35 million Euro.

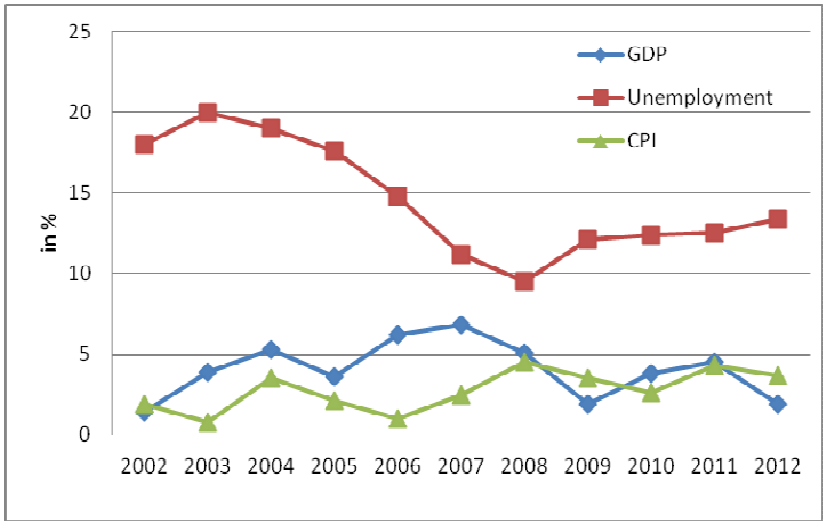
Additionally, the macroeconomic variables were included in the study. These variables are shown in Figure 2. It is worth mentioning that in the models using discriminant functions these variables were static (applied at a certain moment of time, e.g. at the date of bankruptcy). However, in the models using survival analysis, these variables were dynamic. All values were available at the time of FS date (end of calendar year).

For the purpose of this analysis only three variables were selected: GDP (Gross Domestic Product – dynamics), unemployment rate (in %) officially registered, CPI (Consumer Price Index). It is expected that high and increasing GDP should positively affect the probability of default. High unemployment rate is characteristic for downturn in economy and should increase the probability of default. The effect of CPI (inflation) is not obvious but higher inflation is rather positive in economy and should decrease the probability of default in enterprises segment indirectly.

An important issue, when estimating the PD parameter, is the fact that the length of the observation period of the data used for the parameter estimation must be at least five years, and come from at least one source, regardless of whether the source is internal, external, or a combination of both (Basel requirements).

[5] FS – financial statement

Figure 2. Macroeconomic variables in the analysed period



Source: GUS (Polish Central Statistical Office) database

THE RESULTS OF THE MODELS

In our empirical analysis we have applied Altman's Z-Score model to our sample. In the next step we estimated the survival Cox model with the same ratios as in Altman's model. In the final step we have included macroeconomic variables. In each step we checked the proportionality assumption in Cox model.

In a first step the Z-Score Altman's model was applied and Z-Score discriminatory power was not very high (AUC=0.699). It is assumed that AUC should be at least 0.75-0.80 to assume the discriminatory power as satisfactory. For the sample of Polish enterprises the effectiveness of this model was rather low-medium. It could be due to high heterogeneity of the sample. The sample represents rather homogenous group of enterprises as far as concerning the size but heterogeneous as far as concerning the type of activity (branch).

In the next step the Cox regression model with original Altman's variables (ratios) was estimated. All five ratios were significant (see Table 1).

The strongest influence was like in Altman's Z-Score for ratio X_3 (Earnings before interest and taxes/Total assets). Increase by 1 unit in this ratio leads to the decrease of default risk by about 90%.

Accuracy power of this model is much higher comparing to Altman's Z-Score and amounts to AUC=0.746. This level of predictive power may be assumed to be on the medium level.

Table 1. Results of the Cox regression model with original Z-Score ratios

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr. > chi^2	Hazard Ratio
X1	1	-0.35351	0.08611	16.8557	.0001	0.702
X2	1	-0.92721	0.20987	19.5195	.0001	0.396
X3	1	-2.42348	0.23393	107.3235	.0001	0.089
X4	1	-0.41582	0.06930	36.0033	.0001	0.660
X5	1	-0.16972	0.03212	27.9141	.0001	0.844

Source: own elaboration

It was necessary to check the proportionality assumption in Cox regression model. One of the simplest ways of checking this assumption is to include the interactions with time for variables X_1-X_5 (see Table 2). Significant interaction effect means lack of proportionality assumption fulfillment. In our model the interaction with time was significant for X_3-X_5.

The interaction with time for variables X_1 and X_5 is significant. For those variables the assumption of proportionality is fulfilled. For variables X_3-X_5 the negative coefficient means that the negative influence on the risk is enhanced with time.

Cox survival model is named non-proportional model when the proportionality assumption is not fulfilled. The accuracy ratio for such a model is much higher AUC=0.827.

The next step was to include the macroeconomic variables in Cox semiparametric model. Results for such a combination are quite promising. The included variables were significant (except inflation). Results are presented in Table 3.

The higher the GDP and unemployment rate levels the lower the risk of enterprises' default. The direction of unemployment rate influence is not obvious. However it is not confirmed in other research results. The accuracy power of this model is even higher than the previous one. Macroeconomic variables increase the effectiveness of the model. The accuracy level may be assessed as satisfactory.

Table 2. Interaction with time in Cox regression model – proportionality assumption

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr. > chi^2
duration*X1	1	0.00631	0.00599	1.1101	0.2921
duration*X2	1	-0.01065	0.01451	0.5381	0.4632
duration*X3	1	-0.06547	0.01536	18.1651	<.0001
duration*X4	1	-0.05530	0.00474	136.2655	<.0001
duration*X5	1	-0.02548	0.00216	138.7192	<.0001

Source: own elaboration

Table 3. Semiparametric Cox model with macroeconomic variables

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr. > chi^2	Hazard Ratio
X1	1	-0.56683	0.08378	45.7800	<.0001	.567
X2	1	-1.39851	0.20301	47.4556	<.0001	.247
X3	1	-2.62719	0.22354	138.1217	<.0001	.072
X4	1	-0.45197	0.07127	40.2140	<.0001	.636
X5	1	-0.17275	0.03288	27.6053	<.0001	.841
GDP	1	-0.16940	0.02839	35.5974	<.0001	.844
Unemployment	1	-0.16784	0.02647	40.2175	<.0001	.845
CPI	1	0.08828	0.06511	1.8383	0.1752	.092

Source: own elaboration

Table 4. Cox regression model with macro variables and interactions with process duration - nonproportional hazards

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr. > chi^2
duration*X	1	-0.00945	0.00906	1.0874	.2971
duration*X2	1	-0.02461	0.02165	1.2925	.2556
duration*X3	1	-0.05049	0.02224	5.1548	.0232
duration*X4	1	-0.01474	0.00489	9.0809	.0026
duration*X5	1	-0.00411	0.00220	3.4828	.0620
duration*GDP	1	-0.03799	0.00185	422.9914	.0001
duration*Unemployment	1	-0.07100	0.00300	559.2011	.0001
duration*CPI	1	-0.12050	0.00540	497.1255	.0001

Source: own elaboration

The interaction with time for CPI is significant; however the estimation for this variable was not significant.

CONCLUSIONS

In recent years there have been many changes in the credit environment. Banks offer a variety of financial products to a wide range of customers, including those who do not know the law. In view of the rapid increase in the volume of information on the applicants, the financial institutions have the ability to seek out and create newer and more sophisticated credit models.

In view of the recent financial crisis, banks have realized the need to take account of macroeconomic variables in these models, since the economy has a huge impact on the ability of customers to settle liabilities. Since the method used does not allow taking into account in the models time-dependent variables (time dependent), there is a need to find such methods. It seems that survival analysis is a technique that is facing these requirements, because it helps to determine when a specific event occurs in the future, and not just to predict whether it occurs at all. In

the case of credit, this event is of course the insolvency of the borrower. Banks want to know when a customer ceases to repay the commitment to be able to prepare in advance for this event, and possibly take action to minimize losses. The possibility of macroeconomic variables, makes these models dynamic, and the banks can observe how these variables affect the level of bad debts.

Currently the most common method used in the default models for SMEs is the discriminant analysis or logistic regression. It seems, however, that more and more importance is put on the survival analysis, due to its properties. From the presented description of the survival analysis it can be noted that there are several reasons why it is worth using it as an alternative method to traditionally used static models (e.g. logistic regression models). First of all, it should be emphasized that the use of survival analysis in modeling the companies bankruptcy risk can extend the standard static approach into the dynamic one.

Other advantages of using the survival analysis method are as follows:

1) possibility to use censored data - event occurs when the company is eliminated from the observation data set before registering the default,
2) avoiding the instability which can appear due to rigidly fixed length of the observation time,
3) event time estimation allows following the risk default intensity,
4) obtaining the "dynamic" probability forecasts of event (forecast value is a function of time), which is very useful when determining the appropriate strategy and policy,
5) ability to implement changes in the economic environment in credit risk assessment using time dependent variables such as macro variables.

Generally, usage of the survival analysis as an alternative approach to model bankruptcy (default) risk, gives the wider chance to use the results of these models than when using the standard statistical methods, gaining also the improvement of the model bankruptcy prediction. In our further research we are going to include more macroeconomic variables with potential influence on the analysed process of bankruptcy. We will apply this model also on more homogeneous sample with the latest available data.

REFERENCES

Allen L.N., Rose L.C., Financial survival analysis of defaulted debtors, "Journal of Research Society", 57/2006.

Allison P., Logistic Regression Using SAS: Theory and Application, Second Edition, 2012,

Allison P., Survival Analysis Using SAS: A Practical Guide, Second Edition, 2010.

Allison P.D., Regression for Longitudinal Event Data, Sage University Paper Series on Quantitative Applications in the Social Sciences, Nr 46, Publications, Sage 1984.

Altman E., Suggitt H.J., Default rates in the syndicated bank loan market: A mortality analysis, "Journal of Banking and Finance", 2000, vol. 24, No 1.

Altman, E. I., Measuring corporate bond mortality and performance, "Journal of Finance" 44(4)/1989.

Blossfeld H.P., G. Rohwer, Techniques of Event History Modeling. New Approaches to Causal Analysis, Lawrence Elbaum Associates Publishers, London, 2002.

Cox D. R., and Oakes D., Analysis of Survival Data, Chapman and Hall, 1984.

Cox D.R., Regression models and life tables, Journal of the Royal Statistical Society (Series B), no 34, 1972, pp.187-202.

Giovannetti G., Ricchiuti G., Velucchi M., Size, innovation and internationalization: a survival analysis of Italian firms, Applied Economics,Volume 43, Issue 12, 2011, pp. 1511-1520

Glennon D.C., Nigro P., Measuring the Default Risk of Small Business Loans: A Survival Analysis Approach, Journal of Money, Credit, and Banking, Volume 37, Number 5, October 2005, pp. 923-947

Lando D., On Cox Processes and Credit Risky Securities, Review of Derivatives Research, 2, 1998.

Malik, M., Thomas, L. Modelling credit risk in portfolios of consumer loans: transition matrix model for consumer credit ratings. Southampton, UK, University of Southampton, 21pp. (Discussion Papers in Centre for Risk Research), 2009.

McDonald R.A., Matuszyk A., Thomas L. C., Application of survival analysis to cash flow modelling for mortgage products, OR Insight (2010) 23.

Moon T.H., Sohn S.Y., Survival analysis for technology credit scoring adjusting total perception, Journal of the Operational Research Society (2011) 62, pp.1159–1168

Narain, B., Survival analysis and the credit granting decision, in: Thomas L. C., Crook J. N., Edelman D. B., eds. Credit Scoring and Credit Control. OUP, 1992, Oxford, U.K.

Nunes P. M., Serrasqueiro Z., da Silva J.V., Family-owned and non-family-owned SMEs: empirical evidence of survival determinants , Economics and Business Letters, 3(1), pp.68-76, 2014

Pierides, Y. A. ,The pricing of credit risk derivatives, Journal of Economic Dynamics and Control 21(10), 1997.

Ptak-Chmielewska A., M. Pęczkowski Analiza dyskryminacji (discriminant analysis), [in:] Wielowymiarowa analiza statystyczna. Teoria - przykłady zastosowań z systemem SAS (Multivariate statistics. Theory and applications in SAS)., ed. E.Frątczak, Oficyna Wydawnicza SGH, Warsaw, 2009.

Stepanova, M., Thomas, L.C., Survival analysis methods for personal loan data, "Journal of Operations Research" 50(2), 2002.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 382 – 391

ESTIMATING THE ROC CURVE AND ITS SIGNIFICANCE FOR CLASSIFICATION MODELS' ASSESSMENT

Krzysztof Gajowniczek, Tomasz Ząbkowski
Department of Informatics, Faculty of Applied Informatics and Mathematics,
Warsaw University of Life Sciences - SGGW
e-mail: krzysztof_gajowniczek@sggw.pl, tomasz_zabkowski@sggw.pl
Ryszard Szupiluk
Warsaw School of Economics
e-mail: rszupi@sgh.waw.pl

Abstract: Article presents a ROC (receiver operating characteristic) curve and its application for classification models' assessment. ROC curve, along with area under the receiver operating characteristic (AUC) is frequently used as a measure for the diagnostics in many industries including medicine, marketing, finance and technology. In this article, we discuss and compare estimation procedures, both parametric and non-parametric, since these are constantly being developed, adjusted and extended.

Keywords: ROC curve, AUC, classification models' assessment

INTRODUCTION

Plotting the ROC curve is a popular way for discriminatory accuracy visualization of the binary classification models and the area under this curve (AUC) is a common measure of its exact evaluation. ROC methodology is derived from signal detection theory developed during the II World War where it was used to determine if an electronic receiver is able to distinguish between the signal and the noise. Nowadays, it has been used for the diagnostics in medical imaging and radiology[Hanley and McNeil 1982], psychiatry, manufacturing inspection systems, finance and database marketing.

The ROC analysis is useful for the following reasons: (1) evaluation of the discriminatory ability of a continuous predictor to correctly assign into a two-group classification; (2) an optimal cut-off point selection to least misclassify the two-group class; (3) compare the efficacy of two (or more) predictors.

Many parametric and non-parametric estimation methods have been proposed for estimating the ROC curve and its associated summary measures. In this study, we focus on three methods which have been mostly employed in practical applications. In the following sections of the article we introduce notation and the basic concepts of the ROC curve and AUC measure. The further sections are devoted to one parametric and two non-parametric methods of ROC and AUC estimation. The paper ends with a simulation study and short discussion in the last section.

MEASURES OF BINARY CLASSIFICATION PERFORMANCE

Determination of the ROC curve and the area under the curve is related to the classification matrix construction (Table 1) and calculation of sensitivity and specificity measures.

Table 1. Classification matrix

		Predicted value	
		Positive (P)	Negative (N)
Real value	Positive (P)	True positive (TP)	False negative (FN)
	Negative (N)	False positive (FP)	True negative (TN)

Source: own preparation

ROC curve is a set of points: $(x, y) : x = 1 - specificity, y = sensitivity$ where for a particular decision threshold value u sensitivity and specificity is determined. Sensitivity is a ratio of true positive cases to all real positive cases:

$$se = \frac{TP}{TP + FN} \tag{1}$$

whilst specificity determines the share of true negatives cases to all real negative cases:

$$sp = \frac{TN}{FP + TN}. \tag{2}$$

The interpretation of these measures is as follows. Sensitivity is the ability of the classifier to detect instances of a given class (the conditional probability of classification for the selected class, provided that the object actually belongs to it). In turn, specificity determines the extent to which the decision classifier of belonging to the selected class is characterized by the class (supplement conditional probability of classification for the selected class, provided that the object of this class should not be).

