METODY ILOŚCIOWE
W BADANIACH EKONOMICZNYCH

QUANTITATIVE METHODS
IN ECONOMICS

Vol. XVI, 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 XVI, No. 2

Warsaw 2015

ISSN 2082 – 792X

Warsaw 2015, Volume XVI, 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. XVI, No. 2, 2015

CONTENTS

Jerzy Korzeniewski – Determining the Number of Clusters for Marketing Binary Data 7

Monika Krawiec, Anna Górska – Granger Causality Tests for Precious Metals Returns 13

Justyna Kujawska – Measurement of Healthcare System Efficiency in OECD Countries 23

Izabela Kurzawa, Jarosław Lira – The Application of Quantile Regression to the Analysis of the Relationships between the Entrepreneurship Indicator and the Water and Sewerage Infrastructure in Rural Areas of Communes in Wielkopolskie Voivodeship 33

Joanna M. Landmesser, Krzysztof Karpio, Piotr Łukasiewicz – Decomposition of Differences Between Personal Incomes Distributions in Poland 43

Jarosław Lira – A Comparison of the Methods of Relative Taxonomy for the Assessment of Infrastructural Development of Counties in Wielkopolskie Voivodeship 53

Rafik Nafkha – Assessment and Selection Model for Management System Supporting Small and Medium-sized Enterprises 63

Maria Parlińska, Iryna Petrovska – The Role of Information Systems in Development of Voivodeships in Poland 73

Marcin Rudź – Precise Estimates of Ruin Probabilities 80

Paweł Sakowski, Robert Ślepaczuk, Mateusz Wywiał – Cross-Sectional Returns from Diverse Portfolio of Equity Indices with Risk Premia Embedded 89

Victor Shevchuk – Determinants of the Demand for International Reserves in Ukraine 102

Iwona Skrodzka – Development of Knowledge-Based Economy in European Union in 2000–2014 113

Emilia Tomczyk – Data Vintage in Testing Properties of Expectations 123

Alexandr Trunov – Modernization of Means for Analyses and Solution of Nonlinear Programming Problems 133

Tomasz Wójtowicz – Macroeconomic Indicators Forecasts Accuracy and Reaction of Investors on the WSE 142

Wojciech Zieliński – A Confidence Interval for Proportion in Finite Population Divided into Two Strata: a Numerical Study 152

Joanna Żyra – The Influence of Public and Private Higher Education in Poland on the Economic Growth of the Country 159

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 7 – 12

DETERMINING THE NUMBER OF CLUSTERS FOR MARKETING BINARY DATA

Jerzy Korzeniewski
Department of Statistical Methods, University of Lodz
e-mail: jurkor@wp.pl

Abstract: In the article a new way of determining the number of clusters was proposed focused on data made up of binary variables. An important application aspect is that the data sets on which the new formula was investigated were generated in the way characteristic for the marketing data following the work of Dimitriadou et al. [2002]. The new formula is a modification of the Ratkowsky-Lance index and proved to be better in some respects than this index, which was the best in the mentioned research. The modification proposed is based on measuring the quality of grouping into the predicted number of clusters and running the same index on the twice smaller set of objects comprising dense regions of the original data set.

Keywords: cluster analysis, binary data, number of clusters index, market segmentation

INTRODUCTION

Predicting of the number of clusters

One of very important parts of cluster analysis (unsupervised learning) is to find out how many clusters there should be in a data set. Obviously, this task is closely related to other cluster analysis tasks e.g. selection of variables and grouping of objects, however, the subject of selecting the proper number of clusters has attracted much interest which resulted in dozens of different proposals of indices or stopping rules. Milligan [1985] was probably the first to carry out a thorough investigation of more than two dozens of different indices but the research was concentrated on continuous variables data sets and it took place 30 years ago. Since that time many new proposals were published and the task has been directed to different targets related to e.g. different variable measuring scales. As far the binary variables are concerned a good examination was carried out by Dimitridou et al.

[2002]. The conclusion from this research is in favour of the Ratkowsky-Lance index which turned out to be better than other indices. Therefore, in order to carry out a new research on similar data sets this index was applied as the reference point. From a couple of newer proposals, the Fang and Wang index [2012] was also used in this article.

Binary marketing data

Binary marketing data specificity consists in a number of variables being correlated (or not) to create separate groups of variables. The whole data set consists of a couple of groups of such variables. In this research we followed the scheme suggested by Dimitraidou et al. [2002] in which every data set is described by twelve binary variables composed into four groups of different or equal numbers of variables. An example of such data pattern is presented in Table 1.

Table 1. An example of binary marketing data pattern, twelve variables in four groups

	Group1			Group2			Group3			Group4		
	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
Cluster1	H	H	H	H	H	H	L	L	L	L	L	L
Cluster2	L	L	L	L	L	L	H	H	H	H	H	H
Cluster3	L	L	L	H	H	H	H	H	H	L	L	L
Cluster4	H	H	H	L	L	L	L	L	L	H	H	H
Cluster5	L	L	L	H	H	H	L	L	L	H	H	H
Cluster6	H	H	H	L	L	L	H	H	H	L	L	L

Source: Dimitriadou et al. [2002]

The idea of this example is to present connections between groups of respondents and groups of questions in a questionnaire. The symbol H stands for the high probability of value 1 on a given variable and the symbol L stands for the low probability of 1. Obviously, the number of variables in each group, their correlation within the group, the level of H and L will be varied (see experiment description for details).

INDICES OF THE NUMBER OF CLUSTERS

Out of the multitude of the number of clusters indices which one can find in the literature we picked up as the reference point one that came the best in the Dimitriadou research i.e. the Ratkowsky-Lance index given by the formula

$$RL = \frac{mean\left(B/T\right)}{\sqrt{k}}, \tag{1}$$

where B stands for the sum of squares between the clusters for each variable, T stands for the total sum of squares for each variable and k stands for the number of groups into which the data has to be previously grouped by means of some grouping method.

The mean in the numerator of formula (1) is taken across all single variables. The value of *k* maximizing *RL* should be selected as the number of clusters prediction.

In order to include in the research some newer proposals we chose the Fang-Wang [2012] index based on the bootstrap method. This index is defined in the following way. We draw independently *B* bootstrap samples

$$X_b, Y_b, \quad b = 1,\ldots,B, \tag{2}$$

With the symbol $\Psi_{Xb,K}$ we denote the grouping of sample X_b nto *k* clusters. Then we define the distance of two groupings/divisions with the formula

$$d(\Psi_{Xb,K}, \Psi_{Yb,K}) = \\ = \frac{1}{n^2} \sum_{i,j=1} \left| I(\Psi_{Xb,K}(x_i) = \Psi_{Xb,K}(x_i)) - I(\Psi_{Yb,K}(x_j) = \Psi_{Yb,K}(x_i)) \right|, \tag{3}$$

where *I* stands for the function assuming value 1 if the condition in the brackets is met. This distance measure has easy and intuitive interpretation. The final step is to define a measure of instability of divisions given by the formula

$$s_B(\Psi, K, n) = \frac{1}{B} \sum_{b=1}^{B} d(\Psi_{Xb,K}, \Psi_{Yb,K}). \tag{4}$$

The value of *k* this time minimizing the right hand side of formula (4) should be selected as the number of clusters prediction. All parameters necessary for the above formulas will be specified in the experiment description in the fourth chapter.

Some interesting recent proposals were given by Tibshirani et al. [2010] but they seem to be dedicated rather for special cases with the number of features being much bigger than the number of objects.

NEW INDEX PROPOSAL

We will try to propose a new index of the number of clusters which consists in the modification of the Ratkowsky-Lance index. The modification will involve two independent steps. One will be devoted to limiting the use of the original Ratkowsky-Lance index to half of the objects of a given data set belonging to “dense regions” of the data set. The other step will consist in measuring the quality of a data set division into a predetermined number of clusters.

Finding “dense regions” is a common concept in cluster analysis. The idea behind it is that limiting ones research to these regions usually gives more pronounced results in comparison with that of the whole data set. A popular technique of defining such regions is a sequential procedure working in the following way. The first object picked up is the one which has the smallest distance to its 20th nearest neighbor. This object is removed from the data set, all pairwise distances are computed again and the second object picked up is the one with the smallest distance to its 20th nearest neighbor. We continue this process until we pick up half of the data objects. Obviously, the number 20 may be changed, for smaller data sets it is usually 5, but for the kind of sets used in our experiment (about 5000 objects) number 20 seems the proper choice.

Measuring the quality of a data set division or grouping is another task which can be performed in a number of ways. In our experiment we will use the following approach. Let us define a measure of the quality of a data set division into two clusters (which we will call the primary division). We choose all objects belonging to the smaller cluster and half of objects belonging to the bigger cluster and we divide these objects once again into two clusters using the same grouping method. The measure of the quality of the primary division will be given by the value of the adjusted Rand index [e.g. Gatnar and Walesiak 2004] as a similarity measure of both divisions. In the number of clusters prediction process, the data set is divided into different numbers of clusters, therefore, to use our measure we will apply it to every pair of clusters into which the data set was divided. For example, if the data set was grouped into 5 clusters we will get 10 measures of the quality of separation of every pair of clusters. Ideally, the value of 1 of the measure is desirable i.e. such value confirms that the division was well done or that the two clusters being assessed are perfectly separable. Formally, if anyone of the 10 values is close to zero i.e. very small it proves that in the division there is at least one pair of clusters which is badly separated. However, it only takes place in the case of very clear cluster structures that all pairs of clusters have division quality close to 1. Therefore, as the final measure of the division of the data set into any number of clusters, we will use a simple arithmetic mean across all pairs of clusters.

The new index formula is a modification of the Ratkowsky-Lance index the idea of which is to apply this index twice. Firstly to the whole data set and, secondly, to half of the data set representing dense regions. Subsequently, if the two instances return different numbers of predicted clusters, we will choose one of them. As we have to decide between from 2 to 10 clusters (see experiment description) we will concentrate our attention on the initial number of clusters i.e. 2, 3, 4 and 5. When the quality of these divisions (of the whole data set) is good we will use the prediction based on the whole data set. If the quality of these initial divisions is bad we will use the prediction based on the denser half of the data set. The logic behind such approach is that when divisions into smaller number of clusters are of bad quality the Ratkowsky-Lance index has a tendency to overestimate the predicted number of clusters. To be precise and not to search for thresholds taken from out of blue, we will use the value of 0.5 as the limiting value deciding about the divisions below this value being judged as bad divisions. Thus, the whole modification can be stated in the form of the following algorithm.

- Divide the whole data set into 2, 3, …, 10 clusters.
- Find the denser half of the data set using the technique of the 20^{th} closest neighbor.
- Divide the denser half into 2, 3, …, 10 clusters.
- If the measure of the quality of the whole data set division into 4 clusters is above 0.5 take the prediction of the Ratkowsky-Lance index based on the whole data set.
- If the measure of the quality of the whole data set division into 4 clusters is below 0.5 take the prediction of the Ratkowsky-Lance index based on the denser half of the data set.

EXPERIMENT DESCRIPTION

In order to evaluate the new index we carried out the following experiment. We generated 162 data sets according to the pattern described in chapter 2. We used the *bindata* package available in R language. The data sets generated were diversified with respect to the following parameters.

- Probability; for H there are 3 variants: 0.9, 0.8, 0.7 and for each variant respectively, for L there are 3 variants: 0.1, 0.2 0.3.
- Correlation inside groups of variables; there are 3 variants: uncorrelated variables, variables correlated with moderate strength (0.4), variables correlated with big strength (0.8).
- Number of clusters; 3 variants: 4, 5, 6.
- Numbers of objects in the clusters: 3 variants: (1000, 1000, 1000, 1000, 1000, 1000), (2000, 500, 1000, 700, 700, 1100), (3000, 300, 1000, 500, 700, 500).
- Number of variables within groups; 2 variants: (3, 3, 3, 3), (5, 4, 2, 1).

We ran the Ratkowsky-Lance index, the Fang-Wang index and the new proposal index using the *k*-means grouping method. The *k*-means grouping was done for a random choice of starting points, repeated 50 times, from which the result with the smallest distance measure was chosen. For the Fang-Wang index we used *B* equal to 50. The number of possible clusters from which the algorithms were choosing ranged from 2 to 10 clusters. In order to assess the efficiency of each index, out of many possible criteria, we used the percentage of properly predicted clusters as well as the percentage of errors equal to 1 and the percentage of bigger errors. In the literature one can find a couple of other criteria e.g. proper cluster recovery or correct dominant recovery. However, if one uses a mish-mash criteria the results are sometimes blurred because some criteria return different results than other criteria and does not get any clear conclusions.

RESULS AND CONCLUSIONS

The Fang-Wang index performed poorly achieving about 25% of correct predictions, therefore we will limit our conclusions to the other two compared indices. The two indices agreed in 50% of cases. Other results are given in Table 2. The new index achieved better overall performance as far as correct predictions are concerned (44% to 32%) with almost equal percentage of going wrong by 1 cluster.

Table 2. Results for the Ratkowsky-Lance index and the new index

	Performance measure	Overall performance	Number of clusters			Probability			Number of group variables	
			4	5	6	0.9	0.8	0.7	(3,3,3,3)	(5,4,2,1)
Ratkowsky-Lance	Correct hits	.32	.43	.20	.35	.41	.44	.12	.49	.16
	Error = 1	.27	.15	.43	.23	.37	.24	.19	.34	.19
New index	Correct hits	.44	.52	.56	.25	.35	.50	.48	.44	.44
	Error = 1	.28	.24	.26	.38	.33	.33	.21	.31	.28

Source: own research

The new index was also better in most subcategories apart from the sets with 6 clusters (the new index lost 25% to 35%), clear cluster structures (the new index lost 35% to 41%) and apart from the group with uniform numbers of variables (the new index lost 44% to 49%). The basic reason for the poorer performance of the Ratkowsky-Lance original index seems to be its poor results (only 16% ! of correct hits) for the data sets in which some groups of variables have much smaller numbers of variables than other groups as well as very poor result (only 12% ! of correct hits) for blurred cluster structures. In conclusion we can state that the new proposal is more robust to unwelcome conditions.

REFERENCES

Dimitriadou E., Dolničar S., Weingessel A. (2002) An examination of indexes for determining the number of clusters in binary data sets, Psychometrika, Vol. 67, Issue 1, pp. 137-159.

Fang Y., Wang J. (2012) Selection of the number of clusters via the bootstrap method, Comput. Statist. Data Anal. 56, pp. 468–477.

Gatnar E., Walesiak M. (2004) Metody Statystycznej Analizy Wielowymiarowej w Badaniach Marketingowych, Wydawnictwo AE we Wrocławiu, pp. 334-36.

Leisch F., Weingessel A., Hornik K. (2015) *bindata* package manual.

Milligan G. W., Cooper M. C. (1985) An examination of procedures for determining the number of clusters in a data set, Psychometrika 50, 159–179, pp. 137-159.

Tibshirani R., Witten D. (2010) A framework for feature selection in clustering, Journal of American Statistical Association, 105(490), pp. 713–726.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 13 – 22

GRANGER CAUSALITY TESTS FOR PRECIOUS METALS RETURNS

Monika Krawiec
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
e-mail: krawiec.monika@gmail.com
Anna Górska
Department of Agricultural Economics and International Economic Relations
Warsaw University of Life Sciences – SGGW
e-mail: anna_gorska@sggw.pl

Abstract: The aim of the paper was examining Granger causality between rates of return of precious metals. The study covers the period from 2008 through 2013 and includes gold, silver, platinum, and palladium. After developing statistical analysis and confirming stationarity of time series under consideration, the Granger causality test was run. Its results revealed a bilateral causation between silver and platinum rates of return. The study also detected causal relationships flowing from gold and palladium rates of return to silver returns.

Keywords: precious metals, stationarity, Granger causality

INTRODUCTION

Since early 2000s, commodity markets have become more like financial markets. The phenomenon of their financialization brings about the need of adopting methods originally designed for investigating financial markets, namely methods of financial econometrics. Beginnings of financial econometrics are often dated back to 1982 when Robert Engle published his paper presenting autoregressive conditional heteroscedasticity (ARCH) model. It opened a door to further development of various models, such as family of generalized ARCH (GARCH) models, autoregressive conditional duration (ACD) model, dynamic conditional correlation (DCC) model etc. [Jajuga 2007]. There were also developed some other concepts of dynamic econometrics, such as cointegration and testing causal relationships between economic variables often referred to as Granger causality.

When applied to commodity markets, Granger causality tests can tell us the nature of inter-relationships between the various markets and categories of commodities. The aim of the paper is to test Granger causality for markets of precious metals. Our study covers rates of return series of four basic precious metals: gold, silver, platinum and palladium. The occurrence of pairwise Granger causality among them would indicate the possibility of improving forecasts by including the lagged values of respective variables in adequate VAR (vector autoregressive) models. Recognizing relationships between precious metals prices and returns is important as on one hand they are considered attractive assets for portfolio investments, and on the other hand all of them have distinct technical uses.

Gold (Au) is found in nature mainly as either high-quality free gold or as finely distributed minerals mixed with silver, copper or mercury. It has seven money properties: it is a luxury good valued by most people; it is dividable in almost any denomination; it is easy to transport; it remains completely stable over time; it can be weighted exactly; it is not easy to forge or artificially producible; and it cannot be multiplied. Gold can also fulfill three money functions: it can be used as a means of exchange or means of payment, it comes in an arithmetic unit, its purchasing power does not diminishes over time [Eller and Sagerer 2008]. Nowadays, gold is used as a monetary commodity, for jewelry, and dental industry, but in fact its use in jewelry production dates back to the 4th millennium BCE.

Silver (Ag), similarly to gold, has been used since the 4th millennium before Christ as both, jewelry and money. It occurs 15-20 times more often than gold, however almost never in pure form. The majority (about 60%) is extracted as a secondary metal during copper, zinc or lead production, 25% comes from pure silver mines, and the smallest part (15%) comes from gold production. Silver, the same as gold, fulfills the three money functions. It is typically used for jewelry, photography, silverware, and in a diverse range of electronic products.

Platinum (Pt) was first used by pre-Columbian South American natives. When Antonio de Ulloa published his report on a new metal of Colombian origin in 1748, it became investigated by scientists. In early 1800's William Wollaston – English chemist became the first one who produced pure, malleable platinum. Platinum occurs with the same frequency as gold and is primarily generated as a byproduct of copper and nickel production. It is used heavily by the dental, chemical, electronics, and auto industries. Because of its chemical qualities, platinum is often used in catalytic converters to reduce emissions.

The last one of the metals we are interested in – palladium (Pd) was discovered in 1803 by William Wollaston during platinum exploration in America. It usually occurs with other platinum metals and it has similar industrial uses as platinum. Very often it substitutes platinum in engines exhaust systems [Balarie 2007].

EMPIRICAL DATA AND RESEARCH METHODS

Numerous researchers analyzing various inter-relationships between precious metals prices or returns base their studies on time series of different length. For example, Wahab et al. [1994] examine the period from 1982 through 1992, Escribano and Granger [1998] – the period from 1971 through 1990, Ciner [2001] – the period: 1992 – 1998, Lucey and Tully [2006]: 1978 – 2002, Kearney and Lombra [2009]: 1985-2006, Tsuchiya [2010]: 2002-2010, Śmiech and Papież [2012]: 2000 – 2011. Our data set covers a 6-year-period from January 2008 to December 2013 and consists of London daily closing prices of four precious metals (gold, silver, platinum and palladium) in USD per ounce. The quotations are available at www.kitco.com.

First, rates of return series were calculated as follows:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \tag{1}$$

where P_t is the price at time *t* and P_{t-1} is the price in the previous period.
Both, prices of the precious metals and their rates of return are displayed in Figure 1.

The rates of return series became the base to evaluate descriptive statistics for considered precious metals. Then, normality of distributions was verified by adopting the Jarque-Bera test. The results are given in Table 1.

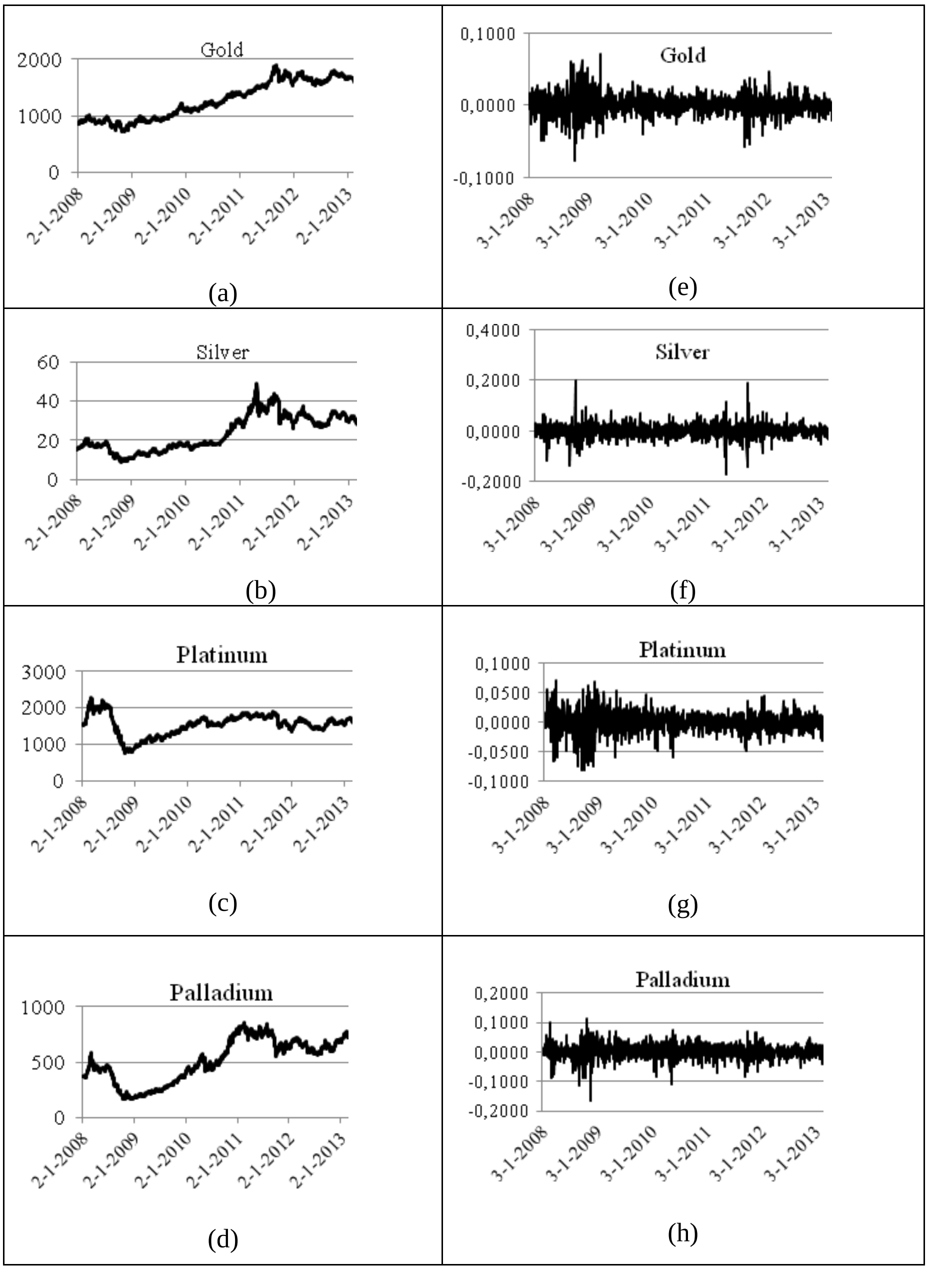
Table 1. Descriptive statistics for daily logarithmic returns of precious metals (2008-2013)

Measure	Metal			
	Gold	Silver	Platinum	Palladium
Minimum	-0.09596	-0.18693	-0.08493	-0.17859
Maximum	0.06841	0.18279	0.06940	0.10920
Mean	0.00023	0.00018	-0.00008	0.00043
Standard deviation	0.01386	0.02676	0.01687	0.02303
Coeff. of variation	59.9911	151.8322	202.1752	53.6307
Skewness	-0.38924	-0.40619	-0.64145	-0.66024
Kurtosis	4.51254	7.39392	3.69171	4.84103
J-B	1312.88	3466.22	956.50	1577.07

Source: own calculations

On the base of data in Table 1, one can notice that mean daily returns range between - 0.008% for platinum and 0.04% for palladium. The maximum of daily returns (18,3%) was observed for silver on September 18, 2008. The minimum of daily returns (-18,7%) was also registered for silver on May 12, 2011. The lowest standard deviation was the one obtained for gold (0.01386), while silver exhibited the highest value of standard deviation (0.02676). However, platinum was the precious metal showing the highest volatility.

Figure 1. Prices of precious metals from 2008 through 2013: gold (a), silver (b), platinum (c), palladium (d) and their returns: gold (e), silver (f), platinum (g), palladium (h)



Source: own elaboration

The lowest volatility was exhibited by palladium (see values of coefficient of variation). In all cases, distributions of returns are negatively skewed. Positive values of kurtosis indicate more acute distributions in comparison to the normal distribution. The Jarque-Bera test confirms the non-normality of daily returns distributions at 0.05 significance level.

Table 2 reports values of Pearson correlation coefficient calculated for various pairs of precious metals. As expected, all values are found to be positive[1] and significant at the 0.05 level. The highest positive correlation was observed for the pair: platinum – palladium (as it is mentioned in the introduction, palladium often substitutes platinum in technical applications), the lowest one for the pair: gold – palladium (they are not close substitutes to each other).

Table 2. Coefficients of correlation between selected precious metals

Metal	Gold	Silver	Platinum	Palladium
Gold	1	0.6061	0.5827	0.4891
Silver	0.6061	1	0.5771	0.5133
Platinum	0.5827	0.5771	1	0.7506
Palladium	0.4891	0.5133	0.7506	1

Source: own calculations

In the next step of research, for answering the question whether past returns of a given precious metal can help better forecast returns of other selected precious metals, Granger causality test will be applied. Generally, since the future cannot predict the past, if variable X Granger-causes variable Y, then changes in X should precede changes in Y. In other words: when we identify one variable as the dependent variable (Y) and another as the explanatory variable (X), we make an implicit assumption that changes in the explanatory variable induce changes in the dependent variable. Therefore, in a regression of Y on other variables (including its own past values) if we include past or lagged values of X and it significantly improves the prediction of Y, we can say that X Granger-causes Y. A similar definition applies if Y Granger-causes X [Gujarati 2003]. If X causes Y and Y causes X, the two variables are jointly determined and there is a bilateral causation.

There are several different procedures for testing Granger causality[2]. In our paper, following Ramanathan [2002], we consider the model:

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{q} \beta_j X_{t-j} + u_t \ , \tag{2}$$

[1] According to Kearney and Lombra [2009], price fluctuations of silver, platinum and palladium seem to follow closely the price of gold over the last two decades.

[2] According to Osińska [2008], in economic practice the most popular are three procedures, differing in construction and in results they provide, that are based on likelihood ratio, Wald test, and Lagrange multiplier.

where u_t is white noise, p is the order of the lag for Y, and q is the order of the lag for X. The null hypothesis that X does not Granger-cause Y is that $\beta_j = 0$ for $j = 1, 2, ..., q$.
Then we have the restricted model:

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + \upsilon_t \ . \tag{3}$$

The test statistic is the standard Wald F-statistic:

$$F = \frac{(ESSR - ESSU)/q}{ESSU/(n-p-q)}, \tag{4}$$

where n is the number of observations used in unrestricted model in equation (2), ESSU is the error sum of squares for equation (2), ESSR is the error sum of squares for the restricted model (3). Under the null hypothesis of X not Granger-causing Y, F has the F-distribution with q d.f. for the numerator and $n-p-q$ d.f. for the denominator. The orders of the lags (p and q) are arbitrary and are usually chosen to be large [Ramanathan 2002].

As a pre-requisite condition for Granger causality testing, time series need to be stationary. Stationarity in the weak sense implies that the mean of the variable, its variance and covariance shell be time invariant. There are several stationarity tests. In the paper we use the augmented Dickey-Fuller test (the ADF-test). The null hypothesis assumes nonstationarity. The first step is to estimate one of the following equations [Witkowska et al. 2008]:

$$\Delta y_t = \alpha_1 y_{t-1} + \sum_{i=1}^{p} c_i \Delta y_{t-i} + \varepsilon_t \ , \tag{5}$$

$$\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \sum_{i=1}^{p} c_i \Delta y_{t-i} + \varepsilon_t \ , \tag{6}$$

$$\Delta y_t = \alpha_0 + \lambda_1 t + \alpha_1 y_{t-1} + \sum_{i=1}^{p} c_i \Delta y_{t-i} + \varepsilon_t \ . \tag{7}$$

The statistic of the test is given by:

$$\tau = \frac{\hat{\alpha}_1}{S(\hat{\alpha}_1)}, \tag{8}$$

where: $\hat{\alpha}_1$ – OLS estimate of α_1 in any of equations (5) – (7), $S(\hat{\alpha}_1)$ – standard error of α_1 estimate.
If the tau value is lower than the critical value, the null hypothesis is rejected. Hamulczuk et al. [2012] note that tau follows the distribution that differs from

other standard distributions, thus it is necessary to use special statistical tables. However, the GRETL software that we use, computes the probability value (p). If $p < 0.05$, H_0 can be rejected.

RESULTS OF GRANGER CAUSALITY TESTS FOR PRECIOUS METALS RETURNS

As it was mentioned in the previous section, when testing Granger causality, it is assumed that the variables are stationary. That is why we start with performing the ADF-test for our data. Its results (values of tau-statistic based on estimates of equation (7)) are presented in Table 3. Since they let us conclude that all considered time series are stationary, the following series of hypotheses can be verified:

H_0: *rates of return of precious metal X are not Granger cause of rates of return of precious metal Y.*

Table 3. The ADF-test results for returns of separate precious metals

Precious metal	Tau-statistic	p-value
Gold	-17.5023	4.37E-056
Silver	-12.6803	1.42E-031
Platinum	-7.3073	3.06E-10
Palladium	-8.4474	5.88E-14

Source: own calculations

Gujarati [2003] suggests the direction of causality may depend critically on the number of lagged terms included. That is why in Table 4 we present the results of the F-test using several lags[3]. Since our interest is testing for causality, we do not show the estimated coefficients of models (2) and (3). In most cases the lag length does not influence test results (the only exception at the 5% rejection rate is relationship silver→platinum). Thus, regardless the lag length, there is causality running from gold returns, platinum returns, and palladium returns to silver returns. One may also notice Granger causality flowing from silver returns to platinum returns, so there is a bilateral causality between them (silver↔platinum). There are no causal relationships at all, between gold and platinum, gold and palladium, and platinum and palladium, although the last pair exhibited the highest value of correlation coefficient. It confirms that correlation does not imply causality.

[3] According to Waściński [2010], the lag length should reflect natural interactions between variables. For example, the recommended number of lags in the case of quarterly data is 4. Our study is based on daily observations, so we start with 1 lag and next we test 5 lags (precious metals quotations are observed on each of 5 weekdays). Finally, taken into account Ramanathan's recommendation to choose large numbers of lags, we consider 10 lags.

Table 4. The Granger causality test results for precious metals returns

Relationship	Number of lags	F-statistic	p-value	Decision at 0.05
gold→silver	1	148.7900	0.0000	Reject
	5	33.7090	0.0000	Reject
	10	17.4740	0.0000	Reject
silver→gold	1	0.4256	0.5142	Do not reject
	5	1.3219	0.2520	Do not reject
	10	1.0048	0.4269	Do not reject
gold→platinum	1	0.3128	0.5760	Do not reject
	5	0.9788	0.4294	Do not reject
	10	1.1091	0.3519	Do not reject
platinum→gold	1	0.1162	0.7332	Do not reject
	5	0.4072	0.8440	Do not reject
	10	0.6358	0.7840	Do not reject
gold→palladium	1	0.2693	0.6039	Do not reject
	5	0.2931	0.9169	Do not reject
	10	0.3557	0.9650	Do not reject
palladium→gold	1	0.1155	0.7340	Do not reject
	5	0.1756	0.9718	Do not reject
	10	0.3468	0.9680	Do not reject
silver→platinum	1	8.8647	0.0030	Reject
	5	2.7424	0.0179	Reject
	10	1.7222	0.0707	Do not reject
platinum→silver	1	70.5640	0.0000	Reject
	5	14.0840	0.0000	Reject
	10	8.1067	0.0000	Reject
silver→palladium	1	2.7338	0.0985	Do not reject
	5	1.0086	0.4111	Do not reject
	10	0.5908	0.8226	Do not reject
palladium→silver	1	57.9310	0.0000	Reject
	5	11.5880	0.0000	Reject
	10	6.3584	0.0000	Reject
platinum→palladium	1	0.3832	0.5360	Do not reject
	5	0.3556	0.8788	Do not reject
	10	1.0521	0.3967	Do not reject
palladium→platinum	1	0.8592	0.3541	Do not reject
	5	0.3675	0.8711	Do not reject
	10	0.6307	0.7885	Do not reject

Source: own calculations

CONCLUDING REMARKS

The paper was aimed at answering the question whether returns of separate precious metals are Granger causes of returns of other precious metals. The study covered the period from 2008 through 2013 and included four precious metals: gold, silver, platinum, and palladium. On the base of their logarithmic returns, there were calculated descriptive statistics and coefficients of correlation. Then tests for normality and stationarity were conducted. Finally, to achieve the purpose of the study, Granger causality test was performed. Our results revealed Granger causality running from gold, platinum, and palladium returns to silver returns, and from silver returns to platinum returns as well. Thus, including lagged values of gold, platinum, and palladium returns improves the prediction of silver returns, whereas including lagged values of silver returns can improve the prediction of platinum returns solely. The analysis presented in the paper is a part of more complex study of precious metals markets the authors have been developing in the last few years (see [Górska and Krawiec 2011, Górska and Krawiec 2013, Górska and Krawiec 2014]).