It should be noted that the output values generated by the model (e.g. neural network, logistic functions) belong to a certain range, therefore, the threshold should be determined on the basis of which the assignment is made of the cases to particular classes. When determining the value of the decision threshold u in the range [0, 1], and setting $f(x)$ such that:

$$f(x) = \begin{cases} 0 \ for \ x < u \\ 1 \ for \ x \geq u \end{cases} \tag{3}$$

a set of points can be obtained, which allows to plot the ROC curve.

In order to present the mechanism of the ROC curve plotting the following example will be shown. Table 2 contains example with 10 observations sorted in descending order of a classifier probability (so-called scoring model) with the actual classification of the observations (1 or 0). The next columns in the table include the settings of the actual and predicted classifications (TP, TP + FN, TN, TN + FP). SE column shows the sensitivity in accordance with formula (1), and the SP column - specificity determined by the formula (2).

Table 2. Mechanism of the ROC curve plotting

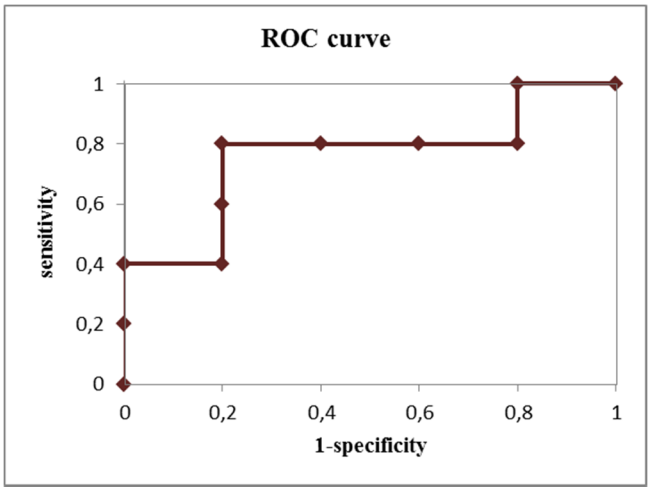
No. obs.	Classifier probability	True class	TP	TP+FN	SE	TN	TN+FP	SP	1-SP
1	0.90	1	1	5	0.2	5	5	1	0
2	0.85	1	2	5	0.4	5	5	1	0
3	0.75	0	2	5	0.4	4	5	0.8	0.2
4	0.70	1	3	5	0.6	4	5	0.8	0.2
5	0.55	1	4	5	0.8	4	5	0.8	0.2
6	0.45	0	4	5	0.8	3	5	0.6	0.4
7	0.40	0	4	5	0.8	2	5	0.4	0.6
8	0.35	0	4	5	0.8	1	5	0.2	0.8
9	0.25	1	5	5	1	1	5	0.2	0.8
10	0.10	0	5	5	1	0	5	0.0	1

Source: own preparation

The ROC curve for the data presented in Table 2 has the following form (Figure 1). The ROC curve was determined based on 10 observations only, therefore this curve has a discrete character. In case of a larger number of observations, the curve would be more smooth.

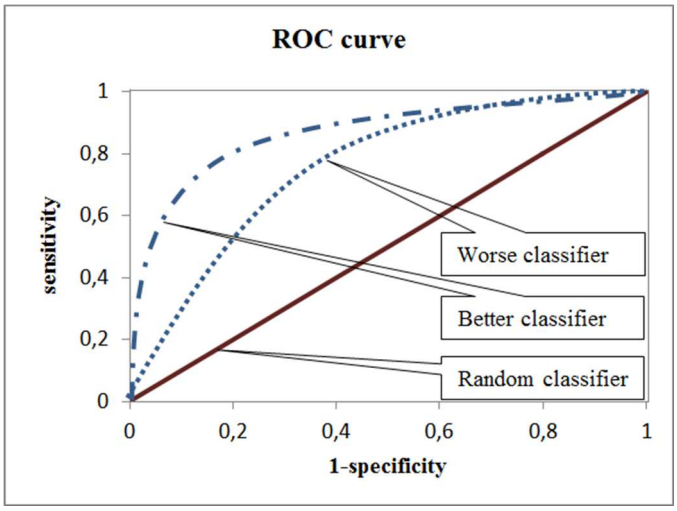
For the purpose of interpretation and comparison of multiple curves, two possible variants of the ROC curve are shown in Figure 2.

Figure 1. ROC curve for the data in Table 2



Source: own preparation

Figure 2. The ROC curve and its possible variants



Source: own preparation

Curve, which coincides with the diagonal curve, has no classification ability. The more the curve is convex and approaching the upper left corner, the better the discrimination has particular model. Highest (perfect) correctness puts classifier in (0,1).

Comparing ROC curves on the graph may be subject to error, especially when comparing a large number of models. Therefore, several ROC curve summary measures of the discriminatory accuracy of a test have been proposed in

the literature, such as the area under the curve (AUC) or the Youden index $\max_c = \{Se(u) + Sp(u) - 1\}$ [Youden 1950].

THE AUC ESTIMATION

One of the main feature associated with the ROC curve, is that curve is increasing and invariant under any monotonic increasing transformation of the considered variables. In general AUC is given by

$$AUC = \int_0^1 ROC(u)\, du \tag{4}$$

Moreover, let X_p and X_n denote the class marker for positive and negative cases, respectively. It could be shown that $AUC = P(X_p > X_n)$. This can be interpreted as the probability that in a randomly selected pair of positive and negative observations the classifier probability is higher for the positive case.

Since the ROC curve measures the inequality between the good and the bad score distributions, it seems reasonable to show a relation between the ROC curve and the Lorenz curve. Twice the area between the Lorenz curve and the diagonal line at 45 degree corresponds to the Gini concentration index. This leads to an interesting interpretation of the AUC measure in terms of the Gini coefficient: $Gini = 2AUC - 1$.

Parametric estimation

A simple parametric approach is to assume the X_p and X_n are independent normal variable with $X_p \sim N(\mu_p, \sigma_p^2)$ and $X_n \sim N(\mu_n, \sigma_n^2)$. Then the ROC curve can be summarized as follow:

$$ROC(u) = \Phi\left(a + b\Phi^{-1}(u)\right) \qquad u \in [0,1] \tag{5}$$

where $a = (\mu_p - \mu_n)/\sigma_p$, $b = \sigma_n/\sigma_p$ and Φ indicates the standard normal distribution function $X \sim N(0,1)$. Furthermore,

$$AUC = \Phi\left(\frac{\mu_p - \mu_n}{\sqrt{\sigma_p^2 + \sigma_n^2}}\right) \text{ or equivalently } AUC = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right) \tag{6}$$

and can be estimated by substituting sample means and standard deviations into all above mentioned formulas.

In practical applications the assumption of normality is untenable, therefore transformation such as the log or the Box–Cox is often suggested [Zou and Hall

$$X^{\lambda} = \begin{cases} X^{\lambda} - 1 \big/ \lambda & for\ \lambda \neq 0 \\ \log(X) & for\ \lambda = 0 \end{cases} \tag{7}$$

2000], and the estimator (6) is then applied to the transformed data. Based on the observations on the positive and negative cases, an appropriate likelihood function can be constructed and maximized giving $\hat{\lambda}$, the maximum likelihood of estimate λ.

Non-parametric estimation

a)

The area under the empirical ROC curve is equal to the Mann–Whitney U statistic [Mann and Whitney 1947] which is usually computed to test whether the levels on some quantitative variable X in one population P tend to be greater than in second population N, without actually assuming how are they distributed in these two population. This measure provides an unbiased non-parametric estimator for the AUC [Faraggi and Reiser 2002]:

$$AUC = \frac{1}{N_p N_n} \sum_{i=1}^{N_p} \sum_{j=1}^{N_n} I\left(X_{pi}, X_{nj}\right) \text{ with } I = \begin{cases} 1 & for\ x_{pi} > x_{nj} \\ \frac{1}{2} & for\ x_{pi} = x_{nj} \\ 0 & otherwise \end{cases} \tag{8}$$

where N_p, N_n are the number of positive and negative cases respectively. Unfortunately, this estimator in some situation is not recommended, because it conceptually requires all $N_p N_n$ comparison and when we are dealing with large number of observations, computational time could be long. Sometimes in (8) sigmoid function is used instead of indicator function [Calders and Jaroszewicz 2007].

b)

When calculating the area under the curve it should be noted that the probabilistic classifiers give the values of the output vector other than the zero and one. Therefore, having m cases classification $o_1,...,o_m$ belonging to a set classes $C = \{C_1, C_2\}$ according to the decision threshold u, sorted so that $0 = se(C_1, o_1) \leq ... \leq se(C_1, o_m) = 1$ and $1 = sp(C_1, o_1) \geq ... \geq sp(C_1, o_m) = 0$ the area under the curve could be calculated via trapezoidal integration:

$$AUC = -\frac{1}{2}\sum_{i=2}^{m}(sp_i se_{i-1} - sp_{i-1} se_i) \tag{9}$$

where $se_i = se(C_1, o_i)$ represents the sensitivity of the classification i-th case to the class C_1, $sp_i = sp(C_1, o_i)$ is the specificity of the classification i-th case to the class C_1. The trapezoidal approach systematically underestimates the AUC, because of the way all of the points on the ROC curve are connected with straight lines rather than smooth concave curves.

To overcome the lack of smoothness of the empirical estimator, [Zou et al. 1997] used kernel methods to estimate the ROC curve, which were later improved by [Lloyd 1998]. Kernel density estimators are known to be simple, versatile, with good theoretical and practical properties.

TESTING DIFFERENCES BETWEEN TWO ROC CURVES

To compare classification algorithms by comparing the area under the ROC curves, is used the following procedure described by [Bradley 1997][Hanley and McNeil 1983]. We consider the following set of hypotheses

$$\begin{aligned} H_0 &: AUC_1 = AUC_2 \\ H_1 &: AUC_1 \neq AUC_2 \end{aligned} \tag{10}$$

to evaluate it, the following test statistic is used

$$z = \frac{A\hat{U}C_1 - A\hat{U}C_2}{\sqrt{SE^2\left(A\hat{U}C_1\right) + SE^2\left(A\hat{U}C_2\right)}} \tag{11}$$

which has the standardize normal distribution $N(0,1)$, and where

$$SE\left(A\hat{U}C\right) = \sqrt{\frac{\theta(1-\theta) + (n_1 - 1)\left(Q_1 - \theta^2\right) + (n_2 - 1)\left(Q_2 - \theta^2\right)}{n_1 n_2}} \tag{12}$$

$$Q_1 = \frac{\theta}{2-\theta},\ Q_2 = \frac{2\theta^2}{1+\theta} \tag{13}$$

where n_1 and n_2 are the number of negative and positive examples respectively and θ is the true area under the ROC curve (but in practice only the estimator $A\hat{U}C$ is used).

SIMULATION STUDY

In order to check the performance of the selected AUC estimator, we conducted the simulations based on the data used for telecom customer churn modelling (the loss of customers moving to some other company). The data is a collection of "Cell2Cell: The Churn Game" [Neslin 2002] derived from the Center of Customer Relationship Management at Duke University, located in North Carolina in the United States. They constitute a representative slice of the entire database, belonging to an anonymous company operating in the sector of mobile telephony in the United States.

The data contains 71047 observations, wherein each observation corresponds to the individual customer. For each observation 78 variables are assigned, of which 75 potential explanatory variables are used for models construction. All explanatory variables are derived from the same time period, except the binary dependent variable (the values 0 and 1) labeled as "churn", which has been observed in the period from 31 to 60 days later than the other variables. In the collection there is an additional variable "calibrat" to identify the learning sample and test sample, comprising 40000 and 31047 observations. Learning sample contains 20000 cases classified as churners (leavers) and 20000 cases classified as non-churners. In the test sample, which is used to check the quality of the constructed model, there is only 1.96% of people who quit. Such a small percentage of the class highlighted can be often found in the business practice.

In this study similar set of modelling techniques has been used as in [Gajowniczek and Ząbkowski 2012]. These were artificial neural networks, classification trees, boosting classification trees, logistic regression and discriminant analysis.

After estimation the $\widehat{\lambda}$ parameter by power transformation, we observed that most of the distributions (Table 3) have not the normal distribution based on Shapiro-Wilk normality test at $\alpha = 0.01$. As stated in [Krzyśko et al. 2008], X_p, X_n may not have a normal distribution, but the reasoning based on the ROC curve built for a normal distribution may give good results, because the ROC curves do not count individual distribution, but the relationship between the distributions.

Table 3. Tests for normality

	Negative cases (churn=1)		Positive cases (churn=0)	
	p-value	$\widehat{\lambda}$	p-value	$\widehat{\lambda}$
Artificial neural network (SANN)	2.88E-18	1.18	0.6694	1.37
Boosting classification trees (Boosting)	1.02E-22	1.16	0.2164	1.41
Logistic regression (Logit)	1.01E-08	0.77	0.0380	0.71
Classification trees (C&RT)	7.93E-51	1.13	1.20E-16	1.55
Discriminant analysis (GDA)	2.27E-08	0.79	0.0181	0.80

Source: own preparation

Very small differences can be seen in Table 4 among non-parametric AUC estimates. The biggest difference in AUC can be observed in case of classification trees. This is due to the fact that C&RT assigns observations to the leafs. Within each leaf there is the same probability of belonging to the positive class. Therefore, when there are only few leafs in the tree then we don't expect the distribution of probabilities to meet the assumption of normality.

Table 4. AUC estimation using different techniques

	Mann-Whitney (non-parametric)	Trapezoidal integration (non-parametric)	Normal assumption (parametric)
Artificial neural network (SANN)	0.6242784	0.6242784	0.6864752
Boosting classification trees (Boosting)	0.6632097	0.6632097	0.7045478
Logistic regression (Logit)	0.6189685	0.6189685	0.6072612
Classification trees (C&RT)	0.6215373	0.6227865	0.752052
Discriminant analysis (GDA)	0.6190384	0.6190384	0.6288627

Source: own preparation

Table 5 show the critical levels (p-values) for testing differences between two ROC curves based on Mann-Whitney estimation. The hypothesis of equality of the areas under the ROC curve could be reject when p-value are smaller than accepted level of significance. It can be observed that, at the significance level $\alpha = 0.05$, the areas under the curves for the SANN, Logit, C&RT, GDA are not significantly different. Only the AUC measures for Boosting significantly differs from the other methods.

Table 5. P-values for the differences between two AUC measure

	SANN	Boosting	Logit	C&RT	GDA
SANN	1.00000000	0.02417606	0.75930305	0.93139089	0.76237611
Boosting		1.00000000	0.01042189	0.01925184	0.01054391
Logit			1.00000000	0.82563864	0.99678138
C&RT				1.00000000	0.82878156
GDA					1.00000000

Source: own preparation

CONCLUSIONS

The aim of this study was to compare the accuracy of commonly used ROC curve estimation methods taking into account different classification techniques. We show that non-parametric methods give convergent results in terms of the AUC measure while parametric approach tends to give the higher values of AUC, except the Logit. In practical applications, for parametric methods of ROC estimation the assumption of normality is untenable, therefore, non-parametric methods should be utilized.

The simulation experiment suggest that the non-parametric ROC estimation using trapezoidal rule is a reliable method when the distributions of the predictive outcome are skewed and that it provides a smooth ROC. Finally, this approach of estimation is not difficult nor computationally time consuming.

Acknowledgments

The study is cofounded by the European Union from resources of the European Social Fund. Project PO KL „Information technologies: Research and their interdisciplinary applications", Agreement UDA-POKL.04.01.01-00-051/10-00.