REFERENCES

Balarie E. (2007) Commodities for Every Portfolio, John Wiley&Sons, Hoboken, New Jersey.

Ciner C. (2001) On the Long run Relationship Between Gold and Silver: a Note, Global Finance Journal, 12, 299-303.

Eller R., Sagerer Ch. (2008) An Overview of Commodity Sectors, The Handbook of Commodity Investing. John Wiley&Sons, Hoboken, New Jersey, 681-711.

Engle R.F. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, Econometrica, 50, 987-1007.

Escribano A., Granger C.W.J. (1998) Investigating the Relationship Between Gold and Silver Prices, Journal of Forecasting, 17 (2), 81-107.

Górska A., Krawiec M. (2011) Zastosowanie narzędzi analizy technicznej w inwestowaniu na rynkach metali szlachetnych, Metody ilościowe w badaniach ekonomicznych, 12, (2), 148-157.

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.

Górska A., Krawiec M. (2014) Analysis of Calendar Effects in Markets of Precious Metals, Metody ilościowe w badaniach ekonomicznych – Quantitative Methods in Economics, 15, (2), 392-402.

Gujarati D.N. (2003) Basic Econometrics, McGraw-Hill, Boston.

Hamulczuk M., Gędek S., Klimkowski C., Stańko S. (2012) Prognozowanie cen surowców rolnych na podstawie zależności przyczynowych, Wydawnictwo Instytutu Ekonomiki Rolnictwa i Gospodarki Żywnościowej, Warszawa.

Jajuga K. (2007) 25 lat ekonometrii finansowej, Zeszyty Naukowe Uniwersytetu Szczecińskiego 462, Finanse, Rynki finansowe, Ubezpieczenia, 6, Szczecin, 91-100.

Kearney A.A., Lombra R.E. (2009) Gold and Platinum: Toward Solving the Price Puzzle, Quarterly Review of Economics and Finance, 49, 884-982.

Lucey B.M., Tully E. (2006) The Evolving Relationship between Gold and Silver 1978-2002: Evidence from a Dynamic Cointegration Analysis – a Note, Applied Financial Economics Letters, 2, 47-53.

Osińska M. (2008) Ekonometryczna analiza zależności przyczynowych, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń.

Ramanathan R. (2002) Introductory Econometrics with Applications, South-Western Thomson Learning, Mason, Ohio.

Śmiech S., Papież M. (2012) A Dynamic Analysis of Causality Between Prices on the Metals Markets, Proceedings of the International Conference "Quantitative Methods in Economics", Multiple Criteria Decision Making, XVI, Bratislava, 221-225.

Tsuchiya Y. (2010) Linkages Among Precious Metals Commodity Futures Prices: Evidence from Tokyo, Economics Bulletin, 30 (3), 1772-1777.

Wahab M., Cohn R., Lashgari M. (1994) The Gold-Silver Spread: Integration, Cointegration, Predictability and Ex-ante Arbitrage, Journal of Futures Markets, 14, 709-756.

Waściński T. (2010) Powiązania cenowe na polskim rynku finansowym, Dom Wydawniczy Elipsa, Warszawa.

Witkowska D., Matuszewska A., Kompa K. (2008) Wprowadzenie do ekonometrii dynamicznej i finansowej, Wydawnictwo SGGW, Warszawa.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 23 – 32

MEASUREMENT OF HEALTHCARE SYSTEM EFFICIENCY IN OECD COUNTRIES

Justyna Kujawska
Department of Economic Analysis and Finance, Gdańsk University of Technology
e-mail: Justyna.Kujawska@zie.pg.gda.pl

Abstract: Increased spending on healthcare systems in many countries tends to attract attention to their efficiency. The aim of this paper is to evaluate the efficiency of healthcare systems in the OECD countries and indicate causes of inefficiency by applying Data Envelopment Analysis (DEA) and using additive and super-efficiency models. The homogeneity of the sample is assessed and outliers are excluded. A ranking is established on the basis of efficiency scores. By means of DEA, fully efficient units are identified, forming a reference set (of best practice) for inefficient countries to follow.

Keywords: healthcare system efficiency, Data Envelopment Analysis

INTRODUCTION

An efficiently operating healthcare system makes an important contribution to increasing the general quality of life. Regularly conducted surveys point out the shortcomings of healthcare services. The most commonly identified problems are: overly expensive healthcare services, excessively long waiting times, and distance to healthcare facilities [OECD 2011]. The data which are most often used for comparisons of different national healthcare systems include total healthcare expenditure as a fraction of gross domestic product (GDP) or GDP per capita [Anell, Willis 2000]. The average healthcare expenditure for all OECD countries amounted to 6,8% of GDP in 1990, 7,8% in 2000 and 9,5% in 2010. In the same years, the corresponding figures for Poland were 4,8%; 5,5% and 7,0% respectively, and for the United States 12,4%; 13,7% and 17,6% [OECD 2012]. Another important factor which affects the performance of healthcare systems is the ageing of populations, which boosts the demand for healthcare services. Life expectancy in the OECD countries has been growing systematically: women's life expectancy (in years) increased from 78 in 1990 to 82,5 in 2010, while men's life expectancy increased from 71 in 1990 to 77 in 2010 [OECD 2012]. Because an average of 72% of healthcare is financed from public funds in the OECD countries, the aforementioned factors justify a need to evaluate the efficiency of these services [OECD 2012]. The aim of this paper is to propose a model for measuring healthcare system efficiency by means of DEA.

DATA ENVELOPMENT ANALYSIS AND THE EFFICIENCY OF HEALTHCARE SYSTEMS

DEA is a nonparametric method for measuring relative efficiency. This method has been undergoing dynamic development since 1978, when Charnes, Cooper and Rhodes published their seminal article entitled "Measuring the efficiency of decision making units" [Charnes et al. 1978]. The growth in the use of DEA is reflected in the statistics of publications registered in the Web of Science database. In the years 1978-1990, 225 such articles were published, but by 2009 this number had reached 4,597. It is foreseen that by 2020 the number of articles on the subject may reach 13,000 [Liu et al. 2013a]. DEA is a data-oriented approach for evaluating the performance of a set of homogeneous entities called decision making units (DMUs), which convert multiple inputs into multiple outputs [Cooper et al. 2011]. This method may be applied to a wide range of sectors (such as banking or healthcare) to identify sources of inefficiencies [Liu et al. 2013b].

Basic DEA models only measure radial efficiency but fail to evaluate the input excesses or output shortfalls (slacks), and hence only detect radial inefficiency. The DEA definition of efficiency is that the performance of a DMU is fully (100%) efficient only when the efficiency score equals one and the input and output slacks equal zero. When the efficiency score is one while one or more of the slacks differ from zero, the DMU is said to be weakly efficient [Cooper et al. 2000, Zhu, Cook 2007]. Unfortunately, the radial efficiency measure does not take into account non-zero slacks. The additive model is free from this flaw because it takes slacks into consideration directly in the computation of the efficiency measure. This was used as the basis for the development of the Slack Based Measure (SBM) model for evaluating efficiency [Cooper et al. 2000].

After choosing the structure of the model, it is important to define its orientation, according to whether the aim is to reduce the inputs and keep the outputs at the same level (input-oriented), or alternatively to maximise the outputs and keep the inputs at the same level (output-oriented) [Ozcan 2008].

Measurement of the efficiency of healthcare systems is not an easy task. The main difficulty is in correctly measuring the outcomes of the system. The most popular approach applies measurable intermediate indicators of services which are assumed to have a fundamental impact on the health status of the population. The outcomes of a healthcare system may be defined as changes in the health of the population attributable to healthcare expenditure, e.g. changes in life expectancy, infant mortality, inequity in access to healthcare, frequency of occurrence of certain diseases, etc. [González et al. 2010]. In spite of controversy over whether some of these variables are appropriate as relevant outcomes of healthcare systems, most analyses at the system level have relied on the use of life expectancy and infant mortality rates to evaluate the outcomes of health systems [e.g. Retzlaff-Roberts et al. 2004, Afonso, Aubyn 2005, Anell, Willis 2000, Hadad et al. 2013]. Nevertheless, some researchers argue that infant mortality in the OECD countries has ceased to be a dramatic problem. Undoubtedly, it does not concern most developed countries, but Mexico, Chile, Turkey and countries of the former eastern bloc still record infant mortality rates above the average. One of the most often quoted studies [Retzlaff-Roberts et al. 2004] adopts the infant mortality rate and life expectancy at birth as outputs. The inputs characterising the resources of the system include the number of physicians and the number of beds per 1,000 residents, the number of magnetic resonance imaging (MRI) devices per million residents and healthcare expenditure as a fraction of GDP.

Sometimes, international comparisons cannot be made due to insufficient data, and consequently some countries must be excluded from the analysis. The use of DEA requires much caution in the selection of the sample because of the rule that the set of objects compared must be homogeneous or almost homogeneous. This may be interpreted as a recommendation not to compare objects which are different in nature (outliers) [Guzik 2009, Haas, Murphy 2003]. An outlier is defined as an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism [Ben-Gal 2010]. For example, Afonso and Aubyn exclude Mexico and Turkey from their study, because their outputs are outliers, in particular their infant mortality rates (25,9 and 40,3 respectively, while the mean value for all OECD countries is 7,1) [Afonso, Aubyn 2005]. Similarly, in the course of another evaluation [Hadad et al. 2013] Chile, Mexico and Turkey are excluded from the analysis because their purchasing power parity-adjusted GDP per capita is below 50% of the OECD average.

PROPOSED MODEL

An output-oriented SBM model with constant returns to scale is adopted here. This is appropriate in this context since healthcare systems desire to maximize health gains, rather than hold health gains constant and minimize inputs, as assumed in an input-oriented model [Hadad et al. 2013]. Let the DMU set consist of *n* objects, each having *m* inputs and *s* outputs. Following Cooper et al. [2011], the output-oriented SBM efficiency ρ_o^* for DMU_o is defined as:

$$\frac{1}{\rho_o^*} = \max_{\lambda, s^-, s^+} \left(1 + \frac{1}{s} \sum_{r=1}^{s} \frac{s_r^+}{y_{ro}} \right) , \tag{1}$$

subject to:

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- \quad (i = 1,\ldots,m)$$
$$y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ \quad (r = 1,\ldots,s) \tag{2}$$
$$\lambda_j \geq 0(\forall j)\ s_j^- \geq 0(\forall i)\ s_r^+ \geq 0(\forall r)$$

where: $\lambda = [\lambda_1,\ldots,\lambda_n]$ are intensity variables,

$s_i^- = [s_1^-,\ldots,s_n^-]$, $s_r^+ = [s_1^+,\ldots,s_s^+]$ are vectors of input and output slacks respectively, and

$x_j = [x_{1j},\ldots,x_{mj}]$, $y_j = [y_{1j},\ldots,y_{sj}]$ are vectors of the inputs and outputs of DMU_j respectively.

In order to rank the SBM-efficient DMUs, the Super-SBM model can be used. An output-oriented super-SBM is defined in Cooper et al. [2011] as:

$$\rho_o^* = \min_{\bar{\mathbf{x}},\bar{\mathbf{y}},\boldsymbol{\lambda}} \frac{1}{(1/s)\sum_{r=1}^{s}(\bar{y}_r / y_{ro})} \tag{3}$$

subject to:

$$\bar{x}_i \geq \sum_{j=1, j\neq o}^{n} x_{ij}\lambda_j \quad (i = 1,\ldots,m)$$
$$\bar{y}_r \leq \sum_{j=1, j\neq o}^{n} y_{rj}\lambda_j \quad (r = 1,\ldots,s) \tag{4}$$
$$\bar{x}_i \geq x_{io}(\forall i)\ \bar{y}_r \leq y_{ro}(\forall r)\ \bar{y}_r \geq 0(\forall r)\ \lambda_j \geq 0(\forall j)\ j \neq o$$

Using an optimal solution of the above equations $(\lambda^*, \boldsymbol{s}^{-*}, \boldsymbol{s}^{+*})$ a projection of DMU$_o$= $(\boldsymbol{x}_o, \boldsymbol{y}_o)$ on the efficient frontier is defined as [Cooper et al. 2011]:

$$(\bar{\boldsymbol{x}}_o, \bar{\boldsymbol{y}}_o) = (\boldsymbol{x}_o - \boldsymbol{s}^{-*}, \boldsymbol{y}_o + \boldsymbol{s}^{+*}) \tag{5}$$

This approach determine the robustness of the efficiency scores by changing the reference set of the inefficient DMUs; rank the efficient DMUs; and estimates the super efficiency of the DMUs. The super efficiency model excludes each observation from its own reference set so that it is possible to obtain efficiency scores that exceed unity [Mogha et al. 2014, Cooper et al. 2011, Zanboori et al. 2014, Hadad et al. 2013].

In this article, the three variables regarded as inputs characterising the financial means invested in a healthcare system and its basic resources are: I1 – total healthcare expenditure expressed as % of GDP; I2 – number of physicians per 1 000 residents; I3 – number of hospital beds per 1 000 residents. Four variables are used to characterise the outputs of healthcare systems: O1 – Infant Mortality Rate (IMR), measured as the number of deaths of children less than one year old per 1,000 live births; Potential Years of Life Lost (PYLL), O2 for men and O3 for women – these indicators are expressed per 100,000 males and females; O4 – Life Expectancy at birth (LE). The efficiency measurement techniques used in this paper imply that the outputs are measured in such a way that "more is better" [Afonso, Aubyn 2005]. In order to maintain this assumption, IMR is converted into ISR (Infant Survival Rate), calculated as the quotient 1,000/IMR, and PYLL is converted into the quotient 100,000/PYLL. The analysis covers 33 OECD countries, data for which are taken from the OECD Health Database for the three years: 2000, 2005 and 2010 [OECD 2012] (Turkey was excluded because PYLL data is missing). Calculations are made using DEA-Solver LV 3.0 by Saitech.

INTERPRETATION OF RESULTS

One of the main advantages of DEA is that it allows the DMUs to have the full freedom to select linear programming weights. However, the efficient frontier can be influenced by outliers. Therefore, it is crucial to check for the presence of atypical DMUs [Bellini 2012]. To ensure the homogeneity of the sample evaluated, the outliers are identified on the basis of the method described earlier [Hadad et al. 2013]. Chile and Mexico are excluded from the sample because the efficient frontier may be influenced by them (when these two countries are considered in the calculations, Chile ranks first and Mexico eighth). Calculations are performed using the SBM model and SBM-efficient countries are ranked using the Super-SBM model (in which the efficiency score may be greater than unity).

Table 1. Efficiency scores

Country	2000	2005	2010	Country	2000	2005	2010
Australia	0.90	1.01	0.91	Japan	1.06	1.02	1.10
Austria	0.55	0.61	0.62	Korea	1.29	1.21	1.09
Belgium	0.66	0.69	0.72	Luxembourg	0.74	1.04	1.01
Canada	1.05	1.09	1.08	Netherlands	0.82	0.83	1.01
Czech Republic	0.73	0.85	1.00	New Zealand	0.68	0.93	1.02
Denmark	0.72	0.72	0.75	Norway	0.85	0.87	0.89
Estonia	0.52	1.01	1.04	Poland	0.66	0.67	1.00
Finland	0.72	0.76	1.00	Portugal	0.66	0.73	0.80
France	0.55	0.63	0.67	Slovak Republic	0.62	0.58	0.53
Germany	0.53	0.61	0.65	Slovenia	0.73	0.83	1.12
Greece	0.74	0.72	0.68	Spain	1.03	1.03	1.01
Hungary	0.40	0.45	0.55	Sweden	1.19	1.35	1.26
Iceland	0.72	1.03	1.06	Switzerland	0.64	0.68	0.70
Ireland	0.74	0.84	0.95	United Kingdom	1.01	1.01	1.01
Israel	0.83	1.00	1.11	United States	1.01	0.71	0.75
Italy	0.85	0.87	0.89				

Source: own calculations

Table 1 shows changes in the efficiency score over the period surveyed. On the basis of the analysis of efficiency scores, several characteristic groups of countries can be distinguished, and on the basis of the analysis of changes in the values of the individual model variables it is possible to determine the potential factors which explain the calculated efficiency level. In all the years considered, the following countries had full efficiency: Canada, Japan, Korea, Spain, Sweden and the United Kingdom. The most dynamic growth in efficiency can be seen for Estonia, Iceland, New Zealand, Poland and Slovenia, while for the Czech Republic, Finland, Israel, Luxemburg and the Netherlands there was growth but to a lesser extent. All these countries were inefficient in 2000 but were fully efficient in 2010. The least changes can be seen for Greece, the Slovak Republic and the United States. In the Australian healthcare system, there is efficiency growth for 2005 and an efficiency drop for 2010. Ireland and Portugal have a significant growth in efficiency, but are inefficient in 2010. The rest of the countries keep the same low efficiency level in all the years surveyed, with small growth trends.

Table 2. Data concerning selected countries which improved their efficiency

Country	I1	I2	I3	O1	O2	O3	O4	Rank	Year
Estonia	5.3	3.3	7.2	119.0	6.9	19.0	70.6	30	2000
	5.0	3.2	5.5	185.2	8.2	23.6	72.7	8	2005
	6.3	3.2	5.3	303.0	11.5	34.7	75.6	8	2010
Iceland	9.5	3.4	7.5	333.3	21.9	43.4	80.1	20	2000
	9.4	3.6	6.0	434.8	32.6	54.0	81.2	5	2005
	9.3	3.6	5.8	454.5	31.5	61.9	81.5	7	2010
New Zealand	7.6	2.2	6.2	158.7	19.2	31.6	78.3	21	2000
	8.4	2.1	4.5	200.0	21.9	35.6	79.8	12	2005
	10.1	2.6	2.7	192.3	22.9	36.4	81	9	2010
Poland	5.5	2.2	4.9	123.5	10.5	25.7	73.8	24	2000
	6.2	2.1	6.5	156.3	11.5	29.0	75.1	26	2005
	7.0	2.2	6.6	200.0	12.9	33.5	76.3	16	2010
Slovenia	8.3	2.2	5.4	204.1	14.1	31.9	75.5	17	2000
	8.3	2.4	4.8	243.9	16.9	36.4	77.7	18	2005
	9.0	2.4	4.6	400.0	21.8	45.7	79.5	2	2010

Source: own calculations

Table 2 presents a juxtaposition of data which explain the increase in efficiency for selected countries. The successive columns contain the inputs and outputs according to the previous description, as well as their ranking for the year shown in the right-hand column. Estonia, New Zealand, Poland and Slovenia considerably improved the efficiency of their healthcare systems, mainly through increased spending. In 2010, growth in spending relative to 2000 amounted to 19% for Estonia, 33% for New Zealand, 27% for Poland and 8,5% for Slovenia. The PYLL ratio improved in the range 15-83% for women and 19-66% for men.

The infant mortality rate improved in the range from 21.1% (New Zealand) to 154.5% (Estonia).

Table 3 describes selected countries whose healthcare systems exhibited full efficiency. In Japan, Korea and Sweden, expenditure on healthcare increased by 25%, 58% and 17%. The number of hospital beds per 1,000 residents decreased in Japan (7%) and Sweden (24%) and increased significantly in Korea (88%). The PYLL ratio improved by 21-57% for men and by 19-46% for women. The IMR improved by 39.1-65.6%.

Table 3. Data concerning selected countries with full efficiency in the period examined

Country	I1	I2	I3	O1	O2	O3	O4	Rank	Year
Japan	7.6	1.9	14.7	312.5	24.0	46.9	81.2	2	2000
	8.2	2.4	14.1	357.1	26.0	50.6	82.0	7	2005
	9.5	2.2	13.6	434.8	29.1	55.7	83.0	4	2010
Korea	4.5	1.3	4.7	188.7	14.4	33.3	76.0	1	2000
	5.7	1.6	5.9	212.8	19.1	40.9	78.5	2	2005
	7.1	2.0	8.8	312.5	22.6	48.6	80.7	5	2010
Sweden	8.2	3.1	3.6	294.1	26.0	42.9	79.7	3	2000
	9.1	3.5	2.9	416.7	29.4	47.1	80.6	1	2005
	9.6	3.8	2.7	400	32.5	53.1	81.5	1	2010

Source: own calculations

Table 4. Data concerning selected countries where efficiency deteriorated

Country	I1	I2	I3	O1	O2	O3	O4	Rank	Year
Greece	8.0	4.3	4.7	169.5	19.0	42.7	78	13	2000
	9.7	5.0	4.7	263.2	20.9	45.7	79.2	21	2005
	10.2	6.1	4.9	263.2	21.9	50.4	80.6	26	2010
Slovak Republic	5.5	3.4	7.9	116.3	10.6	26.9	73.3	26	2000
	7.0	3.0	6.8	138.9	11.9	28.6	74	30	2005
	9.0	3.3	6.4	175.4	13.8	32.5	75.2	31	2010
United States	13.7	2.3	3.5	144.9	15.0	25.7	76.7	7	2000
	15.8	2.4	3.2	144.9	15.4	26.6	77.4	23	2005
	17.6	2.4	3.1	163.9	16.3	27.8	78.7	22	2010

Source: own calculations

The next group, shown in Table 4, consists of countries whose efficiency deteriorated to various extents. The United States is an interesting case. Healthcare expenditure rose by 28% but only led to an improvement in the PYLL ratio of 8.5% for women and men and in the IMR of 13.1%. The Slovak Republic radically increased expenditure – by 64% – which translated into an improvement in the PYLL ratio for women of 21% and for men of 30%, and in the IMR of 51%. Greece increased expenditure by 28% but the changes in outputs are similar to those in the Slovak Republic. For fifteen countries which were inefficient in 2010, a projection is calculated which describes the output levels that need to be achieved in order to attain full efficiency with the current level of inputs. The suggested values of outputs for a given country are calculated using the restrictive conditions (2) when the optimal values of the decision variables (i.e. the optimal slacks and intensity variables) are inserted. Table 5 shows the direction and magnitude of the required change in outputs, according to equation (5).

Table 5. Suggested percentage changes in outputs for countries inefficient in 2010

Country	Score	O1	O2	O3	O4
Australia	0.906	38.7	0.0	2.7	0.0
Austria	0.618	107.3	57.5	59.1	23.5
Belgium	0.718	64.1	26.7	46.5	20
Denmark	0.749	42.4	41.1	43.7	6.7
France	0.670	87.3	40.6	41.7	27.7
Germany	0.651	82.3	39.8	57.5	34.6
Greece	0.675	41.2	79.8	33.2	38.2
Hungary	0.546	66.4	143.1	106.7	17
Ireland	0.948	17.9	3.3	0.7	0.0
Italy	0.894	34.8	9.2	3.6	0.0
Norway	0.893	3.4	15.6	18.8	10.1
Portugal	0.799	4.6	58.5	25.3	12.2
Slovak Republic	0.526	147.0	118.9	84.7	9.5
Switzerland	0.703	92.2	22.3	34.7	19.5
United States	0.753	30.4	56.1	44.4	0.0

Source: own calculations

Australia, Ireland, Italy and Norway are the best performers in this group, with some outputs at the level of the fully efficient countries and with significantly lower suggested changes in outputs compared to the other countries. The least efficient countries, such as Austria, Hungary and the Slovak Republic should improve their outputs to a significantly greater extent. For example the greatest changes required regard the PYLL ratio in Hungary – 143.1% for men and 106.7% for women - and in the Slovak Republic – 118.9% for men and 84.7% for women. These output changes are attainable through the proper management of healthcare systems by following the examples of the reference group of efficient countries. In several countries, surpluses of inputs are observed, e.g. GDP per capita could be reduced in Denmark (11.4%) and the United States (38.2%). The United States is the country which of all the OECD members spends the largest proportion of GDP on healthcare. In 2010, it spent 17.6% of its GDP on healthcare while the average was 9.5%. In these cases the outputs could be achieved with reduced inputs.

CONCLUSIONS

The model which has been proposed for assessing the efficiency of healthcare systems in OECD countries has allowed the goal of this study to be achieved. Changes have been traced over an 11-year period at 5-year intervals. In a large majority of countries a constant upward trend in efficiency can be observed or it is maintained at a constant good level. Of course, there are exceptions, such as Greece, the Slovak Republic and the United States.

In addition to ranking, which as such is valuable information, the DEA method also makes it possible to identify fully efficient units which constitute a reference set (of best practice) for the inefficient countries to follow. Because DEA determines relative efficiency, in subsequent periods it undergoes changes. Moreover, the number of efficient DMUs increases. Hence, application of the super-efficiency model provides additional information about the ranking of fully efficient countries. The use of the SBM model enables all slacks in inputs and outputs to be taken into account, which increases the discriminatory power and let to obtain more accurate measurements of the model applied [Hsu 2014].

Certain general conclusions can be drawn from the analysis in this article. In the group of countries with full efficiency throughout the entire period analysed, there are mainly countries with good established economic conditions. The greatest growths in efficiency are recorded for three countries of the former eastern bloc: Estonia, Poland and Slovenia, the healthcare systems of which are still undergoing transformation. As mentioned by other authors [e.g. González et al. 2010], achieving additional increases in the health status of the population in rich countries is much more expensive because of decreasing returns (flat-of-the-curve hypothesis). This is exemplified by the US healthcare system, where a significant increase in inputs does not translate into a proportional increase in outputs. In poorer countries, a more decisive increase in outputs is attained with relatively lower inputs. By contrast, in rich countries increasingly costly innovations and services can barely lead to modest improvements in the general health level of the population. Hence, it has been suggested that redirecting resources to other programmes promoting healthy lifestyles and habits could perhaps better improve general health in rich countries. On the other hand, in less developed countries, even modest investments in healthcare can be dramatically effective in terms of lives saved, increases in life expectancy and general improvements in living conditions. However, the magnitude of these effects also critically depends on the way in which resources are employed.

Projections which show the directions of desired changes are an important part of DEA efficiency evaluation.

REFERENCES

Afonso A., Aubyn M.St. (2005) Non-parametric approaches to education and health efficiency in OECD countries, Journal of Applied Economics, Vol. VIII, No. 2, pp. 227-246.

Anell A., Willis M. (2000) International comparison of health care systems using resource profiles, Bulletin of the World Health Organization, Vol. 78, No. 6, pp. 770-778.

Bellini T. (2012) Forward search outlier detection in data envelopment analysis, European Journal of Operational Research, Vol. 216, Issue 1, pp. 200–207.

Ben-Gal. I. (2010) Outlier detection, in: Maimon O., Rockach L. (eds.) (2010) Data Mining and Knowledge Discovery Handbook, Springer, New York.

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.

Cooper W.W., Seiford L.M., Tone K. (2000) Data Envelopment Analysis, A Comprehensive Text with Models, Applications, References and DEA-Solver Software, Kluwer Academic Publishers, Boston/Dordrecht/London.

Cooper W.W., Seiford L.M., Zhu J.Z. (2011) Handbook on Data Envelopment Analysis, Springer, New York, Dordrecht, Heildelberg, London.

González E., Cárcaba A., Ventura J. (2010) Value efficiency analysis of health systems: does public financing play a role?, Journal of Public Health, Vol. 18, pp. 337-350.

Guzik B. (2009) Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.

Haas D.A., Murphy F.H. (2003) Compensating for non-homogeneity in decision-making units in data envelopment analysis, European Journal of Operational Research, Vol. 144, No 3, pp. 530–544.

Hadad S., Hadad Y., Simon-Tuval T. (2013) Determinants of healthcare system's efficiency in OECD countries, European Journal of Health Economics, Vol. 14, Issue 2, pp. 253-265.

Hsu Y.C. (2014) Efficiency in government health spending: a super slacks-based model, Quality and Quantity Journal, Vol. 48, Issue 1, pp. 111–126.

Liu J.S., Lu L.Y.Y, Lin B.J.Y. (2013 a) Data envelopment analysis 1978-2010: A citation-based literature survey, Omega-International Journal of Management Science, Vol. 41, Issue 1, pp. 3-15.

Liu J.S., Lu L.Y.Y., Lu W.M., Lin B.J.Y. (2013 b) A survey of DEA applications, Omega-International Journal of Management Science, Vol. 41, Issue 5, pp. 893-902.

Mogha S.K., Yadav S.P., Singh S.P. (2014) New slack model based efficiency assessment of public sector hospitals of Uttarakhand: state of India, International Journal of System Assurance Engineering and Management, Vol. 5, Issue 1, pp. 32–42.

OECD (2011) Health at a Glance 2011: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/health_glance-2011-en, p. 130, accessed 10 May 2013.

OECD Health Data (2012) Frequently Requested Data, http://www.oecd.org/health/health-systems/oecdhealthdata2012-frequentlyrequesteddata.htm, accessed 20 March 2013.

Ozcan Y.A. (2008) Health care benchmarking and performance evaluation. An assessment using Data Envelopment Analysis (DEA), Springer, New York.

Retzlaff-Roberts D., Chang C.F., Rubin R. M. (2004) Technical efficiency in the use of health care resources: a comparison of OECD countries, Health Policy, Vol. 69, Issue 1, pp. 55-72.

Zanboori E., Rostamy-Malkhalifeh M., Jahanshahloo G.R., Shoja N. (2014) Calculating Super Efficiency of DMUs for Ranking Units in Data Envelopment Analysis Based on SBM Model, The Scientific World Journal, Vol. 2014, Article ID 382390.

Zhu J., Cook W.D. (2007) Modelling data irregularities and structural complexities in Data Envelopment Analysis, Springer, New York.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 33 – 42

THE APPLICATION OF QUANTILE REGRESSION TO THE ANALYSIS OF THE RELATIONSHIPS BETWEEN THE ENTREPRENEURSHIP INDICATOR AND THE WATER AND SEWERAGE INFRASTRUCTURE IN RURAL AREAS OF COMMUNES IN WIELKOPOLSKIE VOIVODESHIP

Izabela Kurzawa, Jarosław Lira
Department of Finance and Accounting, University of Life Sciences in Poznan
e-mail: kurzawa@up.poznan.pl; jlira@up.poznan.pl

Abstract: The article presents the usefulness of quantile regression for the analysis of diversification in entrepreneurship in rural areas of communes in Wielkopolskie Voivodeship. The dependence between the entrepreneurship indicator value and the density and availability of the water and sewerage infrastructure was determined for individual quantiles of the entrepreneurship indicator distribution. This approach enables estimation of different quantile functions of the conditional cumulative distribution function of the entrepreneurship indicator. This analysis enables atypical observations when the conditional cumulative distribution function is diversified and does not have a standard form.

Keywords: quantile regression, entrepreneurship indicator, economic infrastructure, rural areas

INTRODUCTION

The demand for infrastructure and its services is related to the degree of socioeconomic development. The greater the growth and development of a county or commune is, the greater the demand for infrastructural services is [Ratajczak 1999].

The development of entrepreneurship in rural areas is significantly related to the distance between a county or commune and a major economic centre. As the distance from the centre increases, the number of business entities per 10,000 rural inhabitants at the working age decreases [Salamon 2009]. The development of economic infrastructure in rural areas significantly influences both the number of business entities in the national economy and the growth of this number per 10,000 inhabitants at the working age. Therefore, increasing the infrastructural equipment in rural areas is a sine qua non, because it is a factor enabling their further growth and development [Lira 2014].

Microenterprises and small enterprises (about 99% of the total number of enterprises) play a key role in rural development. Therefore, the development of entrepreneurship plays a significant role in the development of local economy. It increases the production of goods, employment, people's income, the commune's budget and better satisfies local needs. It is necessary to stress the significance of the development of entrepreneurship for the functioning and socioeconomic development of the commune. It seems significant to research the influence of various traits, such as the influence of the density and availability of the water and sewerage infrastructure in rural areas on entrepreneurship in communes.

The aim of the study was to assess the usefulness of quantile regression for the analysis of dependence between the entrepreneurship indicator and the water and sewerage infrastructure in rural areas in Wielkopolskie Voivodeship. The article analyses communes in 2013 according to the density and availability of the water and sewerage infrastructure and registered business entities of the national economy (the REGON business entity register) in rural areas of Wielkopolskie Voivodeship.

METHOD

The article uses quantile regression because it is a method that enables the use of atypical observations (outlying observations). It is an advantage that the method uses the whole sample, so there is no problem of burdening parameter estimators. This problem would occur if the method of least squares was applied to subsamples identified on the basis of the dependent variable, i.e. the entrepreneurship indicator [Koenker 2005]. The authors of this method, Koenker and Basset [1978], observed that in the case of heteroscedasticity the estimation of a 0.5 quantile regression may prove to be a more effective method of searching for parameter values than the traditional regression based on the expected value of a dependent variable.