REFERENCES

Bradley A.P. (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms, Pattern Recognition, vol. 30, No. 7, pp. 1145-1159.

Calders T., Jaroszewicz S. (2007) Efficient AUC Optimization for Classification, Proceedings of The 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'07), pp. 42-53.

Faraggi D.. Reiser B. (2002) Estimation of the area under the ROC curve, Statistics in Medicine, vol. 21, pp. 3093–3096.

Gajowniczek K., Ząbkowski T. (2012) Problemy modelowania rezygnacji klientów w telefonii komórkowej, Metody Ilościowe w Badaniach Ekonomicznych, vol. 13, No 3, pp. 65-79.

Hanley J. A., McNeil B. J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology vol. 143, pp. 29-36.

Hanley J. A., McNeil B. J. (1983) A method of comparing the areas under receiver operating characteristic curves derived from the same cases, Radiology vol. 148, pp. 839-843.

Krzyśko M., Wołyński W., Górecki T., Skorzybut M. (2008) Systemy uczące się, Wydawnictwo Naukowo-Techniczne.

Lloyd C. J. (1998) Using smoothed receiver operating characteristic curves to summarize and compare diagnostic systems, Journal of the American Statistical Association, vol. 93, pp. 1356–1364.

Mann H. B., Whitney D. R. (1947) On a test of whether one of two random variables is stochastically larger than the other, The Annals of Mathematical Statistics; vol. 18, pp. 50–60.

Neslin S. (2002) Cell2Cell: The churn game. Cell2Cell Case Notes, Hanover, NH: Tuck School of Business, Dartmoth College, Downloaded from: http://www.fuqua.duke.edu/centers/ccrm/datasets/cell/

Youden W. J. (1950) An index for rating diagnostic tests, Cancer, vol.3, pp. 32–35.

Zou K. H.; Hall W. J., Shapiro D. E. (1997). Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests, Statistics in Medicine, vol. 16, pp. 2143–2156.

Zou K. H., Hall W.J. (2000) Two transformation models for estimating an ROC curve derived from continuous data, Journal of Applied Statistics, vol. 27, pp. 621–631.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 392 – 402

ANALYSIS OF CALENDAR EFFECTS IN MARKETS OF PRECIOUS METALS

Anna Górska
Department of Agricultural Economics and International Economic Relations
Warsaw University of Life Sciences – SGGW
e-mail: anna_gorska@sggw.pl
Monika Krawiec
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
e-mail: krawiec.monika@gmail.com

Abstract: Calendar effects are anomalies in the behavior of asset prices that may disprove the efficient market hypothesis. The well recognized are: day-of-the-week effect, month-of-the-year effect, holidays effect and turn-of-the-month effect. These anomalies are observed in many financial markets, most often on stock exchanges, thus studies on calendar effects usually focus on stock markets. However, the aim of the paper is searching for the anomalies in precious metals markets (the empirical data covers London daily spot prices from 2008 through 2013). This is the continuation of authors' prior research aimed at testing weak market efficiency hypothesis for precious metals markets.

Keywords: precious metals, calendar effects, linear regression, GARCH models

INTRODUCTION

Numerous financial analyses are based on the efficient market hypothesis, which may be also applied to commodity markets. According to the classical Fama's definition: a market in which prices always fully reflect available information is called efficient [Fama 1970]. Park and Irwin [2007] present a more detailed definition proposed by Jensen [1978]: a market is efficient with respect to information set θ_t if it is impossible to make economic profits by trading on the basis of information set θ_t. Jensen also subdivides the efficient market hypothesis

into three types based on definitions of the information set θ_t: weak-form efficiency, where the information set θ_t is limited to the information contained in the past price history of the market as of time t; semistrong-form efficiency, where the information set θ_t is all information that is publicly available at time t; strong-form efficiency, where the information set θ_t is all public and private information available at time t [Jensen 1978]. One should note that the weak form is a restricted version of the semistrong form, and weak and semistrong forms are restricted versions of the strong form. The weak-form market efficiency hypothesis is tested the most often. To do this, technical analysis tools and statistical tests veryfing random changes in prices are applied. These are for example: unit root tests, autocorrelation tests, variance ratio test, runs test, as well as analysis of long-run relationships and correlations, and calendar effects analysis [Witkowska at al. 2008].

Calendar effects are anomalies in the behavior of asset prices that make the market inefficient. The best known are the following: the day-of-the-week effect, the month-of-the-year effect, holidays effect, and the turn-of-the-month effect. These and other anomalies[1] are observed in many financial markets, most often on stock exchanges, thus studies on calendar effects usually focus on stock markets. The aim of the paper is searching for the anomalies in precious metals markets. We focus on the day-of-the-week and the month-of-the-year effects. The paper is the continuation of authors' prior research aimed at testing weak market efficiency hypothesis for precious metals markets by the use of runs test, variance ratio test and autocorrelation test. The results obtained then were not homogeneous [Górska, Krawiec 2013].

EMPIRICAL DATA AND METHODOLOGY

Data

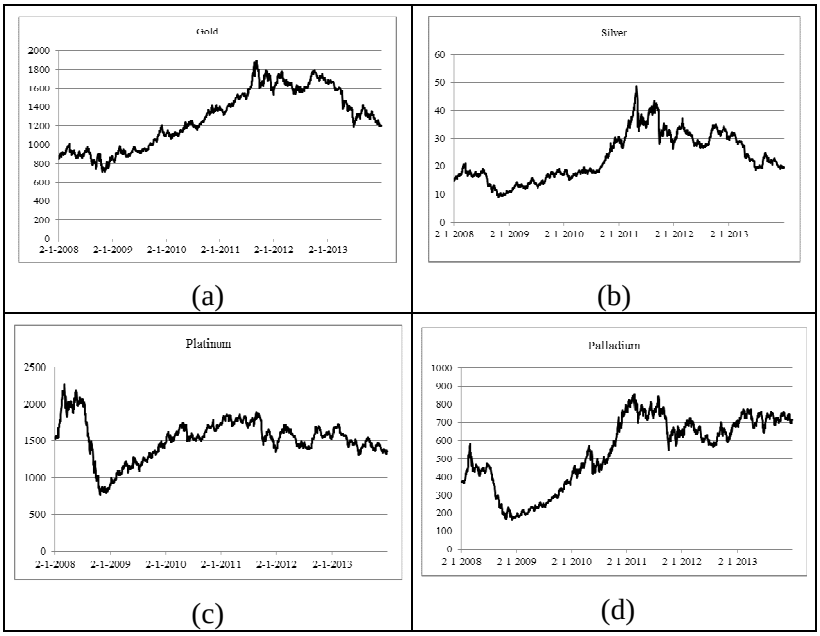
The empirical data used for the purpose of the analysis covers London daily closing prices of four precious metals (gold, silver, platinum and palladium) from 2008 through 2013. The quotations in USD per ounce are available at www.kitco.com. The prices of precious metals in the period under consideration are displayed in figure 1. First of all, there were calculated the returns:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}, \tag{1}$$

[1] There are also more advanced studies on lunar (new moon and full moon) effects in stock markets. Findings of Yuan at al. [2006], based on stock market returns of 48 countries, indicate that stock returns are lower on the days around a full moon than on the on the days around a new moon and the lunar effect is independent of other calendar-related anomalies such as the January effect, the day-of-the-week effect, the month effect, and the holiday effect. Keef and Khaled [2011] provide further evidence on international lunar effects.

where P_t is the price at time t and P_{t-1} is the price in the previous period. These rates of return became the base to evaluate some descriptive statistics for considered metals. The results are given in tables 1 and 2.

Figure 1. Prices of precious metals from 2008 through 2013: gold (a), silver (b), platinum (c), palladium (d)



Source: own elaboration

On the base of data given in tables 1 and 2, one can notice that expected rates of return (means) calculated for all days were positive. The highest of them was the one produced by palladium, the lowest – by platinum. However, analyzing weekday returns separately, we can notice negative returns in the case of gold: Tuesdays and Thursdays, in the case of silver: Fridays (this is the highest negative return of all weekdays), in the case of platinum: Tuesdays and Fridays, and in the case of palladium: Fridays. The highest positive expected rate of return is Monday silver return. When considering separate months, the highest positive expected rate of return is January platinum return, while the highest negative expected rate of return is September palladium return. There are also other negative monthly expected rates of return (in the case of gold: March, April, June, October, December, in the case of silver: May, June, October December, in the case of platinum: April, May, June, July, September, October, December, in the case of palladium: March, May, June, August).

Table 1. Descriptive statistics for daily returns on separate weekdays (2008-2013)

	Measure	All days	Mondays	Tuesdays	Wednesdays	Thursdays	Fridays
Gold	Mean	0,00033	0,00150	-0,00010	0,00036	-0,00042	0,00039
	std. dev.	0,01383	0,01519	0,01232	0,01373	0,01481	0,01298
	minimum	-0,09150	-0,09150	-0,04901	-0,05650	-0,05846	-0,03938
	maximum	0,07081	0,06198	0,03817	0,04586	0,07081	0,04946
	skewness	-0,25893	-1,01961	-0,25902	-0,19897	0,17142	0,10798
	kurtosis	4,25636	8,98234	1,24253	2,49095	3,92464	1,32139
Silver	Mean	0,00053	0,00316	0,00010	0,00167	0,00104	-0,00317
	std. dev.	0,02668	0,02802	0,02396	0,02444	0,02951	0,02691
	minimum	-0,17050	-0,14407	-0,10452	-0,07613	-0,17050	-0,13728
	maximum	0,20056	0,11111	0,18963	0,07232	0,20056	0,11385
	skewness	-0,03458	-0,69031	1,32845	-0,11617	0,24043	-0,64854
	kurtosis	7,47039	5,66200	13,41601	0,57967	11,21311	4,63205
Platinum	Mean	0,00006	0,00033	-0,00004	0,00061	0,00020	-0,00081
	std. dev.	0,01679	0,01722	0,01840	0,01511	0,01671	0,01643
	minimum	-0,08143	-0,07534	-0,08143	-0,05405	-0,07277	-0,07407
	maximum	0,07186	0,06853	0,07186	0,05200	0,05499	0,05657
	skewness	-0,50923	-0,36834	-0,60776	-0,48588	-0,34239	-0,70027
	kurtosis	3,44222	3,05964	5,08753	1,73355	2,40617	3,33568
Palladium	Mean	0,00069	0,00185	0,00015	0,00146	0,00077	-0,00070
	std. dev.	0,02289	0,02306	0,02189	0,02258	0,02601	0,02057
	minimum	-0,16355	-0,08498	-0,11355	-0,08591	-0,16355	-0,08353
	maximum	0,11538	0,07399	0,07084	0,11538	0,09958	0,05799
	skewness	-0,44769	-0,14178	-0,60218	0,05349	-0,90905	-0,43579
	kurtosis	4,08381	1,68553	3,42433	3,28852	6,94054	1,25974

Source: own calculations

The highest value of standard deviation calculated for all daily observations was observed for silver, while the lowest – for gold. When considering separate weekdays, the highest standard deviation is that for Thursday silver returns, the lowest for Tuesday gold returns. In the case of separate months, the highest standard deviation is that for September silver returns, and the lowest for July gold returns. For the whole period (all days) we can observe negative skewness, however for Tuesday silver returns, Wednesday palladium returns, Thursday gold and silver returns, and for Friday gold returns, skewness is positive. In the case of monthly returns, we have positive skewness for January (gold, silver, platinum, palladium), February (silver, platinum, palladium), March (gold), May (gold), June (platinum and palladium), July (gold, silver), September (gold and silver), October (palladium), November (gold and silver), December (gold, silver, platinum, palladium). Almost all values of kurtosis are positive (the only exception is negative kurtosis for May gold returns). This indicates more acute distributions in comparison to a normal distribution.

Table 2. Descriptive statistics for daily returns in separate months (2008-2013)

	Month	Measure					
		Mean	Std. dev.	Min	Max	Skewness	Kurtosis
Gold	January	0,0014	0,0121	-0,0255	0,0467	0,5803	1,5574
	February	0,0012	0,0115	-0,0429	0,0311	-0,3482	1,7237
	March	-0,0008	0,0131	-0,0479	0,0708	0,6516	7,6488
	April	-0,0005	0,0147	-0,0915	0,0232	-2,6992	12,5589
	May	0,0003	0,0127	-0,0266	0,0351	0,0267	-0,1347
	June	-0,0008	0,0125	-0,0585	0,0308	-0,7499	3,5530
	July	0,0010	0,0110	-0,0364	0,0426	0,1117	2,0572
	August	0,0014	0,0137	-0,0565	0,0342	-0,6431	2,9281
	September	0,0006	0,0163	-0,0539	0,0615	0,0893	3,0237
	October	-0,0005	0,0169	-0,0766	0,0574	-0,5644	3,9535
	November	0,0018	0,0145	-0,0315	0,0620	1,0086	2,6309
	December	-0,0012	0,0156	-0,0448	0,0520	0,1699	1,6472
Silver	January	0,0032	0,0228	-0,0514	0,0688	0,1333	0,4391
	February	0,0034	0,0190	-0,0595	0,0668	0,0537	1,1501
	March	0,0000	0,0258	-0,1182	0,0678	-0,6434	3,2383
	April	0,0002	0,0250	-0,1409	0,0751	-1,4655	8,2851
	May	-0,0010	0,0347	-0,1705	0,1138	-0,7502	6,1163
	June	-0,0026	0,0227	-0,0638	0,0723	-0,0670	0,8031
	July	0,0012	0,0201	-0,0543	0,0721	0,3171	0,9050
	August	0,0018	0,0266	-0,1373	0,0553	-1,3073	5,5573
	September	0,0004	0,0391	-0,1441	0,2006	1,2385	10,3227
	October	-0,0009	0,0288	-0,0963	0,1084	-0,3423	2,4793
	November	0,0016	0,0236	-0,0419	0,0949	1,0035	2,3791
	December	-0,0009	0,0255	-0,0869	0,0772	0,1034	1,1604
Platinum	January	0,0043	0,0155	-0,0447	0,0566	0,4300	1,5047
	February	0,0029	0,0153	-0,0478	0,0516	0,0842	1,5712
	March	0,0002	0,0194	-0,0664	0,0719	-0,3657	3,2863
	April	-0,0006	0,0147	-0,0598	0,0249	-1,3360	3,4248
	May	-0,0011	0,0172	-0,0602	0,0548	-0,0041	1,9272
	June	-0,0009	0,0138	-0,0384	0,0385	0,2945	0,5847
	July	-0,0001	0,0128	-0,0504	0,0304	-0,5586	1,5249
	August	0,0001	0,0170	-0,0741	0,0393	-0,9195	3,4445
	September	-0,0032	0,0207	-0,0814	0,0550	-1,2299	3,4482
	October	-0,0008	0,0193	-0,0728	0,0629	-0,7780	3,0803
	November	0,0007	0,0153	-0,0670	0,0520	-0,1054	3,6059
	December	-0,0007	0,0178	-0,0753	0,0685	0,0223	4,4839

(continued)

Table 2. (continued)

	Month	Measure					
		Mean	Std. dev.	Min.	Max.	Skewness	Kurtosis
Palladium	January	0,0027	0,0206	-0,0649	0,0708	0,1557	1,1193
	February	0,0038	0,0237	-0,0835	0,0996	0,1548	2,9462
	March	-0,0007	0,0255	-0,0859	0,0725	-0,9164	1,8687
	April	0,0006	0,0185	-0,0575	0,0521	-0,4259	0,8257
	May	-0,0005	0,0250	-0,1102	0,0740	-0,6173	3,5662
	June	-0,0009	0,0204	-0,0550	0,0584	0,1028	0,5325
	July	0,0013	0,0175	-0,0719	0,0544	-0,2398	1,9966
	August	-0,0007	0,0191	-0,0619	0,0469	-0,4550	0,8013
	September	-0,0033	0,0259	-0,1136	0,0730	-0,8787	3,5731
	October	0,0020	0,0257	-0,0863	0,1154	0,2556	3,5867
	November	0,0019	0,0278	-0,1636	0,0787	-1,2060	9,2073
	December	0,0022	0,0219	-0,0695	0,0674	0,0153	2,3167

Source: own calculations

Methodology

In order to study calendar effects we use econometric models. The study focuses on the-day-of-the-week and the month-of-the-year effects. The day-of-the-week effect is one of the most frequently investigated seasonal anomalies in capital markets [Cross 1973, French 1980, Gibbons and Hess 1981, Keim and Stambaugh 1984, Lakonishok and Smidt 1988, Lakonishok and Maberly 1990]. Researchers revealed that Monday and Friday returns differ from returns on other weekdays: Monday returns are statistically significantly negative, while Friday returns – positive. Another well documented anomaly is January effect. It is proved that returns on stock markets often are much higher in January than in other months [Rozeff and Kinney 1976, Keim 1983, Haugen and Lakonishok 1988, Bouman and Jacobsen 2002, Fountas and Segredakis 2002, Lucey and Zhao 2008]. Another monthly effects are: May effect (low returns) and September effect (high returns). There were also several studies in Poland investigating calendar effects for Warsaw Stock Exchange. Different results were presented by Szyszka [1999], Papla [2000], Osińska [2006], Landmesser [2006], Rozkrut [2006], Witkowska and Kompa [2007], Kluth [2007] etc.