The quantile regression model can be presented in the following form [Koenker 2005,Trzpiot 2012, Davino at al. 2014]:

$$y_i = \beta_0^{(p)} + \sum_{j=1}^{J} \beta_j^{(p)} x_{ij} + \varepsilon_i^{(p)} \tag{1}$$

where:

$$Q_p(y_i|x_{ij}) = \beta_0^{(p)} + \sum_{j=1}^{J} \beta_j^{(p)} x_{ij} \tag{2}$$

conditional p-th quantile of dependent variable Y with known values of X_j variables,
y_i – dependent variable values,

x_{ij} – values of independent variables (j=1, 2, …, J),
$0 < p < 1$ – the index defining regression parameters for p-th quantile of variable Y distribution.

Each time the estimation is carried out on the total sample. However, a different beta parameter is estimated for each quantile of the dependent variable. The estimation of quantile regression parameters consists in minimisation of the weighted sum of absolute remainder values, where appropriate weights are assigned to them[1]:

$$min \sum_{i=1}^{n} \rho_p \left(\left| y_i - \left(\beta_0^{(p)} + \sum_{j=1}^{J} \beta_j^{(p)} x_{ij} \right) \right| \right) \tag{3}$$

where:

$$\rho_p(z) = \begin{cases} pz & for\, z \geq 0 \\ (1-p)z & for\, z < 0 \end{cases} \tag{4}$$

$$z = y_i - \left(\beta_0^{(p)} + \sum_{j=1}^{J} \beta_j^{(p)} x_{ij} \right) \tag{5}$$

The estimated quantile regression parameters are interpreted similarly to classical regression estimators, i.e. parameter $\beta_j^{(p)}$ indicates variation in a particular quantile p of dependent variable Y as a result of unit variation of j-th independent variable X_j, where we assume that the other variables do not change. This enables us to show the diversified influence of independent variables on individual quantiles of dependent variable distribution. On the other hand, an absolute term can be interpreted as an approximate conditional distribution of the quantile function for dependent variable Y, where we assume that the values of independent variables equal zero.

The following measures are usually used to assess the quality of estimated quantile regression:

1. The Wald test is used to measure the significance of parameter assessment (the zero hypothesis assumes the insignificance of each individual parameter in the model) [Koenker, Machado 1999]:

[1] Atypical observations receive lower weights and thus, the problem of including them in the model is solved. Depending on the phenomenon character and data distribution, in empirical applications 3-9 different quantile regressions are usually estimated (these regressions correspond to consecutive quantiles or deciles in the distribution). The phenomenon is analysed according to all the models obtained. The bootstrap method is usually applied to obtain the estimators of standard errors of coefficients for the quantile regression. The STATA 12 software was used to estimate models in the article.

$$\begin{cases} H_0: \beta_j^{(p)} = 0 \\ H_1: \beta_j^{(p)} \neq 0 \end{cases} \qquad z = \frac{\hat{\beta}_j^{(p)}}{D(\hat{\beta}_j^{(p)})} \quad \text{(j=0,1,...,J)} \tag{6}$$

2. Pseudo-R^2 [Davino et al. 2014]:

$$\text{pseudo-}R^2 = 1 - \frac{\sum_{y_i \geq \hat{y}_i} p \cdot |y_i - \hat{y}_i| + \sum_{y_i < \hat{y}_i} (1-p) \cdot |y_i - \hat{y}_i|}{\sum_{y_i \geq \hat{y}_i} p \cdot |y_i - \hat{Q}| + \sum_{y_i < \hat{y}_i} (1-p) \cdot |y_i - \hat{Q}|} \tag{7}$$

and [Koenker, Machado 1999]:

$$R^1 = 1 - \frac{\sum_{y_i \geq \hat{y}_i} \tau \cdot |y_i - \hat{y}_i| + \sum_{y_i < \hat{y}_i} (1-\tau) \cdot |y_i - \hat{y}_i|}{\sum_{y_i \geq \hat{y}_i} \tau \cdot \left|y_i - \hat{\beta}_0^{(p)}\right| + \sum_{y_i < \hat{y}_i} (1-\tau) \cdot \left|y_i - \hat{\beta}_0^{(p)}\right|} \tag{8}$$

where:
y_i – dependent variable values,
$\hat{y}_i$ – theoretical values of dependent variable,
$0 < p < 1$ – the index defining regression parameters for p-th quantile of the distribution of the variable Y – it is used as a weight,
$\hat{Q}$ – estimated quantile from the sample,
$\hat{\beta}_0$ – quantile for dependent variable Y from the estimated model, where we assume that the values of independent variables equal zero.
In theory these measures assume values from the interval [0,1], but they cannot be interpreted as coefficients of determination from classical linear regression. They are only a local measure of goodness of fit between the model and a particular quantile rather than the global measure of goodness of fit in the total conditional distribution. The higher the value of the measures is, the better the model was estimated.
In this study the authors suggest that the assessment of the model should be additionally supplemented with a quantile coefficient of determination and quantile coefficient of variation, adapted to quantile regression. They prove the goodness of fit between the model and empirical data [Rousseeuw, Leroy 1987]:

$$\text{quantile } R^2 = 1 - \left[\frac{Med\left|r_i^{(p)}\right|}{Med\left|y_i - \hat{\beta}_0^{(p)}\right|}\right]^2 \tag{9}$$

and

$$v = \frac{\hat{\sigma}}{\hat{\beta}_0^{(p)}} \tag{10}$$

$$\hat{\sigma} = \left\{ {\sum_{i=1}^{n} w_i \cdot r_i^{(p)}}^2 \Big/ \left(\sum_{i=1}^{n} w_i - 2\right) \right\}^{0,5} \tag{11}$$

$$w_i = \begin{cases} 1 & if \ \left|\frac{r_i^{(p)}}{s^0}\right| \leq 2.5 \\ 0 & in\ other\ case \end{cases} \tag{12}$$

$$s_0 = 1.4826 \cdot (1 + \frac{5}{n-2}) \cdot \sqrt{Med_i {r_i^{(p)}}^2} \tag{13}$$

where:

$Med\left|r_i^{(p)}\right|$ – median of absolute values of residuals,

$Med\left|y_i - \hat{\beta}_0^{(p)}\right|$ – median of deviations of real values of dependent variable from quantile $\hat{\beta}_0^{(p)}$ for dependent variable Y from the estimated model, where we assume that the values of independent variables equal zero,

$\hat{\sigma}$ – quantile standard deviation,

w_i – weights.

The higher the value of the quantile coefficient of determination is and the lower the value of the quantile coefficient of variation is, the better the model was estimated.

DATA

The study was based on data from 207 rural communes and isolated rural areas of urban-rural communes in Wielkopolskie Voivodeship in 2013 [Local Data Bank, Central Statistical Office, Warsaw].

The entrepreneurship indicator in the communes was determined by calculating the number of business entities in the REGON business entity register per 10,000 inhabitants at the working age[2]. The indicator was assumed as dependent variable *Y*. In the first variant of the model independent variables were related with the density of the water and sewerage infrastructure and they were expressed as follows:

X_1 – density of water supply distribution network (km/100 km^2),
X_2 – density of sewerage distribution network (km/100 km^2),
X_3 – percentage of rural inhabitants[3] with access to sewage treatment plants (%).

In the second variant independent variables were related with access to the water and sewerage infrastructure and they were expressed as follows:

X_4 – percentage of rural inhabitants with access to water supply network (%),
X_5 – percentage of rural inhabitants with access to sewerage network (%),
X_3 – percentage of rural inhabitants with access to sewage treatment plants (%).

Table 1 shows basic descriptive statistics of the independent variables and dependent variable under analysis. The sewerage network density was characterised

[2] Inhabitants at the working age: men (15-64 years), women (15-59 years).
[3] The actual population residing in the area was taken into consideration (as of 31 December 2013).

by the greatest diversification (about 108%). Apart from that, the network density was also characterised by relatively high right-sided asymmetry (3.39). Some of the communes under analysis did not have a sewerage network (about 7% of the total number of communes under study). This resulted in diversification between the communes in terms of the percentage of inhabitants with access to sewage treatment plants (the standard deviation was 71% of the mean value of the variable). As far as the entrepreneurship indicator is concerned, there were rather considerable differences between the communes. The lowest value of the indicator was noted in the commune of Wysoka in Piła County (439.9 business entities per 10,000 inhabitants at the working age). High values of the indicator were noted in the enterprising rural communes of Suchy Las, Tarnowo Podgórne and Komorniki in Poznań County, where the number of business entities exceeded 2,500).

Table 1. Descriptive statistics for components of the water and sewerage infrastructure and the entrepreneurship indicator in rural areas of communes in Wielkopolskie Voivodeship in 2013

Descriptive statistics	Density of distribution network (km/100 km^2)		Percentage of rural inhabitants with access to (%)			Entrepreneur-ship indicator
	water supply	sewerage	sewage treatment plants	water supply network	sewerage network	
minimum	8.80	0.00	0.00	52.80	0.00	439.90
0.25 quantile	68.00	10.20	15.50	84.00	17.80	874.20
0.50 quantile	92.20	18.70	32.80	89.60	34.40	1 042.40
0.75 quantile	116.00	33.80	52.80	93.30	47.60	1 210.50
maximum	236.20	222.50	96.80	98.90	77.00	3 224.60
coefficient of variation based on mean (%)	41.22	107.64	70.57	8.49	59.62	34.70
skewness	0.55	3.39	-1.36	-0.07	0.46	2.20

Source: own calculation based on Local Data Bank, Central Statistical Office, Warsaw

SELECTED RESEARCH FINDINGS

The analysis of individual estimated quantile regression models was started with statistical verification of the models. Particular attention was paid to the significance of structural parameters because only these parameters can be interpreted. Apart from that, the goodness of fit between the model and empirical data was determined by analysing pseudo-R^2, R^1, quantile R^2 and quantile coefficient of variation v, which can be respectively treated as the local equivalents of the coefficient of determination and the coefficient of variation of the random component in a classical regression analysis estimated with the method of least squares. Tables 2 and 3 show the estimated parameter values and their errors for each model as well as probability *p* and coefficients describing the goodness of fit. It is noteworthy that quantile R^2 is characterised by much greater values than pseudo-R^2 and R^1. The values of pseudo-R^2 and R^1 are similar when the quartiles of the entrepreneurship indicator distribution are similar to the quartiles estimated on the basis of the model (the intercept).

The estimated quantile regression models enable determination of the diversification of the influence of individual traits referring to the density and availability of the water and sewerage infrastructure in rural areas of communes for the identified entrepreneurship indicator quantiles.

Table 2 shows the estimated 0.25, 0.5 and 0.75 quantile regression parameters of the entrepreneurship indicator. The analysis of the data in the table reveals that the density of the water supply distribution network had negative influence on the entrepreneurship indicator value in rural communes. For example, in 0.75 quantile of the entrepreneurship indicator distribution when the density of the water supply distribution network was increased by one ceteris paribus unit (i.e. 1 km per 100 km^2), the entrepreneurship indicator decreased by about 2.359. This means that in more entrepreneurial communes increasing the density of the water supply distribution network will not increase entrepreneurship, because this network is sufficient and it is treated as the basic network. Apart from that, in the 0.75 quantile, which referred to the communes with a relatively high entrepreneurship indicator, the influence was statistically significant ($p < 0.05$). On the other hand, the density of the sewerage network had statistically significant positive effect on entrepreneurship. The higher the value of this indicator was, the greater the entrepreneurship was. For example, an increase by one ceteris paribus unit in the communes with the highest entrepreneurship indicator values caused the entrepreneurship indicator to increase by about 11.705. This observation leads us to think that in more entrepreneurial communes the density of the sewerage distribution network is still unsatisfactory and the extension of the network causes a considerable increase in entrepreneurship. Usually this network is being developed in rural areas and it still is not sufficient. Apart from that, in each quantile of the distribution the influence was stronger by 7.757, 9.592 and 11.705, respectively. Thus, we can suppose that by increasing the density of the sewerage distribution network communes with high entrepreneurship indicator values will have greater chances for further significant development than less entrepreneurial communes and they will attract new investors. On the other hand, the percentage of rural inhabitants with access to sewage treatment plants had negative influence on the indicator under analysis and it proved to be statistically significant only in the 0.75 quantile.

Table 2. The dependence between the entrepreneurship indicator and the density of the water and sewerage infrastructure in rural areas of communes in Wielkopolskie Voivodeship in 2013

Conditional quantile for the entrepreneurship indicator	Constant	Density of distribution network (km/100 km^2)		Percentage of rural inhabitants with access to sewage treatment plants (%)
		water supply network	sewerage network	
0.25 quantile	**858.483**	**-0.957**	**7.757**	**-0.671**
standard error	71.084	0.747	3.662	1.952
p - value	0.000	0.202	0.035	0.732
goodness of fit (%)	7.8$^{a)}$	7.9$^{b)}$	28.7$^{c)}$	29.8$^{d)}$
0.50 quantile	**978.521**	**-1.073**	**9.592**	**-0.973**
standard error	90.706	0.883	2.274	1.235
p - value	0.000	0.226	0.000	0.432
goodness of fit (%)	14.2$^{a)}$	16.1$^{b)}$	32.7$^{c)}$	20.5$^{d)}$
0.75 quantile	**1,256.938**	**-2.359**	**11.705**	**-2.923**
standard error	124.291	1.012	1.816	1.360
p - value	0.000	0.021	0.000	0.033
goodness of fit (%)	20.5$^{a)}$	20.9$^{b)}$	44.0$^{c)}$	22.9$^{d)}$

$^{a)}$ pseudo-R^2, $^{b)}$ R^1, $^{c)}$ quantile R^2,$^{d)}$ quantile coefficient of variation υ.

Source: as in Table 1

Table 3. The dependence between the entrepreneurship indicator and the availability of the water and sewerage infrastructure in rural areas of communes in Wielkopolskie Voivodeship in 2013

Conditional quantile for the entrepreneurship indicator	Constant	Percentage of rural inhabitants with access to (%)		
		water supply network	sewerage network	sewage treatment plants
0.25 quantile	**529.759**	**2.739**	**3.190**	**0.736**
standard error	187.754	2.129	3.359	2.197
p - value	0.005	0.200	0.344	0.738
goodness of fit (%)	6.1$^{a)}$	42.3$^{b)}$	88.6$^{c)}$	45.6$^{d)}$
0.50 quantile	**724.186**	**1.991**	**6.397**	**-1.700**
standard error	346.362	4.165	2.782	2.612
p - value	0.038	0.633	0.022	0.516
goodness of fit (%)	7.2$^{a)}$	41.4$^{b)}$	77.6$^{c)}$	30.5$^{d)}$
0.75 quantile	**1,297.940**	**-3.604**	**11.485**	**-3.556**
standard error	289.171	3.486	2.370	1.836
p - value	0.000	0.302	0.000	0.054
goodness of fit (%)	8.3$^{a)}$	10.3$^{b)}$	39.1$^{c)}$	25.0$^{d)}$

$^{a)}$ pseudo-R^2, $^{b)}$ R^1, $^{c)}$ quantile R^2,$^{d)}$ quantile coefficient of variation υ.

Source: as in Table 1

Table 3 shows the estimated 0.25, 0.5 and 0.75 quantile regression parameters of the entrepreneurship indicator depending on the availability of the water and sewerage infrastructure. The analysis of the data in the table revealed that the percentage of the population with access to the sewerage network had positive influence on entrepreneurship in the communes. However, the influence increased for consecutive quantiles of the distribution of variable Y. In turn, there was low statistical significance between the entrepreneurship indicator and the percentage of the population with access to the water supply network. On the other hand, the percentage of the population with access to sewage treatment plants had positive influence on the increase in the entrepreneurship indicator only in communes with low entrepreneurship indicator values. Its negative influence was observed in more entrepreneurial communes. For example, in 0.75 entrepreneurship distribution quantile an increase in the percentage of rural population with access to sewage treatment plants by one ceteris paribus unit (by one per cent) caused the entrepreneurship indicator to decrease by 3.556, on average.

CONCLUSIONS

1. The quantile regression method complements the classical regression of least squares and it diversifies the influence of independent variables in the conditional quantiles of the dependent variable.
2. The quantile R^2 and quantile coefficient of variation v complement the possibilities to verify quantile regression models statistically.
3. The density of the water supply distribution network (km/100 km^2) had negative influence on the entrepreneurship indicator in the communes. The influence decreased in more entrepreneurial communes.
4. The density of the sewerage distribution network (km/100 km^2) had positive influence on the entrepreneurship indicator in the communes. The influence was the greatest in the most entrepreneurial communes.
5. The percentage of rural population with access to sewage treatment plants (%) had negative influence on the entrepreneurship indicator in the communes. Most likely, it indicates certain saturation and further increase in the percentage will decrease entrepreneurship.
6. As far as entrepreneurship is concerned, the density of the water and sewerage infrastructure is more significant than rural inhabitants' access to it.

REFERENCES

Davino C. Furno. M. Vistocco. D. (2014) Quantile Regression. Theory and Applications, John Wiley & Sons. Ltd. Oxford.

Koenker R. & Bassett G. Jr. (1978) Regression quantiles, Econometrica, Vol. 46. No. 1. pp. 33-50.

Koenker R. (2005) Quantile Regression, Econometric Society Monographs, No. 38, Cambridge University Press.

Koenker R. & Machado J. (1999) Goodness of fit and related inference processes for quantile regression, Journal of the American Statistical Association, 94. pp. 1296-1310.

Lira J. (2014) Rozwój infrastruktury gospodarczej a wskaźniki przedsiębiorczości na obszarach wiejskich województwa wielkopolskiego w latach 2004-2012, Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu, Vol. XVI, Issue 6, Warszawa-Poznań-Lublin, pp. 320-235.

Ratajczak M. (1999) Infrastruktura w gospodarce rynkowej, Akademia Ekonomiczna w Poznaniu, Poznań.

Rousseeuw P. J. Leroy A. M. (1987) Robust regression and outlier detection. John Wiley and Sons. New York.

Salamon J. (2009) Przestrzenne zróżnicowanie wartości wskaźnika przedsiębiorczości na obszarach wiejskich województwa świętokrzyskiego, Infrastruktura i Ekologia Terenów Wiejskich, No. 5, PAN, Oddział Kraków, pp. 231-239.

Trzpiot G. (2012) Ekstremalna regresja kwantylowa, Studia Ekonomiczne Uniwersytetu Ekonomicznego w Katowicach, Zeszyty Naukowe Wydziałowe, No. 91, Katowice 2012, pp. 11-20.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 43 – 52

DECOMPOSITION OF DIFFERENCES BETWEEN PERSONAL INCOMES DISTRIBUTIONS IN POLAND

Joanna Małgorzata Landmesser
Department of Econometrics and Statistics, WULS-SGGW
joanna_landmesser@sggw.pl
Krzysztof Karpio
Department of Informatics, WULS-SGGW
krzysztof_karpio@sggw.pl
Piotr Łukasiewicz
Department of Informatics, WULS-SGGW
piotr_lukasiewicz@sggw.pl

Abstract: In this paper we study differences between personal incomes distributions in Poland in 2002 and 2012. The empirical data have been collected within the Household Budget Survey project. We used the Machado & Mata decomposition, which utilizes quantile regression. This method allowed us to investigate differences between income distributions in the whole range of values, going beyond simple average value decomposition. We evaluated influence of person's attributes on the differences of incomes distributions in 2002 and 2012. By decomposing the differences into the explained and unexplained components we got information about their causes. The differences described by the explained part are caused by different characteristics of samples. The unexplained part shows differences caused by the changes of attribute importance.

Keywords: quantile regression, Machado & Mata decomposition, counterfactual distribution

INTRODUTION

Nowadays one can observe significant development of various microeconomic decomposition methods. Based on the works [Oaxaca 1973] and [Blinder 1973] one elaborated techniques which went far beyond simple comparison of average values, for example decomposition of the variances or the whole distributions. New techniques allowed to discover various factors influencing incomes distributions, as minimal wage [DiNardo et al. 1996]. They also have been useful in studying differences of incomes distributions for various group of people [Albrecht et al. 2003].

During the decomposition of differences between the distributions one utilizes so called counterfactual distributions. They are a mixture of an conditional distribution of the dependent variable and various distributions of the explanatory variables [see Juhn et al. 1993, DiNardo et al. 1996]. One of them, proposed in Machado & Mata [2005] decomposes differences of distributions using a quantile regression.

In this paper we compared incomes of the employees running the one-person households in 2012 with those in 2002. The data have been collected in the Household Budget Survey project in Poland. The aim of the work is to study differences between income distributions in year 2002 and 2012. By use of the [Machado & Mata 2005] decomposition method we investigated differences in the whole range of income values. The past studies in Poland were mostly focused on the decomposition of the average values by using the Oaxaca & Blinder method [e.g. Śliwicki & Ryczkowski 2014]. On the other hand, the studies of [Newell & Socha 2005] showed that many factors influence only high wages, localized in the high quatiles on the wages distribution. Similarly [Rokicka & Ruzik 2010] showed that differences between wages of men and women are the biggest in the right part of the distributions.

DECOMPOSITION METHODS

Oaxaca & Blinder decomposition of average incomes differences

We consider two groups of one-person households. The first one contains data for 2002, the second one – for 2012, denoted by T_1 and T_2 respectively. We also deal with the outcome variable *y*, and a set of predictors *X*. The variable *y* is individual income and predictors *X* are individual sociodemographic characteristics of households (people) such as sex, age, education and others. The idea of Oaxaca & Blinder decomposition can be applied whenever we need to explain the differences between the expected values of dependent variable *y* in two comparison groups [Oaxaca 1973, Blinder 1973]. We assume that the expected value of *y* conditionally on *X* is a linear function of *X*:

$$y_i = X_i \beta_i + v_i, \qquad i = T_1,\, T_2, \tag{1}$$

where X_i are characteristics of objects in the year *i* and β_i is the vector of parameters. The equation (1) can be estimated for both years:

$$\hat{y}_i = X_i \hat{\beta}_i, \qquad i = T_1,\, T_2\,. \tag{2}$$

The difference between the expected values of *y* in both years is as follows:

$$\Delta^{\mu} = E(y_{T_2}) - E(y_{T_1}) = E(X_{T_2})\beta_{T_2} - E(X_{T_1})\beta_{T_1} \,. \tag{3}$$

Based on (2) and (3) the decomposition of the difference Δ^{μ} has the following form:

$$\hat{\Delta}^{\mu} = \overline{X}_{T_2}\hat{\beta}_{T_2} - \overline{X}_{T_1}\hat{\beta}_{T_1} = \underbrace{(\overline{X}_{T_2} - \overline{X}_{T_1})\hat{\beta}_{T_2}}_{\hat{\Delta}^{\mu}_{\text{explained}}} + \underbrace{\overline{X}_{T_1}(\hat{\beta}_{T_2} - \hat{\beta}_{T_1})}_{\hat{\Delta}^{\mu}_{\text{unexplained}}} \tag{4}$$

The expression (4) is named Oaxaca & Blinder decomposition of average incomes differences [Oaxaca 1973, Blinder 1973]. The first component gives the effect of characteristics and expresses the difference of the potentials of both groups. The second component represents the effect of coefficients, typically interpreted as discrimination in numerous studies.

Decomposition of differences between distributions

The mean decomposition analysis may be extended to the case of differences along the whole distribution. Let $f^{T_1}(y)$ and $f^{T_2}(y)$ be the density functions for the variable y in 2002 and 2012, respectively. The distribution $f^{i}(y)$, $i = T_1, T_2$, is the marginal distribution of the joint distribution $\varphi^{i}(y, X)$:

$$f^{i}(y) = \int_{C(X)} \ldots \int \varphi^{i}(y, X)dX \,, \tag{5}$$

where X is a vector of individual characteristics observed [Bourguignon & Ferreira 2005]. Let $g^{i}(y|X)$ be the conditional distribution of y. Then one can (5) express as:

$$f^{i}(y) = \int_{C(X)} \ldots \int g^{i}(y|X)h^{i}(X)dX \,, \tag{6}$$

where $h^{i}(X)$ is the joint distribution of all elements of X in year i. The difference between the two distributions may be decomposed onto

$$f^{T_2}(y) - f^{T_1}(y) = [f^{T_2}(y) - f^{C}(y)] + [f^{C}(y) - f^{T_1}(y)], \tag{7}$$

where $f^{C}(y)$ is the counterfactual distribution, which can be constructed as

$$f^{C}(y) = \int_{C(X)} \ldots \int g^{T_2}(y|X)h^{T_1}(X)dX \,. \tag{8}$$

The first component in (7) gives the effect of the unequal different personal characteristic’s distributions in 2012 and 2002. The second component describes the inequalities between two distributions of y conditional on X. The difference with respect to the Oaxaca & Blinder decomposition is that this decomposition refer to full distributions, rather than just to their means.

Quantile regression

The linear regression assumes the relationship between the regressors and the outcome variable based on the conditional mean function. This gives only a partial insight into the relation. The quantile regression allows the description of the relationship at different points in the conditional distribution of y [Koenker & Bassett 1978].

We consider the relationship between the regressors and the outcome using the conditional quantile function:

$$Q_\theta(y|X) = \Phi^{-1}_{y|X}(\theta, X) = X\beta(\theta), \tag{9}$$

where $Q_\theta(y|X)$ – the θ^{th} quantile of a variable y conditional on covariates X, $\theta \in (0,1)$; $\Phi_{y|X}$ – the cumulative distribution of the conditional variable $y|X$. We assume that all quantiles of y conditional on X are linear in X. The quantile regression estimator for quantile θ minimizes the sum:

$$\sum_{i=1}^{n} \rho_\theta(y_i - X_i\beta) \to \min, \tag{10}$$

where $\rho_\theta(u) = \begin{cases} \theta u & \text{dla} \quad u \geq 0 \\ (\theta-1)u & \text{dla} \quad u < 0 \end{cases}$.

The sum (10) gives asymmetric penalties for over and under prediction. For each quantile other parameters are estimated. We interpret these coefficients as the returns to different characteristics X at given quantiles of the distribution of y. The standard errors of parameters are calculated using bootstrap method [Gould 1992].

Machado & Mata decomposition of differences in distributions

Machado & Mata [2005] have used quantile regression in order to estimate counterfactual unconditional income distributions. The unconditional quantile is not the same as the integral of the conditional quantiles. Therefore, authors provide a simulation based estimator where the counterfactual distribution is constructed from the generation of a random sample. The approach is the as follows:

(1) generate a random sample of size m from a $U[0,1]$: $\theta_1, \theta_2, \ldots, \theta_m$;

(2) using the dataset for T_1 estimate m different quantile regression $Q_{\theta_i}(y|X_{T_1})$, obtaining coefficients $\hat{\beta}_{T_1}(\theta_i), i = 1,\ldots,m$;

(3) generate a random sample of size m with replacement from X_{T_1}, denoted by $\{X^*_{T_1 i}\}, i = 1,\ldots,m$;

(4) $\{y^*_{T_1 i} \equiv X^*_{T_1 i}\hat{\beta}_{T_1}(\theta_i)\}, i = 1,\ldots,m$ is a random sample from the unconditional distribution $f^{T_1}(y)$.

Alternative distributions could be estimated by drawing X from another distribution and using different coefficient vectors. To generate a random sample from the income density that would prevail in group T_2 and covariates were distributed as $h^{T_1}(X)$, we follow the steps (1), (2) from the previous procedure for T_2 and then:

(3) generate a random sample of size m with replacement from X_{T_1}, denoted by $\{X^*_{T_1 i}\},\ i=1,...,m$;

(4) $\{y^{C*}_{T_2 i} \equiv X^*_{T_1 i}\hat{\beta}_{T_2}(\theta_i)\},\ i=1,...,m$ is a random sample from the counterfactual distribution $f^C(y)$.

The Machado & Mata decomposition of the difference between the income densities in two years for each quantile is as follows:

$$\begin{aligned}\hat{\Delta}^{\theta} &= Q_\theta(y^*_{T_2}|X^*_{T_2}) - Q_\theta(y^*_{T_1}|X^*_{T_1}) = \\ &= Q_\theta(y^*_{T_2}|X^*_{T_2}) - Q_\theta(y^{C*}_{T_2}|X^*_{T_1}) + Q_\theta(y^{C*}_{T_2}|X^*_{T_1}) - Q_\theta(y^*_{T_1}|X^*_{T_1}) = \\ &= \underbrace{(X^*_{T_2} - X^*_{T_1})\hat{\beta}_{T_2}(\theta)}_{\hat{\Delta}^{\theta}_{\text{explained}}} + \underbrace{X^*_{T_1}(\hat{\beta}_{T_2}(\theta) - \hat{\beta}_{T_1}(\theta))}_{\hat{\Delta}^{\theta}_{\text{unexplained}}}\end{aligned} \tag{11}$$

To estimate standard errors for the estimated densities we repeat the Machado & Mata procedure many times and generate a set of estimated densities.

EMPIRICAL DATA

The data were collected in the Household Budget Survey project for 2002 and 2012, group T_1 and T_2 respectively. The analyzed data regards households running by one person whose main source of earning comes from a work as an employee. The annual disposable incomes (variable *INC*, in thousands of PLN) in 2012 were compared with those obtained in 2002. The incomes in 2002 during the analysis were expressed in prices in 2012 (variable *INCREAL*). The sample consisted of 834 and 1594 persons in 2002 and 2012 respectively. For each person: sex, age, education, place of residence, type of labor position have been obtained. Based on the obtained attributes one defined the following describing variables:

SEX (0 – woman, 1 – man),
AGE (years),
EDU (education, 1–9, 1 – primary, . . ., 9 – tetriary),
RES (residence, 1–6, 1 – village, . . ., 6 – town ≥ 500k of inhabitants),
POS (0–1, 0 – manual labor position, 1 – non-manual labor position).
Features of the variables have been collected in the Table 1.

Table 1. The mean values and the standard deviations for the selected variables

	Whole sample	2002	2012
Number of observation	2428	834	1594
INC	27.92 (21.32)	19.76 (17.92)	32.19 (21.71)
INCREAL	30.09 (22.58)	26.07 (23.64)	32.19 (21.71)
SEX (% men)	40.28	38.49	41.22
AGE	41.65 (12.12)	40.65 (11.29)	42.16 (12.51)
POS (% non-manuals)	67.26	64.51	68.70

Source: own calculations

RESULTS

We compared the personal incomes distributions for years 2002 and 2012. In the first step of the analysis the Oaxaca & Blinder decomposition has been applied for the average values. The results are listed in the Table 2.

Table 2. The Oaxaca & Blinder decomposition of the average incomes differences

Average *INCREAL* in 2002			26.072
Average *INCREAL* in 2012			32.189
Raw gap			6.117
Aggregate decomposition			
Explained effect			1.302
Unexplained effect			4.816
% explained			21.3
% unexplained			78.7
Detailed decomposition			
explained component		unexplained component	
SEX	0.217	*SEX*	6.251
AGE	0.199	*AGE*	0.445
EDU	1.038	*EDU*	10.852
RES	–0.369	*RES*	1.462
POS	0.216	*POS*	–6.804
const	0.000	*const*	–7.390
Total	1.302	Total	4.816

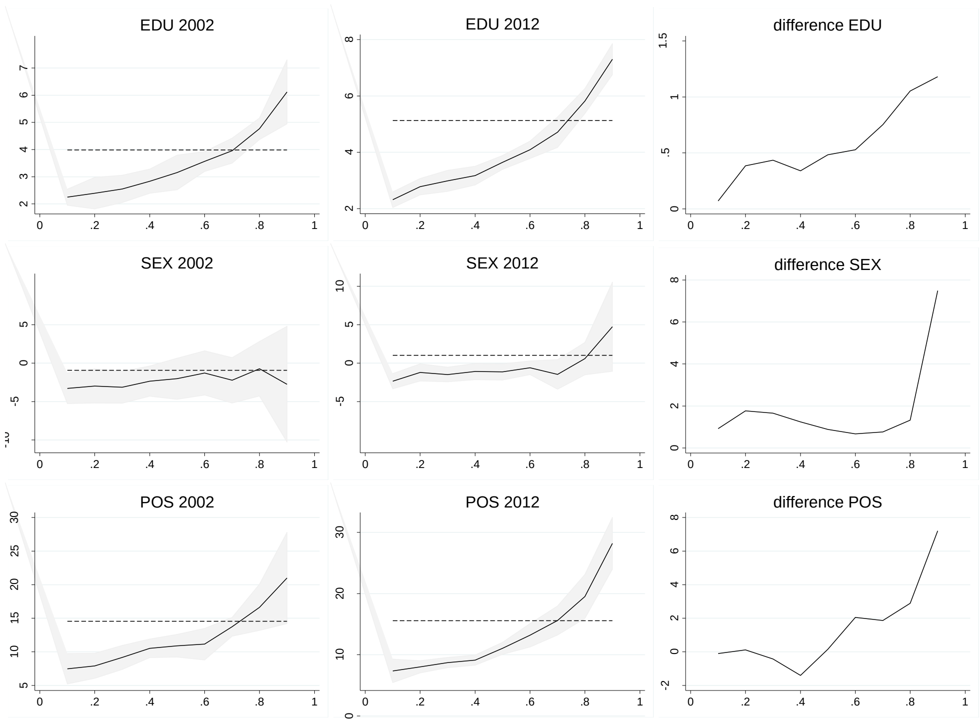
Source: own calculations

One can observe the positive difference between average values of the real incomes in 2012 and 2002. In the next step one tried to explain the observed difference. Using the decomposition method, one evaluated strength of the influence of the analyzed factors onto the average incomes. Generally, the differences are explained by the factors being studied in 21.3%. The *SEX*, *AGE*, *EDU*, *POS* were positively correlated with the change of the average value of incomes. However the biggest influence exhibited the *EDU* attribute. The increase of the average incomes are explained the most by the big increase of the education level in 2002 to 2012. On the other hand the *RES* exhibits negative correlation with the change of the average income.

The remaining 78% of average income changes are unexplained by the regression models being used. The unexplained part is assigned to the changes of the estimated parameters' influence onto the average income between 2002 and 2012. Such a different "labor market value" of parameters in the two years is the main source of the unexplained part of the model.

In the next step of the analysis the quantile regression models have been estimated. The influence of the selected attributes (measured by the model coefficients) onto the various quantiles of the income distribution are summarized in the Figure 1. The strength of the influence is presented as a function of the quantile range for both years. The differences of the results for both years are also presented.

Figure 1. The influence (vertical axis) of the selected attributes on the income distribution as a function of the quantile range (horizontal axis)



* The shaded areas represent 95% confidence intervals. The dashed lines represent results of the classical linear regression model.

Source: own preparation

The education has the positive impact on the income distribution in the whole range of values. However its importance increases with the quantile range, being the biggest for the highest incomes. The same behavior is observed for both years. The bigger influence on the income distribution is exhibited by the *POS* variable. This is directly related to the bigger incomes gained by the persons on the non-manual positions. However, instead of such a big overall influence we observe its rise along with the income distribution. On the other hand for the *SEX* variable we observe constant, negligible influence on the income for both years, especially in the right part of income distributions.

Studying the unexplained components of the quantile differences, related to the different parameters values in both years, one can conclude that for *EDU* as well as for *POS* they are quite similar. They increase with the values of the quantiles. However the differences for the *POS* parameters are relatively small for the left part of the distribution rising strongly for the higher quantiles.

On the other hand the unexplained part of the *SEX* parameters is negligible. The big rise of the differences for the last quantile is within the statistical error.

In the last step of the analysis one performed the decomposition of differences between income distributions. The differences are expressed as the sum of the explained and unexplained components along the whole income distributions. The Machado & Mata method has been used to estimate quantile regressions for 19 percentiles. The results for deciles are presented in the Table 3. The errors have been evaluated using the bootstrap method.

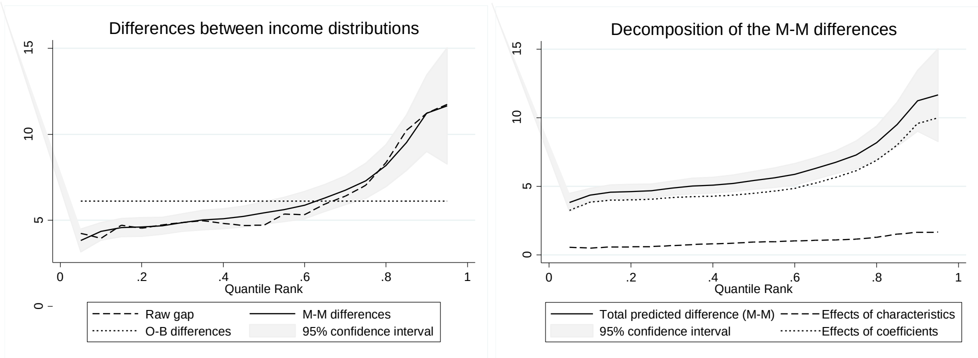
Table 3. The results of the Machado & Mata decomposition of the differences of the incomes distributions for 2002 and 2012

Quantile	Raw gap	Diference M-M	Explained part	Unexplained part	Explained %	Unexplained %
0.10	3.9567	4.3506 (0.2590)	0.5032 (0.3057)	3.8473 (0.2559)	12%	88%
0.20	4.5382	4.6094 (0.2784)	0.5934 (0.2674)	4.0160 (0.2707)	13%	87%
0.30	4.8726	4.8689 (0.2608)	0.6819 (0.2748)	4.1870 (0.2473)	14%	86%
0.40	4.8204	5.0937 (0.2955)	0.8079 (0.3076)	4.2858 (0.2825)	16%	84%
0.50	4.7140	5.4314 (0.3305)	0.9372 (0.3436)	4.4942 (0.3103)	17%	83%
0.60	5.3322	5.8808 (0.4041)	1.0257 (0.4118)	4.8551 (0.3791)	17%	83%
0.70	6.4007	6.7484 (0.4282)	1.0990 (0.4992)	5.6494 (0.4517)	16%	84%
0.80	8.3482	8.1739 (0.6231)	1.2908 (0.6968)	6.8831 (0.5822)	16%	84%
0.90	11.2275	11.228 (1.1327)	1.6515 (1.1115)	9.5765 (1.0202)	15%	85%

Source: own calculations

The results in Table 3 are also presented in the Figure 2. The left plot contains the raw differences between income distributions and the model predictions. The right plot shows the decomposition of the differences onto the explained and unexplained parts.

Figure 2. The results of the Machado & Mata decomposition of the differences of the incomes distributions for 2002 and 2012



Source: own preparation

The whole model approximates the data well. The differences between income distributions are rising with incomes. Their decomposition onto the explained and unexplained parts indicate low share of its explained part (12% to 17%). The share of the model's unexplained part is relatively high (83% to 88%) and increasing what indicates on the increase of the "labor market value" of the households' attributes. However, the explained part of the model also increases with income level.

SUMMARY

In this paper one studied differences between personal's income distributions in Poland in 2002 and 2012. The households of the single employers have been taken into account. The Oaxaca & Blinder and Machado & Mata decompositions of the average values and the whole incomes distributions respectively have been used. The Oaxaca & Blinder decomposition showed the positive influence of the most analyzed variables (*SEX*, *AGE*, *EDU*, *POS*) on the average income differences. The only variable with negative impact was *RES* what indicates on a "shift of big incomes towards smaller town". The Machado & Mata decomposition showed the increase of the differences between income distributions with the value of income. The differences were mostly caused by change of "labor market values" of the households' characteristics, described by the unexplained part of the model. These changes were greater when going towards big incomes.

The observed change might be caused by global processes in the European economy. It is known that such events took place in the past (e.g. financial crisis 2007-2009). Of course such a hypothesis needs to be confirmed through further studies. The future studies can also cover a detailed decomposition, which may exhibit the influence of the attributes on the whole income distribution.

REFERENCES

Albrecht J., Björklund A., Vroman S. (2003) Is There a Glass Ceiling in Sweden? Journal of Labor Economics, 21, pp. 145-177.

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

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

Gould W.W. (1992) Quantile Regression with Bootstrapped Standard Errors, Stata Technical Bulletin, 9, pp. 19-21.

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

Koenker R., Bassett G. (1978) Regression Quantiles, Econometrica, 46, pp. 33-50.

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

Newell A., Socha M. (2005) The Distribution of Wages in Poland, IZA Discussion Paper, 1485.

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

Rokicka M., Ruzik A. (2010) The Gender Pay Gap in Informal Employment in Poland, CASE Network Studies and Analyses, 406.

Śliwicki D., Ryczkowski M. (2014) Gender Pay Gap in the micro level – case of Poland, Quantitative Methods in Economics, Vol. XV, 1, pp. 159-173.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 53 – 62

A COMPARISON OF THE METHODS OF RELATIVE TAXONOMY FOR THE ASSESSMENT OF INFRASTRUCTURAL DEVELOPMENT OF COUNTIES IN WIELKOPOLSKIE VOIVODESHIP

Jarosław Lira
Department of Finance and Accounting
University of Life Sciences in Poznan
e-mail: jlira@up.poznan.pl

Abstract: The study analyses the possibility to measure the scale of disproportion in the development of the synthetic feature between spatial objects over a period of time on the basis of relativised values of diagnostic features. The study also proposes the construction of taxonomically relative indices of development according to the approach based on spatial median and it compares this approach with the classic approach proposed by Wydymus (2013). Both approaches are illustrated with a numerical example referring to the economic infrastructure in rural areas of the counties in Wielkopolskie Voivodeship.

Keywords: relative taxonomy, economic infrastructure, rural areas

INTRODUCTION

Relative taxonomy considers relativised indices, which are defined as quotients of the values of individual diagnostic features describing the synthetic feature for each spatial object relative to the values of other objects. Methods used in relative taxonomy enable us to determine the position of a particular spatial object relative to other objects. In the dynamic approach, they allow us to determine the process of levelling developmental disproportions in terms of development of the synthetic feature. Wydymus [2013] proposed the construction of taxonomically relative indices of development in the dynamic approach based on mean values of indices of relativised diagnostic features for individual spatial objects. On the other hand, Lira et al. [2014] proposed the identification of relative typological classes based on relative indices of development and their analysis in the dynamic approach.

The study presents the construction of taxonomically relative indices of development according to the approach based on Weber's spatial median and it compares this approach with the classic approach proposed by Wydymus [2013]. Both approaches were applied to measure the position of a particular county relative to other counties in terms of rural inhabitants' access to infrastructural services in Wielkopolskie Voivodeship in 2013. The research material was based on data provided in the electronic form by the Central Statistical Office in Warsaw – Local Data Bank [2015].

RESEARCH METHODOLOGY

The synthetic feature is composed of simple features, i.e. diagnostic features, which can be expressed with values in the form of structure or intensity indices. Let us assume that simple features are stimulants. Otherwise, it is necessary to unify their character (by converting destimulants and nominants into stimulants). Next, the values of individual diagnostics for each object and each time period were relativised according to the formula:

$$d_{(b/c)jt} = \begin{cases} x_{bjt}/x_{cjt} & x_{cjt} \neq 0 \\ 0 & x_{cjt} = 0 \end{cases}$$

where: $b \neq c$, $b = 1, 2, \cdots, n$, $c = 1, 2, \cdots, n$,

x_{ijt} – denoted the observation in the *i-th* object ($i = 1, 2, \cdots, n$) of the *j-th* diagnostic feature ($j = 1, 2, \cdots, m$) in time period *t* ($t = 1, 2, \cdots, k$).

Thus relativised values of diagnostic features in an object relative to other spatial objects for diagnostic *j* and time period *t* could be presented in the following form:

$$\mathbf{D}_{jt} = \begin{bmatrix} 1 & d_{(2/1)jt} & \cdots & d_{(\mathrm{n}/1)jt} \\ d_{(1/2)jt} & 1 & & d_{(\mathrm{n}/2)jt} \\ \vdots & & \ddots & \vdots \\ d_{(1/\mathrm{n})jt} & d_{(2/\mathrm{n})jt} & \cdots & 1 \end{bmatrix}.$$

Matrices $\mathbf{D}_{jt}$ make the basis for the construction of taxonomically relative indices of development of the synthetic feature in the classic and positional approach based on Weber's spatial median.

Classic approach proposed by Wydymus [2013]

In order to classify the objects with respect to all diagnostic features simultaneously the subsequent matrices were calculated [Wydymus 2013]:

$$\mathbf{D}_{jt}^{*} = \mathbf{A} \cdot \mathbf{D}_{jt}$$

where the matrix **A** was defined as:

$$\mathbf{A}=\begin{bmatrix} 0 & \cdots & \frac{1}{(n-1)} \\ \vdots & \ddots & \vdots \\ \frac{1}{(n-1)} & \cdots & 0 \end{bmatrix}$$

The diagonal elements of $\mathbf{D}_{jt}^{*}$ matrices formed matrices $\mathbf{W}_t$ for each time period *t*:

$$\mathbf{W}_t=\begin{bmatrix} w_{11t} & w_{12t} & \cdots & w_{1mt} \\ w_{21t} & w_{22t} & \cdots & w_{2mt} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1t} & w_{n2t} & \cdots & w_{nmt} \end{bmatrix}$$

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 [Wydymus 2013]:

$$\mathrm{S}_{it}=\frac{1}{m}\sum_{j=1}^{\mathrm{m}}\frac{1}{w_{ijt}},$$

but if $w_{ijt}=0$, we assume that $1/w_{ijt}$ is an arbitrarily determined value close to 0, e.g. 0.001.

Positional approach based on Weber's spatial median

Matrices $\mathbf{D}_{jt}$ make the basis for the construction of each spatial object *i* at period in time *t* of matrix $\mathbf{\Delta}_{it}$, which adopt the following form for individual objects:

$$\mathbf{\Delta}_{1t}=\begin{bmatrix} d_{(1/2)1t} & d_{(1/2)2t} & \cdots & d_{(1/2)mt} \\ d_{(1/3)1t} & d_{(1/3)2t} & & d_{(1/3)mt} \\ \vdots & & \ddots & \vdots \\ d_{(1/\mathrm{n})1t} & d_{(1/\mathrm{n})2t} & \cdots & d_{(1/\mathrm{n})mt} \end{bmatrix},$$

$$\mathbf{\Delta}_{2t}=\begin{bmatrix} d_{(2/1)1t} & d_{(2/1)2t} & \cdots & d_{(2/1)mt} \\ d_{(2/3)1t} & d_{(2/3)2t} & & d_{(2/3)mt} \\ \vdots & & \ddots & \vdots \\ d_{(2/\mathrm{n})1t} & d_{(2/\mathrm{n})2t} & \cdots & d_{(2/\mathrm{n})mt} \end{bmatrix},$$

$$\cdots$$

$$\mathbf{\Delta}_{nt}=\begin{bmatrix} d_{(\mathrm{n}/1)1t} & d_{(\mathrm{n}/1)2t} & \cdots & d_{(\mathrm{n}/1)mt} \\ d_{(\mathrm{n}/2)1t} & d_{(\mathrm{n}/2)2t} & & d_{(\mathrm{n}/2)mt} \\ \vdots & & \ddots & \vdots \\ d_{(\mathrm{n}/\mathrm{n}-1)1t} & d_{(\mathrm{n}/\mathrm{n}-1)2t} & \cdots & d_{(\mathrm{n}/\mathrm{n}-1)mt} \end{bmatrix}.$$

Next, Weber's spatial median

$$L_1\_med_{it}=(L_1\_med_{i1t},L_1\_med_{i2t},\cdots,L_1\_med_{imt})'$$

was calculated for each spatial object *i* at period in time *t* for the data gathered in matrices $\mathbf{\Delta}_{it}$, which can be treated as *n*-l observation vectors of *m*-feature objects. Matrix $\mathbf{\Omega}_t$ was used to classify spatial objects according to all diagnostic features.

$$\mathbf{\Omega}_{\mathrm{t}} = \begin{bmatrix} \omega_{11t} & \omega_{12t} & \cdots & \omega_{1mt} \\ \omega_{21t} & \omega_{22t} & \cdots & \omega_{2mt} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{n1t} & \omega_{n2t} & \cdots & \omega_{nmt} \end{bmatrix},$$

where:

for object $i = 1$:

$$\omega_{11t} = L_1\_med_{1t}\{d_{(1/2)1t}, d_{(1/3)1t}, \cdots, d_{(1/n)1t}\} = L_1\_med_{11t}$$
$$\omega_{12t} = L_1\_med_{1t}\{d_{(1/2)2t}, d_{(1/3)2t}, \cdots, d_{(1/n)2t}\} = L_1\_med_{12t}$$
$$\cdots$$
$$\omega_{1mt} = L_1\_med_{1t}\{d_{(1/2)mt}, d_{(1/3)mt}, \cdots, d_{(1/n)mt}\} = L_1\_med_{1mt}$$

for object $i = 2$:

$$\omega_{21t} = L_1\_med_{2t}\{d_{(2/1)1t}, d_{(2/3)1t}, \cdots, d_{(2/n)1t}\} = L_1\_med_{21t}$$
$$\omega_{22t} = L_1\_med_{2t}\{d_{(2/1)2t}, d_{(2/3)2t}, \cdots, d_{(2/n)2t}\} = L_1\_med_{22t}$$
$$\cdots$$
$$\omega_{2mt} = L_1\_med_{2t}\{d_{(2/1)mt}, d_{(2/3)mt}, \cdots, d_{(2/n)mt}\} = L_1\_med_{2mt}$$

for object $i = n$:

$$\omega_{n1t} = L_1\_med_{nt}\{d_{(n/1)1t}, d_{(n/2)1t}, \cdots, d_{(n/n-1)1t}\} = L_1\_med_{n1t}$$
$$\omega_{n2t} = L_1\_med_{nt}\{d_{(n/1)2t}, d_{(n/2)2t}, \cdots, d_{(n/n-1)2t}\} = L_1\_med_{n2t}$$
$$\cdots$$
$$\omega_{nmt} = L_1\_med_{nt}\{d_{(n/1)mt}, d_{(n/2)mt}, \cdots, d_{(n/n-1)mt}\} = L_1\_med_{nmt}.$$

Matrix $\mathbf{\Omega}_t$ was used to calculate the Φ_{it} matrix of relative indices of development for individual spatial objects at consecutive periods in time *t*:

$$\Phi_{it} = med\left\{\frac{1}{\omega_{i1t}}, \frac{1}{\omega_{i2t}}, \cdots, \frac{1}{\omega_{imt}}\right\},$$

but if $\omega_{ijt} = 0$, we assume ${}^{1}/_{\omega_{ijt}}$ is an arbitrarily determined value close to 0, e.g. 0.001.

The construction of taxonomically relative indices of development was extended for the case when a particular diagnostic feature assumes the value of 0 for a specific spatial objects. It can be observed especially with smaller objects, e.g. when rural inhabitants do not have access to a gas network in their communes.

This approach applies the median vector defined according to Weber's criterion. Therefore, it is called Weber's spatial median or it is defined as L_1 median or spatial median in reference publications [Lira 1999]. Let $\mathrm{K}_n^m = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n\} \in \boldsymbol{\mathcal{R}}^m$ be a set of *n* vectors of observation of *m*-feature objects and let $\widehat{\mathbf{\Theta}} \in \boldsymbol{\mathcal{R}}^m$ be the vector solving the optimisation problem

$$\mathrm{T}(\widehat{\mathbf{\Theta}}, \mathrm{K}_n^m) = \min_{\mathbf{\Theta} \in \boldsymbol{\mathcal{R}}^m} \mathrm{T}(\mathbf{\Theta}, \mathrm{K}_n^m),$$

where the objective function of this problem assumes the following form:

$$\mathrm{T}(\mathbf{\Theta}, \mathrm{K}_n^m) = \sum_{i=1}^{n}\left[\sum_{j=1}^{m}\left(\mathrm{x}_{ij} - \theta_j\right)^2\right]^{1/2},$$

where $\mathbf{X}_i = (\mathrm{x}_{i1}, \mathrm{x}_{i2}, \cdots, \mathrm{x}_{im})'$, $i = 1, 2, \cdots, n$ and $\mathbf{\Theta} = (\theta_1, \theta_2, \cdots, \theta_m)'$.

The measurement of relative synthetic feature variations is based on the construction of taxonomically relative indices of development and it consists of the following four stages:

stage 1 – proposing a system of diagnostic features which can be determinants of development of the synthetic feature,

stage 2 – relativising the values of diagnostic features for each spatial object at a period in time,

stage 3 – constructing taxonomically relative indices of development of the synthetic feature,

stage 4 – identifying relative typological classes of spatial objects according to relative synthetic feature variations and describing them.

Stage 1. The selection of diagnostic features is based on substantive premises and statistical analysis of diagonal elements in the inverse matrix of correlation matrix **R** in order to avoid excessive correlation of features[1] [Lira, Wysocki 2004].

Stage 2. We calculate individual indices, assuming that individual spatial objects at a period in time under analysis are the basis for comparisons of each diagnostic feature. Next, we construct appropriate relative matrices for these features at a particular period in time.

Stage 3. It is possible to apply the classic approach proposed by Wydymus to construct taxonomically relative indices of development. When the set of diagnostic features includes strongly asymmetric features or features with outlying observations, we can apply the positional approach based on Weber's spatial median. The lesser the value of relative index is than 1, the greater the relative advantage of a particular spatial object is over all other spatial objects in terms of synthetic evaluation of the period in time under investigation.

Stage 4. The values of relative indices are used for linear ordering of spatial objects and identification of relative typological classes. Grouping spatial objects from a (very) low to a (very) high level of relative development can be based e.g. on analysis of differences in the relative indices. When we have ordered spatial objects according to decreasing values of relative indices, we can calculate differences between its values for neighbouring objects, i.e. for the first and second object, for the second and third object, etc. When we analyse consecutive differences, starting with the first difference (between the first and second object), and when we find that the value of this difference is much greater than the others,

[1] If a feature is excessively correlated with other features, diagonal elements of inverse matrix $\mathbf{R}^{-1}$ are much greater than 10, which is a symptom of wrong numerical conditioning of matrix **R** [Malina & Zeliaś 1997].

we can identify a relative typological class with low development. The other differences enable us to identify more classes.

RESEARCH RESULTS

We conducted an empirical analysis in four stages, as described in the research methodology. In order to measure relative variations in the economic infrastructure in rural areas of the counties in Wielkopolskie Voivodeship in 2013 we selected five continuous quantitative diagnostic features, whose values were expressed as stimulant structure indices (the intensity index, except roads):

- length of public communal roads of improved hard surface in km per 100 km^2 of county's rural areas (roads),
- users of water supply network as percentage of total rural population[2] (water supply),
- users of sewerage network as percentage of total rural population (sewerage),
- users of gas network as percentage of total rural population (gas),
- users serviced by sewage treatment facilities as percentage of total rural population (sewage treatment).

The calculations presented in Table 1 let us draw the following conclusions:

1. There is strong right-sided asymmetry in rural inhabitants' access to gas networks (1.47). This situation is strongly influenced by the outlying observation for Poznań County (66.6%), which is over 22 p.p. greater than in Grodzisk County, which is in the second position. Apart from that, this feature is characterised by relatively high dispersion (98.64% in the classic approach or 72.48% in the positional approach) among the features for which the coefficient of variation was calculated.
2. The coefficients of variation based on Weber's spatial median resulted in lower values for all the features than the classic coefficient of variation with a corresponding value.

Table 2 shows the results of research conducted at stages 3 and 4. In the classic approach the taxonomically relative indices of development resulted in excessively high values in Turek County (14.622) and Czarnków-Trzcianka County (2.425). Apart from that, we studied the similarity of orderings obtained by means of taxonomically relative indices based on the classic approach and on Weber's spatial median. We assumed Spearman's rank correlation coefficient as a measure of similarity between the orderings. Although the value of this coefficient was relatively high $r = 0.84; p = 0.000$, there were some differences between the positions of counties in terms of their economic infrastructure. The greatest difference was observed in Konin County, because in the positional

[2] taken as the number of actual inhabitants as of December 31, 2013

approach the position of this county was 10 positions worse than in the classic approach. Only two counties, i.e. Śrem and Turek, occupied the same positions in both approaches.

Table 1. Descriptive statistics characterising the components of economic infrastructure in rural areas of the counties in Wielkopolskie Voivodeship (as of 31 December 2013)

Characteristics	Roads	Water supply	Sewerage	Gas	Sewage treatment
minimum	10.24	75.57	16.60	0.03	19.19
lower quartile	22.13	85.76	27.21	3.64	27.02
marginal median	41.55	89.03	34.29	13.04	39.59
Weber's spatial median	38.90	88.59	35.69	14.64	38.58
upper quartile	52.03	91.50	43.04	22.05	51.41
maximum	75.23	94.13	53.68	66.64	72.47
coefficient of variation (%)	44.68	4.92	29.33	98.64	37.21
median absolute deviation[1)]	15.52	3.28	8.41	10.61	12.43
coefficient of variation based on Weber's spatial median[2)] (%)	39.91	3.70	23.57	72.48	32.23
kurtosis	-0.97	0.78	-0.81	2.32	-0.60
skewness	-0.01	-0.95	-0.13	1.47	0.42
diagonal elements of the inverse correlation matrix $\mathbf{R}^{-1}$	1.68	1.64	5.30	1.31	5.10

[1)] Median absolute deviation is defined by $mad_j = med_{ij}|x_{ij} - L_1\_med_j|$.

[2)] Coefficient of variation based on Weber's spatial median is given by $v_{pj} = \frac{mad_j}{L_1\_med_j} \cdot 100\%$.

Source: own calculations based on Local Data Bank, Central Statistical Office, Warsaw 2015

In both approaches we identified 4 relative typological classes (stage 4) and we made the following observations:

- 54.8% of the total number of counties belonged to the same relative typological class, whereas the other counties were grouped one class higher or lower,
- the class with the highest degree of relative infrastructural development (class I) was less numerous in the positional approach (4 counties) than in the classic approach (7 counties),
- class II in the positional approach (12 counties) was more numerous than in the classic approach (7 counties),
- there were similar numbers of counties in classes III and IV.

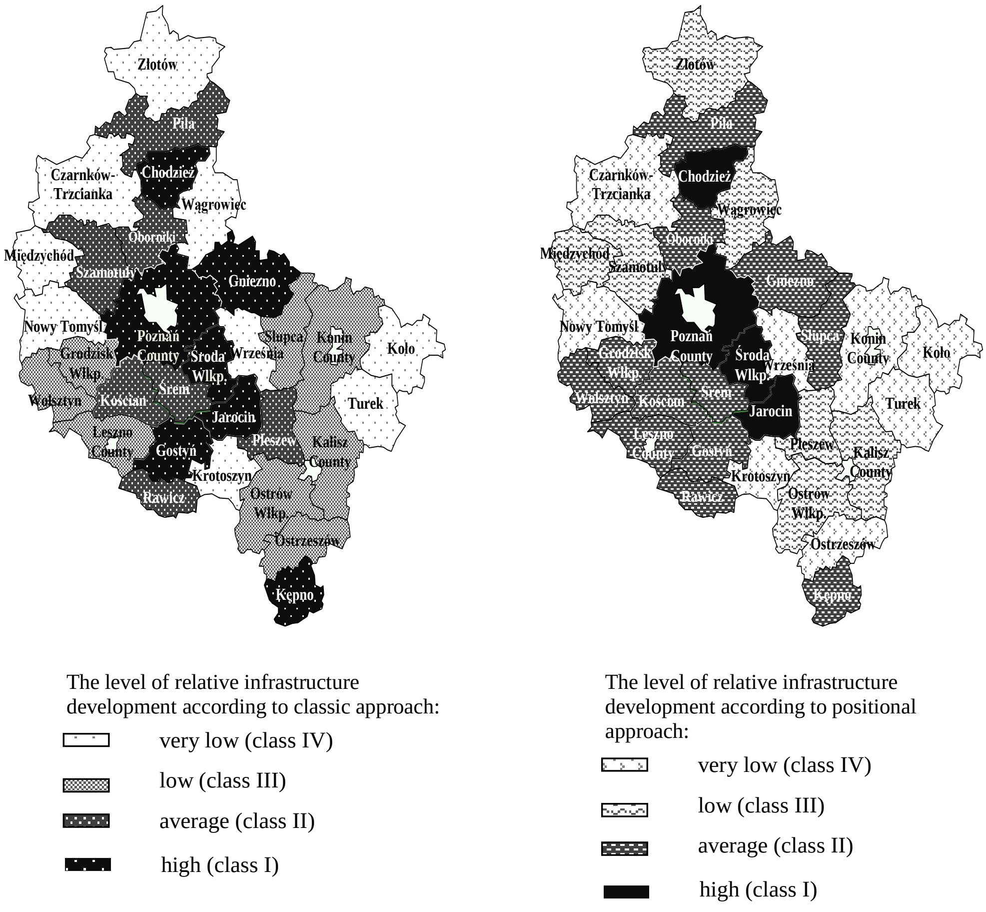
Table 2. The ordering and classification of rural areas of the counties in Wielkopolskie Voivodeship according to the economic infrastructure by means of the taxonomically relative indices of development

No.	Counties	Relative indices based on mean			Relative indices based on Weber's spatial median		
		ranks	indices	typological classes	ranks	indices	typological classes
1	Chodzież	3	0.571	I	1	0.664	I
2	Czarnków-Trzcianka	30	2.425	IV	25	1.307	IV
3	Gniezno	5	0.624	I	6	0.815	II
4	Gostyń	6	0.643	I	11	0.933	II
5	Grodzisk Wlkp.	22	0.889	III	16	0.984	II
6	Jarocin	1	0.542	I	2	0.668	I
7	Kalisz	15	0.808	III	21	1.084	III
8	Kępno	7	0.656	I	8	0.871	II
9	Koło	24	1.025	IV	30	1.730	IV
10	Konin	16	0.818	III	26	1.373	IV
11	Kościan	12	0.750	II	5	0.800	II
12	Krotoszyn	25	1.044	IV	29	1.618	IV
13	Leszno	19	0.836	III	13	0.950	II
14	Międzychód	27	1.067	IV	22	1.086	III
15	Nowy Tomyśl	29	1.132	IV	27	1.432	IV
16	Oborniki	13	0.782	II	9	0.879	II
17	Ostrów Wlkp.	21	0.889	III	17	1.006	III
18	Ostrzeszów	18	0.835	III	24	1.252	IV
19	Piła	8	0.693	II	7	0.863	II
20	Pleszew	11	0.730	II	20	1.078	III
21	Poznań	2	0.546	I	4	0.762	I
22	Rawicz	9	0.711	II	12	0.948	II
23	Słupca	20	0.853	III	15	0.979	II
24	Szamotuły	10	0.728	II	18	1.010	III
25	Środa Wlkp.	4	0.601	I	3	0.693	I
26	Śrem	14	0.784	II	14	0.954	II
27	Turek	31	14.622	IV	31	1.822	IV
28	Wągrowiec	26	1.046	IV	19	1.045	III
29	Wolsztyn	17	0.828	III	10	0.882	II
30	Września	23	1.001	IV	28	1.532	IV
31	Złotów	28	1.069	IV	23	1.088	III

Source: as in Table 1

The identification and analysis of the values of the taxonomically relative indices of development of the economic infrastructure according to the classic and positional approach revealed considerable spatial diversification in rural areas of the counties in Wielkopolskie Voivodeship. The diversification is illustrated in Figure 1, where the counties with the highest level of relative infrastructural development are marked with the darkest colour, whereas the counties with the lowest level of relative infrastructural development are marked with the lightest colour.

Figure 1. The delimitation of rural areas in Wielkopolskie Voivodeship according to the relative development of availability of infrastructural services in individual counties in 2013



Source: the Author's compilation based on the information in Table 2

CONCLUSIONS

The application of multidimensional methods of relative taxonomy enables analysis of relative developmental disproportions between individual counties and all the others in terms of the economic infrastructure in rural areas. The article presents two methods of construction of taxonomically relative indices of development according to the classic and positional approach in the dynamic aspect. The comparison of both approaches was based on one period of time that was arbitrarily selected for presentation.

The positional approach based on Weber's spatial median, which was applied for the construction of taxonomically relative indices of development, let us draw the following conclusions:

- the application of the positional approach is justified in a situation when the set of diagnostic features includes strongly asymmetric features or features with outlying observations,
- the taxonomically relative indices of development based on Weber's spatial median better reflects developmental disproportions between individual counties and all the others than the classic approach,
- the counties with sustainable development of economic infrastructure are positioned higher in the positional approach,
- the analysis of relative developmental disproportions can be conducted in a dynamic aspect, observing whether a particular county or relative typological class increases or decreases the developmental advantage over the others within the period of time under investigation.

REFERENCES

Lira J. (1998) Zastosowanie wybranych metod statystycznych opartych na medianie w doświadczalnictwie rolniczym, Praca doktorska, AR Poznań.

Lira J., Głowicka-Wołoszyn R., Wołoszyn A. (2014) The application of relative taxonomy methods to the study of technical infrastructure development in rural areas across the provinces of Poland, Quantitative Methods in Economics, Vol. XV, No. 2, pp. 330-338.

Lira J., Wysocki F. (2004) Zastosowanie pozycyjnego miernika rozwoju do pomiaru poziomu zagospodarowania infrastrukturalnego powiatów, Wiadomości Statystyczne, 9, GUS PTS, Warszawa, str. 39-49.

Local Data Bank, www.stat.gov.pl

Malina A., Zeliaś A. (1997) Taksonomiczna analiza przestrzennego zróżnicowania jakości życia ludności w Polsce w 1994 r., Przegląd Statystyczny 1, str. 11-27.

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, str. 631-645.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 63 – 72

ASSESSMENT AND SELECTION MODEL FOR MANAGEMENT SYSTEM SUPPORTING SMALL AND MEDIUM-SIZED ENTERPRISES

Rafik Nafkha
Department of Informatics
Warsaw University of Life Sciences – SGGW
e-mail: rafik_nafkha@sggw.pl

Abstract: The article presents analysis of selected ERP (Enterprise Resource Planning) systems in term of the application of the "Production" module in a manufacturing enterprise. Conditions for system evaluation and selection have been expressed in classes of criteria. In each class several practical criteria were considered and rated with great detail. The final results are presented on a point scale allowing the comparison of same class systems.

Keywords: assessment ERP systems, ERP system selection criteria, ERP evaluation model

SO IS USUALLY IS!