One of possible methods for examining calendar effects is estimating the following equation:

$$R_t = \beta_1 D_{1t} + \beta_2 D_{2t} + \beta_3 D_{3t} + \beta_4 D_{4t} + \beta_5 D_{5t} + \varepsilon_t , \quad (2)$$

where:

R_t - is the daily return of the asset,

D_{it} - are dummy variables which take on the value of 1 if the corresponding return for day *t* is Monday, Tuesday, Wednesday, Thursday or Friday respectively, and 0 otherwise.

β_i capture the mean daily return for each of the days of the week, but Borges [2009] points out that the *t*-test for those coefficients only inform us if they are significantly different from zero. According to her, if the time period under study is sufficiently long, it is to be expected that mean daily return is positive, however a very small number. Therefore, the significance of the *t*-tests is biased in favor of accepting positive excess returns, and against accepting negative excess returns. She suggests to estimate five equations separately, each aiming at detecting a specific day of the week effect:

$$R_t = \alpha + \beta_{i1} D_{it} + \varepsilon_t . \tag{3}$$

Here, if we include only the dummy variable for Mondays, α captures the mean daily return of non-Mondays and β_1 is the excess return of Mondays, relative to non-Mondays. The *t*-test of β_1 tells us if this effect is significant. The same is for β_2, β_3, β_4 and β_5, for detecting other weekdays effects.

The above discussion can be transferred to month effect analysis as well. The only difference is that we need twelve different dummies: M_i (i=1,…,12). Each takes the value of 1 if the corresponding return for day *t* is of January, February etc., and 0 otherwise.

In order to investigate calendar effects GARCH models may be applied as well. GARCH (p, q), developed by Bollerslev in 1986, is of the following form:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \lambda_i \sigma_{t-i}^2 . \tag{4}$$

In the paper here, following Borges [2009], after performing a test of ARCH effects on our data, we evaluate the most simple form - GARCH (1,1), including only one lag both in the ARCH term (last period's volatility) and in the GARCH term (last period's variance). Dummies indicate days and months of observation.

RESULTS

In order to detect calendar effects in daily returns generated by four precious metals, following procedures described in the previous section, we start with estimating individual regressions for the day-of-the-week and month effects. The results for the β_i coefficients are given in table 3. They reveal significant the day-of-the-week effect – Friday effect in the case of silver only. Here, Friday silver returns are statistically significantly negative (not positive as one could have expected). Significant monthly effects are: January effect (platinum) and September effect (platinum and palladium). September returns are statistically significantly negative.

Table 3. Calendar effects for precious metals (2008-2013) – linear simple regressions

Effect	Metal			
	Gold	Silver	Platinum	Palladium
Monday (D_1)	0,0014	0,0032	0,0003	0,0014
Tuesday (D_2)	-0,0005	-0,0005	-0,0001	-0,0007
Wednesday (D_3)	0,0000	0,0014	0,0007	0,0010
Thursday (D_4)	-0,0009	0,0006	0,0002	0,0001
Friday (D_5)	0,0001	-0,0046*	-0,0011	-0,0017
January (M_1)	0,0012	0,0029	0,0046*	0,0022
February (M_2)	0,0010	0,0031	0,0030	0,0034
March (M_3)	-0,0012	-0,0005	0,0001	-0,0016
April (M_4)	-0,0009	-0,0003	-0,0007	0,0000
May ($M_{5)}$	0,0000	-0,0016	-0,0012	-0,0013
June (M_6)	-0,0013	-0,0035	-0,0010	-0,0017
July (M_7)	0,0007	0,0007	-0,0002	0,0006
August (M_8)	0,0012	0,0014	0,0001	-0,0015
September (M_9)	0,0003	-0,0002	-0,0035*	-0,0043*
October (M_{10})	-0,0010	-0,0016	-0,0010	0,0015
November (M_{11})	0,0016	0,0011	0,0008	0,0013
December (M_{12})	-0,0017	-0,0015	-0,0008	0,0016

Source: own calculations Note: *significance at the 0,05 level

In the second step of research, after confirming the presence of ARCH effects in our data, we estimate GARCH (1, 1) models. The adequate results are presented in table 4. They confirm the January effect in the case of platinum and detect another month effect – September effect in the case of gold and silver. However, the GARCH (1,1) approach gives no evidence of any day-of-the-week effect.

CONCLUDING REMARKS

The aim of the paper was to study calendar effects in markets of precious metals. Usually researchs on calendar effects focus on stock markets rather than on commodity markets and if they cover commodities the most often are limited to oil or gold and sometimes silver. Our study extends the investigation to the four most important precious metals: gold, silver platinum and palladium and is based on London daily closing prices from 2008 through 2013.

First of all, we find almost no evidence of calendar effects on London gold market (the only exception is September effect under GARCH methodology), that implies its efficiency. Our results confirm prior findings of Smith [2002], who stated that London gold closing prices follow a random walk. In his opinion the closing price is more efficient than morning and afternoon fixings, because it is determined by additional information during the day and involves many more market participants. Moreover, researchers broadly agree that the U.S. gold market

is also efficient (see Tschoegl [1980], Solt, Swanson [1981], Aggarwal, Soenen [1988]). Charlie at al. [2007] generalize and conclude that the gold market itself is relatively efficient. However, Qi and Wang [2013] provide evidence of monthly effects in Chinese gold market.

Table 4. Calendar effects for precious metals (2008-2013) – GARCH (1,1) approach

Effect	Metal			
	Gold	Silver	Platinum	Palladium
Monday (D_1)	0,000916	0,001787	0,000339	0,001255
Tuesday (D_2)	0,000711	0,001525	0,000713	0,001196
Wednesday (D_3)	0,000373	0,002177	0,000951	0,000930
Thursday (D_4)	-0,000017	0,001126	-0,000384	0,000429
Friday (D_5)	0,000561	-0,002125	-0,000319	0,000785
α_0	0,000003	0,000031	0,000004	0,000008
α_1	0,052945	0,088835	0,062163	0,065413
λ_1	0,930780	0,868468	0,921607	0,919585
January (M_1)	0,000529	0,003189	0,004316*	0,001855
February (M_2)	0,000776	0,002076	0,001068	0,002425
March (M_3)	-0,000418	0,000218	-0,000059	-0,000865
April (M_4)	0,001225	-0,002897	-0,000492	-0,000088
May (M_5)	0,000579	-0,000495	-0,001731	-0,000040
June (M_6)	-0,000634	-0,002145	-0,001506	-0,001113
July (M_7)	0,000721	0,000757	0,000899	0,001919
August (M_8)	0,001530	0,003474	0,001047	0,000907
September (M_9)	0,001948*	0,006581*	0,000489	-0,000049
October (M_{10})	0,000355	-0,000668	0,000200	0,001622
November (M_{11})	0,001173	0,001380	0,000041	0,001959
December (M_{12})	-0,001733	-0,001806	-0,000778	0,001679
α_0	0,000003	0,000031	0,000004	0,000008
α_1	0,0558847	0,096919	0,057289	0,062837
λ_1	0,928998	0,859957	0,928591	0,922777

Source: own calculations Note: *significance at the 0,05 level

According to Christian [2007] silver shares some characteristics with gold, so one could expect its efficiency as well, but our study for silver detects both the-day-of-the-week (Friday) effect and the-month-of-the-year (September) effect. In the case of platinum and palladium the only observable calendar anomalies are monthly effects: January and September for platinum and September for palladium. However, considering only the effects that are significant under both applied methodologies, we should recognize the January effect for platinum solely.

REFERENCES

Aggarwal R., Soenen L. (1988) The Nature and Efficiency of the Gold Market, Journal of Portfolio Management, 14 (3), 18-21.

Bollerslev T. (1986) Generalized Autoregressive Conditional Heteroscedasticity, Journal of Econometrics, 31, 307-327.

Borges M.R. (2009) Calendar Effects in Stock Markets: Critique of Previous Methodologies and Recent Evidence in European Countries, Working Papers, WP 37/2009/DEUECE, School of Economics and Management, Technical University of Lisbon.

Bouman S., Jacobsen B. (2002) The Halloween Indicator , "Sell in May and Go Away": Another Puzzle, The American Economic Review, 92, 1618-1635.

Charlie X.C., Clacher I., Faff R., Hillier D. (2007) A Practical Guide to Gold as an Investment Asset, The Handbook of Commodity Investment, John Wiley&Sons, Hoboken, New Jersey, 712-735.

Christian T. (2007) Fundamental Analysis of the World Silver Market, The Handbook of Commodity Investment, John Wiley&Sons, Hoboken, New Jersey, 763-802.

Cross F. (1973) The Behavior of Stock Prices on Fridays and Mondays, Financial Analysts Journal, 29, 67-69.

Fountas S., Segredakis K.N. (2002) Emerging Stock Markets Return Seasonalities: the January Effect and the Tax-loss Selling Hypothesis, Applied Financial Economics, 12, 291-299.

French K.R. (1980) Stock Returns and the Weekend Effect, Journal of Financial Economics, 8, 55-80.

Fama E.F. (1970) Efficient Capital Markets: a Review of Theory and Empirical Work, Journal of Finance, 25 (May), 383 – 417.

Gibbons M., Hess P. (1981) Day of the Week Effects and Asset Returns, Journal of Business, no 54, pp. 579-596.

Górska A., Krawiec M. (2013) Badanie efektywności informacyjnej w formie słabej na rynku metali szlachetnych, Zeszyty Naukowe Uniwersytetu Szczecińskiego, 768, Finanse, Rynki Finansowe, Ubezpieczenia, 63, Rynek Kapitałowy – Skuteczne Inwestowanie, Szczecin, 143-156.

Haugen R., Lakonishok J. (1988) The Incredible January Effect, Dow-Jones Irwin, Homewood.

Jensen M.C. (1978) Some Anomalous Evidence Regarding Market Efficiency, Journal of Financial Economics, 6 (June-September), 95-101.

Keef P.K., Khaled M.S. (2011) Are Investors Moonstrack? Further International Evidence on Lunar Phases and Stock Returns, Journal of Empirical Finance, 18, 56-63.

Keim D. (1983) Size-related Anomalies and Stock Return Seasonality: Further Empirical Evidence, Journal of Financial Economics, 12, 13-32.

Keim D., Stambaugh F. (1984) A Further Investigation of Weekend Effects in Stock Returns, Journal of Finance, 39, pp. 819-840.

Kluth K. (2007) Analiza efektów sezonowych na Giełdzie Papierów Wartościowych w Warszawie, Metody ilościowe w badaniach ekonomicznych, VII, Warszawa, 163-172.

Lakonishok J., Maberly E. (1990) The Weekend Effect: Trading Patterns of Individual an Institutional Investors, Journal of Finance, 45(1), 231-243.

Lakonishok J., Smidt S. (1988) Are Seasonal Anomalies Real?: a Ninety Year Perspective, Review of Financial Studies, 1, 403-425.

Landmesser J. (2007) Efekt dnia tygodnia na Giełdzie Papierów Wartościowych w Warszawie, Ekonomika i Organizacja Gospodarki Żywnościowej, 60, Warszawa, 187-196.

Lucey B.M., Zhao S. (2008) Halloween or January? Yet Another Puzzle, International Review of Financial Analysis, 17, 1055-1069.

Papla D. (2000) Badanie efektywności rynku, Metody ekonometryczne i statystyczne w analizie rynku kapitałowego, red. K. Jajuga, Wydawnictwo AE im. Oskara Langego we Wrocławiu, Wrocław, 21-69.

Osińska M. (2006) Ekonometria finansowa, PWE, Warszawa.

Park Ch.-H., Scott I.H. (2007) The Profitability of Technical Analysis in Commodity Markets, The Handbook of Commodity Investing, John Wiley&Sons, Hoboken, New Jersey, 909 – 947.

Qi M., Wang W. (2013) The Monthly Effects in Chinese Gold Market, International Journal of Economics and Finance, 15 (10), 141-146.

Rozeff M. S., Kinney W. R., (1976) Capital Market Seasonality: the Case of Stock Returns, Journal of Financial Economics, 3, 379-402.

Rozkrut D. (2006) Searches for the Day of the Week Effect on the Polish Stock Exchange, Folia Oeconomica, 5 (13), Szczecin, 93-102.

Smith G. (2002) Tests of the Random Walk Hypothesis for London Gold Prices, Applied Economic Letters, 9 (10), 671-674.

Solt M., Swanson P. (1981) On the Efficiency of the Markets for Gold and Silver, Journal of Business 54 (3), 453-478.

Szyszka A. (1999) Efektywność rynku a anomalie w rozkładach stóp zwrotu w czasie, Nasz Rynek Kapitałowy, 12.

Tschoegl A. (1980) Efficiency in the Gold Market – a Note, Journal of Banking and Finance, 4 (4), 371-379.

Witkowska D., Kompa K. (2007) Analiza własności stóp zwrotu akcji wybranych spółek. Rynek Kapitałowy – skuteczne inwestowanie, Zeszyty Naukowe US, 462, Finanse, Rynki finansowe, Ubezpieczenia, 6, Szczecin, 255-266.

Witkowska D., Matuszewska A., Kompa K. (2008) Wprowadzenie do ekonometrii dynamicznej i finansowej, Wydawnictwo SGGW, Warszawa.

Yuan K., Zheng L., Zhu Q. (2006) Are Investors Moonstrack? Lunar Phases and Stock Returns, Journal of Empirical Finance, 13, 1-23.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 403 – 413

CLASSIFICATION OF POLISH HOUSEHOLDS BASED ON THEIR INCOMES BY MEANS OF DECISION TREES

Krzysztof Karpio, Piotr Łukasiewicz, Grzegorz Koszela, Arkadiusz Orłowski
Faculty of Applied Informatics and Mathematics
Warsaw University of Life Sciences – SGGW
e-mails: krzysztof_karpio@sggw.pl, piotr_lukasiewicz@sggw.pl,
grzegorz_koszela@sggw.pl, arkadiusz_orlowski@sggw.pl

Abstract: Classification trees included in SQL Server 2008R2 Analysis Services package have been used to classify Polish households based on their incomes. The analysis has been performed by means of the three algorithms and their effectiveness has been measured. Using the best algorithm a groups of households with the lowest and the largest incomes have been distinguished. The most important attributes describing households with the lowest and the largest incomes were identified and discussed.