Taking part in a dozen or so of ERP systems implementation projects, from my observations, the method to choose an appropriate system by customers frequently consists in comparing functionality of selected primary a couple of systems. According to them, the best system is the one that has better functionality and a wider scope, fits to all company departments, has a nice interface, easy to implement, is reliable, etc. To avoid partiality suspicions during the system selection, customers compare them on their own responsibility. In this case, they often invite producers and their representatives to present their solutions, they ask for testing different tasks and sometimes they make reference visits in order to achieve users opinions about offered systems. From the statements of individuals taking parts when system selection, follows that the more facts they know about systems, then the less difference between them they note. Finally they feel confused. This shows that comparative method raises problems that are often difficult to deal. This happens because comparative methods assumes, that one

should be familiar with presented systems because only in this way they can be objectively compared. So the question is whether there is an alternative solution to the comparative method? and how to choose the simplest and better?
The main method that can be helpful when choosing an ERP system is to carry out an exact technical analysis based on a mathematical model, in which detailed selection criteria will be evaluated by individual users. Each of the criteria is given an indicator rating based on survey results (Table 1). The resulting indicators are corrected by other factors such as individual user preferences, risk or timeout of the project. On the basis of partial results, total assessment indicator for each considered system will be determined. Obtained result will not be definitive or incontestable because it depends on company type , its size and the preferences of groups taking part in the selection of the new system.

MODEL ASSESSMENT AND SELECTION OF INFORMATION SYSTEM

Consider i $\in$ (1..n) - baseline information system evaluation criteria. Basic rating indicator system selection (R) for adopted criteria can be defined as:

$$R = \sum_{i=1}^{n} x_i \tag{1}$$

Where:
x_i - assessing indicator for i-th selected criterion.
Considering that any single criterion consists of m particular criteria j $\in$ (1..m), then the sum of indicators at a given assessment level gives the evaluation indicator system selection:

$$R_{SI} = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} \tag{2}$$

x_{ij} - assessing indicator for n considered criteria.
The system selection indicator usually should be supplemented by individual user factors such as: the effect of personal preference arising from operated system attribute, the weight of each evaluation criteria, project timeout and/or budget exceeding ect.
Let consider the following - a_1 ,b_1……. z_m individual user factor, then the evaluation indicator system selection after correction will take the form,

$$R_{SI} = \sum_{i=1}^{n} ((a_1 b_1 \ldots z_1) x_{i1} + (a_2 b_2 \ldots z_2) x_{i2} + \ldots + (a_m b_m \ldots z_m) x_{im}) \tag{3}$$

One of the main criteria when selecting a new ERP system remains the price and system implementation cost. On the other hand, most of companies are also interested in the greatest functionality of the system. To ensure a balance between financial indicator criteria (Table 2 shows that users assign fewer points to the more expensive system) and the rest indicator criteria, the model should be optimized to meet Pareto-optimal point.

$$Max\sum_{j=1}^{k}(x_{ij}) \quad Min\sum_{j=k+1}^{m}(x_{ij}) \tag{4}$$

$\sum_{j=1}^{k}(x_{ij})$ – financial criteria indicators

$\sum_{j=k+1}^{m}(x_{ij})$ – rest of criteria indicators.

Such objective function approach allows the calculation of Pareto-optimal point [Dixit, Nalebuff 2008], by maximizing financial indicators criteria relative to the rest of criteria indicators.

MODEL APPLICATION FOR EVALUATION ERP SYSTEM FOR SMALL AND MEDIUM-SIZED MANUFACTURING ENTERPRISES

Browsing the statement available on the Polish market of ERP systems [Computerworld 2010], we can see how much competition is in this class solutions and how wide range of options are available to potential users. Tables [Computerworld 2010] show that, ERP market leaders belong to three leading brands. One of the leading suppliers of this class of systems include companies such: BPSP, SAP and Microsoft, then there is a middle class group with quite variable composition over time including TETA, IFS, COMARCH and others. The last group serves niche market segments. ERP systems ranking is a very virtual product sets and their suppliers. Data for these rankings are given frequently by stakeholders themselves that's why in this paper I used a subjective criterion when choosing ERP systems to investigate in the study: about whom the most it is said and heard among system users. The analysis concerns three selected ERP systems in management Small and Medium-Sized manufacturing enterprises (SME): Comarch ERP XL (shortly Comarch), Sage ERP X3 (Sage) and SAP Business One (SBO). The presented analysis has a task, rather, to be aware of aspects that one should pay attention during the selection process and system evaluation, than to specify absolutely the best system or confirm these ratings because these ratings look like they want those responsible for PR and advertising from suppliers.

To start the creation of the analysis firstable we should think about the main criteria to evaluate and appropriate ordering them according to the degree of importance [Grudzewski, Hejduk 2004]. Both the price and the software implementation value still remain as the main selection criteria. Sectorial companies opt for more expensive systems, more scalable and better suited to their industry. Financial companies or those acting in the regulated markets expect a system that is in accordance with the legislation in given sector. Regardless of any business specifics, system functionalities, which must have the system will provide the basis for the selection. Right from the start it is necessary to find answers to questions such as: what is the probability of system failure and data loss? The time required to restore the software? How data loss can be prevented in case of server damage or virus attack?

Besides the problem of choosing the suitable system, SME are faced with the problem of choosing the appropriate company that will supply, implement and provide the necessary level of technical support. Keeping in mind that ERP system usually is a project for many years, then from the selection step, it is necessary to check out the system upgrade features and the possibilities of making change in the system. At such moments, it becomes very important the scalability of the system and its integrity. The validation of specified criteria in the literature [Doradcy-IT 2014] are the result of the survey presented in Table 1 where companies show a high value of functionality, fault tolerance and cost of the system.

Table 1. System selection criteria according to respondents

1 – least important 5 – most important	1	2	3	4	5	Number of respondent
System functionality	0	0	0	8	22	30
System failure	1	0	4	5	20	30
Software price and implementation cost	0	3	6	4	17	30
Software implementation duration	2	2	5	9	12	30
Ability to add new modules	2	1	2	18	7	30
System flexibility	0	2	4	16	8	30
Technical support after implementation	0	2	5	8	15	30
System technology	1	1	10	16	2	30
Hardware requirement	2	6	6	12	4	30
Others	8	3	12	5	2	30

Source: [Doradcy-IT 2014]

For each criterion weight is determined on the basis of respondents answers. At the beginning we define for each criterion the weight of importance.

$$p_i = \frac{i}{\sum_{i=1}^{n}(i)} \tag{5}$$

p_i – i-th weight criterion validity, i $\in$ (1,2,3,4,5), n = 5

$$\sum_{i=1}^{n}(p_i) = 1$$

Then weights are determined for each criterion separately (single criterion consists of m particular criteria).

$$p_k = \sum_{i=1}^{m} p_i * l_i \tag{6}$$

l_i – number of respondents.

The weight of criterion can be used to determine the coefficient of respondents preferences. In this case, and to differentiate the systems relative to each other, it can be necessary to allocate additional points, e.g. 0,25 point to each investigated attribute over the standard feature [Chmielarz 2003]. The evaluation and selection of ERP system supporting production in SME's is going to be carried out according to the following criteria: financial, production attributes functionality, vendor support, system implementation and training, system integrity, scalability, and technological environment. Each of the criteria will be prescribed an evaluation of five-point scale. Points are designed for the existence or lack of particular features (lack = 0, exist but highly insufficient = 1, function partially complied = 2, fully satisfied = 3, exceeding the basic functionality = 4).

Financial criterion

One of the main criteria for the selection of a new ERP system is the software price and implementation cost [Grudzewski, Hejduk 2004]. The ERP Total Cost of Ownership (TCO) goes beyond these two components. The majority of SME companies, incorrectly interpret system value evaluating the software value on the basis of its standard functionality. It should, however, take into account other additional costs including installation, implementation and employees training costs. If the company does not have appropriate hardware infrastructure may also occur replacement computers, network equipment and other related services. In the investigated sector, SAP AG, a big provider of business software, proposes license (Standard/Start) depending on the enterprise needs. SAP AG together with its partners offer a special package with fixed price including 5 licenses with implementation service for 25 000 PLN [VisaCom 2015]. Comarch SA, a Polish business software supplier, provides a license package for 10 posts, including production feature for a price of 2000 PLN per module. Sage (The Sage Group, plc) is a general ERP solution provider for distributors and manufacturers, in accordance with the analysis "Sage ERP X3 vs Epicor", Sage costs for 10 users – including but not limited to, manufacturing: around $45,000 1 year (perpetual license).

Table 2. Financial criterion evaluation

Investigated criteria	SAP BO	Sage	Comarch
License cost	2	1	3
Implementation cost	2	2	3
Training cost	1	2	3
Technical Support cost	2	3	3
Client opinion	2	3	3
Total	9	11	15

Source: own based on survey result

Functional criterion

Each of the three analyzed systems have a built-in tools with similar functionality. SAP BO includes products tree and trees assembling products that serve as production routing. In the other hand it is not possible to indicate the work centers and machine responsible for a given stage production. The situation looks different in Comarch, where by defining the production routing, the workstations, machines and individual time operations are taken into consideration. Sage, relative to Comarch is enriched with service cooperation routing. The system has also weight station interface so that one can manage weighing materials during the manufacturing process. Material management in all considered systems covers the functionality required to purchase goods and services, manage inventory, and inspect incoming materials. Inventory Management includes issuing and transferring inventory, inventory restocking, and the inventory count and adjustment processes. Planning control functions for series-type production, make-to-order, single-item production and planning to warehouse are a fundamental part of production and control components of SAP BO. Comarch supports in addition variant production process, including mixed planning form (simplified, detailed or roughing). Sage supports finite and infinite capacity requirements planning. Both Sage and Comarch provide for Material Planners an interactive drag-and-drop scheduling tool in GANTT format for manual viewing, simulation, and updating of outstanding work orders and routing operations.

Table 3. Production module functionality evaluation

Investigated criteria	SAP BO	Sage	Comarch
Production management	3	4	3
Production planning	2	3	4
Production scheduling	2	3	4
Customer satisfaction	3	3	3
Total	10	13	14

Source: own based on survey result

Support system criterion

The most transparent and functional support features given by ERP system suppliers is to help client on-line. SAP AG provides a full description of the functionality of all modules of the system often supported by practical examples. On-line service is available in 26 languages. The supplier website provides an interactive guide, demonstration videos and tutorials. SAP Consultant can be contacted by phone and chat from 8 am to 4 pm. The Sage website contains information on different system modules and solutions which have been applied. There are brochures and documents describing the system functionalities. Sage inserted a hyperlink to its channel on YouTube platform, where there are instructional videos. Help and support is available directly from the system, but it is necessary to buy a subscription. Phone support is available 24 hours a day. On the Comarch SA website general description of individual system modules and their functionalities are presented. Detailed information about the system and benefits that arise from their use are well described in the documentation. Selected issues are illustrated and explained. The website has a database of instructional videos, tutorials for users, discussion forum and advices. In order to use these helps you have to be logged on the “Individual customer site”. Phone support operate from 9 am to 5 pm.

Table 4. Support system criterion evaluation

Investigated criteria	SAP BO	Sage	Comarch
On-line	4	2	2
Documentation access	3	3	3
Instructional videos	3	3	4
Technical phone support	2	4	2
Client opinion	4	3	3
Total	16	15	14

Source: own based on survey result

System implementation and training criterion

To standardize the ERP systems implementation, several methods were created. Almost every large supplier has developed its own methodology. SAP AG and its Partners have developed: GoForWard, STI (Short Time of Implementation), MASAP and others. These methods are usually developed during numerous implementation, carried out in various organizations adapting them to industries specificities. System deployment lasts between 2 and 8 weeks. Implementation costs depend on the number of purchased licenses. In addition to traditional on-site installation, SAP AG offers cloud solution and mobile deployment. Comarch SA employees have created a tool that assists and synthesizes the work of group involved in the realization of both partner and client implementation. It is an action plan, on the basis of which the implementation is proceeding smoothly. System deployment lasts from 2 to 8 weeks and depends on company size and business processes complexity. SIGMA (Sage Implementation Global Methodology Approach) by means of which Sage carries out implementation projects. The methodology includes project life cycle, adapted each time to the specificities of organization. The SIGMA technique developed over thousands of implementations carried out in different organizational cultures in many countries. Sage deployment may take about 3 months and depends on the number of purchased modules. The last Sage ERP X3 version 7 may be sold as cloud-based business software.

Table 5. Implementation and training criterion

Investigated criteria	SAP BO	Sage	Comarch
Implementation - Quickness	4	4	3
System fitting to customer need	4	3	3
Training access	4	2	3
Client opinion	4	3	3
Total	16	12	12

Source: own based on survey result

Integrity and system scalability criterion

SAP BO has an open architecture and uses Microsoft SQL Server, which became the standard, especially for SME. Main processes enabling the functioning of any company are fully integrated thanks to user and data interfaces. Data can be exported into Microsoft Excel or Word document. There is a standard interface for CAD (Computer Aided Design) applications and PDA (Personal Digital Design) applications to collect production data and then their storage. Data synchronization , contacts, tasks and e-mail between SAP BO and Microsoft Outlook improves system performance. Comarch is characterized by its flexible modular structure, thanks to which System can be enlarged for new users, new fields, modules and functionalities. Thanks to Comarch EDI service (Electronic Data Interchange), data exchange with any business partner becomes automatically and transparency. A characteristic feature of Comarch in term of logistic and production processes is the electronic exchange of different types of documents during customer orders execution including: storage state report, transshipment of consignment or goods, order status, etc.). The integration of Sage technology and Microsoft environment enables IT (Information Technology) sharing process, resources and increase company business effectiveness. Electronic Document Management (EDM) solution, ensures information exchange across the enterprise. Within the production scope, where supply chain is a very important process for effective planning, Sage and Preactor have conclude cooperation, which resulted in a global and fully integrated solutions for planning and scheduling production.

Table 6. System integrity and scalability evaluation

Investigated criteria	SAP BO	Sage	Comarch
Scalability	3	4	4
Integrity	1	3	4
Client opinion	2	4	4
Total	7	11	12

Source: own based on survey result

Systems technology criterion

SAP BO is based on a single server integrated with Windows NT network. It uses, based on Win 32 bilayer architecture client-server. Custom development (called Add-ons) are done using the SAP BO SDK (Software Development Kit). SDK is a"toolbox" that contains interfaces, sample code, documentation and development tools. It provides application programming interfaces (APIs) that allow developers to enhance SAP BO. The system uses the following databases: MS SQL Server 2008 or 2012, Sybase Adaptive Server Enterprise and IBM Universal Database Express Edition. Sage in turn is built on SAFE X3 (Sage Application Framework for Enterprise) platform. This platform provides users with best-in-class collaboration capabilities in either client/server or web mode, as well as an integrated business. Thanks to this solution, the system is available on the following Operation Systems (OS): Microsoft Windows, Linux Red Hat and Unix. Data can be stored and processed in both MS SQL Server or Oracle technologies. Comarch operates in a client-server mode. Thanks to its modern technology - Microsoft SQL Server - provides efficient, reliable work, data security and application integration with Microsoft Office. The system uses Microsoft SQL Server 2008 / 2008R2 / 2012 / databases.

Table 7. System technology evaluation

Investigated criteria	SAP BO	Sage	Comarch
Modularity	3	3	3
Openness	3	3	3
OS technology	4	4	2
Data Base	4	3	2
User opinion	3	2	2
Total	17	15	12

Source: own based on survey result

Systems overall rating

Considered system analysis is not intended to identify the best system but rather to pay attention to the main features that can be adopted during ERP systems selection. After summing main criteria total points and in order to calculate the final indicator system selection rating, only individual user preference factor is added in this case. Other measures such risk coefficient, budget overruns or project time exceeding factors and others may also be used.

Table 8. Total weighted evaluation criteria

System selection criteria	User reference	SAP BO	Sage	Comarch
Financial criterion	1.06	9	11	15
Functionality	1.21	10	13	14
System support	1.07	16	15	14
Implementation and training	1.06	16	12	12
Integrity and system scalability	1	6	11	12
System technology	0.91	17	15	12
Total		74	77	79
Indicator system rating		77.19	80.8	83.46

Source: own calculations

After determining the coefficient of respondents preferences basis on Table 1 and (6), and after summing all individual given points for basic and associated criteria, the analysis result shows that the described in this paper systems are almost on the same level. SAP BO scored the fewest points due to the weaker production management and system integrity. It makes however up for effective implementation, training and support system. Sage took second place. It differs from its predecessor a higher level of production management feature and greater number of built-in tools. Thanks to good production management support and, above all, a greater number of integrated tools that translate into large functionality, Comarch ERP scored the highest number of points. This result explains the higher license sales of the software on the Polish market.

SUMMARY

As follows from experts experience, the comparative method gives raise to various problems that are often difficult to resist. It assumes a good knowledge of examined systems, because only then they can be objectively compared. A casual systems knowledge will not lead to a reasonable evaluation and system selection. It seems that the presented in this paper method can avoid the subjective features of the comparative method as better functionality, greater flexibility, etc. and provides measurable indicators that can be used to evaluate and compare selected criteria for choosing an ERP system for SME's. The simplicity of the methods is that scores are assigned to each criterion according to user own discretion, which was included in the evaluation systems model. The presented model can be modified and used to evaluate other systems class in accordance with the needs of the company.

REFERENCES

Chmielarz W. (2003) Teoria i praktyka oceny informatycznych systemów finansowo-księgowych, Wydawnictwo Naukowe Wydziału Zarządzania UW.

Comarch (2015) Comarch ERP XL Light, downloaded April 30th, 2015 from http://www.comarch.pl/files-pl/file_64/Comarch-ERP-XL-Light-10.pdf

Computerworld (2010) Przegląd systemów do wspomagania zarządzania.

Dixit A. K, Nalebuff B. J. (2008) Sztuka Strategii, MT Biznes 2008.

Doradcy-IT (2014) Metody i kryteria wyboru systemu ERP, downloaded April 30th, 2015.

Hejduk I. K., Grudzewski, W. M. (2004) Metody Projektowania systemów zarządzania. Difin S.A.

Sage (2015) Sage vs Epicor, downloaded April 30th, 2015 from http://www.rklesolutions.com/sage-x3-vs-epicor/

SAP (2014) downloaded 30th March, 2015 from demo systemu ERP | Business One | SAP: sap.com/poland/solution/sme/ software/erp/small-business-management/demos.html

Visacom released (2014) and downloaded April 08th, 2015. sap.visacom.pl/index.php?option=com_content&view=article&id=223&Itemid=1,

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 73 – 79

THE ROLE OF INFORMATION SYSTEMS IN DEVELOPMENT OF VOIVODESHIPS IN POLAND

Maria Parlińska, Iryna Petrovska
Department of Agricultural Economics and International Economic Relations
Warsaw University of Life Sciences – SGGW
e-mail: maria_parlinska@sggw.pl, iryna_petrovska@sggw.pl

Abstract: Role of information and Information Systems becomes more and more important for company activity. Their influence on economy of Poland was not researched yet. Usage of Information and Communication technology rapidly increased for the past years. The research goal of the article was evaluation of dependency between development of ICT and economic situation of voivodeships in Poland.

Keywords: information, economics of information, Information Systems (IS)

ECONOMICS OF INFORMATION AND INFORMATION SYSTEMS

Economics of information is one of groundbreaking change in economy. Digital revolution became new economic revolution, sometimes called the third industrial revolution [Varian 2002].

Economics of information took first beginning from 1960 by Stigler[1] and and from 1962 by Machlup.[2] Stiglitz worked with the problem of imperfect information on the market. If information is not perfect, market equilibrium does not exist. According to Stiglitz it is possible to decrease the level of imperfect information by proper expenses [Stiglitz 2013, Żelazny 2014, Petrovska 2014]. These expenses are connected with possibility to get right-time, clear and truthful information.

1 Stigler G. (1961) The Economics of Information, The Journal of Political Economy, Vol. 69, pp. 213-225.
2 Marchlup F. (1962) The Production Distribution of Knowledge in the United States, Princeton University Press.

New institutional economy put transaction costs on the first place at the analysis, to which are also included costs of information [Daniłowska 2007].

One of the way to decrease the transaction costs by company can be introduction of proper Information System.

Based on nowadays development of Information Technologies, it is not possible to skip evaluation of their influence on economy.

Plenty of articles were already written about the benefits from the using of Information Systems. Some of these researches confirmed that Information Systems have positive influence on company development, some of them do not agree with it. For example Oliner and Sicher[3] said that computers have no influence on economic development. Some economists say that investments in IS are not always profitable. There are still continued discussions if companies which use IS gain comparative advantage [Carr 2003, Masli et al. 2011]. Lin in his work said that IS has no positive influence on business value [Lin 2009]. Based on results from works of Barua, Kriebel and Mukhopadhyay[4], it can be concluded that company, which invests in IS gains economic efficiency [Romero 2014].

Among the newest research works it should be mentioned the book of Schiller "Digital depression: information technologies and financial crisis", where author demonstrates digital technology's central role in the global political economy development processes. Especially author connected it to the financial, production and military networks [Schiller 2014].

Company introduces Information System in purpose to increase its competitiveness. It can be defined the following advantages such as increasing production flow, shortening time delivering, increasing sales and so on [Parlińska et al. 2013].

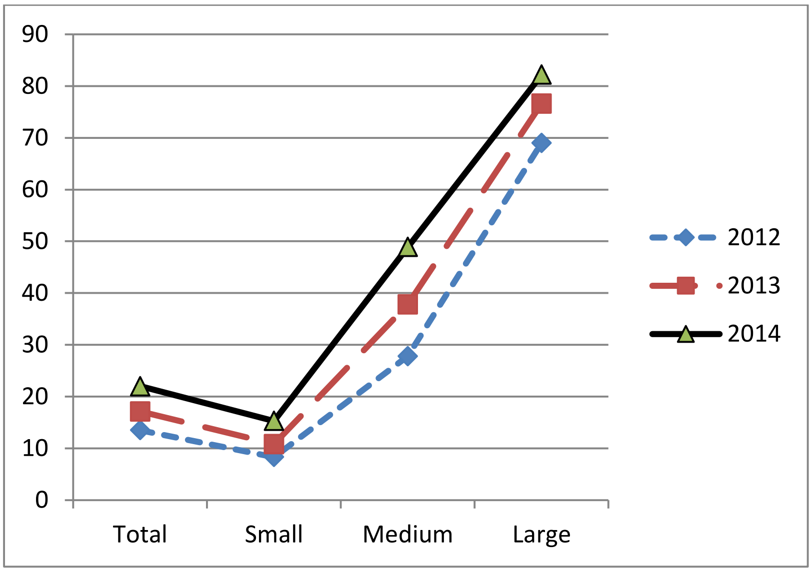
THE USE OF INFORMATION SYSTEMS BY ENTERPRISES IN POLAND

Information is the main element of Information System. In economic sciences the Information System is defined as set of actions to collect, store and disseminate information in purpose to succor decision making process [Borkowski 2003].

[3] Oliner S., Sichel D. (1994) Computers and output growth revisited: how big is the puzzle?, Brookings Paper on Economic Activity, Vol. 25, No. 2, pp. 273-334.

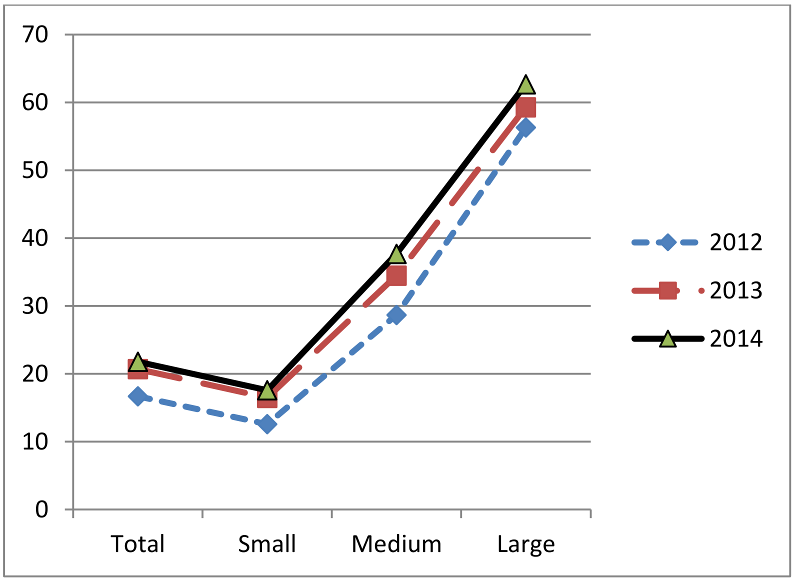
[4] Barua A., Kriebel C., Mukhopadhyay T. (1995) Information technology and business value: an analytic and empirical investigation, Information System Research, Vol. 6, No. 1, pp. 3-24.

Figure 1. Usage of ERP in Poland by company size [%]



Source: made by Author using data from statistical reports "Społeczeństwo informacyjne w Polsce"

Figure 2. Usage of CRM in Poland by company size [%]



Source: made by Author using data from statistical reports "Społeczeństwo informacyjne w Polsce"

In general in Poland there are used two systems which allow automatic share of information between different departments within the enterprise: Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM). ERP is used for planning and managing company activity such as accounting, production, marketing and so on. Customer Relationship Management is used for managing the information about clients.[5]

Table 1. Usage of ERP and CRM in Poland by types of economic activity

Specification	Enterprises using system %					
	ERP			CRM		
	2012	2013	2014	2012	2013	2014
Manufacturing	15.9	20.6	25.2	16.0	20.2	20.6
Electricity, gas, steam and air conditioning supply	31.7	35.8	48.4	28.2	24.2	31.4
Water supply; sewerage, waste management and remediation activities	14.5	21.2	30.5	23.3	24.1	29.1
Construction	6.2	6.7	10.8	8.0	9.1	9.6
Trade; repair of motor vehicles	17.0	20.0	24.1	21.6	26.5	26.4
Transportation and storage	9.2	12.5	18.0	14.8	17.4	18.7
Accommodation and catering	6.2	6.7	12.4	6.8	14.6	17.3
Information and communication	19.0	36.4	45.5	34.8	52.1	57.2
Financial and insurance activities	19.0	24.9	28.0	54.3	54.7	55.0
Real estate activities	9.5	12.5	23.2	17.8	16.9	21.4
Professional, scientific service activities	11.4	16.6	20.9	17.5	20.4	24.4
Administrative and support service	10.7	14.4	19.9	17.3	20.5	21.6
Repair of computer and communication equipment	27.7	35.5	33.8	40.0	48.4	46.8

Source: Statistical reports „Społeczeństwo informacyjne w Polsce"

Based on data from Statistical Office, which is shown in Figure 1 and 2, it can be noticed growing tendency of using ERP and CRM by enterprises in Poland from 2012 to 2014. The biggest share of using these systems belongs to large companies. In 2014 it reached more than 80 % for ERP and more than 60% for CRM.

According to type of economic activity, ERP and CRM are used the most in information, communication, insurance and finance activity (Table 1). Mostly in each type of activity it is observed notable growth of using these systems (especially if to compare year 2012 with 2014).

---

[5] Statistical reports (2012, 2013, 2014) "Społeczeństwo informacyjne w Polsce", Central Statistical Office of Poland, Warszawa.

RANKING OF VOIVODESHIPS ACCORDING TO ECONOMIC AND ICT DEVELOPMENT USING HELLWIG METHOD

In the empirical part of the article, it was made comparison of voivodeships' economic and ITC development and evaluated relation between them.

Based on background of new institutional economics, which says that costs of information belongs to transaction costs and can be reduced by proper expenses (in our research – expenses for ICT), as dependent variable it was chosen voivodeship's economic development of enterprises, as independent – development of ICT. Number of observations is equal to 16.

In purpose to make ranking of voivodeships' economic situation of enterprises there were chosen the following variables, connected with their economic activity:

- Level of unemployment in voivodeship,
- Number of enterprises in voivodeship,
- Average salary in voivodeship,
- Average enterprise's income in voivodeship.

In purpose to make ranking of ICT usage in voivodeships there were chosen the following variables, connected with using of information and telecommunication systems by enterprises:

- Using of ERP by enterprises,
- Using of CRM by enterprises,
- Enterprises receiving orders via computer network,
- Enterprises receiving/sending e-factures,
- Share of enterprises, which use computers,
- Share of enterprises with Internet access,
- Enterprises providing portable devices to the employers.

Hellwig methodology allows to make ranking of observations based on synthetic measure of object's distance from theoretical standard of development [Parlinska et al. 2014].

Using Hellwig method, there was made a ranking of voivodeships according to their economic and ICT development. All data for calculations was taken from statistical reports of year 2013.

Table 2. Hellwig measures

Voivodeship	Hellwig measure of ICT development	Hellwig measure of economic development
Mazowieckie	0.785695	0.935069
Dolnośląskie	0.64392	0.516134
Śląskie	0.605156	0.525213
Pomorskie	0.448798	0.506155
Podlaskie	0.434099	0.255387
Małopolskie	0.412305	0.487596
Opolskie	0.400786	0.281285
Łódzkie	0.390058	0.378291
Kujawsko-pomorskie	0.356221	0.289714
Zachodniopomorskie	0.347572	0.278421
Wielkopolskie	0.292489	0.556026
Podkarpackie	0.285461	0.273039
Lubuskie	0.266336	0.257155
Lubelskie	0.248223	0.282421
Warmińsko-mazurskie	0.133952	0.133591
Świętokrzyskie	0.012142	0.246839

Source: made by Author using data from statistical reports of Central Statistical Office of Poland

After making regression analysis, it was calculated that relation between economic development of the region and the use of ICT is significant (p<0,05, model is $y = 0{,}08 + 0{,}81 \times x$, R = 0,79, F = 22,87). Based on this results it can be made conclusion that together with the rise of using ICT, economic indicators of enterprises in the region also increase.

CONCLUSIONS

With rapid development of Information Systems in business it became not possible to make evaluation of economic situation of enterprises in the region, without taking them into account. In this research the background was taken from new institutional economics, that is why as the independent variable it was chosen ICT development in the region, which has influence on economic situation of enterprises in the region.

Using Hellwig method the measures of development were calculated. As a result of regression analysis, it was obtained significant dependence between dependent and independent variables, which means that using of information and communication technologies has positive influence on economic development of enterprises in the voivodeships of Poland.

REFERENCES

Barua A., Kriebel C., Mukhopadhyay T. (1995) Information technology and business value: an analytic and empirical investigation. Information System Research, Vol. 6, No. 1, pp. 3-24.

Borkowski B. (2003) Systemy informacyjne na potrzeby Wspólnej Polityki Rolnej, Wydawnictwo SGGW, Warszawa.

Carr N. (2003) IT doesn’t matter, Harvard Business Review, Vol. 81, No. 5, pp. 41-49.

Daniłowska A. (2007) Poziom, zróżnicowanie oraz uwarunkowania kosztów transakcyjnych kredytów i pożyczek rolniczych, Wydawnictwo SGGW, pp. 13-55, Warszawa.

Lin W. (2009) The business value of information technology as measure by technical efficiency: evidence from country-level data, Decision Support Systems, No. 46, pp. 865-874.

Marchlup F. (1962) The Production Distribution of Knowledge in the United States, Princeton University Press.

Masli A., Richardson V., Sanchez J., Smith R. (2011) The business value of IT: a synthesis and framework of archival research, Journal of Information Systems, Vol. 25, No. 2, pp. 81-116.

Oliner S., Sichel D. (1994) Computers and output growth revisited: how big is the puzzle?, Brookings Paper on Economic Activity, Vol.25, No. 2, pp. 273-334.

Parlińska M., Petrovska I., Pietrych L. (2014) Ocena zróżnicowania produkcji mlecznej w Ukrainie z zastosowaniem wielowymiarowych metod statystycznych, Quantitative Methods in Economics, Vol. 15, No. 4, pp. 229-235.

Parlińska M., Petrovska I. (2013) Role of information systems at the logistic enterprises, Quantitative Methods in Economics, Vol. 14, No 2, pp. 139-147.

Petrovska I. (2014) The use of AHP for decision making process in organization, Zeszyty Naukowe Politechniki Śląskiej. Organizacja i zarządzanie 2014, Issue 1921(74), pp. 229-238.

Romero J. (2014) Business Value of Information Technology: measuring performance and sources of profitability, Approaches and Processes for Managing the Economics of Information Systems, pp. 82-90.

Schiller D. (2014) Digital Depression: Information Technology and economic crisis, University of Illinois Press, Urbana, Chicago and Springfield.

Statistical reports (2012, 2013, 2014) Społeczeństwo informacyjne w Polsce, Central Statistical Office of Poland, Warszawa.

Stigler G. (1961) The Economics of Information, The Journal of Political Economy, Vol. 69, pp. 213-225.

Stiglitz J. (2013) Selected works of Joseph E. Stiglitz: Information and Economic Analysis, Vol. II, Oxford University Press, Oxford.

Varian Hal R. (2002) Mikroekonomia, Wydawnictwo Naukowe PWN, Warszawa.

Żelazny R. (2014) Ekonomika informacji – przedmiot i zakres badawczy, Studia Ekonomiczne/Uniwersytet Ekonomiczny w Katowicach, Kategorie i teorie ekonomiczne oraz polityka gospodarcza, No. 176, pp. 168-178.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 80 – 88

PRECISE ESTIMATES OF RUIN PROBABILITIES

Marcin Rudź
Institute of Mathematics, Technical University of Lodz
e-mail: marcin.rudz@p.lodz.pl, marcin_rudz@interia.pl

Abstract: In this paper we investigate a sequence of accurate approximations of ruin probabilities in discrete time models. We prove its convergence to the exact ruin probability without any restrictive assumptions on the claim distribution. Numerical studies show that the sequence, from the first term on, accurately approximates ruin probabilities. A formula for ruin probabilities in the finite horizon is also proposed.

Keywords: discrete time risk models, ruin probabilities, approximations, Solvency II

INTRODUCTION

Changes taking place on the worldwide financial market cause reforms of the supervision systems. The forthcoming EU directive Solvency II gives each insurance company a possibility to submit a motion (to the relevant supervision authority) for approving the internal model for determining the solvency capital requirement (SCR).

In this paper we investigate estimates of ruin probabilities in such a natural risk model. The discrete time setup, where the financial situation is reported on to the supervisor at the end of each fixed time period, has several advantages [Cheng et al. 2000], [Jasiulewicz 2013], [Gajek and Rudź 2013], [Rudź 2015] and others.

There are numerous papers investigating approximations of ruin probabilities in various risk models [De Vylder 1978], [Rolski et al. 1999], [Asmussen 2000], [Grandell 2000], [Čížek et al. 2005], [Dickson 2005], [Asmussen and Albrecher 2010] and others. Some of the approximations (for instance the De Vylder method) are based upon purely empirical grounds[1].

[1] See Asmussen S. [2000] Ruin Probabilities, World Scientific, Singapore, p. 80.

In this paper we investigate, along theoretical as well as numerical lines, some precise estimates of ruin probabilities in the following risk model.

A discrete time risk model

In this section we summarise, after [Bowers et al. 1997], [Klugman et al. 1998], [Rolski et al. 1999], [Gajek 2005], [Gajek and Rudź (2013)] and [Rudź 2015], the relevant material on a discrete time risk model.

All stochastic objects considered in the paper are assumed to be defined on a probability space $(\Omega, \Im, P)$. Let N denote the set of all positive integers and $\boldsymbol{R}$, the real line. Set $N^0 = N \cup \{0\}$, $R_+ = (0, \infty)$, $R_+^0 = [0, \infty)$, and $\overline{R_+} = (0, \infty]$.

Let a non-negative random variable X_i denote the aggregated sum of the claims in the i th time period; a positive real γ, the amount of aggregated premiums received each period; and a non-negative real u, the insurer's surplus at 0. We assume that $X_1, X_2, \ldots$ are i.i.d. random variables with a common distribution function F. Let $\{S(n)\}_{n\in N^0}$ denote the insurer's surplus process defined by

$$S(0) = u \tag{1}$$

and

$$S(n) = u + \gamma n - \sum_{i=1}^{n} X_i, \ n \in N. \tag{2}$$

The moment of ruin is defined by

$$\tau = \tau(u) = \inf\{n \in N: \ S(n) < 0\} \tag{3}$$

under the convention that $\inf \varnothing$ means ∞. The probability of ruin in a finite horizon n is defined by

$$\Psi_n(u) = P(\tau(u) \le n) \tag{4}$$

and the infinite-horizon probability of ruin by

$$\Psi(u) = P(\tau(u) < \infty). \tag{5}$$

Fix $\gamma \in \boldsymbol{R}_+$. Define $M: \ R \to \overline{R_+}$ by

$$M(r) = Ee^{-r(\gamma - X_1)}, \ r \in R. \tag{6}$$

The positive real solution r_0 of the following equation

$$M(r_0) = 1, \tag{7}$$

if it exists, is called adjustment coefficient. Denote $R_0(u) = e^{-r_0 u}$, $u \ge 0$ and $M_0 = \{r \in R: \ M(r) < \infty\}$.

The following result provides a sufficient condition for the existence of r_0.

Lemma 1. [Gajek 2005, p. 15]. Assume that $EX_1 < \gamma$, $F(\gamma) < 1$ and the set M_0 is open. Then there exists a unique adjustment coefficient $r_0 > 0$.

Under the assumptions of Lemma 1 the following equality holds [Bowers et al. 1997]

$$\Psi(u) = \frac{R_0(u)}{E[R_0(S(\tau)) | \tau < \infty]}. \tag{8}$$

Research objectives

There are some difficulties when working with the probability of ruin in the finite horizon. Let us quote at least one opinion of eminent researchers in this field [Rolski et al. 1999, p. 148]:

"Unfortunately, in many cases it is difficult to express the finite-horizon ruin probabilities in a closed form. The infinite-horizon ruin probabilities are mathematically simpler".

Thus, the research objectives are:

1. Finding a formula for the probability of ruin Ψ_n in the finite horizon.
2. Construction of a sequence of Ψ's approximations by means of Ψ_n and some other functionals.
3. Investigating the convergence of the sequence of Ψ's approximations.
4. Finding a close connection between the approximations and Formula (8).
5. Numerical studies which illustrate precision of the approximations.

MAIN RESULTS

From now on, we will use the convention that $\sum_{m=1}^{0} x_m = 0$. For simplicity of notation, we write $\int_a^\infty$ instead of $\int_{(a,\ \infty)}$ and $\int_0^b$ instead of $\int_{[0,\ b]}$, where $a,\ b \in R_+^0$.

In the present section we prove that

$$\Psi_n(u) = \sum_{k=1}^{n} \overbrace{\int \underset{A_k(u)}{\dots} \int}^{k} dF(x_k)\dots dF(x_1),\ n \in N,\ u \ge 0, \tag{9}$$

where the set $A_k(u),\ k \in \{1, \dots, n\}$ consists of ordered tuples $(x_1, \dots, x_k) \in R^k$ satisfying the following conditions:

- $x_k > u + k\gamma - \sum_{m=1}^{k-1} x_m$,
- $x_i \le u + i\gamma - \sum_{m=1}^{i-1} x_m$ for every $i \in \{1, \dots, k-1\}$.

Denote $\Gamma = R_+ \setminus (0, \gamma)$ and $d_0 = \sup\{d \in \Gamma :\ F(d) < 1\}$, if such a $d_0 \in \overline{R_+}$ exists. We will assume that $d_0 = \infty$, $\gamma > EX_1$ and the set M_0 is open.

Formula (9) will be used to define a sequence $\{\tilde{\Psi}_n\}_{n \in N}$ given by

$$\tilde{\Psi}_n(u) = \frac{\Psi_n(u)}{D_n(u)},\ n \in N, u \ge 0, \tag{10}$$

where

$$D_n(u) = \sum_{k=1}^{n} \overbrace{\int \underset{A_k(u)}{\dots} \int}^{k} e^{-r_0(k\gamma - \sum_{m=1}^{k} x_m)} dF(x_k)\dots dF(x_1). \tag{11}$$

Throughout the paper we will investigate the approximations of Ψ given by (10).

Theorem 1.

i) for all $n \in N$ and $u \ge 0$, it holds

$$\Psi_n(u) = \sum_{k=1}^{n} \overbrace{\int \underset{A_k(u)}{\dots} \int}^{k} dF(x_k)\dots dF(x_1). \tag{12}$$

Assume that $d_0 = \infty$, $\gamma > EX_1$ and the set M_0 is open. Then:

ii) for all $n \in N$ and $u \ge 0$

$$\tilde{\Psi}_n(u) = \frac{R_0(u)}{E[R_0(S(\tau)) | \tau \le n]}, \tag{13}$$

iii) the sequence $\{\tilde{\Psi}_n\}_{n \in N}$ converges pointwise, as $n \to \infty$, to the exact probability of ruin Ψ.

Proof: Observe that $\{\tau=1\}, \ldots, \{\tau=n\}$ are disjoint events whose union is $\{\tau \le n\}$. Moreover, under the condition $\tau=1$: $S(\tau)=S(1)=u+\gamma-X_1$ and

$$\tau=1 \Leftrightarrow S(1)<0 \Leftrightarrow X_1>u+\gamma. \tag{14}$$

In much the same way as above, under the condition $\tau=2$: $S(\tau)=S(2)=u+2\gamma-(X_1+X_2)$ and

$$\tau=2 \Leftrightarrow (S(2)<0,\ S(1)\ge 0) \Leftrightarrow (X_2>u+2\gamma-X_1,\ X_1\le u+\gamma). \tag{15}$$

In general, under the condition $\tau=n,\ n\in N: S(\tau)=S(n)=u+\gamma n-\sum_{m=1}^{n} X_m$ and

$$\begin{aligned}\tau=n &\Leftrightarrow (S(n)<0, \underset{1\le i<n}{\forall} S(i)\ge 0)\\ &\Leftrightarrow (X_n>u+n\gamma-\sum_{m=1}^{n-1}X_m, \underset{1\le i<n}{\forall} X_i\le u+i\gamma-\sum_{m=1}^{i-1}X_m).\end{aligned} \tag{16}$$

i) Thus,

$$\begin{aligned}\Psi_n(u)&=P(\tau\le n)=P(\tau=1)+P(\tau=2)+\ldots+P(\tau=n)\\ &=\int_{u+\gamma}^{\infty}dF(x_1)+\int_{0}^{u+\gamma}\int_{u+2\gamma-x_1}^{\infty}dF(x_2)dF(x_1)+\ldots\\ &+\int_{0}^{u+\gamma}\int_{0}^{u+2\gamma-x_1}\cdots\int_{0}^{u+(n-1)\gamma-\sum_{i=1}^{n-2}x_i}\int_{u+n\gamma-\sum_{i=1}^{n-1}x_i}^{\infty}dF(x_n)\ldots dF(x_1)\\ &=\sum_{k=1}^{n}\overbrace{\int\ldots\int}^{k}_{A_k(u)}dF(x_k)\ldots dF(x_1).\end{aligned} \tag{17}$$

ii) Furthermore, under the assumptions, $\int_{u+\gamma}^{\infty}dF(x)>0$ for every $u\ge 0$. Thus, by (17), $\Psi_n(u)>0,\ \ u\ge 0$. Following Lemma 1, there exists a unique adjustment coefficient $r_0>0$. Therefore, for all $n\in N$ and $u\ge 0$

$$E[R_0(S(\tau))|\tau\le n]=\frac{R_0(u)D_n(u)}{P(\tau\le n)}=\frac{R_0(u)D_n(u)}{\Psi_n(u)}, \tag{18}$$

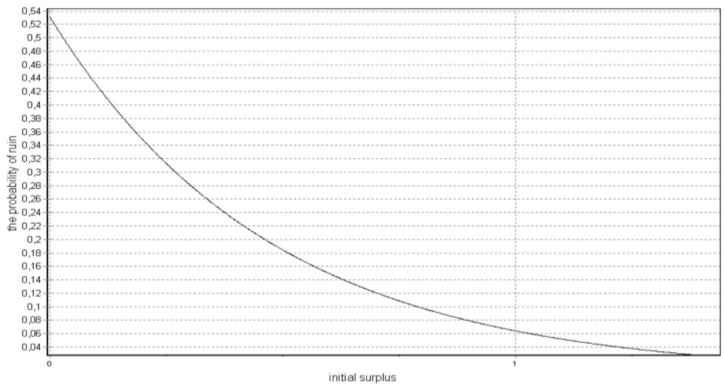
where the above-mentioned observations and some properties of the conditional expectation [Jakubowski and Sztencel 2001, p. 124] were used. Thus, ii) and iii) hold as well. ■

NUMERICAL EXAMPLES

In this section we present the results of numerical studies carried out by the author. We compare the approximations discussed in the previous section with the exact ruin probabilities obtained by means of the method described in [Rudź 2015]. The bisection method was also used.

We consider: exponentially distributed claims with the expected value $EX_1 \approx 0.22$ (Figure 1), a gamma claim distribution with the shape parameter 2 and the scale one 5.5 (Figure 2), a mixture of two gamma distributions with the shape parameters $(2,\ 2)$, the scale ones $(3,\ 7.5)$ and weights $(0.85,\ 0.15)$ respectively (Figure 3), a mixture of five exponential distributions with the scale parameters $(1,\ 2,\ 3,\ 7,\ 13)$ and weights $(0.3,\ 0.2,\ 0.3,\ 0.1,\ 0.1)$ respectively (Figure 4) and a mixture of five exponential distributions with the scale parameters $(1,\ 3,\ 5,\ 8,\ 10)$ and weights $(0.1,\ 0.3,\ 0.2,\ 0.2,\ 0.2)$ respectively (Figure 5).

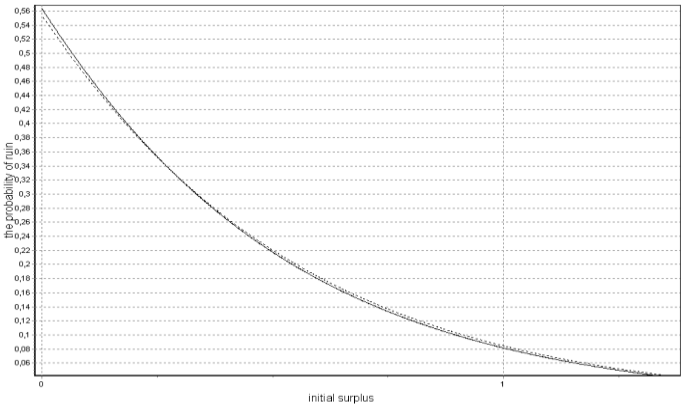
Figure 1. Exponential distribution with the scale parameter 4.5 (i.e. the expected value $EX_1 \approx 0.22$). The amount of aggregated premiums $\gamma = 0.3.$ The $\tilde{\Psi}_1$ approximation coincides with the exact probability of ruin Ψ (the solid line)



Source: own computations

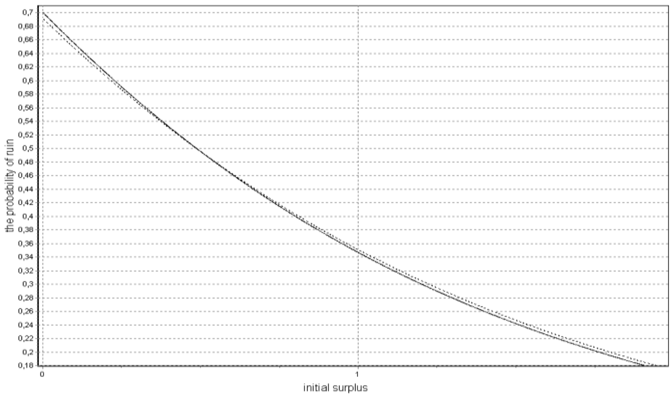
The approximation $\tilde{\Psi}_1$ might be faultless (Figure 1) or almost faultless (Figures 2 and 3). We have confirmed the above observation by means of a number of simulation experiments.

Figure 2. Gamma distribution with the shape parameter 2 and the scale one 5.5 (i.e. the expected value $EX_1 \approx 0.36$). The amount of aggregated premiums $\gamma = 0.45$. The $\tilde{\Psi}_1$'s graph (the dotted line) almost coincides with the exact ruin probability's one (the solid line)



Source: own computations

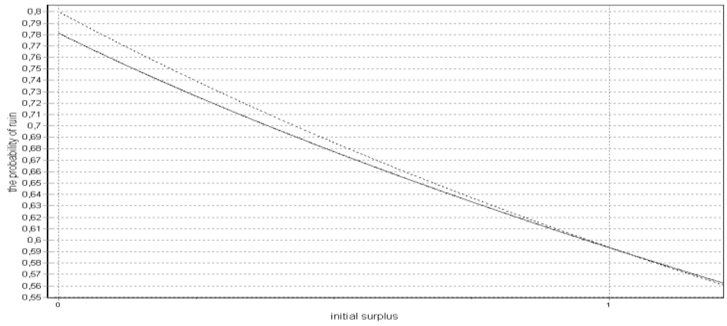
Figure 3. A mixture of two gamma distributions with the shape parameters $(2,\ 2)$, the scale ones $(3,\ 7.5)$, weights $(0.85,\ 0.15)$ respectively and the amount of aggregated premiums $\gamma = 0.7$. The $\tilde{\Psi}_1$'s graph (the dotted line) almost coincides with the exact ruin probability's one (the solid line)



Source: own computations

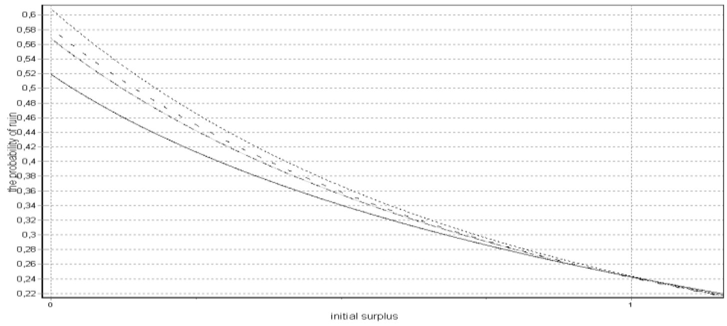
One distinguishes the graph of $\tilde{\Psi}_1$ from the Ψ's one in Figures 4 and 5. The latter one illustrates the approximations $\tilde{\Psi}_1$, $\tilde{\Psi}_2$ and $\tilde{\Psi}_3$.

Figure 4. A mixture of five exponential distributions with the scale parameters $(1,\ 2,\ 3,\ 7,\ 13)$, weights $(0.3,\ 0.2,\ 0.3,\ 0.1,\ 0.1)$ respectively and the amount of aggregated premiums $\gamma = 0.6$. The $\tilde{\Psi}_1$ approximation (the dotted line) is presented in relation to Ψ (the solid line)



Source: own computations

Figure 5. A mixture of five exponential distributions with the scale parameters $(1,\ 3,\ 5,\ 8,\ 10)$, weights $(0.1,\ 0.3,\ 0.2,\ 0.2,\ 0.2)$ respectively and the amount of aggregated premiums $\gamma = 0.41$. The approximations: $\tilde{\Psi}_1$ (the dotted line), $\tilde{\Psi}_2$ (the dashed line) and $\tilde{\Psi}_3$ (the dashed-dotted line) are presented in relation to Ψ (the solid line)



Source: own computations

CONCLUSION

In this paper we investigated Sequence (10) of Ψ's approximations in discrete time models. All research objectives have been achieved.

Theorem 1 i) gives a formula for the probability of ruin in the finite horizon. Theorem 1 ii) and iii) shows a close connection between Sequence (10) and Formula (8). In particular, it is shown that Sequence (10) converges pointwise to the exact probability of ruin without any restrictive assumptions on the claim distribution. The results of a numerical study lead to the conclusion that Sequence (10), from the first term on, accurately approximates ruin probabilities in the considered cases of claim distributions.

ACKNOWLEDGEMENTS

The author would like to thank Professor Lesław Gajek for helpful discussions.

REFERENCES

Asmussen S. (2000) Ruin Probabilities, World Scientific, Singapore, reprinted 2001.

Asmussen S., Albrecher H. (2010) Ruin probabilities, 2nd ed., World Scientific, Singapore.

Bowers N. L., Gerber H. U., Hickmann J. C., Jones D. A., Nesbitt C. J. (1997) Actuarial Mathematics, 2nd ed., The Society of Actuaries, Schaumburg, IL.

Cheng S., Gerber H. U., Shiu E. S. W. (2000) Discounted probabilities and ruin theory in the compound binomial model, Insurance: Mathematics and Economics, Vol. 26, pp. 239 – 250.

Čížek P., Härdle W., Weron R. (2005) Statistical tools for finance and insurance, Springer-Verlag, Berlin Heidelberg.

De Vylder F. E. (1978) A practical solution to the problem of ultimate ruin probability, Scandinavian Actuarial Journal, pp. 114 – 119.

Dickson D. C. M. (2005) Insurance Risk and Ruin, Cambridge University Press, Cambridge.

Gajek L. (2005) On the deficit distribution when ruin occurs–discrete time model, Insurance: Mathematics and Economics, Vol. 36, pp. 13 – 24.

Gajek L., Rudź M. (2013) Sharp approximations of ruin probabilities in the discrete time models, Scandinavian Actuarial Journal, pp. 352 – 382.

Grandell J. (2000) Simple approximations of ruin probabilities, Insurance: Mathematics and Economics, Vol. 26, pp. 157 – 173.

Jakubowski J., Sztencel R. (2001) Wstęp do teorii prawdopodobieństwa [An introduction to probability theory], 2nd ed. (additional printing), SCRIPT, Warsaw.

Jasiulewicz H. (2013) Dyskretny proces ryzyka z uwzględnieniem reasekuracji i losowej stopy procentowej [Discrete risk process with reinsurance and random interest rate], Roczniki Kolegium Analiz Ekonomicznych [Annals of Collegium of Economic Analyses], Warsaw School of Economics, Warsaw, Vol. 31, pp. 11 – 26.

Klugman S. A., Panjer H. H., Willmot G. E. (1998) Loss models. From data to decisions, Wiley, New York.

Rolski T., Schmidli H., Schmidt V., Teugels J. (1999) Stochastic processes for insurance and finance, Wiley, New York.

Rudź M. (2015) A method of calculating exact ruin probabilities in discrete time models, Roczniki Kolegium Analiz Ekonomicznych [Annals of Collegium of Economic Analyses], Warsaw School of Economics, Warsaw, Vol. 37, pp. 307 – 322.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 89 – 101

CROSS-SECTIONAL RETURNS FROM DIVERSE PORTFOLIO OF EQUITY INDICES WITH RISK PREMIA EMBEDDED[1]

Paweł Sakowski
Faculty of Economic Sciences, Warsaw University
e-mail: sakowski@wne.uw.edu.pl
Robert Ślepaczuk
Faculty of Economic Sciences, Warsaw University
and Union Investment TFI S.A.
e-mail: rslepaczuk@wne.uw.edu.pl (corresponding Author)
Mateusz Wywiał
Faculty of Economic Sciences, Warsaw University
and Quedex Derivatives Exchange
e-mail: wywialm@quedex.net

Abstract: The main purpose of this article is to extend evaluation of classic Fama-French and Carhart model for global equity indices. We intend to check the robustness of models results when used for a wide set of equity indices instead of single stocks for the given country. Such modification enables us to estimate equity risk premium for a single country. However, it requires several amendments to the proposed methodology for single stocks.

Our empirical evidence reveals important differences between the conventional models estimated on single stocks, either international or US-only, and models incorporating whole markets. Our novel approach shows that the divergence between indices of the developed countries and those of emerging markets is still persistent. Additionally, research on weekly data for equity indices presents rationale for explanation of equity risk premia differences between variously sorted portfolios.

Keywords: cross-sectional models, asset pricing models, equity risk premium, equity indices, new risk factors, sensitivity analysis, book to market, momentum, market price of risk, emerging and developed equity indices,

[1] The support of NCN grant number 2014/13/B/HS4/03209 is gratefully acknowledged. The views presented in this text are those of the authors and do not necessarily represent those of Union Investment TFI S.A.

INTRODUCTION

The phenomenon of equity risk premium and stock market returns fluctuations is thoroughly described in financial literature. The discussion started with the seminal papers introducing CAPM of Sharpe [1964], Lintner [1965] and Black [1972]. Then, it has greatly evolved with the three-factor model of Fama [1992] and four-factor model of Carhart [1997]. Nowadays, it concentrates around many other modifications which propose various set of risk factors in order to fully explain the variability of stock market returns. This paper aims to introduce several new ideas to this debate.

At the beginning we would like to stress the most popular effects revealed in financial literature, which were indicated as the most important risk factors explaining outperformance of the given groups of stocks:

- the robustness of outperformance of the value investing strategy (i.e. investing in stocks that have high book to market, dividends yield, earnings ratio, etc.) produce higher risk adjusted returns [Fama 1992], [Lakonishok 1994], [Arshanapalli et al. 1998], [Bondt and Thaler 1985] and [Bondt and Thaler 1987],
- size effects (i.e. small minus big stocks effect) [Fama and French 2012],
- momentum and reversal effect (i.e. winners minus losers effect) captured for many different time frames [Wu 2002], [Jegadeesh and Titman 1993], and [Asness 1995],
- liquidity effect [Rahim and Noor 2006], [Liu 2004],
- market factor, investment factor and return on equity factor [Chen et al. 2011],
- five factors, profitability and investment on the top of standard three-factor model [Fama and French 2015],
- betting against beta, i.e. long leveraged low-beta assets and short high-beta assets produce significant positive risk-adjusted returns [Frazzini and Pedersen 2014],
- accounting manipulation factor performs better for New EU Member states [Foye et al. 2013],
- cash-flow-to-price factor, momentum and market factor analyzed for 49 countries [Hou et al. 2011].

At the same time many authors claimed that CAPM still works, arguing that deviations due to missing factors are difficult to detect and it is relatively difficult to reject data-snooping bias in case of multifactor models [MacKinlay 1995]. Other kinds of biases which can be encountered while performing stock returns analysis include among others look ahead bias [Lo and MacKinlay 1990].

Based on the current state of the art for stock returns and the fact that very few papers covered the problem for equity indices returns so far, we want to better explain the diversity of equity indices returns and hence follow the conclusion of Griffin [2002] who stated that Fama-French factors are country-specific rather than global.

Therefore, the main aim of this paper was to present a cross sectional analysis for global indices with special attention to equity risk premium. We want to find an answer to a question whether based on combination of well-known asset pricing models we are able to pick these equity indices which are relatively cheap (or expensive), at the same taking into account all other important risk factors.

Our main research questions are as follows:

1. Can models of [Fama and French 1992] and [Carhart 1997] be used for explanation of equity risk premium for global indices? Our intention is to answer this question on single equity indices basis and on aggregated level as well.
2. Can we say that beta coefficients are rather similar or do they differ among countries?
3. Are signs for beta coefficients coherent with the results for single stocks?
4. Can we say that the model of Carhart fully explain the variability of equity risk premium for worldwide indices?
5. Is it possible to distinguish countries with consecutively high beta sensitivities?
6. Which risk factor was the most important in portfolio construction?

Above mentioned questions helped us to plan the methodology section of this research.

METHODOLOGY AND DATA

Methodology

The methodology is based on the seminal paper of [Carhart 1997], who proposed the four-factor model for mutual funds analysis. One of the reason that we prefer the model of Carhart over the methodology of [Fama and French 1992] (the three-factor model for stocks return analysis) are the results of [Fama and French 2012] and comprehensive results obtained for emerging markets by [Cakici et al. 2013]. They focused on 18 emerging markets treating each of them separately and they revealed the significance of value and momentum everywhere except Eastern Europe and additionally noted that value factors and momentum factors were negatively correlated.

Taking into account that our research is intended for equity indices we have to introduce several amendments to the initial methodology. Necessary modifications include:

- converting monthly to weekly data in order to reveal dynamics during shorter time intervals,
- introducing lags to the original models in order to use them for forecast purposes,

- including new risk factors that explain the diversity of returns more deeply,
- necessary conversion of well-known risk factors from the country level to the worldwide level,
- creating adequate zero investment portfolios that fully reflect the influence of particular risk factor on equity risk premium.

Before we present our model it is crucial to define the equity risk premium as the expected excess return of equities over the risk free rate. The point here is that current literature proposes many alternative ways to measure it, depending on what we want to focus on:

- historical returns approach:

$$ERP = \sum_{t=t_0}^{n} R_t - Rf_t \tag{1}$$

where R_t -Rf_t is excess return at time t over risk-free rate.

- earnings yield approach:

$$ERP = \frac{E}{P} - Rf_t \tag{2}$$

where $\frac{E}{P}$ is earnings to price ratio

- dividend yield approach:

$$ERP = \frac{D}{P} + g - Rf_t \tag{3}$$

where $\frac{D}{P}$ is dividend to price ratio and g is dividend growth rate.

- regression- and factor-based approach which can be characterized by point-in-time estimates instead of long-term estimates only, not dependent on e.g. tax policy, and which allows dynamic forecasts:

$$ERP = \alpha + \sum_{i=1}^{n} \beta_i X_{i,t} + \varepsilon_t \tag{4}$$

where $X_{i,t}$ is the *i*-th risk factor at moment t and β_i is sensitivity to this factor.

- survey-based approach which is often systematically biased, negatively correlated with future returns, and positively with previous returns.

The selection of particular definition can certainly affect final results but before we focus on that we describe factor models used in this research. We start with the Fama-French three-factor model:

$$(R_i - Rf_t) = \alpha + \beta_{MKT,i} * (R_m - Rf_t) + \beta_{HML,i} * HML_t + \beta_{SMB,i} * SMB_t \tag{5}$$

where $(R_i$-$Rf_t)$ is weekly return of equity index in excess to weekly risk free rate, $(R_m$-$Rf_t)$ is equally weighted equity index minus risk free rate, HML_t represents the monthly premium of the book-to-market factor, and SMB_t is the monthly premium of the size factor. We assumed 3m Libor USD as risk free rate measure in our research.

Then we concentrate on the four-factor model of Carhart, which additionally introduces the WML factor:

$$(R_i - Rf_t) = \alpha + \beta_{MKT,i} * (R_m - Rf_t) + \beta_{HML,i} * HML_t + \beta_{SMB,i} * SMB_t + \beta_{WML,i} * WML_t \quad (6)$$

The WML factor is the monthly premium on winners minus losers (WML) and can be calculated by subtracting the equal weighted average of the highest performing firms from the equal weighed average of the lowest performing firms [Carhart, 1997]. The detailed procedure of calculating HML, SMB and WML risk factors is summarized below.

The HML is a zero-investment portfolio that is long on the highest decile group of book-to-market (B/M) equity indices and short on the lowest decile group. The difference of returns of these extreme decile groups is calculated in each weekly interval, which finally constitutes HML factor. Based on these returns we created cumulative returns for HML and then LMH zero investment portfolio (where LMH was created as the difference between lowest and highest decile group of book-to-market).

The SMB is a zero-investment portfolio that is long on the highest decile group of small capitalization (cap) equity indices and short on the lowest decile group. The difference of returns of these extreme decile groups is calculated in weekly interval as well. Similarly, based on these returns we created cumulative returns for SMB and then BMS zero investment portfolio (where BMS was created as the difference between lowest and highest decile group of small capitalization (cap) equity indices).

Lastly, the WML is a zero-investment portfolio that is long on the highest decile group of previous 2-month return winner equity indices and short on its lowest decile group (loser equity indices). The difference of returns of these extreme decile groups is calculated again for each weekly interval and based on that we create cumulative returns for WML and then LMH zero investment portfolio (where LMW was created as the difference between lowest and highest decile group of previous 2-month return winner equity indices).

Data and descriptive statistics

We used the data for the most comprehensive set of investable equity indices[2] covering the period between 1990 and 2015[3]. The detailed list of all equity

[2] For practical purposes we used only these indices which can be easily invested through options, futures or ETFs.

[3] For practical purposes the study was limited to 2000-2015 because of unavailability of longer time series for several of our risk factors.

indices and their descriptive statistics can be obtained upon request. Descriptive statistics for risk factor used in the study are presented in Table 1.

Table 1. Descriptive statistics for risk factors: R_m-R_f, HML, SMB and WML

	R_m	R_f	R_m-R_f	HML	HML top	HML bottom	LMH	SMB	SMB top	SMB bottom	BMS	WML 12m	WML top	WML bottom	LMW
nobs	795	795	795	795	795	795	795	795	795	795	795	795	795	795	795
NAs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Minimum	0.000	-0.17	-0.171	-0.072	-0.174	-0.182	-0.143	-0.129	-0.151	-0.140	-0.079	-0.115	-0.106	-0.185	-0.106
Maximum	0.001	0.071	0.071	0.143	0.113	0.099	0.072	0.079	0.118	0.096	0.129	0.106	0.097	0.097	0.115
1.Q	0.000	-0.01	-0.006	-0.011	-0.009	-0.013	-0.014	-0.012	-0.009	-0.006	-0.011	-0.011	-0.005	-0.011	-0.019
3. Q	0.000	0.011	0.011	0.014	0.016	0.014	0.011	0.011	0.014	0.011	0.012	0.019	0.015	0.014	0.011
Mean	0.000	0.001	0.001	0.001	0.002	0.007	-0.001	-0.000	0.001	0.001	0.000	0.003	0.004	0.000	-0.003
Median	0.000	0.004	0.003	0.000	0.004	0.006	-0.000	-0.001	0.002	0.002	0.001	0.004	0.005	0.001	-0.004
Sum	0.336	1.426	1.089	1.455	1.994	0.535	-1.455	-0.198	1.306	1.108	0.198	2.942	3.631	0.688	-2.942
SEMean	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.001
LCLMean	0.000	0.000	0.000	0.000	0.000	-0.009	-0.003	-0.001	0.000	0.000	-0.001	0.001	0.003	-0.001	-0.005
UCLMean	0.000	0.003	0.002	0.003	0.004	0.003	-0.000	0.001	0.003	0.002	0.001	0.005	0.006	0.002	-0.001
Variance	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Stdev	0.000	0.017	0.017	0.022	0.026	0.021	0.022	0.019	0.021	0.019	0.019	0.026	0.020	0.026	0.026
Skewness	0.726	-1.82	-1.843	0.551	-0.813	-1.108	-0.551	-0.260	-0.523	-0.887	0.260	-0.279	-0.432	-0.510	0.279
Kurtosis	-0.90	13.61	13.68	3.100	5.294	6.079	3.100	3.139	5.539	6.242	3.139	1.336	3.050	4.460	1.336
IR	7.845	0.771	0.582	0.619	0.742	0.219	-0.564	-0.092	0.568	0.537	0.093	1.092	1.826	0.244	-0.905
cum_ret	0.390	3.047	1.911	3.162	6.060	0.690	-0.760	-0.177	2.601	1.964	0.215	16.840	33.985	0.964	-0.944

The data cover the period between 1990-2015. Detailed time frames for every risk factor are summarized in the table. HML, SMB and WML represent differences in returns between extreme decile groups, top and bottom represent returns of extreme decile groups, LMH, BMS and LMW represent anti-factors.