Keywords: income distribution, high incomes, classification trees, entropy, SQL Server

INTRODUCTION

In this paper classification trees included in a SSAS package have been used for classification of Polish households based on their incomes. We search for households with the lowest or the highest incomes. The main aim of this paper is to identify attributes of households with the lowest or the highest incomes. The extreme incomes are defined by the first and the last deciles. We are going to answer the following questions: (1) what is a hierarchy of the attributes based on their influence on the classification results (whether a household has the lowest or the highest income); (2) what attributes and their values best describe households with the lowest or the highest incomes? Studying incomes is preceded by investigation of all algorithms and a choice of the most effective one.

A multidimensional analysis of poverty exists widely in literature about incomes and classification of households. This technique was used to search for households with incomes below social minimum by [Kozera et al. 2013]. The

authors used classification trees based on the CART algorithm with Gini index. The same method was adopted by [Anioła et al. 2012] to analyze factors influencing savings management by Polish households. The authors use also a logistic regression and cluster analysis. Decision trees were also used by [Dziechciarz–Duda et al. 2012]. Factors influencing tendency of households to benefit from social care were identified by means of the CHAID algorithm. The study of factors that determine wages in Poland were performed by [Kompa et al. 2013]. The authors utilized the decision trees based on the QUEST algorithm [Lim et al. 2000].

Noteworhty studies of households performer by [Beckel et al. 2013]. The authors use and compare four types of the classifiers: the *k*-Nearest Neighbor classifier, the Linear Discriminant Analysis classifier, the Mahalanobis classifier, and the Support Vector Machine (SVM). The new classification approach is presented by [D'Ambrosio 2001]. It is based on the index of social distance defined in the paper.

The problem of an identification of households with high incomes is not widely present in literature. However this problem seems to be interesting because of a distribution of high incomes differs from the distribution of remaining incomes. It is clearly visible in the case of personal incomes, where there is an exponential distribution for ca. 95% of incomes, while 5% of the highest incomes follow a power law. This behavior was showed for example for incomes in US and UK [Dragulesku et al. 2001] as well as for EU countries [Jagielski et al. 2013]. It is also known that commonly used economics models of incomes distributions e.g. Dagum and Shing-Maddala [Łukasiewicz & Orłowski 2004] do not describe the highest incomes although they are characterized by a high overall precision and fat tails [Łukasiewicz et al. 2012]. In this paper we deal with those issues and study characteristics of households with high incomes.

Classification trees (decision trees) are one of the methods of multidimensional data analysis, whose beginnings were around 60'ties of the XX century [Morgan & Sonquist 1963]. A very fast development of algorithms used in classification trees took place in eighties and nineties [Breiman et al. 1984, Quinlan 1993, Lim et al. 2000]. Nowadays, classification trees are included in many statistical packages and widely used in biology, sociology, medicine and, economy [Chrzanowska et al. 2009]. They are one of the statistical learning methods. One randomly chooses learning sample from a set of objects characterized by independent variables (attributes). Values of dependent variable (classes) must be known for each selected object. A hierarchy of attributes is determined and rules of splitting objects among subsets of homogeneous class composition are being set out. Based on results of the calculations a tree is constructed and its parameters are evaluated. The hierarchical structure is created, which is often presented graphically as an inverted tree with a root, nodes and leaves (terminal nodes). At this moment one can use the tree to classify other

objects into the classes. For more information see [Koronacki & Ćwik 2008, Gatnar 1998].

Classification trees have been implemented in the biggest decision making system in Business Intelligence (BI) infrastructure based on e.g. Oracle Data Mining, SQL Server Analysis Services (SSAS), SAS Enterprise Miner. In this paper classification trees included in a SSAS package have been used for classification of Polish households. There are implemented three algorithms: *Entropy (E), Bayesian with K2 Prior (BK2P), Bayesian Dirichlet Equivalent with Uniform Prior (BDEUP)* into a SSAS package. The first one is based on a Shannon's Entropy, which acts as a measure of classes' homogeneity in nodes and leaves. Algorithms based on entropy are also present in many other statistical packages. The other two algorithms - *BK2P* and *BDEUP* are newer technics, based on a Bayesian analysis [Cooper & Herskovits 1992, Heckerman 1995]. These algorithms are based on the probability theory, which is used to construct probabilistic networks called Bayesian belief networks. Generally, a Bayesian belief network is a pair (G, P), where G is a directed acyclic graph, and P is a conditional probability distribution of vertices of the graph [Jensen 1996, Olbryś 2007]. It might be interesting that each algorithm built in SSAS has three options: *Binary* (nodes are split into two subsets only, binary tree), *Complete* (nodes are split into maximum number of subsets based on the all possible values of the attribute) and, *Both* (during each split of a node a decision is being made, based on an effectiveness, which of the previous options to use).

DATA

In these studies microdata regarding budgets of households in 2008 have been analyzed. There were 37,107 households in the data set. Households were classified based on their 10 attributes (independent variables) in three groups: 1. variables describing a head of the household (a person with the biggest income); 2. variables describing a household as a whole; 3. variables describing location of a household. All attributes and their possible values are summarized in the Table 1. At this moment we only point out that the socio-economic group is defined as the main source of household's income.

A majority of households are employee's households (about 50%), on the other hand the smallest group consists of households maintained from non-earned sources (about 3.5%).

Table 1. The attributes of the households and their values

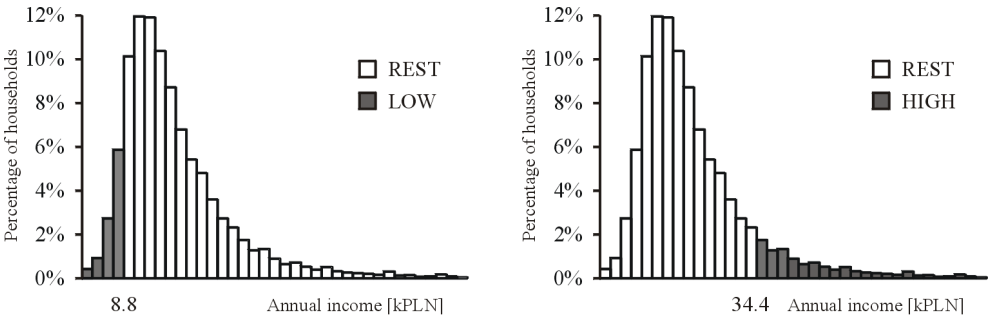
Group	Attribute	Attribute values
1	SEX (Sex of a household's head)	male (1), female (2)
	EDU (Education of a household's head)	tertiary (1), post-secondary (2), upper secondary vocational (3), upper secondary general (4), basic vocational (5), lower secondary (6), primary (7), no formal education (8)
	AGE (Age of a household's head)	16 - 102 (years)
	EGROUP (Economic group of a household)	employed in manual labor position (11), employed in non-manual labor position (12), farmer (2), self-employed (3), retired (41), pensioner (42), maintained from non-earned sources (5)
2	FTYPE (Family type)	marriage without children (1), marriage with 1 to 4 children (2 - 5), mother with children (6), father with children (7), marriage with children and other persons (8), mother with children and other persons (9), father with children and other persons (10), other persons with children (11), singles (12), others (13)
	NPER (Number of persons in a household)	1 – 15
	NCHIL (Number of children)	0 – 9
	NEAR (Number of earners)	1 – 10
3	PRES (Place of residence)	town >= 500 (1), town 200 - 499 (2), town 100 - 199 (3), town 20 - 99 (4), town < 20 (5), village (6) (thousands of residents)
	VOI (Voivodeship)	1 – 16

Source: own study

An annual income of a household is a dependent variable. We study household's income per number of earners, not per number of persons. We consider two cases with two distinct income classes:
(i) variable LOW-REST: LOW (10% of households with the lowest incomes), REST (remaining households);
(ii) variable REST-HIGH: HIGH (10% of households with the highest incomes), REST (remaining households).

Deciles groups (LOW, HIGH) and income limits are shown in the Figure 1.

Figure 1. Income distribution in Poland in 2008 and deciles groups



Source: own study

ANALYSIS OF SSAS ALGORITHMS

The full classification trees have been built based on the 30% random sample for each of the algorithms and options. The remaining part of the data was a validation set. The trees were constructed for each of the dependent variable. The effectiveness of the trees was measured by two parameters obtained for the validation set. The first one is a percentage of properly classified objects. The second one is a percentage properly classified LOW or HIGH objects. Results are in Table 2.

Table 2. The attributes of the households and their values

	Algorythm		
	Entropy	BK2P	BDEUP
Mode	Variable LOW-REST		
Binary	90.6 (9.9)	90.4 (7.0)	90.5 (9.2)
Complete	90.3 (13.8)	90.2 (11.5)	90.2 (11.8)
Both	90.7 (17.3)	90.3 (6.3)	90.5 (8.8)
Mode	Variable REST-HIGH		
Binary	90.9 (17.7)	90.6 (16.7)	90.6 (16.7)
Complete	90.4 (13.0)	90.2 (6.7)	90.0 (0.0)
Both	91.2 (28.9)	90.7 (15.6)	90.7 (15.6)

Source: own study

The first parameter is high (exceeds 90%) and has roughly the same values for all algorithms and options. This behavior is a result of proper classifications of REST objects. Thus the effectiveness of the trees was determined based on the second parameter. The best results were obtained for the Entropy algorithm with Both option for LOW-REST (17.3%) as well as for REST-HIGH (28.9%) variables. The tree for the second variable has a significantly better effectiveness that for the first one. It seems to be related to a shape the income distribution – high incomes, located in the tail of the distribution have much higher dispersion than low incomes. The algorithms based on bayesian networks have relatively low

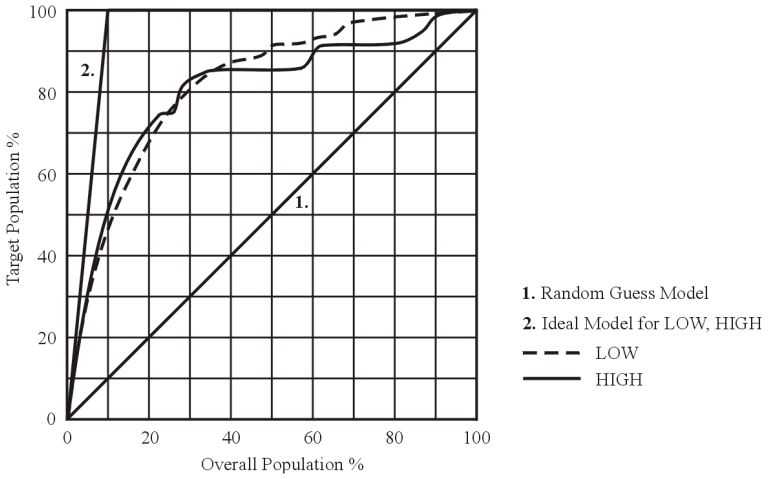
effectiveness. The second parameter has similar values for Binary and Both options, what has not been observed for the Entropy algorithm. All the algorithms with Complete option (bayesian networks in particular) have higher effectiveness for LOW than for HIGH objects.

The results for the Entropy algorithm were compared to the effectiveness of binary entropy based trees implemented in SAS Enterprise Miner package. The first parameter was about 90.6%, the second one was 11.3% for LOW-REST and 19.0% for REST-HIGH variable. The results of SAS are slightly better than for SSAS. However, in SSAS for entropy based trees one can obtain significant improvement of effectiveness by using the Both option.

ANALYSIS OF INCOMES

The following analysis concerns entropy based trees with Both option. Because of a large complexity of obtained trees they will not be presented in graphical form. We will present their global characteristic, attribute rank and characteristic of selected nodes and leaves. The shares of LOW and HIGH objects in a sample as a function of sample size are presented in Figure 2. The random and ideal models are also added to the plot.

Figure 2. Lift chart for LOW-REST and REST-HIGH trees



Source: own study

Using the constructed trees we analyze hypothetic sample data being a part of the data set (horizontal axis). The sample data set will contain percentages of LOW or HIGH objects indicated on the vertical axis. The effectiveness of identification is slightly better for HIGH than for LOW objects for the sample sizes from 0% to 25%. Above 25% the situation is reversed but from 30% increase of sample size causes only a small increase of the effectiveness. Both models become

less effective than random model. The best effectiveness is for small sample sizes. The ideal model reaches 100% for a sample size 10%, what means that all the LOW or HIGH objects are identified. On the other hand random model gives 10%. The trees identify 46% of LOW objects and 50% of HIGH objects in a 10% sample.

During the next stage of analysis the trees were trimmed by setting a node minimal support to 50. The attributes of objects ordered based on their significance are in Table 3. We observe the same three most important attributes for *LOW-REST* and *REST-HIGH* variables. The *EGROUP (Economic group of a household)* is the most important factor in the LOW or HIGH subgroup membership. Based on this attribute the groups of objects are split only on the first levels of the trees (complete split). The remaining two most important attributes are: *EDU (Education of a household's head)* and *NEAR (Number of earners)*. The attributes *EDU* and *NCHIL (Number of children)* have bigger influence on the identification of HIGH than LOW objects. The attributes *PRES (Place of residence)* and *FTYPE (Family type)* are more important in the case of the *LOW-REST* variable. The attribute *AGE (Age of a household's head)* is insignificant for both variables.

After the trimming the tree for *LOW-REST* variable had 58 leaves on 6 levels and the tree for *REST-HIGH* variable had 66 leaves on 8 levels. The majority of the nodes and leaves were of the REST type. A few but the most interested nodes and leaves with a majority of HIGH or LOW objects are described in Tables 4 and 5.

Table 3. The attributes of the households and their values

Variable *LOW-REST*		Variable *REST-HIGH*	
Attribute	Tree level	Attribute	Tree level
EGROUP	1	*EGROUP*	1
NEAR	2, 3, 4, 5	*EDU*	2, 3, 5, 6
PRES	2, 3, 5	*NEAR*	2, 3, 4, 5
EDU	3, 4, 5, 6	*NCHIL*	3, 4, 7
FTYPE	3, 4	*VOI*	3, 5
VOI	4	*SEX*	3, 4, 5, 6, 7
SEX	4, 5	*PRES*	2, 4, 7, 8
NPER	5	*NPER*	4, 6
NCHIL	3	*FTYPE*	6, 7
AGE	–	*AGE*	–

Source: own study

Table 4. Nodes and leaves with majority of HIGH objects

Node / Leaf	Attribute value	% HIGH
N 1	*EGROUP* = 3 & *NEAR* = 1 & *EDU* = 1	70.2
L 1.1	& *LOS* = 1	59.3
L 1.2	& *LOS* ≠ 1	80.8
N 2	*EGROUP* = 12 & *EDU* = 1 & *VOI* = 14 & *PRES* = 1	68.3
L 2.1	& *NEAR* = 1	83.8
N 2.2	& *NEAR* ≠ 1	60.8
L 2.2.1	& *SEX* ≠ 1	50.6
N 2.2.2	& *SEX* = 1	67.7
L 2.2.2.1	& *NCHIL* = 0	57.6
L 2.2.2.2	& *NCHIL* ≠ 0	76.8
L 3	*EGROUP* = 12 & *EDU* = 1 & *VOI* = 14 & *PRES* ≠ 1 & *SEX* = 1	50.1
L 4	*EGROUP* = 12 & *EDU* = 1 & *VOI* = 24 & *NEAR* = 1	50.1
L 5	*EGROUP* = 12 & *EDU* = 1 & *VOI* = 22 & *SEX* = 1	56.2

Source: own study

Table 5. Nodes and leaves with majority of LOW objects

Node / Leaf	Attribute value	% LOW
L 1	*EGROUP* = 5 & *PRES* = 5	54.3
N 2	*EGROUP* = 5 & *PRES* = 6	59.2
N 2.1	& *NEAR* ≠ 2	65.1
L 2.1.1	& *NEAR* = 1	56.4
L 2.1.2	& *NEAR* ≠ 1	73.4
N 3	*EGROUP* = 5 & *PRES* = 4	58.9
L 3.1	& *NCHIL* = 0	52.1
L 3.2	& *NCHIL* ≠ 0	68.1
L 4	*EGROUP* = 42 & *NEAR* ≠ 1 & *PRES* = 6	50.1

Source: own study

In the case of the *REST-HIGH* variable the first group (N 1) was obtained for households with family heads having a tertiary education, self-employed and being the only earner in the household (N 1). More than 70% of households in this group have a high income. We can distinguish two household's models: singles (L 1.1) and households with at least two members, but one earner (L 1.2).