Source: own calculations

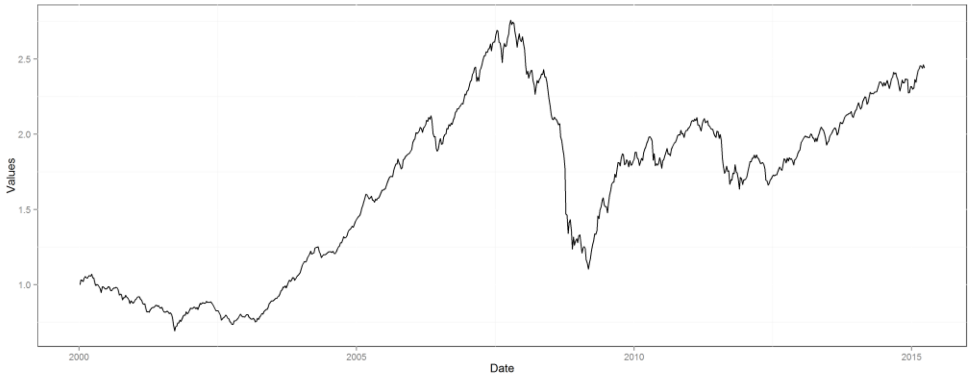
Analysis of risk factors

Detailed analysis of dynamics of the standard fours factors from the Carhart model helped us to define the final specification of the model. Our observation concerning these factors dynamics are shortly summarized below.

Figure 1 shows the dynamics of the first factor $(R_m\text{-}R_f)_t$. It does not significantly differ from equally weighted index path. This actually informs us that we analyzed the period of exceptionally low rates which have not a crucial impact on the value of this factor.

Figure 2 presents fluctuations of the second factor (HML_t). It reveals two distinctive periods. The first one (between 2000 and the beginning of 2012) shows a strong HML effect revealing much better behavior of equity indices with high book-to-market characteristics (Figure 2). Similar phenomenon was heavily presented in the literature for stock returns. Starting from 2012, the HML effect disappeared and totally transformed into the LMH effect what is quite surprising and requires some additional research.

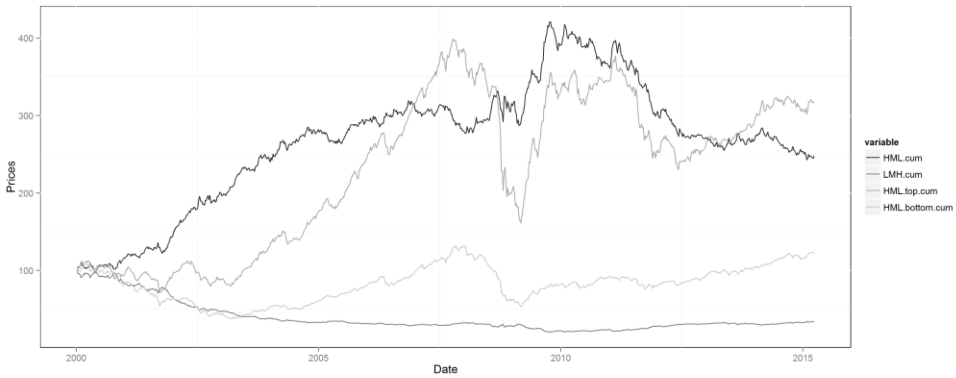
Figure 1. Dynamics of cumulative R_m-R_f factor



R_m-R_f factor was calculated on data between 1990-2015 and presents cumulative returns for R_m-R_f factor calculated on weekly data

Source: own calculations

Figure 2. Cumulative returns of HML factor with top/bottom 20% percentiles

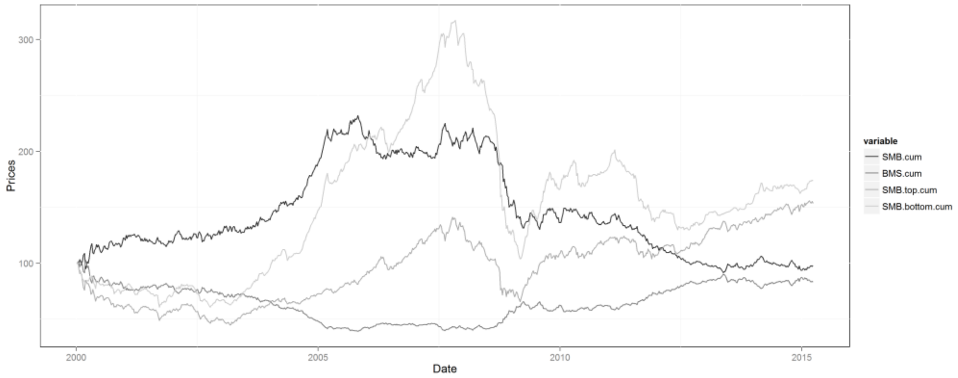


HML factor was calculated on data between 2000-2015 and presents cumulative returns for HML factor calculated on weekly data

Source: own calculations

Fluctuations of third risk factor (SMB_t) are presented in the Figure 3. It can be divided into two differing periods as well. The first period ends in 2006 and is characterized by outperformance of small capitalization equity indices. In the second period (between 2006 and 2015) this effect is totally reversed and we can see high outperformance of big capitalization equity indices.

Figure 3. Cumulative returns of SMB factor with top/bottom 20% percentiles

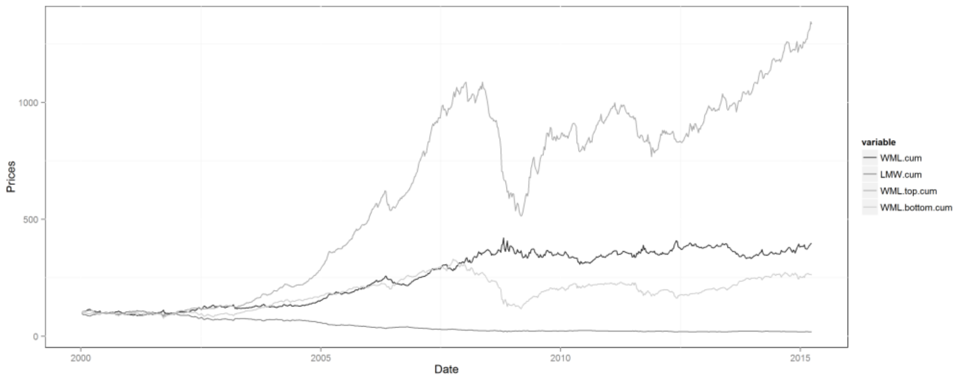


SMB factor was calculated on data between 2000-2015 and presents cumulative returns for SMB factor calculated on weekly data

Source: own calculations

Finally, the fourth risk factor (WML_t) reveals the strongest WML effect (Figure 4) which is stable during the whole period and exactly confirms the short-term momentum effect observed in financial literature.

Figure 4. Cumulative returns of WML factor with top/bottom 20% percentiles



WML factor was calculated on data between 2000-2015 and presents cumulative returns for WML factor calculated on weekly data for the last two months

Source: own calculations

RESULTS

Detailed results of regression for the Carhart model are presented in Table 2.

Table 2. Regression results for the Carhart model

index	country	ticker	n	a	t.a.	Rm-Rf	t.Rm-Rf	HML	t.HML.	SMB	t.SMB.	WML	t.WML.	Rsq	F
Developed markets															
Euro STOXX 50	Europe	SX5E	795	-0.002	-2.808	1.466	41.480	-0.204	-7.617	-0.054	-1.741	-0.110	-5.148	0.727	527.10
CAC 40	France	CAC	795	-0.002	-2.652	1.426	41.001	-0.195	-7.407	-0.076	-2.472	-0.093	-4.459	0.723	514.69
AEX-Index	Netherlands	AEX	795	-0.002	-2.808	1.473	40.674	-0.227	-8.307	-0.018	-0.563	-0.093	-4.249	0.715	494.57
DAX	Germany	DAX	795	-0.001	-0.939	1.495	37.928	-0.206	-6.925	-0.105	-3.020	-0.109	-4.606	0.694	448.82
FTSE 100	UK	UKX	795	-0.001	-2.255	1.122	37.665	-0.165	-7.324	-0.070	-2.657	-0.098	-5.472	0.694	447.84
FTSE MIB	Italy	FTSEMIB	795	-0.002	-3.317	1.524	36.100	-0.130	-4.067	0.001	0.036	-0.060	-2.348	0.656	376.68
OMX Stockholm 30	Sweden	OMX	795	-0.001	-1.245	1.359	33.616	-0.189	-6.182	-0.102	-2.840	-0.083	-3.411	0.639	349.14
OMX Helsinki 25 Index	Finland	HEX25	795	-0.001	-1.894	1.508	33.903	-0.117	-3.484	-0.090	-2.300	-0.073	-2.709	0.637	346.76
Vienna Stock Exch	Austria	ATX	795	-0.001	-1.154	1.423	33.371	0.058	1.809	-0.019	-0.508	-0.082	-3.191	0.633	341.17
S&P500	USA	SPX	795	-0.001	-1.224	1.065	31.884	-0.142	-5.638	-0.144	-4.871	-0.089	-4.418	0.629	335.10
BEL 20	Belgium	BEL20	795	-0.001	-1.552	1.255	32.850	-0.194	-6.703	0.057	1.672	-0.092	-3.987	0.616	317.48
S&P/ASX 200	Australia	AS51	795	-0.001	-1.202	0.927	32.605	-0.032	-1.468	-0.038	-1.519	-0.016	-0.962	0.611	309.98
Oslo SE OBX Index	Norway	OBX	795	0.000	0.064	1.432	31.935	-0.040	-1.168	-0.054	-1.374	-0.056	-2.068	0.606	303.26
IBEX 35 Index	Spain	IBEX	795	-0.001	-1.986	1.372	31.487	-0.068	-2.063	-0.043	-1.113	-0.095	-3.624	0.605	301.91
Luxembourg SE	Luxembourg	LUXXX	795	-0.002	-2.178	1.355	32.646	-0.002	-0.066	0.063	1.728	-0.002	-0.087	0.600	296.82
OMX Kopenhagen 20	Denmark	KFX	795	0.000	0.618	1.227	31.417	-0.112	-3.781	0.016	0.460	-0.090	-3.834	0.598	294.00
S&P/TSX 60	Canada	SPTSX	795	-0.001	-1.074	1.048	30.106	-0.085	-3.220	-0.080	-2.598	-0.028	-1.331	0.578	270.55
Swiss Market Index	Switzerland	SMI	795	-0.001	-0.817	1.076	28.347	-0.228	-7.960	0.021	0.617	-0.140	-6.127	0.567	258.67
Hang Seng Index	Hong Kong	HSI	795	-0.001	-1.401	1.291	28.285	-0.071	-2.070	-0.143	-3.533	-0.026	-0.954	0.552	243.46
Irish SE 20 Price Index	Ireland	ISEQ20P	524	-0.001	-1.112	1.306	21.173	0.083	1.478	-0.179	-2.664	-0.056	-1.379	0.549	157.64
Prague SE Index	Czech Rep	CTXEUR	795	-0.001	-0.803	1.444	27.652	0.156	3.952	0.023	0.490	-0.083	-2.632	0.548	239.89
MSCI Singapore Index	Singapore	MXSG	795	-0.001	-2.058	1.141	28.154	-0.030	-0.967	-0.075	-2.087	-0.006	-0.234	0.540	232.25
PSI20 Index	Portugal	PSI20	795	-0.002	-3.355	1.080	26.739	0.009	0.298	-0.018	-0.502	-0.063	-2.586	0.523	216.73
Nikkei 225 Index	Japan	NKY	795	-0.001	-1.810	1.143	23.789	-0.038	-1.054	-0.167	-3.933	-0.004	-0.127	0.469	174.74
NZE 50 Gross Index	Zealand	NZSE50FG	742	0.001	1.261	0.523	17.762	-0.003	-0.136	0.010	0.335	-0.012	-0.661	0.329	90.34
ICEX Main Index	Iceland	ICEXI	795	-0.001	-0.594	0.451	6.035	0.114	2.021	0.082	1.244	-0.025	-0.557	0.060	12.54
Malta SE Index	Malta	MALTEX	795	0.000	-0.694	0.182	4.370	0.031	0.977	0.038	1.038	0.049	1.937	0.027	5.58
Emerging markets															
MEX BOLSA IPC	Mexico	MEXBOL	795	0.001	1.003	1.179	25.069	0.028	0.796	-0.255	-6.139	-0.068	-2.407	0.525	106.66
M1HU Iran	Iran	M1HU	742	-0.001	-0.965	1.893	24.477	0.265	4.388	-0.165	-2.182	-0.130	-2.727	0.521	16.44
Budapest SE	Hungary	BUX	795	0.000	-0.549	1.339	26.069	0.063	1.632	-0.006	-0.128	-0.111	-3.590	0.520	190.81
KOSPI 200 Index	South Korea	KOSPI2	795	-0.001	-1.102	1.184	21.807	0.231	5.619	-0.534	-11.131	-0.035	-1.083	0.519	40.44
FTSE JSE Namibia	Namibia	FTN098	588	0.000	0.082	1.098	20.162	0.097	2.018	-0.341	-6.033	-0.001	-0.037	0.511	135.66
Bovespa	Brazil	IBOV	795	-0.001	-0.632	1.389	23.098	0.229	5.046	-0.307	-5.769	-0.021	-0.572	0.491	23.19
FTSE/JSE Top40	South Africa	TOP40	795	0.001	1.073	1.106	24.507	-0.028	-0.826	-0.154	-3.857	-0.035	-1.298	0.490	69.82
Russian Trad System	Russia	RTSI	795	0.000	-0.172	1.699	20.993	0.498	8.145	-0.291	-4.059	-0.070	-1.436	0.466	76.58
WIG20 Index	Poland	WIG20	795	-0.001	-1.181	1.191	22.441	0.071	1.774	-0.127	-2.695	-0.097	-3.043	0.460	38.21
S&P CNX Nifty	India	NIFTY	795	0.001	0.832	1.184	21.860	-0.123	-3.004	-0.098	-2.050	0.045	1.377	0.408	118.96
Ipsa	Chile	IPSA	795	0.000	0.173	0.783	19.133	0.046	1.498	-0.285	-7.869	0.036	1.462	0.407	3.00
Athens SE General	Greece	FTASE	795	-0.005	-3.468	1.498	18.806	0.196	3.262	-0.079	-1.120	-0.129	-2.680	0.377	119.60
OMX sTallinn Index	Estonia	TALSE	795	0.001	1.166	1.017	20.418	0.005	0.135	0.460	10.441	0.065	2.174	0.376	213.92
MSCI Taiwan Index	Taiwan	TAMSCI	795	-0.002	-1.650	1.036	18.124	0.069	1.589	-0.228	-4.514	-0.011	-0.318	0.360	136.39
Cyprus SE General	Cyprus	CYSMMAPA	551	-0.006	-2.601	1.536	12.345	0.971	8.590	0.172	1.302	0.068	0.842	0.359	106.83
Jakarta LQ45 Index	Indonesia	LQ45	795	0.001	0.880	1.027	15.887	0.014	0.286	-0.514	-8.986	0.052	1.342	0.351	200.59
Merval Buenos Aires	Argentina	MERVAL	795	0.002	1.502	1.627	18.892	0.112	1.714	-0.021	-0.275	-0.032	-0.616	0.351	0.94
OMX Vilnus Index	Lithuania	VILSE	794	0.001	0.716	0.881	18.478	0.057	1.587	0.448	10.632	0.060	2.100	0.346	65.92
SET50 Index	Thailand	SET50	795	0.000	-0.331	1.083	18.286	0.083	1.849	-0.133	-2.538	0.054	1.499	0.339	6.27
Colcap	Colombia	COLCAP	662	0.002	1.581	0.777	13.310	0.149	3.108	-0.298	-5.116	0.032	0.880	0.298	5.38
Ljubljana SE	Slovenia	SBITOP	625	-0.001	-1.049	0.822	15.790	-0.082	-1.820	0.180	3.396	0.000	-0.011	0.297	3.65
Belgrde SE	Serbia	BELEX15	494	-0.001	-0.730	1.110	13.353	-0.084	-1.090	0.484	5.178	0.164	2.970	0.276	21.74
Bucharest SE	Romania	BET	795	0.002	1.510	1.121	16.186	0.140	2.680	0.136	2.216	0.076	1.820	0.271	104.56
FTSE Bursa Malaysia	Malaysia	FBMKLCI	795	0.000	-0.313	0.592	15.895	0.000	-0.003	-0.054	-1.638	0.063	2.808	0.264	24.51
Istanbul SE 30 Index	Turkey	XU030	795	0.001	0.483	1.449	15.174	0.100	1.383	-0.038	-0.445	-0.006	-0.096	0.257	70.97
PSEi Index	Philippines	PCOMP	795	0.000	0.456	0.840	15.078	0.041	0.981	-0.030	-0.610	0.001	0.031	0.253	31.19
Tel Aviv 25 Index	Israel	TA-25	795	0.000	0.297	0.774	15.672	-0.036	-0.959	-0.017	-0.392	0.072	2.407	0.250	218.25
Ukraine PFTS Index	Ukraine	PFTS	795	0.002	1.054	1.214	13.848	0.093	1.411	0.203	2.612	0.107	2.030	0.206	3.30
FTSE Nasdaq Dubai 20	UAE	DUAE	458	0.000	0.138	1.128	10.692	-0.342	-3.341	0.429	3.465	0.022	0.311	0.205	25.82
SOFIX	Bulgaria	SOFIX	752	0.001	0.841	0.866	11.747	-0.040	-0.706	0.491	6.867	0.090	2.012	0.178	6.17
Egyptian EGX30	Egypt	EGX30	795	0.002	1.124	1.023	12.244	-0.038	-0.601	0.252	3.412	0.100	1.986	0.162	152.09
MBI 10	Macedonia	MBI	534	0.000	0.255	0.899	9.669	-0.080	-0.936	0.388	3.847	0.128	2.080	0.156	7.07
MONEX20	Montenegro	MONEX20	629	0.003	1.607	0.845	8.721	0.062	0.737	0.559	5.672	0.166	2.749	0.142	21.07
Mauritius SE	Mauritius	SEMDEX	795	0.001	1.414	0.401	10.891	0.012	0.420	0.084	2.589	0.093	4.184	0.136	12.75
FTSE China A50	China	XIN9I	609	0.000	0.310	0.706	8.102	-0.385	-5.031	-0.056	-0.630	0.071	1.277	0.133	66.83
KSE 15 Index	Kuwait	KSX15	149	0.000	-0.018	0.251	1.816	0.053	0.463	-0.218	-1.819	-0.145	-2.179	0.130	167.91
FTSE NSE Kenya 25	Kenya	FNKEN2	204	0.002	1.465	0.581	4.529	-0.061	-0.530	-0.057	-0.428	0.171	2.376	0.112	22.53
Bahrain Bourse	Bahrain	BHSEASI	559	-0.001	-0.884	0.239	6.634	-0.095	-2.897	0.231	6.036	0.033	1.429	0.106	73.33
Quatar Exchange Index	Quatar	DSM	795	0.002	1.734	0.652	9.371	-0.105	-1.994	0.207	3.356	0.075	1.787	0.102	172.11
Vietnam HO Chi Minh	Vietnam	VNINDEX	765	0.001	0.732	0.664	7.509	0.138	2.016	-0.278	-3.258	0.103	1.901	0.101	20.67
OMX Riga Index	Latvia	RIGSE	794	0.001	0.776	0.547	8.375	0.072	1.458	0.259	4.483	0.057	1.446	0.099	46.70
Muscat Securities	Oman	MSM30	795	0.000	0.496	0.448	8.811	-0.129	-3.355	0.106	2.359	0.095	3.085	0.096	18.32
TASI Tadawul All	SaudiArabia	SASEIDX	795	0.002	1.328	0.605	8.556	-0.216	-4.046	0.152	2.425	0.033	0.780	0.095	65.51
Slovak Share Index	Slovakia	SKSM	795	0.001	1.065	0.323	5.922	-0.050	-1.210	0.351	7.280	0.042	1.265	0.085	189.38
Colombo SE All Share	Sri Lanka	CSEALL	795	0.002	2.321	0.459	7.998	0.033	0.763	0.166	3.267	0.084	2.425	0.083	213.47
Laos SE Composite	Laos	LSXC	219	0.002	1.053	0.395	2.025	0.277	1.630	0.219	1.065	0.220	2.096	0.064	17.76
Karachi SE Index	Pakistan	KSE100	795	0.003	2.846	0.448	6.452	-0.029	-0.548	-0.095	-1.544	0.048	1.147	0.061	111.23
Ghana SE Composite	Ghana	GGSECI	221	0.004	3.714	0.081	0.869	0.080	0.991	0.162	1.650	-0.063	-1.243	0.053	0.55
Tunisia SE	Tunisia	TUSI20	430	0.002	2.093	0.158	3.330	0.098	2.084	0.060	1.067	0.013	0.396	0.051	101.28
NGSE All Share Index	Nigeria	NGSEINDX	795	0.001	1.331	0.321	4.846	0.041	0.811	0.022	0.384	0.113	2.817	0.035	5.69
Casablanca SE	Morocco	MCSINDEX	795	0.000	0.363	0.205	4.841	-0.041	-1.274	0.008	0.223	0.027	1.074	0.030	68.28
MSE Top20 Index	Mongolia	MSETOP	795	0.006	2.272	0.497	2.970	0.000	-0.002	0.310	2.097	0.193	1.914	0.016	51.32
Iraq SE Index	Iraq	ISXGI	290	-0.001	-0.542	0.122	0.765	-0.114	-0.792	0.242	1.397	-0.127	-1.367	0.013	29.20
Tanzania Share Index	Tanzania	DARSTSI	434	0.004	4.498	0.027	0.622	-0.050	-1.181	0.017	0.320	-0.027	-0.895	0.005	21.35

The data cover the period between 2000-2015. Results are sorted from the highest to the lowest Rsquared values, for developed and emerging markets separately. Additionally, all significant *t*-Statistics are marked with grey background. Factors are based on top/bottom decile groups. WML factor is based on returns form last 2 months

Source: own calculations

In the results section we refer only to the four-factor model because its explanatory power was higher than three-factor model. Our results for equity indices are in many ways similar to well know studies for stock returns ([Lieksnis 2010], [Davis et al. 2000]), however they do not reveal so strong effects as was presented in the literature before. Therefore our main conclusions can be summarized as follows:

1. The highest explanatory power of the four-factor model we observe mainly for developed equity indices. In this group almost all Rsquared values are higher that 50%. On the other hand, for emerging markets they get much lower values.
2. The results of regressions for developed countries with highest Rsquared coefficients have negative (but close to zero) alpha coefficients (significant in 50% of cases) which informs us that there was no any additional returns which were not explained by four-factor model.
3. On the other hand, most alpha coefficients for emerging equity indices are positive but still rather insignificant.
4. Stability of risk factors effects can be observed only with regards to WML factor, which reveals a strong short-term momentum effect stable over 15 years period of research.

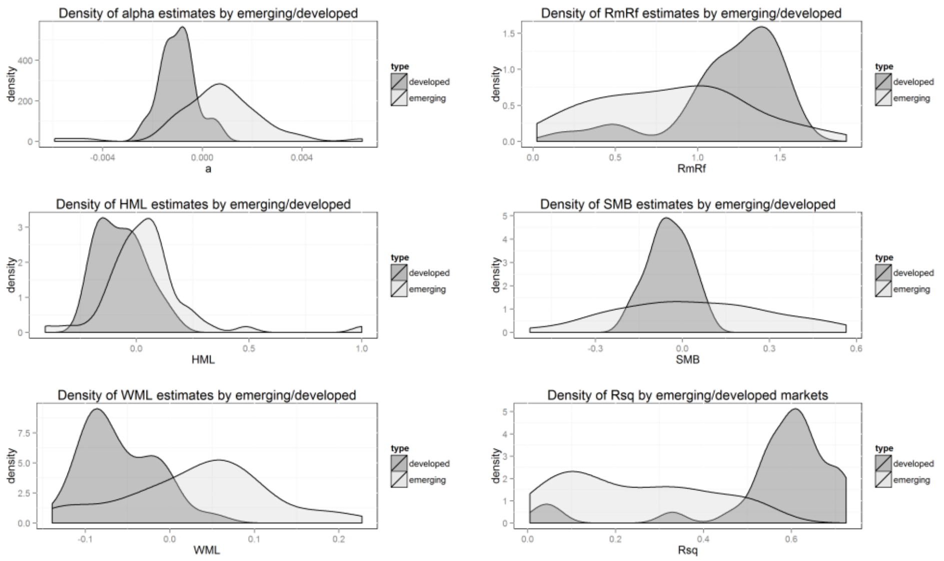
In order to draw more conclusions with regards to different results for developed and emerging markets, we analyzed the densities of parameters estimates and Rsquared values separately for these two types of equity indices (Figure 5).

Conclusions can be summarized as follows:

1. We observe significant difference between positive alpha for emerging and negative alpha for developed markets.
2. Beta for (R_m-R_f) factor is substantially higher for developed countries and additionally less diversified across countries.
3. The sensitivity to HML factor is higher for emerging markets, however again it is much more diversified for emerging equity indices.
4. The means of SMB beta estimates are almost equal, however their diversity is much higher for emerging market as well.
5. Characteristics of WML beta estimates is very similar to HML but the difference between means of beta and the diversity of betas is even larger.
6. Separate histograms for Rsquared for developed and emerging markets confirmed previous observations that regression for developed markets have higher explanatory power than these for emerging markets.

Above mentioned observations suggest that the four-factor model of Carhart can be quite robust approach for developed markets with high explanatory power. However, it should be amended and enhanced with additional risk factors for emerging markets.

Figure 5. Kernel density of parameter estimates and Rsquared values using Gaussian kernel function, separately for developed and emerging equity indices



The data cover the period between 2000-2015 (from 1990 only for R_m-R_f factor).

Source: own calculations

SUMMARY

It is not easy to summarize results which are only partly in line with other studies already presented in the literature and are only the first part of rather larger attempt to fully understand cross-section of equity indices returns. The most intriguing part revealing some light on equity indices returns is the difference of results for developed and emerging markets.

The main result is that using the well-known four-factor model we can only explain the variability of developed markets returns. On the other hand, emerging market equity indices require further investigation and research should be focused mainly on additional factors and on novel model implementation.

Further research should address the following questions:

- Are sensitivities to risk factors stable during various phases of economic cycles?
- Do correlations among international equity markets differ between high and low volatility periods?
- Can we build a zero investment portfolio with positive alpha based on analyzed risk factors?
- Which risk factor is the most important in portfolio construction?

REFERENCES

Arshanapalli B. G., Coggin T. D., Doukas J. (1998) Multifactor Asset Pricing Analysis of International Value Investment Strategies, The Journal of Portfolio Management, 24(757), pp. 10–23.

Asness C. S. (1995) The power of past stock returns to explain future stock returns, Unpublished working paper Applied Quantitative Research.

Black F., Jensen M. C., and Scholes M. (1972) The Capital Asset Pricing Model: Some Empirical Tests, Vol. 81.

Bondt W. F., Thaler R. (1985) Does the stock market overreact?, The Journal of Finance, 40(3), pp. 793–805.

Bondt W. F., Thaler R. H. (1987) Further evidence on investor overreaction and stock market seasonality, The Journal of Finance, 42(3), pp. 557–581.

Cakici N.,. Fabozzi F. J, Tan S. (2013) Size, value, and momentum in emerging market stock returns, Emerging Markets Review, 16, pp. 46–65.

Carhart M. M. (1997) On persistence in mutual fund performance, The Journal of Finance, 52(1), pp. 57–82.

Chen, L., Novy-Marx R., Zhang L. (2011) An Alternative Three-Factor Model, SSRN Electronic Journal.

Davis J. L., Fama E. F., French K. R. (2000) Characteristics, covariances and average returns: 1929 to 1997, The Journal of Finance, 55(1), pp. 389–406.

Fama, E. F., French K. R. (1992) The Cross-Section of Expected Stock Returns, Journal of Finance, XLVII(2).

Fama, E. F., French K. R. (2012) Size, value, and momentum in international stock returns,Journal of Financial Economics, 105(3), pp. 457–472.

Fama, E. F., French K. R. (2015) A five-factor asset pricing model, Journal of Financial Economics, 116(1), pp. 1–22.

Foye J., Mramor D., Pahor M. (2013) A Respecified Fama French Three-Factor Model for the New European Union Member States, Journal of International Financial Management & Accounting, 24(1), pp. 3–25.

Frazzini A., Pedersen L. H. (2014) Betting against beta, Journal of Financial Economics, 111(312417), pp. 1–25.

Griffin J. M. (2002) Are the Fama and French Factors Global or Country Specific?, Review of Financial Studies, 15, pp. 783–803.

Hou K., Karolyi G. A., Kho B. C. (2011) What factors drive global stock returns?, Review of Financial Studies, 24, pp. 2527–2574.

Jegadeesh N., Titman S. (1993) Returns to buying winners and selling losers: Implications for stock market efficiency, The Journal of finance, 48(1), pp. 65–91.

Lakonishok J., Shleifer A., Vishny R. W. (1994) Contrarian Investment, Extrapolation and Risk, The Journal of Finance, 49(5), pp. 1541–1578.

Lieksnis R. (2010) Multifactor Asset Pricing Analysis of the Baltic Stock Market, Ekonomika, 89(4), pp. 85–95.

Lintner J. (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, The review of economics and statistics, pp. 13–37.

Liu W. (2004) Liquidity premium and a two-factor model, EFA 2004 Maastricht Meetings Paper, 44 (September 2003), pp. 0–52.

Lo A. W., MacKinlay A. C. (1990) When are contrarian profits due to stock market overreaction?, Review of Financial Studies, 3(2), pp. 175–205.

MacKinlay A. C. (1995) Multifactor models do not explain deviations from the CAPM, Journal of Financial Economics, 38(1), pp. 3–28.

Rahim R. A., Noor A. H. S. M. (2006) A Comparison Between Fama and French Model and Liquidity-Based Three-Factor Models, Asian Academy of Management Journal and Finance, 2(2), pp. 43–60.

Sharpe W. F. (1964) Capital asset prices: A theory of market equilibrium under conditions of risk, The journal of finance, 19(3), pp. 425–442.

Wu X. (2002) A conditional multifactor analysis of return momentum, Journal of Banking and Finance, 26 (February 1997), pp. 1675–1696.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 102 – 112

DETERMINANTS OF THE DEMAND FOR INTERNATIONAL RESERVES IN UKRAINE

Victor Shevchuk
Institute of Economy, Sociology and Philosophy
Cracow University of Technology
e-mail: victorshevchuk@netscape.net

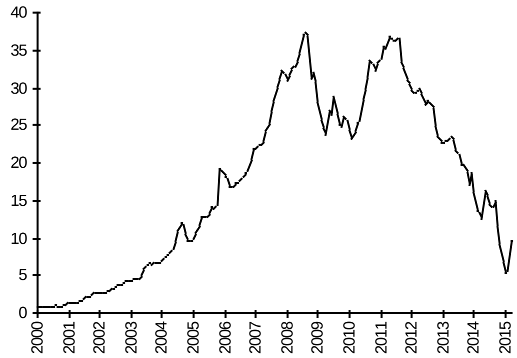
Abstract: This study investigates the demand for international reserves in Ukraine and its structural change in the wake of the 2008-2009 financial crisis in the context of a univariate error correction model (ECM). We find a time-invariant demand for international reserves in the short-run, with the inverse relationship with the volatility of international transactions, exchange rate depreciation and the excessive money stock and the positive link to imports and crisis developments. However, the long-term relationships are not stable over time, except for the effects of the money disequilibrium and crisis disturbances. The exchange rate depreciation and electoral cycle contribute to a depletion of reserves during the post-crisis period only. The adjustment of actual reserves to their long-run relationships is quite rapid.

Keywords: international reserves, imports, exchange rate, money stock

INTRODUCTION

As of August 2011, Ukraine held international reserves totaling 38.4bn USD but the level of reserves has shrunk to just 5.2bn USD by February 2015 (Figure 1). Following a short-lived depletion of reserves in the wake of a 2004 presidential campaign, international reserves rose sharply between 2005 and the middle of 2008, being bolstered mainly by significant capital inflows into the banking sector. However, a third of the international reserves has been lost in 2009–2010 as a result of the sudden stop in capital flows. Then the level of reserves had recovered to its pre-crisis level by the middle of 2011, but it was mainly the result of intense external borrowings. In what followed, a combination of the current account deficit, the burden of foreign debt service payments and political turmoil has resulted in a gradual loss of 85 percent of Ukraine's international reserves.

Figure 1. The level of international reserves in Ukraine (bn of USD), 2000–2015



Source: The IMF on-line *International Financial Statistics*

Empirical studies of the demand for international reserves are usually based on the buffer stock model or the monetary approach to the balance of payments. According to Frenkel and Jovanovic [1981], reserves serve as a buffer stock, with optimal reserves balancing the adjustment costs with the opportunity cost of holding reserves. As suggested by Johnson [1965], accumulation of international reserves is proportional to the net demand for money. For a stable demand for money, Edwards [1984] provided with a synthesis of both theoretical approaches. Recent self-insurance models explain the build-up of reserve holdings by exposure to volatile capital flows thus shifting focus of reserve adequacy assessments from flows of goods to flows of assets [Aizenman and Genberg 2012].