The second group (N 2) are households with family heads employed in non-manual labor position, having a tertiary education and living in Warsaw (*VOI* = 14: mazowieckie voivodeship & *PRES* = 1). About 68% of households in this group have HIGH income. The split of this group into the smaller subgroups is showed in the Table 4. The groups of households characterized by *NEAR* = 1 or *SEX* = 1 (male) are created.

The L 3 group are households with male family heads employed in non-manual labor position, having a tertiary education, living in mazowieckie outside Warsaw. The remaining groups of households (L 4, L 5) are showed in Table 4. In this case we have the same split criteria: household's head employed in a non-manual labor position, has tertiary education and is the only earner or is a male. But those households are located in pomorskie (VOI = 22) and śląskie (VOI = 24) voivodeships. Note at this point that mazowieckie, pomorskie and, śląskie are the voivodeships with the highest mean income in Poland.

In the case of the *LOW-REST* variable there are tree distinguished groups of households with a predominance of LOW objects (L 1, N 2, N 3, see Table 5). The households are maintained from non-earned sources and are located in small towns: 20k–99k (*PRES* = 4), less than 20k (*PRES* = 5) and villages (PRES = 6). The split of N 2 and N 3 nodes into leaves is shown in Table 5. The forth group (L 4) are households of pensioners living in villages who do not maintain their households by themselves.

SUMMARY

In this paper a SQL Server 2008R2 Analysis Services package has been used for the classification of households based on their income. Households with the lowest (LOW) or highest (HIGH) incomes were identified. The effectiveness of the three classification tree algorithms was investigated. For each algorithm the results for three options were compared and discussed. The most effective was algorithm based on a Shannon's entropy with Both option. The effectiveness of identification of HIGH objects (28.9%) turned out to be better than for LOW objects (17.3%). The algorithms based on bayesian networks had lower effectiveness than other methods.

The Economic group of a household (*EGROUP*) was a major classification attribute for both dependent variables. Next most important attributes were: Education of a household's head (*EDU*), Number of earners (*NEAR*), Number of children (*NCHIL*) and Voivodeship (*VOI*) for *REST-HIGH* variable and Number of earners (*NEAR*), Place of residence (PRES), Education of a household's head (*EDU*) and Family type (*FTYP*) for *LOW-REST* variable. Similarly, the high importance of an economic group of households and education was observed by [Kozera et al. 2013]. However the authors also reported a number of persons in the households as an important attribute, what was not a case in our studies. In the other studies [Anioła et al. 2012]: education, family type and a place of residence were proved to be most important attributes.

Despite of relatively low effectiveness of obtained classification trees one distinguished some groups of households with a majority of LOW or HIGH objects. Groups of the biggest share of HIGH incomes are characterized by: incomes coming from a self-employment or non-manual labor position, tertiary

education of a household's head, one earner, mazowieckie, pomorskie or śląskie voivodeships. On the other hand groups of households with the highest share of LOW incomes are characterized by: incomes from non-earned sources or pension, location in small towns or villages.

Noteworthy is high significance of the Number of earners (*NEAR*) attribute. The studied incomes are the incomes per number of earners, so the *NEAR* attribute shall not be so important. Households in the three groups with a majority of HIGH incomes (N 1, L 2.1, L 4) are maintained only by their family heads. It seems to be a characteristic for at least a part of households with high incomes. On the other hand many of the households with low incomes are maintained by more than one person (L 2.1.1, L 4). However a quantitative evaluation of significance of *NEAR* attribute requires further studies.

REFERENCES

Anioła P., Gołaś Z. Zastosowanie wielowymiarowych metod statystycznych w typologii strategii oszczędnościowych gospodarstw domowych w Polsce, NBP, 2012.

Beckel Ch., Sadamori L., Santini S. (2013) Automatic Socio–Economic Classification of Households Using Electricity Consumption Data, e-Energy '13 Proceedings of the fourth International Conference on Future Energy Systems, 75–86.

Breiman L., Friedman J.H., Olshen R.A., Stone C.J. Classification and Regression Trees, Wadsworth, Belmont, CA, 1984.

Chrzanowska M., Alfaro E., Witkowska D. (2009) The individual borrowers recognition: Single and ensemble trees, Expert Systems with Applications 36, 6409-6414.

Cooper G.F., Herskovits E. (1992) A Bayesian Method for the Induction of Probabilistic Networks from Data, Machine Learning 9, 309-347.

D'Ambrosio C. (2001) Household Characteristics and the Distribution of Income in Italy: an Application of Social Distance Measures, Review of Income and Wealth 47, No 1, 43–64.

Dragulescu A.A., Yakovenko V.M. (2001) Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States, Physica A, 299, 213–221.

Dziechciarz-Duda M., Król A., Przybysz K. (2012) Minimum egzystencji a czynniki warunkujące skłonność do korzystania z pomocy społecznej. Klasyfikacja gospodarstw domowych, Taksonomia Nr 19 (242), 144–152.

Gatnar E. Symboliczne metody klasyfikacji danych, PWN, Warszawa, 1998.

Heckerman D. (1997) Bayesian Networks for Data Mining, Data Mining and Knowledge Discovery 1, 79-119.

Heckerman D., Geiger D., Chickering D.M. (1995) Learning Bayesian networks: the combination of knowledge and statistical data, Machine Learning 20, 197-243.

Jagielski M., Kutner R. (2013) Modelling of income distribution in the European Union with the Fokker–Planck equation, Physica A 392, 2130–2138.

Kompa K., Witkowska D. (2013) Application of Classification Trees to Analyze Income Distribution in Poland, Quantitative Methods in Economics, Vol. XIV, No. 1, 265–275.

Koronacki J., Ćwik J. Statystyczne systemy uczące się, Exit, Warszawa, 2008.

Kozera A., Stanisławska J., Wysocki F. (2013) Klasyfikacja gospodarstw domowych ze względu na stopień zaspokojenia ich potrzeb mierzony kategorią minimum socjalnego, Marketing i Rynek 11, 31-38.

Lim T., Loh W., Shih Y. (2000) A Comparison of Prediction Accuracy, Complexity, and Training Time of Thirty-three Old and New Classification Algorithms, Machine Learning, Vol. 40, 203–229.

Łukasiewicz P., Karpio K., Orłowski A. (2012) The Models of Personal Incomes in USA, Acta Physica Polonica A, Vol. 121, B-82–B-85 (2012).

Łukasiewicz P., Orłowski A. (2004) Probabilistic models of income distributions, Physica A 344, 146.

Morgan J.N., Sonquist J.A. (1963) Problems in the Analysis of Survey Data, and a Proposal, Journal of the American Statistical Association 58 (302), 415-434.

Olbryś J. (2007) Sieć bayesowska jako narzedzie pozyskiwania wiedzy z ekonomicznej bazy danych, Zeszyty Naukowe Politechniki Białostockiej, Informatyka 2, 93-107.

Quinlan J. R. (1993) C4.5: Programs for Machine Learning, Morgan Kaufmann, Los Altos.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XV, No. 2, 2014, pp. 414 – 428

RETURNS FROM THE ART MARKET. PRICE INDEX EVALUATED FOR THE MOST-TRADED POLISH PAINTERS[1]

Krzysztof Kompa
Department of Econometrics and Statistics,
Warsaw University of Life Sciences – SGGW
e-mail: krzysztof_kompa@sggw.pl
Dorota Witkowska
Department of Finance and Strategic Management
University of Lodz
e-mail: dorota.witkowska@uni.lodz.pl

Abstract: Relatively new approach toward investing is to concern pieces of art as investment instrument or collateral. In order to decide whether art is a good investment, it is necessary to evaluate expected returns, which might be obtained from such investment thus an art price index should be developed. The aim of the paper is to discuss artworks as investment assets and evaluate price index of paintings produced by 11 Polish artists whose artworks were traded the most often on auctions that held in Poland in the years 2007-2010. In our research, employing data concerning 750 objects, we apply hedonic index methodology to estimate returns from the paintings market. Our results justify the opinion that art can be treated as safe asset class, especially in comparison to equity market.

Keywords: investment, art market, hedonic price index

INTRODUCTION

Situation at financial markets tends investors to look for alternative investments to diversify their portfolio. Relatively new approach toward investing is to concern pieces of art as investment instrument or collateral. There is growing liter-

[1] Investing in Paintings on the Financial Market**,** research conducted under the National Science Centre Grant No. 2012/05/B/HS4/04188

ature concerning artworks as investment opportunity for investors: Frey, Pommerehne (1988), (1989a), Pesando (1993), Mei, Moses (2002), Worthington, Higgs (2003), (2004), Campbell (2004), (2008), Adamowska (2008), Higgs (2012), Hsieh, et al. (2010), Kraeussl, Wiehenkamp (2012), and Frey, Cueni (2013), to mention some.

In the year 2013, the global art market equaled 47.4 billion of Euro. It means that the global art and antiques market was almost back to the extraordinary heights of the pre-recession boom years powered by buyers in America and by rising prices for major Post-War and Contemporary artists[2]. The Polish art market is very small i.e. about 0.2% of the world sales - its value was estimated for 300-350 million PLN, while auction sales was 60.5 million PLN in 2012[3].

The aim of the paper is to evaluate the returns from the investment in Polish paintings market on the basis of the art price index. Commonly used methods to construct art price indexes are repeat-sales regression and hedonic regression [Ginsburgh, et al., 2006, p. 947]. Limited scale of the art market in Poland does not allow applying the repeated sale approach, therefore the hedonic method is employed. The research is provided on the basis of data concerning the selected sample of paintings sold at auctions held by auction houses and foundations in the years 2007-2010.

ART AS INVESTMENT ASSET

Artworks as investment assets are quite special. The current market value of art piece is difficult to evaluate since there is no "natural value", which could be used as reference or fair value. The price of artwork is limited only by the amount of money that collectors are willing and able to pay for [Goetzmann, et al. 2011]. Investment in art is undivided and illiquid in comparison to "classical" financial assets. Artworks are often expensive, also costs of purchase and sale may be high (from 10 to 25 percent of hammer price while costs on financial markets are about 1% of the price) and they seem to be long-lasting investments.

There are also several risks (regardless the risk of price variability of investment instruments) special for the art market [Frey, Cueni, 2013]. (1) Buyer can never be certain whether the purchased object is original (i.e. not a copy or forgery). And even if the artwork is an original piece, one never knows who was

[2] In the year 2013, USA accounted for 38% of the global market by value, China - 24%, the UK - 20%, France - 6%, Switzerland - 2%, Italy, Germany, Austria and Sweden - for 1% each of them, and the rest of the world – 6%. Taking into account Europe as a whole, United Kingdom created 63% of the market value, France - 19%, Germany - 5%, Italy – 3%, Austria – 2%, Sweden – 2%, Spain -1%, the Netherlands 1% and the rest of European Union 4%. The European Union as a whole kept 32% of the world market [McAndrew 2014].

[3] Deloitte 2013.

a painter: the master himself, the circle, the school or the painting is made only in style of a grand master[4]. (2) Quality of paintings, which have been repainted, damaged, not properly renovated or stored, may cause the declining of their values. (3) Owners of art pieces can be afraid that the masterpieces (that they poses) might be stolen, destroyed (by fire, revolution, etc.) or seized by government as a part of "national heritage". (4) Possession of valuable art can cause the necessity to pay additional taxes (sale or property taxes) and government may impose new export restrictions. (5) Tastes and fashions changes over a time and art market is characterized by extreme heterogeneity thus one never knows if the certain artist will not be "fallen from fashions" in the future. Renneboog and Spaenjers (2013) report 220 "fallen from fashions" artists who were included in the 1926, 1959 and 1980 edition of Gardner's *Art through the Ages*[5], but not in the 1996 or 2004 edition. (6) Art belongs to the group of luxury goods and it's price is very sensitive on general economic situation and income changes. Goetzmann et al. (2011) present the evidence that personal income of the highest earners determines price of art - one percent fall in income of the earners in the top 0.1 percent income distribution in the UK triggers a decline in art prices of nearly 10 percent. (7) Behavioral anomalies seem to play an important role in the art market since collectors are usually not willing to sell pieces of art from their collections and they tend to buy art produced by domestic artists.

In order to decide whether art is a good investment, it is necessary to evaluate expected returns which might be obtained from such investment thus an art price index should be developed. The are three major goals behind construction of art price indexes is ([Ashenfelter and Graddy 2006], [Ginsburgh et al., 2006]): (1) to measure financial performance of art, relative to other alternative forms of investment, (2) to check whether adding art to a diversified market portfolio can lower the overall risk and/or increase the rate of return, and (3) to outline general trends on the art market.

Construction of the art price index dedicated to the certain market requires several decisions concerning: selection of the index methodology and selection of the sample i.e. its size and the way of the object selection. There are several methods and indexes which can be used to analyze changes of prices or returns from the art market, such as naive price index, repeat-sales, average price (geometric mean), composite price (basket) index, and hedonic index. Pesando (1993), Pesando, Shum (1999), Mei, Moses (2002) employ repeat sales, Mok et al. (1993), Landes (2000) – geometric repeat sales, Ginsburgh, Schwed (1992), Kraussl, van Elsland (2008), Higgs, Warthington (2005) – hedonic index, Candela et al. (2004) – quality adjusted price, Renneboog, van Houtte (2002) –basket index.

[4] See (Frey and Pommerehne, 1989b) describing the story of the painting "Daniel in the Lion's Den".

[5] *Gardner's Art Through the Ages* is an American textbook on art history that was written by Helen Gardner (1878 – 1946) and published for the first time in 1926.

The repeat sales methodology considers only those artworks that were sold at least twice in the investigated period, and their price were registered. Thus it requires transparent and liquid market. Ginsburgh et al. (2006) argue that repeat-sales regression should not be applied to periods shorter than 20 years, since the number of observations may be too small to obtain reliable results. The main advantage of hedonic index approach is that it may consider all traded objects.

According to the Deloitte (2013) report, average annual return from 800 repeat sales, that took place in Poland during last 20 years, was 25.7% while in the same time equity returns measured by Warsaw Stock Exchange Index WIG20 was only 8.7%[6]. Annual return from artworks hold longer than 15 years was 46.6% while investments with the horizon shorter than 5 years gave only 0.2% profit. Thus the time span of investments is crucial in obtained returns.

Naive art price indexes are constructed using average and median auction prices. In this method, a basket of representative paintings is created and tracked over time. This approach avoids the repeated sales restrictions but requires instead aggregation on some a priori criteria in order to construct so-called "average painting" [Candela, Scorcu 1997].

The problem of the sample selection is crucial in construction the fixed basket of representative artworks that must be made by experts. However it appears also in application other methods and it is known from literature since many different criteria of the artwork selection are employed. In the literature the most often selected criteria are: (1) mediums (for example Locatelli-Biey, Zanola (2002) consider sculptures, Pesando (1993) – prints, Candela et al. (2004) – paintings, Ginsburgh, Schwed (1992) – drawings); (2) style of the piece of art or the period when the artwork came into being (for instant Pesando (1993) concerns modern prints and Picasso prints, Candela et al. (2004) – Modern and contemporary, 19-th century, Old Master paintings, Ginsburgh, Schwed (1992) - Flemish-Dutch, French, Italian Old Master drawings); (3) authors selected by the experts (as an example is a paper by Pesando, Shum (1999) who analyze Picasso prints); (4) nationality of artists or place where they created their artworks (for instance: Mok et al. (1993) consider Modern Chinese paintings, a Kraussl, van Elsland (2008) – German paintings in general, Ginsburgh, Schwed (1992) - Flemish-Dutch, French, Italian drawings, Higgs, Warthington (2005) – Australian paintings), and (5) belonging to the certain collection (as it is done by Landes (2000) who takes Ganz collection).

Such arbitrary choice may be criticized because selected artworks may not be representative for the whole market. The size of the sample depends on the investigated period i.e. time span when artworks are sold.

[6] See [Gajewski, Potocki, 2013].

HEDONIC ART PRICE INDEX

Art price index can be evaluated employing hedonic regression, which is a method for estimating an approximate value of a piece of art adjusting the average price of the artist's works for the qualitative characteristics which are incorporated into the hedonic model. Hedonic price functions are used to predict prices of new goods, adjust for quality change in price indexes and to measure consumer and producer valuations of different products. They describe the relationship between characteristic of a product and its price. The analytic function is determined by distributions of buyers and sellers and their preferences as well as the structure of competition in the market [Neiheim 2006].

Artworks are heterogeneous assets, with a variety of physical and non-physical characteristics that make them unique, including artist reputation, materials used, the period of production and subjective traits like quality. Therefore the price of an artwork depends on these characteristics. In hedonic approach[7], the value attached to each one of the attributes, that are deemed to be significant in the determination of the price, is estimated and the price index is evaluated employing so called hedonic quality adjustment (HQA). The basic formula for hedonic index (*HI*) is as following:

$$HI_{t+1} = \frac{NI_{t+1}}{HQA_{t+1}} = \frac{\prod_{i=1}^{n}(P_{i,t+1})^{1/n} \Big/ \prod_{i=1}^{m}(P_{i,t})^{1/m}}{HQA_{t+1}} \tag{1}$$

where at time *t*: NI_t – naive price index, P_{it} - the price of the *i*-th artwork, HQA_t – hedonic quality adjustment.

The naive price index (NI) describes "average painting" from the aggregation of all artworks that create the sample representing the art market or it's segment. The hedonic index allows the tracking of short-term price movements and returns in this market. It let us a better understanding of how this market performed over the investigated period. The performance of this market can then be compared to investments in traditional financial assets.

Changes in quality, style, mode and type of the artworks is described by the hedonic quality adjustment which is defined as following:

$$HQA_{t+1} = \exp\left[\sum_{j=1}^{k}\hat{\alpha}_j\left(\sum_{i=1}^{n}\frac{X_{ij,t+1}}{n} - \sum_{i=1}^{m}\frac{X_{ij,t}}{m}\right)\right] \tag{2}$$

[7] Hedonic price indexes are discussed by Nesheim (2006), Triplett (2006) while their application on the art market by Candela, et al. (2004), Kraeussl, van Elsland (2008), Kraeussl, Wiehenkamp (2012) to mention some research provided for developed art markets. However the first attempt to construct hedonic art price indexes for emerging markets was made by Kraeuss, Logher (2008) who consider art markets in China, Russia and India.

where $X_{ij,t}$ - observations of the *j*-the feature characterizing the *i*-th artwork at time *t*, *m* and *n* are numbers of lots sold in the period *t* and *t*+1 respectively, $\hat{\alpha}_j$ - parameter estimates of hedonic regression (pooled regression) that usually takes the following form:

$$\ln P_{i,t} = \alpha_0 + \sum_{j=1}^{k} \alpha_j X_{ij,t} + \sum_{t=1}^{\tau} \beta_t Z_{i,t} + \varepsilon_{i,t} \tag{3}$$

where α_j, β_t - the coefficient values of the characteristic X_j and Z_t, Z_t - time dummy variable, which takes the value 1 if the painting *i* is sold in the period *t* and takes the value 0 otherwise, and ε_{it} - the disturbance term.

One of the underlying assumptions is that the price of an artwork depends essentially on its quality, which is to a great extent quantifiable. The explanatory variables represent characteristics of the objects, such as the artist, size, format, technique, materials, period, subject of the artwork, signature and artist's living status or related to the sale, including auctioneer, location and date of sale. These attributes are usually qualitative so they are represented by dummy variables in (3). The dependent variable in hedonic models is usually the natural logarithm of the sales price. All auctions relating to an artist are included in the calculation in order to avoid selection bias. The time dummy variables can be annual, semi-annual, quarterly or even monthly depending on the frequency of trading. The hedonic approach essentially entails running an Ordinary Least Squares (OLS) regression.

The hedonic regression method therefore controls for quality changes by attributing implicit prices to a set of value-adding characteristics. In other words, hedonic regressions strip the observable characteristics from the artworks to obtain an index reflecting the price of a "standard artwork" [Renneboog, Van Houtte 2002].

Hedonic models require knowledge of the artworks' characteristics and mechanisms driving art prices. The main weakness of this methodology is limitation of attributes used as explanatory variables and arbitrariness of their selection. Therefore application of hedonic index methodology requires decision about:

a) the form of the price index[8], for example Kraeussl, van Elsland (2008) use formula (1) while Higgs (2012) defines hedonic price index in different way;
b) the selection of explanatory variables in the regression model (3) that is connected with the information describing artworks being subject of transaction,
c) the selection of the artwork sample that will be used for art price index evaluation in formulas (1)-(2).

EMPIRICAL RESULTS

The art market in Poland practically did not exist under the communist regime. Although in that time several art galleries and artist associations were operat-

[8] See [Widłak 2010] who presents different forms of the price index.

ing buying and selling art pieces, mostly currently produced by domestic artists or the one from other soviet bloc countries. Art market in Poland started to normalize and develop at the beginning of political and economic transformation in 1989 when new art galleries and foundations together with auction market have been created. In years 1989-2012, number of art auctions increased from 8 to 122. It is worth mentioning that last two decades were characterized by changes in the income distribution and the structure of consumption that caused the increasing of the interest on art market in the Polish society.

Table 1. Structure of Polish art market in the first half of 2012 by mediums and segments

Mediums	Lots [%]	Value [%]	Segments	Lots [%]	Value [%]
Sculpture	2	1	Post-war and contemporary	30	31
Photography	1	0	Ultra-contemporary	44	8
Arts & Crafts	10	5	Art before 1945	26	61
Graphic	14	2			
Drawing	17	20			
Painting	56	72			

Source: own elaboration on the basis of data from [Gajewski, Potocki, 2013, pp. 18-19].

Structure of the art market in Poland in 2012 is presented in Table 1. Paintings are the most popular medium of the Polish art market both in number of lots (56%) and value of transaction (72%). Therefore paintings may be used as representative segment of the Polish art market. "Ultra-contemporary art", which describes young artists (under 40 years old) is distinguished since it creates the majority (44%) of artworks sold but they characterize by low prices, and creates only 8% of total value.

Table 2. Structure of the database: paintings sold on auctions in years 2007-2010

Structure of transactions in years	2007	2008	2009	2010
Number of lots sold	2493 (24%)	2548 (25%)	2427 (23%)	2932 (28%)
Value of transactions [mln PLN]	39.22 (24%)	58.71 (37%)	36.71 (23%)	25.68 (16%)

Source: own elaboration

Hedonic models are estimated employing data[9] from auctions of paintings that took place in Poland in the years 2007-2010. The whole sample contains 10,400 objects produced by nearly 3,000 artists who represent different periods and styles, as well as a great variety of techniques. Taking into account number of lots sold, one should notice that situation is stable i.e. every year these numbers are similar (Table 2). However the highest value of transaction was observed in 2008, and in 2010 it equaled less than 50% of the sales in 2007.

[9] The basic data base from auctions of paintings was constructed by Lucińska (2012).

The range of prices is huge: from 20 PLN for a piece created by J. Jakóbowska (born in 1984) to 1.1 million PLN for an artwork by W. Czachórski (1850 - 1911), with average price for a single lot 8,691 PLN and standard deviation 33,698 PLN. Therefore here a question arises which traded objects should be taken into account and what is the minimal price for a work to treat is as a piece of art.

Description of data and variables

We decided to consider only artworks, painted by the Polish artists who are selected according to the biggest number of lots sold in the investigated period[10], and we assume that minimal average price for an artist should be 2,000 PLN. Employing these criteria the sample of paintings is constructed. It consists of 750 paintings created by 11 artists (see Table 3).

Table 3. List of Polish painters whose artworks created the sample

No.	Author	Count	Value [PLN]	Average value [PLN]	Year of	
		of artworks sold in 2007 - 2010			birth	death
1	Chmieliński Stachowicz Wł.	55	648 200	11 786	1911	1979
2	Dominik Tadeusz	46	608 000	13 217	1928	-
3	Dwurnik Edward	63	431 300	6 846	1943	-
4	Erb Erno	58	816 500	14 078	1890	1943
5	Kossak Wojciech	60	2 027 500	33 792	1856	1942
6	Wyczółkowski Leon	61	3 848 300	63 086	1852	1936
7	Hofman Wlastimil	85	1 817 050	21 377	1881	1970
8	Kossak Jerzy	91	1 261 000	13 857	1886	1955
9	Malczewski Jacek	71	9 401 300	132 413	1854	1929
10	Nikifor Krynicki	79	196 400	2 486	1895	1968
11	Nowosielski Jerzy	81	5 706 700	70 453	1923	2011
	Sum	750	26 762 250			

Source: own elaboration

The biggest number of lots sold in analyzed period were produced by Jerzy Kossak (91) while the highest value of transactions concerns artworks by Malczewski (more than 1 million PLN). In our sample, the lowest average value for the single artwork obtained paintings by Nikifor (2486 PLN). The selected sample covers 7.2% of all lots and 16.2% of the turnover registered in the database. Table 3 contains basic information about "the most liquid" painters whose artworks are used in our research.

In our investigation we select several explanatory variables which are usually applied in hedonic models constructed for the art price. They describe artist's

[10] In the paper [Kompa, Witkowska 2013] sample selection provided due to different criteria is discussed.

and exhibitor's reputation, type and quality of the artwork as well as conditions of the transaction. *Auction house* describes the reputation of auctioneer and this variable is specified as a number of dummies defined name of auctioneer. In our models there are 9 variants of this variable, and the reference variant is: "other auctioneers". Artist reputation is defined by the name of a painter that is represented by the variable *artist* (see Table 3), and "Wyczółkowski" is the reference painter, among 11 artists. Artist's living status is often incorporated to hedonic models since when an artist dies the production stops and prices may rise. We assume that variable *living status* equals 0 if artist is still living when and auction takes place.

Type and quality of the art piece is described by several variables such as: *signature*, *technique* and *surface* of the painting. Technique and materials characterize type of work and this variable is specified as a number of dummies that indicate whether the art piece represents certain type of work. We distinguished 9 variants of the variable and the reference variant is: "other techniques". *Signature* is one of the artworks attributes, it equals 1 if signature is visible. *Surface* [cm^2] of the artwork is the most commonly used variable that describes the physical characteristics of painting. In general the parameters estimates for this variable should be positive however larger works may be difficult to display thus in some models squared surface is applied. In the model we use natural logarithms of surface area or of squared surface.

Conditions of the transaction is represented by two variables: *year* and *price relation*. Year of sale is a set of binary variables defined the year of transaction. In our research we consider 4 years, and the reference variant of this variable is "Year_2010". *Price relation* between reserve and hammer price, this variable equals 1 if the former is bigger than the letter since in such a case sale might not take place (so-called conditional sale).

Hedonic regression

In this paper we present four selected models that are estimated employing OLS method on the basis of described above sample. The models' specification is presented in Table 4, where symbol ● denotes variables present in the certain model, numbers in parenthesis informs about number of variants that are statistically significant for the significance level $\alpha=0.05$, + informs that qualitative or binary variables are significant with positive sign of the parameter estimates.

Models M3 and M4 are characterized by the highest adjusted R^2 and negative value of Akaike's information criterion, which make them the best models from the point of view their statistical characteristics. The specification of the models M1 and M2 is nearly the same, except presence of the variable: *price relation* thus their characteristics are similar. One may notice that the models M2 and M3 differs one from another only by representation of the painting's size (logarithm *surface area* or logarithm *surface area*^{2}) but their statistical properties are completely different that is especially visible comparing Akaike's criterion. While

the models M3 and M4 differ by two variables and the constant term but their characteristics are nearly the same.

Table 4. Comparison of models' specification

Models	M1	M2	M3	M4
const	● +	● +	● +	
year	● (1)	● (1)	●	●
auction house	● (4)	● (4)	● (8)	● (7)
artist	● (9)	● (9)	● (9)	● (9)
signature	●	●	●	●
technique	● (5)	● (5)	● (2)	● (2)
price relation	●	●		
surface area	● +	● +		
surface area^2		● +	● +	
living status		● +		
R^2 adjusted	0.8114	0.8115	0.9953	0.9953
F	101.68	105.07	5071.98	4910.11
Akaike	1269.10	1267.30	-1494.40	-1492.90

Source: own elaboration

Table 5 contains parameter estimates of selected models, stars denote significance level: * α=0.10, ** α=0.05 and *** α=0.01. Artist's *living status* is significant variable with positive sign in model M4. In all models variable *signature* is not significant, as well as nearly all variants of the variable: *year*. Parameter estimates standing by names of all painters, except Malczewski, are negative that is correct since only Malczewski's artworks obtained higher prices than Wyczółkowski. Positive parameters standing by names of auctioneers is also proper because selected auction houses are well-known and seem to be trustworthy thus they organize auction which are selected more often when valuable artworks are subjects of transactions. Taking into consideration sign and significance of selected techniques we notice that oil paintings are usually more expensive than the ones prepared using other techniques.

Table 5. Parameter estimates

Name of variable	Variable variants	Model M1	Model M2	Model M3	Model M4
		Parameter estimates			
const.		**2,788 *****	**2,793 *****	**5,040 *****	
year	YEAR_2007	0,093	0,097	0,008	0,007
	YEAR_2008	0,076 **	0,077 **	0,001	0,001
	YEAR_2009	0,015	0,015	0,002	0,002
auction	AGRAART	0,295 **	0,299 **	0,084 ***	0,083 **

cont. on the next page

Table 5. (cont.) Parameter estimates

house	Desa	0,199	0,201	0,085 ***	0,084
	Desa Unicum	0,408 ***	0,403 ***	0,054 **	0,055 ***
	Okna Sztuki	0,480 ***	0,482 ***	0,071 **	0,070 **
	Ostoya	0,100	0,103	0,065 ***	0,064 ***
	Polswiss Art	0,805 ***	0,797 ***	0,070 ***	0,072 ***
	Rempex	0,090	0,081	0,058 ***	0,061 ***
	Rynek Sztuki	0,017	0,008	0,048 **	0,050 **
signature		-0,044	-0,046	-0,004	-0,004
price relation		-0,027			**-0,007**
surface area		0,565 ***	0,564 ***	**0,048 *****	**0,048 *****
surface area^2				0,564 ***	
artist's living status					**5,041 *****
artist	Kossak_J	**-1,591 *****	**-1,418 *****	**-0,057 *****	**-0,057 *****
	Kossak_W	-0,877 ***	-0,686 ***	-0,032 *	-0,032
	Chmieliski	-1,227 ***	-1,020 ***	-0,060 ***	-0,060 ***
	Dwurnik	-2,282 ***	-2,174 ***	-0,141 ***	-0,141 ***
	Erb	-1,091 ***	-0,812 ***	-0,041 **	-0,042 **
	Hofman	-1,088 ***	-0,848 ***	-0,048 **	-0,048 **
	Malczewski	0,312 ***	0,535 ***	-0,100 ***	-0,101 ***
	Nikifor	-1,333 ***	-0,874 ***	-0,255 ***	-0,256 ***
	Nowosielski	-0,119	0,072	-0,047 ***	-0,047 ***
	Dominik	-1,905 ***	-1,854 ***	-0,060 ***	-0,059 ***
Technique	watercolour	0,197	0,553	-0,015	-0,016
	acrylic	0,698 ***	0,891 ***	0,045	0,045
	gouache	0,285	0,641	-0,002	-0,004
	oil	0,886 ***	1,043 ***	0,056 **	0,056 **
	pencil	-0,246	0,235	-0,070 **	-0,070 **
	pastel	0,450 **	0,746 **	0,034	0,034
	tempera	0,635 ***	0,997 ***	0,030	0,030
	drawing ink	-0,598 **	-0,275 **	-0,017	-0,017

Source: own elaboration

Hedonic art price indexes

In our research we evaluate naive and hedonic price indexes. Parameter estimates of the models (3), presented in Table 5, are used to evaluate hedonic quality adjustment (2), and, finally, price index (1). In Table 6 we present obtained results. Naïve price index is a nominator in relation (1), and describes "average" changes of prices regarding selected artworks in every investigated year. Hedonic quality adjustments (HQA) is evaluated separately for the estimated models therefore also hedonic price indexes depends on the parameter estimates.

Table 6. Hedonic art price indexes

Type of index	Model	Year	HQA	Art price Index	Equity market index*	
					WIG	WIG20
Naïve		2007		1.0000	55 648.54	3 456.05
		2008		1.4984	27 228.64	1 789.73
		2009		0.6163	39 985.99	2 388.72
		2010		0.9441	47 489.91	2 744.17
Hedonic	M1	2008	1.4137	1.0599		
		2009	0.6867	0.8975		
		2010	0.9860	0.9575		
	M2	2008	1.4160	1.0582		
		2009	0.6862	0.8981		
		2010	0.9887	0.9549		
	M4	2008	1.3256	1.1303		
		2009	0.8552	0.7206		
		2010	0.9088	1.0389		

Source: own elaboration and *http://www.gpw.pl/indeksy_gieldowe

Table 7. Percentage returns from different investment assets

Investment in art			Returns			Returns		
Index type	Model	Year	comparison to the		Average annual	comparison to the		Average annual
			(t-1)	2007		(t-1)	2007	
Naïve		2008	49.8	49.8	-4.5	Warsaw Stock Exchange Index WIG		
		2009	-38.4	-7.7				
		2010	-5.6	**-12.8**				
Hedonic	M1	2008	6.0	6.0	-3.1	-51.1	-51.1	-5.1
		2009	-10.3	-4.9		46.9	-28.1	
		2010	-4.3	**-8.9**		18.8	**-14.7**	
	M2	2008	5.8	5.8	-3.2	Warsaw Stock Exchange Blue Chip Index WIG20		
		2009	-10.2	-5				
		2010	-4.5	**-9.2**				
	M4	2008	13.0	13.0	-5.4	-48.2	-48.2	-7.4
		2009	-27.9	-18.6		33.5	-30.9	
		2010	3.9	**-15.4**		14.9	**-20.6**	

Source: own elaboration

Last two columns in Table 6 contain quotations of the market indexes from the Warsaw Stock Exchange (WIG is total return index while WIG20 is blue chip index). Having quotations of stock indexes we may calculate returns that could be

obtained investing on equity market and compare it to the investment on art market.

Therefore we calculate percentage returns for each year (Table 7) as: (1) annual return i.e. in comparison to the previous year (t-1), (20 total return i.e. in comparison to the year 2007, and (3) average annual return as geometric mean from the total return obtained in the year 2010.

It is visible that both equity indexes generated losses in every year of investigation while negative returns on the art market appear in 2009 and 2010. It worth mentioning that naïve price index and hedonic indexes evaluated on the basis of the models M1 and M2 show smaller average annual losses than stock index WIG. While all of them have generate smaller losses than index WIG20.

CONCLUSIONS

Analysis provided in the paper shows that the art market in Poland has been developing since 1989. This market is still immature however it may be attractive for the investors. Presented results justify the opinion that artworks, created by the well-known Polish painters, can be treated as safe asset class, although in the investigated period the returns from treasury bonds were higher (- annually from 4% to 5.75% which depends on mature of the bonds that is from 2 to 10 years). Comparing the returns from the investment in art, represented by hedonic or naive indexes, to returns from WIG and WIG20 we notice that losses from investment in artworks are smaller. We also notice that financial crisis is visible in the Polish equity market in 2008 while in the art market a year later.

Taking into consideration the construction of hedonic regression models, it is visible that the model specification essentially influences hedonic quality adjustment hence it affects evaluation of the art price index which is the artworks' prices proxy. Thus it might be convenient to employ aggregated indexes as it is proposed by Witkowska 2014.

REFERENCES

Adamowska W. (2008) Is Art. such a good investment? Investing in fine art. on the international and Polish action market. Master thesis written under the supervision of F. Vermeylen, Erasmus University, Rotterdam.

Ashenfelter O., Graddy K. (2006) Art Auctions, [in:] Ginsburgh, V., Throsby, D. (eds.), Handbook of the Economics of Arts and Culture, Elsevier, Amsterdam: 909-945.

Campbell R.A.J. (2004) The Art of Portfolio Diversification, LIFE Working Paper, No. 04-009, Maastricht University. Retrieved May 19, 2011, from http://www.google.pl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CCoQFjAB&url=http%3A%2F%2Fweb.econ.ku.dk%2Ffru%2Fconference%2FProgramme%2FSaturday%2FD4%2FCampbell_Art%2520InvestmentCampbell.pdf&ei=K7ShVJPuJ5PU7Ab8xYHwDg&usg=AF

QjCNGVAT2ofx_F0N76v2nGuDg8bA18tw&sig2=CexJGPxmHTwxqWOmMRNaCg&bvm=bv.82001339,d.ZGU

Campbell R. A. J. (2008) Art as a Financial Investment, Journal of Alternative Investments, 10(4), 64-81.

Candela G., Scorcu A. E. (1997) A Price index for Art Market Auctions. An Application to the Italian Market of Modern and Contemporary Oil Paintings, Journal of Cultural Economics, 21, 175-196.

Candela G., Figini P., Scorcu A. E. (2004) Price indices for artists - A proposal, Journal of Cultural Economics, 28(4), 285-302.

Culture in 2012 (2013) Wydawnictwo GUS, Warsaw.

Deloitte (2013) Rynek sztuki. Sztuka rynku, http://www.deloitte.com/assets/ DcomPoland/Local%20Assets/Documents/Raporty,%20badania,%20rankingi/pl_Art&Banking_PL.pdf (last accessed 12.11.2013).

Frey B.S., Cueni R. (2013) Why invest in art?, The Economist's Voice 2013, 1-6.

Frey B.S., Pommerehne W.W. (1988) Is Art Such a Good Investment?, The Public Interest, 91, 79-86.

Frey B.S., Pommerehne W.W. (1989a) Art Investment: An Empirical Inquiry, Southern Economic Journal, 56(2), 396-409.

Frey B.S., Pommerehne W.W. (1989b) Muses and Markets. Explorations in the Economics of the Arts, Oxford: Basil Blackwell, Cambridge, MA.

Gajewski M., Potocki T. (2013) Skate's Focus: Poland's Art Market The Rising Star of Central Europe, report created by Skate's Art Market Research Art & Business Magazine S.A., Skate's, LLC, New York, NY, USA, skatesart.files.wordpress.com/2013/03/skates-focus-polands-art-market.pdf., last accessed 29.01.2014.

Ginsburgh V., Mei J. and Moses M. (2006) The Computation of Prices Indices, Handbook of the Economics of Art and Culture, chap. 27 (947-979).

Ginsburgh V., Schwed N. (1992) Price Trends for Old Masters' Drawings: 1980-1991, [in:] The Art Newspaper, September.

Goetzmann W.N., Renneboog L., Spaenjers C. (2011) Art and money, American Economic Review, 101, 222-226.

Higgs H. (2012) Australian Art Market Prices during the Global Financial Crisis and two earlier decades, Australian Economic Papers, 51(4), 189-209.

Higgs H., Worthington A. (2005) Financial Returns and Price Determinants in the Australian Art Market 1973–2003, Economic Record, Vol. 81(253), pp. 113-123.

Hsieh S., Lee J., Tzeng L. Y. (2010). Art as an Investment: Empirical Study of Asian Contemporary and Chinese 20th Century Modern Art, International Conference on Finance Department of Finance, National Taiwan University, discussion paper, http://www.fin.ntu.edu.tw/~conference/conference2010/proceedings/proceeding/11/11-1(A40).pdf (viewed 7.11.2013).

Kompa K., Witkowska D. (2013) Art Price Index. Preliminary Investigation for Selected Polish Painters (in Polish), Zarządzanie i Finanse, Vol. 11, No. 3, Part 2, 33-50.

Kraeussl R., van Elsland N. (2008) Constructing the True Art Market Index - A Novel 2-Step Hedonic Approach and its Application to the German Art Market, CFS working paper, Center for Financial Studies Frankfurt University, No. 2008/11.

Kraeussl R., Logher R. (2010) Emerging Art Markets, Emerging Markets Review, 11(4), 301-318.

Kraeussl R., Wiehenkamp Ch. (2012) A call on Art investments, Review of Derivatives Research, 15(1), 1-23.

Landes W.M. (2000) Winning the Art Lottery: The Economic Returns to the Ganz Collection, John M. Olin Law and Economics Working Paper No. 76, Second series, The Law School, University of Chicago, Chicago.

Locatelli-Biey M., Zanola R. (2002) The Sculpture Market: An Adjacent Year Regression Index, „Journal of Cultural Economics, vol. 26 (1), 65-78.

Lucińska A. (2012) Rozwój artystyczny malarzy a ceny obrazów na polskim rynku sztuki, [w:] Zarzecki D. (ed.), Finanse, Rynki Finansowe, Ubezpieczenia 51, Zeszyty Naukowe Uniwersytetu Szczecińskiego, 715-725.

Mc Andrew C. (2013) The global art market, with a focus on China and Brazil, TEFAF Art market report 2013, The European Fine Art Foundation, TEFAF Maastricht.

Mc Andrew C. (2014) TEFAF Art market report 2014, The European Fine Art Foundation, TEFAF Maastricht.

Mei J., Moses M. (2002) Art as an Investment and the Underperformance of Masterpieces, American Economic Review 92 (5), 1656-1668.

Mok H.M.K., Ko V.W.K., Woo S.S.M., Kwok K.Y.S. (1993) Modern Chinese Paintings: An Investment Alternative, [in:] Southern Economic Journal, vol. 59(4), 808-816.

Nesheim L. (2006) Hedonic Price Functions, CEMMAP, London.

Pesando J.E. (1993) Art as an Investment: The Market for Modern Prints, American Economic Review, 83, 1075-1089.

Pesando J.E. and Shum, P.M. (1999) The Returns to Picasso's Prints and to Traditional Financial Assets, 1977 to 1996, Journal of Cultural Economics, 23(3): 183-192.

Renneboog L. & Van Houtte T., (2002). The monetary appreciation of paintings: From realism to Magritte, Cambridge Journal of Economics, No. 26, 331-357.

Renneboog L., Spaenjers C. (2013) Buying Beauty on Prices and Returns in the Art Market, Management Science 59(1), 36 – 53.

Triplett J. (2006) Handbook On Hedonic Indexes And Quality Adjustments In Price Indexes: Special Application To Information Technology Products, OECD Publishing, Paris.

Witkowska D. (2014) Sensitivity of Hedonic Price Indexes an Example of the Art Price Indexes Evaluated for Well-Known Polish Painters (in Polish), Ekonometria, 2(44), Wrocław.

Worthington A. C., Higgs H. (2003) Art as an Investment: Short and Long-Term Comovements in Major Paintings Markets, Empirical Economics, 28, 649-668.

Worthington A.C., Higgs H. (2004) Art as an Investment: Risk, Return and Portfolio Diversification in Major Painting Markets, Accounting and Finance, 44, 257-271.

REVIEWERS COOPERATING WITH THE JOURNAL IN 2014

Hanna Adamkiewicz-Drwiłło
Aleksandra Baszczyńska
Jarosław Becker
Jacek Bednarz
Piotr Bilski
Zbigniew Binderman
Dariusz Błaszczuk
Lucyna Błażejczyk-Majka
Bolesław Borkowski
Jadwiga Bożek
Ryszard Budziński
Mariola Chrzanowska
Wiesław Dębski
Robert W. Dimand
Hanna Dudek
Marcin Dudziński
Gema Fernández-Avilés Calderón
Jennifer P. Foo
Peter Friedrich
Konrad Furmańczyk
Paolo Gajo
Stanislaw Gędek
Vasile Glavan
Marek Gruszczyński
Urszula Grzybowska
Mariusz Hamulczuk
Stanisław Jaworski
Andrzej Karpio
Krzysztof Karpio
Marek Karwański
Stanisław Kasiewicz
Joanna Kisielińska
Stephan Klasen
Paweł Kliber
Marek Kociński
Krzysztof Kompa
Yuriy Kondratenko
Vassilis Kostoglou
Grzegorz Koszela
Stanisław Kot
Robert Kragler
Monika Krawiec
Tadeusz Kufel
Justyna Kujawska
Karol Kukuła
Tadeusz Kwater
Joanna Landmesser
Barbara Z. Liberda
Andrzej Łodziński
Piotr Łukasiewicz
Wanda Marcinkowska-Lewandowska
Jerzy Marzec
Aleksandra Matuszewska-Janica
Jose-María Montero Lorenzo
Rafik Nafkha
Magdalena Osińska
Maria Parlińska
Mariola Piłatowska
Marian Podstawka
Iwona Pomianek
Artur Prędki
Alexander Prokopenya
Mariusz Próchniak
Ryszarda Rempała
Janina Sawicka
Yochanan Shachmurove
Mirbulat B. Sikhov
Agnieszka Sompolska-Rzechuła
Stanisław Stańko
Aleksander Strasburger
Jacek Strojny
Ewa Syczewska
Wiesław Szczesny
Waldemar Tarczyński
Grażyna Trzpiot
Beata Turyna–Pułaska
Ewa Wędrowska
Andrzej Wiatrak
Marek Wierzbicki
Dorota Witkowska
Bartosz Witkowski
Michał Zasada
Jan Zawadzki
Jarosław Zawadzki
Tomasz Ząbkowski
Beata Zduniak
Monika Zielińska-Sitkiewicz
Wojciech Zieliński
Krzysztof Zmarzłowski
Dorota Żebrowska-Suchodolska