The purpose of this paper is to estimate the determinants of Ukraine's international reserves taking into account both the long-run and short-run relationships. The remainder of the paper proceeds as follows. Section THEORETICAL ISSUES provides a review of the theoretical approaches for the demand of international reserves. In Section DATA AND STATISTICAL METHODOLOGY, the Johansen cointegration test and the Engle-Granger two-step procedure are presented. Section EMPIRICAL RESULTS contains the econometric estimates of the determinants of international reserves in Ukraine. The last Section concludes.

THEORETICAL ISSUES

The buffer stock model implies that the level of international reserves is a stable function of just a few variables – a scale of international transactions, the adjustment cost, the opportunity cost and reserve volatility [Aizenman and Marion 2004]. Average reserves depend negatively on the exchange rate flexibility and depend positively on GDP and adjustment costs [Aizenman and Genberg 2012]. The monetary model of international reserves is based on the assumption that changes in the reserves reflect the difference between the demand for money and the domestic money supply [Frenkel and Johnson 1976; Edwards 1984].

Accounting for the monetary factors, the extended demand function for international reserves could be presented as follows:

$$\ln\left(\frac{R_t}{X_t}\right) = \alpha_0 + \alpha_1 \ln S_t + \alpha_2 \sigma_t + \alpha_3 \ln C_t + \alpha_4 (r_t^* - r_t) + \alpha_5 \ln E_t + \alpha_6 (\ln M_t - \ln \overline{M}_t) + \varepsilon_t, \tag{1}$$

where R_t is the actual reserve holdings, valued in US dollars and expressed as a ratio of X, where X could be the US GDP deflator, the domestic GDP, the total foreign debt or the money stock, S_t is a scaling variable, σ_t is the volatility of international transactions, C_t is for adjustment costs, r_t and r_t^* are domestic and foreign interest rates, respectively, E_t is the exchange rate (defined as units of domestic currency per unit of foreign currency), M_t and $\overline{M}_t$ are the actual and equilibrium stocks of money, ε_t is the stochastic factor.

The scaling variable (S_t) reflects a positive relationship ($\alpha_1 > 0$) between the reserves and the size of international transactions, usually proxied by real GDP, population size or the level of imports. The volatility of international receipts and payments (σ_t) as measured by the standard deviation of the trend-adjusted changes in reserves or the volatility of export receipts or macroeconomic fundamentals should positively correlate with the level of international reserves ($\alpha_2 > 0$).

The effect of the marginal (average) propensity to imports (C_t) is ambiguous ($\alpha_3 <> 0$). Initially, the central bank financing of the external deficit with its international reserves was considered as a standard measure of adjustment costs and thus the alternative for a loss in output. Assuming implementation of expenditure-reducing policies, the coefficient on imports is expected to be negative. However, a positive relationship between the propensity to import and international reserves in many empirical studies suggested that imports reflect greater external vulnerability.

Interpreting an interest rate differential between foreign and domestic rates as the opportunity cost of holding international reserves, an increase in $r_t^* - r_t$ should contribute to accumulation of reserves, because earnings on the liquid reserves reduce the opportunity costs of holding those reserves ($\alpha_4 > 0$).

As exchange rate depreciation is expected to improve the balance-of-payments, it is likely to reduce the need for international reserves ($\alpha_5 < 0$). The negative elasticity of the reserve demand with respect to the real) exchange rate is found for 13 industrial countries by Bahmani-Oskooee and Niroomand [2008]. Among country studies, a statistically significant relationship between international reserves and exchange rate is found for Turkey [Kasman and Ayhan 2008].

If the excessive money stock increases, the demand for international reserves would decrease ($\alpha_6 < 0$), though for a different reason. Assuming that M_{t-1} stands for the money supply and $\overline{M}$ is the proxy for money demand, excessive spending by domestic residents leads to a loss in international reserves; if the demand for money exceeds the money supply, there is an increase in reserves. As suggested by Edwards [1984], the term $M_{t-1} - \overline{M}_t$, henceforth denoted by M_t^{DE}, captures the effect the monetary disequilibrium on international reserves.

Among other factors, the demand for international reserves depends on the level of external debt [Alfaro and Kanczuk 2014], political business cycle [Dreher and Vaubel 2009], concerns about competitiveness [Delatte and Fouquau 2012], attempts to diminish real exchange rate volatility [Hviding, Nowak and Ricci 2004]. Obstfeld, Shambaugh and Taylor [2010] explain hoarding of international reserves by (i) the "fear of floating", (ii) expansion of domestic banking and financial system relative to GDP and (iii) an increase in the financial integration with international markets. Finally, the "keeping up with the Joneses" (regional imitation) motive implies that a country tries to build-up its reserves in order not to be seen to have lower reserves than a neighboring country and thereby become more susceptible to a loss of investor confidence [Aizenman and Genberg 2012, Cheung and Qian 2009]. Cheung and Ito [2009] found that the relationship between international reserves and their determinants is significantly different between developed and developing economies and is not stable over time. For Ukraine, the presence of precautionary and mercantilist motives of reserve holdings is found for short-term relationships, but the former does not hold over the long term [Makarenko and Gordieieva 2015].

DATA AND STATISTICAL METHODOLOGY

The sample comprises monthly data from 2000 to 2014. All data are obtained from the IMF *International Financial Statistics* online database. The variables used in the reserve demand function are: the value of imports, $IMPORT_t$ (in millions of 2000 US dollars), the volatility of reserves, σ_{t-1}, the nominal effective exchange rate, E_t (index, 2010=100), the excess money stock, M_{t-1}^{DE}, as specified in Eq. (1). Except for the excess money stock and reserves volatility, the variables are transformed into natural logarithms. The electoral dummy, PBC_t, meant to capture electoral effects, equals 1 in six pre-election months and 0 otherwise. Another dummy, $CRISIS_t$, controls for crisis developments.

Since the Ukraine's domestic interest rate contains a significant risk premium, it is rather difficult to estimate the opportunity cost of holding international reserves. Similar to other studies, for example, Badinger [2004], the average propensity to imports as measured by scaling the value imports with the level of income is excluded from the analysis, as there is a potential multicollinearity problem. Finally, the reserve demand model is reduced to a function of scale (the value of real import), uncertainty (volatility of reserves), competitiveness (exchange rate) and the monetary disequilibrium.

The volatility applied is the estimated conditional variance of the reserves from a univariate GARCH(1,1) model:

$$\Delta \ln R_t = \eta + \xi_t, \qquad \xi_t / \Omega_{t-1} \approx N(0, \sigma_t), \tag{2}$$

$$\sigma_t = \omega + \alpha \xi_{t-1}^2 + \beta \sigma_{t-1}, \quad \omega > 0, \quad \alpha \geq 0, \quad \beta \geq 0, \tag{3}$$

where Δ is the operator of first differences, η is the mean $\Delta \ln R_t$ conditional on past information (Ω_{t-1}), and ε_t is the stochastic factor.

The estimated σ_t (conditional variance) from the GARCH(1,1) model is applied in the estimation of the demand function for reserves as a measure of the volatility of international transactions. Table 1 presents the result from the GARCH(1,1) model for the international reserves, indicating a significant ARCH process. Compared to other studies [Choudhry and Hasan 2008], the ARCH effect is not very large, but the coefficient on α is statistically significant at the 1% level.

Table 1. Univariate GARCH results

Coefficients			
η	ω	α	β
0.466 (6.22***)	0.0003 (3.29***)	0.289 (3.71***)	0.684 (10.96***)

Note: z-statistic in parenthesis; ***, **, * imply statistical significance at the 1, 5 and 10% level, respectively.

Source: own calculations

Given that all series are nonstationary in levels (these results are not provided in order to save space and are available on request), the cointegration tests are conducted. The Johansen cointegration tests are presented in Tables 2 and 3. As indicated by the likelihood ratio test and the Akaike Information Criterion (AIC) test, eight and three lags were used in the cointegration test for the baseline model and the extended model, respectively. Both the trace test and the eigenvalue test indicate that changes in Ukraine's international reserves do form a cointegrating relationship with changes in its volatility, imports, nominal exchange rate and excessive money stock in either baseline or extended models (the cointegration tests are conducted with the inclusion of a linear trend with intercept).

As implied by the Engle-Granger two-step methodology, cointegration of the data containing unit roots in the individual time series allows to estimate the long-run relationship (in levels) by standard least-squares techniques and then use the lagged residuals to estimate a short-run dynamics (in first differences). The standard error-correction starts with positing long-run relationships between a dependent variable, and one or more independent variables, with lag structures to be empirically determined. In the second step, the short-run dynamics is estimated using the long-run result as an error-correction mechanism.

Table 2. Johansen Test Statistics for international reserves, its volatility and imports

Number of cointegrating equations		Trace statistic	0.05 Critical value	Prob.	Max-Eigen Statistic	0.05 Critical value	Prob.
H0: $r = r_0$	$r = 0$	53.74***	49.91	0.00	33.95***	25.82	0.00
	$r = 1$	19.79	25.87	0.23	14.65	19.38	0.21
	$r = 2$	5.13	12.51	0.57	5.12	12.51	0.57

Note: * denotes rejection of the null hypothesis at the 10 percent level (** at the 5 percent level, *** at the 1 percent level).
Source: own calculations

Table 3. Johansen Test Statistics for international reserves, its volatility, imports, nominal effective exchange rate and excessive money stock

Number of cointegrating equations		Trace statistic	0.05 Critical value	Prob.	Max-Eigen Statistic	0.05 Critical value	Prob.
H0: $r = r_0$	$r = 0$	105.10***	88.80	0.00	41.80***	38.33	0.01
	$r = 1$	63.29*	63.87	0.06	33.13**	32.11	0.03
	$r = 2$	30.16	42.91	0.49	15.39	25.82	0.59
	$r = 3$	14.76	25.87	0.59	7.93	19.39	0.82
	$r = 4$	6.83	12.51	0.36	6.83	12.51	0.36

Source: own calculations

Following suggestions from economic theory, the long-run relationship between international reserves and its key determinants may be written as follows:

$$\ln R_t = \alpha_0 + \alpha_1 \ln R_{t-1} + \beta_i \mathbf{X}_{it} + \nu_t, \tag{4}$$

where $\mathbf{X}_{it}$ is the vector of explanatory variables and ν_t is the stochastic factor.

The short-run dynamics around the long-run relationship (4) is defined as:

$$\Delta \ln R_t = \gamma_0 + \gamma_1 \Delta \ln R_{t-1} + \eta_i \Delta \mathbf{X}_{it} + \phi \nu_{t-1} + \varepsilon_t, \tag{5}$$

where ε_t is the error term.

The parameter ϕ on z_{t-1} is the error-correction coefficient, which reflects the speed of short-run adjustment. According to the Engle-Granger specification, if the lagged error-correction term carriers a negative and statistically significant coefficient, all variables are converging towards their long-run equilibrium.

EMPIRICAL RESULTS

Table 4 presents the results of the estimated long-run relationships, representing Eq. (3). We estimate two different time periods: the pre-crisis period (2000M1:2008M8) and the post-crisis period (2009M5:2015M2). For both the baseline and extended models, the ADF test indicates the stationarity of residuals, as required for the validity of the results.

Table 4. Determinants of international reserves (long-run coefficients)

Variables	Baseline model		Extended model	
	Pre-crisis	Post-crisis	Pre-crisis	Post-crisis
$\ln R_{t-1}$	0.927 (52.79***)	1.106 (38.70***)	0.919 (36.23***)	1.125 (39.77***)
$\ln IMPORT_{t-1}$	0.090 (4.33***)	–0.127 (–3.73***)	0.091 (1.78*)	–0.039 (–0.93)
σ_{t-1}	–3.093 (–3.10***)	–3.079 (–2.54**)	–3.342 (–3.36***)	1.340 (0.74)
$CRISIS_t$	0.267 (3.89***)	0.104 (3.16***)	0.256 (3.75***)	0.101 (3.31***)
$\ln E_{t-6}$	–	–	0.014 (0.29)	–0.203 (–3.01***)
M_{t-1}^{DE}	–	–	–0.329 (–1.78*)	–0.968 (–2.75***)
PBC_t	–	–	0.006 (0.37)	–0.075 (–3.05***)
R^2	0.98	0.97	0.98	0.97
ADF	–5.31***	–8.06***	–5.34***	–8.12***

Note: the numbers in parenthesis are *t* statistics.
Source: own calculations

The significant coefficient on the lagged imports is found to be positive for the pre-crisis period in both specifications, but it is not the case for the post-crisis period. The coefficient on $\ln IMPORT_{t-1}$ is negative and statistically significant in the baseline model, while it is not statistically different from zero in the extended model. Contrary to what is predicted by the buffer-stock model, a higher variability in international reserves reduces the level of reserves in the pre-crisis period. If control for a wider set of macroeconomic variables, the coefficient on σ_t becomes positive (but not significant) for the post-crisis period. The demand for reserves is stimulated by the crisis developments. In absolute term, the coefficient on $CRISIS_t$ is much larger in specifications for the pre-crisis period.

The nominal effective exchange rate has no effect on the demand for international reserves in the pre-crisis period, but a moderate and long-delayed (five months) impact is found for the post-crisis period. As depreciation of the exchange rate contributes to a depletion of the reserves, it is in line with the argument linking a weakening of the exchange rate and improvement in the balance-of-payments. On the other hand, it is not ruled out that a negative link between the exchange rate depreciation and international reserves reflects a strong demand for foreign exchange on the domestic market.

There is the very strong negative effect that the lagged excessive money stock has on the level of international reserves, especially in the post-crisis period. Much higher value of the estimated coefficient on M_{t-1}^{DE} may be due to the fact that the money variable reflects also the impact of a very low inflation over the 2011–2013 period. An exchange rate depreciation is neutral in respect to reserves during the pre-crisis period, but it clearly contributes to a depletion of the reserves in the pre-crisis period. It is worthwhile to note that the money lag is much shorter than the exchange rate lag. Given the results from regression for the post-crisis period, appreciation of the *hryvna* and a restrictionary monetary stance could be recommended as policy instruments for accumulation of international reserves.

Specifications of a dynamic error-correction model follow closely those of the long-run regressions. Both regression models explain about 20% of the variation in the dependent variable for the pre-crisis period but the value of R^2 becomes twice as large for the post-crisis period (Table 5). Considering that regressions explain the rate of changes in the reserve demand, the R^2 value can be considered quite good. The extended model performs slightly better than the baseline model but the differences in the value of R^2 are rather marginal.

Table 5. Determinants of international reserves (short-run coefficients)

Variables	Baseline model		Extended model	
	Pre-crisis	Post-crisis	Pre-crisis	Post-crisis
$\Delta \ln R_{t-1}$	0.779 (5.60^{***})	0.963 (6.47^{***})	0.818 (5.96^{***})	0.875 (7.07^{***})
$\Delta \ln IMPORT_t$	0.162 (1.86^{*})	0.176 (1.72^{*})	0.169 (1.99^{*})	0.216 (2.19^{**})
$\Delta\sigma_{t-1}$	–2.500 (-1.83^{*})	–4.291 (-2.24^{**})	–2.596 (-1.96^{*})	–3.707 (-1.99^{*})
$CRISIS_t$	0.246 (3.91^{***})	0.056 (2.01^{**})	0.244 (3.99^{***})	0.085 (2.75^{***})
$\Delta \ln E_t$	–	–	–0.644 (-1.71^{*})	–0.662 (-2.69^{***})
ΔM_{t-1}^{DE}	–	–	–0.722 (-1.80^{*})	–1.061 (-1.88^{*})
Error-correction term	–0.843 (-5.05^{***})	–0.808 (-3.95^{**})	–0.873 (-5.34^{***})	–0.918 (-5.28^{***})
R^2	0.18	0.38	0.20	0.40
ADF	-9.81^{***}	-8.98^{***}	-10.61^{***}	-7.53^{***}

Source: own calculations

While the long-run reserve response to the level of imports is heterogeneous across specifications and time span, the short-run effects are about the same for the pre- and post-crisis periods. An increase in imports has a somewhat stronger (and significant) effect on international reserves in the post-crisis period. Also, short-run import elasticity has slightly increased in the extended model.

Contrary to the estimates of the long-run coefficients, the inclusion of the monetary variables to the dynamic model has not changed the impact of reserves volatility in the post-crisis period. However, the coefficient on the lagged volatility remains negative which is not consistent with the buffer stock model.

Other variables perform as expected. It is confirmed that crisis developments are associated with an increase in the international reserves. Demand for reserves seems to react symmetrically to exchange rate and money stock changes. The size of the coefficients on $\Delta \ln E_t$ shows large and immediate effect of a change in the exchange rate on reserves demand. The same negative effect is obtained for the changes in the excessive money stock, being in full accordance with the estimates of the long-run coefficients on the monetary variable. This suggests that the dynamics of international reserves is exchange rate and money stock sensitive, similar to the long-run relationships.

The coefficient of the lagged error correction term is very large and highly significant at the 1% level, pointing to a very fast adjustment of the short-run dynamics of international reserves to the deviation from its long-run equilibrium. Our results are in sharp contrast to those by Makarenko and Gordieieva [2015], who obtained that the speed of adjustment is rather slow (just 8% over a quarter).

CONCLUSIONS

Using monthly data for the period 2000:1 to 2014:12, this paper analyzes the dynamic relationship between the value of Ukraine’s international reserves and its determinants in the context of a multivariate error-correction model. The results reported in this paper indicate that international reserves maintain a stable time-invariant long-run relationship with crisis developments and excessive money stock, suggesting that the monetary disequilibrium plays a significant role in reserve movements even in the long-run. Contrary to the theoretical explanations, the variability of international transactions is negatively correlated with the reserves. The relationship between the levels of international reserves and imports seems to be asymmetrical for the pre- and post-crisis periods. The exchange rate depreciation and political business cycle contribute to a depletion of international reserves during the post-crisis period only.

Our results indicate a rapid correction of disequilibrium situations between desired and actual reserves. Ukraine’s international reserves maintain a uniform short-run positive relationship with changes in imports and the negative relationship with changes in variability of reserves, exchange rate depreciation and excess money stock. It is further confirmed that crisis disturbances lead to a loss of international reserves. All short-run relationships are robust across specifications and sub-periods chosen.

REFERENCES

Aizenman J., Genberg H. (2012) Research on the Demand for International Reserves: Developments in Academia, the Contribution of IMF Researchers, and Influence on IMF Surveillance, IEO Background Paper BP/12/01, Washington, International Monetary Fund.

Aizenman J., Marion N. (2004) International reserve holdings with sovereign risk and costly tax collection, The Economic Journal, Vol. 114, Issue 497, pp. 569–591.

Alfaro L., Kanczuk F. (2014) Debt Redemption, Reserve Accumulation, and Exchange-Rate Regimes, Working Paper 13-074, Cambridge, Mass., Harvard Business School.

Badinger H. (2004) Austria's Demand for International Reserves and Monetary Disequilibrium: The Case of a Small Open Economy with a Fixed Exchange Rate Regime, Economica, Vol. 71, Issue 281, pp. 39–55.

Bahmani-Oskooee M., Niroomand F. (1998) On the Exchange-Rate Elasticity of the Demand for International Reserves: Some Evidence from Industrial Countries, Weltwirtschaftliches Archiv (Review of World Economics), Vol. 124, No. 1, pp. 161–168.

Cheung Y.-W., Ito H. (2009) A Cross-Country Empirical Analysis of International Reserves, CESifo Working Paper No. 2654, Munich, Ifo Institute.

Cheung Y.-W., Qian X. (2009) Hoarding of International Reserves: Mrs Machlup's Wardrobe and the Joneses, Review of International Economics, Vol. 17, No. 4, pp. 824–843.

Choudhry T., Hasan M. (2008) Exchange Rate Regime and Demand for Reserves: Evidence from Kenya, Mexico and Philippines, Open Economy Review, Vol. 18, Issue 2, pp. 167–181.

Delatte A.-L., Fouquau J. (2012) What Drove the Massive Hoarding of International Reserves in Emerging Economies? A Time-varying Approach, Review of International Economics, Vol. 20, No. 1, pp. 64–176.

Dreher A., Vaubel R. (2009) Foreign exchange intervention and the political business cycle: A panel data analysis, Journal of International Money and Finance, Vol. 28, Issue 5, pp. 755–775.

Edwards S. (1984) The Demand for International Reserves and Monetary Equilibrium: Some Evidence from Developing Countries, Review of Economics and Statistics, Vol. 66, No. 3, pp. 495–500.

Frenkel J., Johnson H. (1976) The Monetary Approach to the Balance of Payments, London, Allen & Unwin.

Frenkel J., Jovanovich B. (1981) Optimal international reserves: a stochastic framework, The Economic Journal, Vol. 91, Issue 362, pp. 507–514.

Hviding K., Nowak M., Ricci L. (2004) Can Higher Reserves Help Reduce Exchange Rate Volatility? IMF Working Paper No. WP/04/189, Washington, International Monetary Fund.

Johnson M. (1965) International trade and economic growth: studies in pure theory, 2nd Edition, London, Allen and Unwin.

Kasman A., Ayhan D. (2008) Foreign exchange reserves and exchange rates in Turkey: Structural breaks, unit roots and cointegration, Economic Modelling, Vol. 25, No. 1, pp. 83–92.

Makarenko M., Gordieieva D. (2015) Determining key determinants of demand for international reserves in Ukraine: Cointegration analysis, Financial and credit activity: problems of theory and practice, Vol. 1, No. 18, pp. 182–194.

Obstfeld M., Shambaugh J., Taylor A. (2010) Financial stability, the trilemma, and international reserves, American Economic Journal: Macroeconomics, Vol. 2, No. 2, pp. 57–94.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 113 – 122

DEVELOPMENT OF KNOWLEDGE-BASED ECONOMY IN EUROPEAN UNION IN 2000-2014

Iwona Skrodzka
Faculty of Economics and Management, University of Bialystok
e-mail: iwona.skrodzka@uwb.edu.pl

Abstract: Knowledge-based economy (KBE) is an economy where knowledge is created, acquired, transmitted and used effectively by businesses, organizations, individuals and communities. The concept of KBE was emphasised in the EU programmes such as the Lisbon Strategy and the Europe 2020 Strategy. One of the three priorities of the Europe 2020 Strategy is to promote smart growth, understood as developing an economy based on knowledge and innovation. The aim of the paper is to analyze the development of KBE in European Union in period 2000-2014. The concept of KBE measurement is based on Knowledge Assessment Methodology and the soft modeling method.

Keywords: knowledge-based economy, economic development, soft modeling, KAM methodology

INTRODUCTION

Recent years have witnessed numerous changes in economic theories, especially with reference to the following concepts:

- information society, i.e. the one which uses teleinformation technologies intensively,
- knowledge-based economy (KBE), including "the new economy" (teleinformation technologies it promotes), issues in education (knowledge society), as well as innovation systems, and the institutional system, which is considered indispensable for the development of the above-mentioned elements [Piech 2009, pp. x-xi].

These concepts were emphasised in the European Union (EU) programmes such as the Lisbon Strategy and the Europe 2020 Strategy. The Lisbon Strategy stated that "knowledge and innovation will be the beating heart of the European growth" [European Commission 2005, pp. 4]. The Europe 2020 Strategy, a new long-term European growth programme, which replaced the Lisbon Strategy, stresses the need for a greater coordination of the EU member states in order to overcome the crisis and implement the reforms which will enable us to face the challenges of globalization, ageing societies and a growing need for resource efficiency. Therefore, three priorities were determined:

- smart growth – developing an economy based on knowledge and innovation,
- sustainable growth – promoting a more resource efficient, greener and more competitive economy,
- inclusive growth – fostering a high-employment economy delivering economic, social and territorial cohesion [European Commission 2010, p. 8].

The paper focuses on the issue of measuring of knowledge-based economy in European Union. KBE is difficult to measure due to its complexity, multidimensionality, unobservability. Its measurement requires prior solution of various problems such as: the imprecise and unquantifiable definition of KBE, the choice of method, the choice of indicators referring to different aspects of KBE, the choice of an optimal set of indicators, data availability. The aim of the paper is to analyze the development of KBE in European Union in period 2000-2014. The concept of KBE measurement is based on Knowledge Assessment Methodology (KAM) and the soft modeling method.

DEFINING AND MEASURING KNOWLEDGE-BASED ECONOMY

Knowledge-based economy is on one hand perceived in a narrow sense as a part of economy dealing with knowledge industry, mainly science. However, in a broader sense, it is understood as the economy whose one production factor is knowledge [Piech 2009, p. 214]. The classical definition of KBE is the one proposed by Organisation for Economic Co-operation and Development (OECD), which defines it as an economy directly depending on knowledge and information production, distribution and use [OECD 1996, p. 7]. The Asia-Pacific Economic Co-operation (APEC) Economic Committee defined KBE as an economy in which the production, distribution, and use of knowledge is the main driver of growth, wealth creation and employment across all industries [APEC Economic Committee 2000, p. vii]. According to the definition coined by the OECD and the World Bank Institute, KBE is an economy where knowledge is created, acquired, transmitted and used effectively by enterprises, organizations, individuals and communities. It does not focus narrowly on high-technology industries or on information and communications technologies, but rather presents a framework for analyzing a range of policy options in education, information infrastructure and innovation systems that can help usher in the knowledge economy [OECD, World Bank 2001, p. 3]. It is also assumed that KBE consists of four pillars:

- human capital, in whom some knowledge is stored,

- innovation system with entrepreneurship, more focused on businesses but also on cooperation with science, which also creates new knowledge,
- teleinformation technologies, which facilitate knowledge exchange, also abroad,
- institutional and legal environment, which creates conditions for the development of the above-mentioned areas [Piech 2009, p. 217].

KBE was first measured by F. Machlup, who regroupped economic branches and created a brand new sector – knowledge. The vital work on KBE was the OECD report published in 1996, where the notion of the "knowledge economy" was used for the first time. In 1998, the World Bank created Knowledge Assessment Methodology (KAM). In the same year, the Progressive Policy Institute presented the index of the new economy. A year later, the APEC initiated a project called: "Towards Knowledge-based Economies in APEC". At the beginning of the year 2000, the Australian Statistical Office started research into the knowledge-based economy and society ("Measuring a Knowledge-based Economy and Society"). In the same year, the Harvard University Center for International Development published a report: "Readiness for the Networked World". It presented the ranking list of countries based on the criterion of the readiness. In 2002, the UNECE published its own knowledge-based economy model "Regional Assessment Report" [Dworak 2014, pp. 11-12].

Although during the last 20 years multiple studies have been conducted and numerous works have been written on KBE, one widely accepted measurement method has not been achieved. We can only list a few dominant measurement methods, such as the KAM, drawn up by the World Bank, or the methodology proposed by the OECD. The methodologies have constantly been developed and each of them is a subject to constant criticism [Piech 2009, p. 315].

The KAM, which was developed within the framework of "The Knowledge for Development" (K4D) programme, is regarded as the most developed way of measuring KBE. It distinguishes four key pillars:
- Economic Incentive and Institutional Regime; indicators: tariff and non-tariff barriers, regulatory quality, rule of law.
- Education and Human Resources; indicators: adult literacy rate (% age 15 and above), latest version – average years of schooling, secondary enrollment, tertiary enrollment.
- Innovation System; indicators: researchers in R&D, per million population or in the latest version: payments and income from licence fees, patents applications granted by the US Patent and Trademark Office, per million population, scientific and technical journals articles, per million population.
- Information Infrastructure; indicators: telephones per 1000 persons (telephone mainlines and mobile phones), computers per 1000 persons, Internet users per 10000 persons.

The pillars are used to construct two global indexes:
- Knowledge Index (KI), which determines the knowledge potential of a country; this indicator is calculated as an arithmetic average of three subindexes, which

represent three pillars of KAM (except the Economic Incentive and Institutional Regime);
- Knowledge Economy Index (KEI), which illustrates a general development level of a knowledge-based economy; this indicator is calculated as an arithmetic average of four subindices, which represent the four pillars of KAM [Chen, Dahlman 2005, pp. 9-13].

The advantages of this method are its simplicity, clarity and versatility. It enables comparison of the KI and KEI indicators and their components in both dimensions: intertemporal and international. The method is criticised *inter alia* for: insufficient theoretical background, the tendency to repeat information by indicators, the lack of differentiated weights for indicators, insufficient information about many of the analysed economies, inaccessibility of indicators in the systems of international statistics, incomparability of data due to a variety of data sources [Becla 2010, pp. 56-70].

In this study the concept of KBE measurement is based on KAM methodology and the soft modeling method. In the literature description of the soft modeling method can be found in [Wold 1980], its generalization in [Rogowski 1990] and examples of application in [Perło 2004, Skrodzka 2015].

THE CONCEPT OF SOFT MODEL

Soft model[1] enables to research unobserved variables (latent variables). The values of these variables cannot be directly measured because the lack of a generally accepted definition or the absence of a clear way of measuring them. Soft model consists of two sub-models: the internal sub-model – a system of relationships among latent variables, which describes the relationship arising from the theory and the external sub-model – defines the latent variables based on observed variables, known as indicators.

Indicators enable indirect observation of latent variables and are selected following a chosen theory or the researcher's intuition. In soft modeling, a latent variable can be defined by indicators in two ways: inductively – this approach is based on the assumption that indicators create latent variables (formative indicators) or deductively – this approach is based on the assumption that indicators reflect their theoretical notions (reflective indicators). In both approaches, latent variables are estimated as weighted sums of their indicators [Rogowski 1990, pp. 25-26].

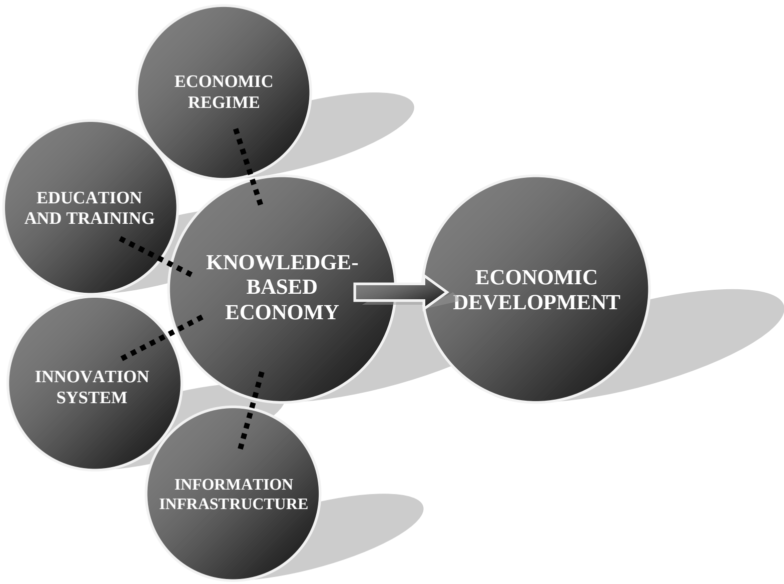
A soft model is constructed similarly to classical econometric models, with the following stages:
Stage I: describing relationships among latent variables in an internal model (specification of an internal model).

[1] Soft modeling method was created by H. Wold [1980].

Stage II: describing latent variables by indicators (specification of an external model).
Stage III: estimating model parameters (the internal one and the external one simultaneously) with the Partial Least Square – PLS method.
Stage IV: content-based and statistical verification of a model (Stone-Geisser test and "2s" rule[2]).

As a result of using the PLS method, we obtain estimates of latent variables, which can be regarded as synthetic measures. These quantities depend not only on external relations but also on relations among latent variables assumed in the internal model. It means that cognition depends not only on the definition of a given notion but also on the theoretical description. Soft modeling makes full use of the theoretical and empirical knowledge. This is what among other things distinguishes the presented method from most of commonly applied methods of multidimensional comparative analysis[3].

Figure 1. The concept of internal sub-model



Source: own elaboration

Figure 1 presents the concept of internal sub-model. The concept assumes relationships between two unobserved categories: the level of development of knowledge-based economy and the level of economic development. KBE is defined by four pillars (according to KAM methodology): economic regime,

[2] Parameter is statistically significant when value of double error is higher than value of estimator.
[3] This is also characteristic of structural models.

education and training, innovation system and information infrastructure. They are also unobserved. Hence KBE is the second-order latent variable.

Estimated model consists of two following equations

$$KBE = \alpha_1 REG + \alpha_2 EDU + \alpha_3 INN + \alpha_4 ICT + \alpha_0 + \varepsilon \quad (1)$$

$$ED = \beta_1 KBE + \beta_0 + \xi \quad (2)$$

where:
KBE – the level of development of knowledge-based economy,
REG – economic regime,
EDU – education and training,
INN – innovation system,
ICT – information infrastructure,
ED – the level of economic development,
α_0, α_1, α_2, α_3, α_4, β_0, β_1 – structural parameters,
ε, ξ – error terms.

Table 1. Indicators of latent variables

Latent variable		Indicator	Meaning	Type of indicator
KBE	*REG*	REG01	Gross capital formation (% of GDP).	stimulant
		REG02	Exports of goods and services (% of GDP).	stimulant
		REG03	Imports of goods and services (% of GDP).	stimulant
	EDU	EDU04	Persons with tertiary education attainment (%).	stimulant
		EDU05	Employees with tertiary education attainment (%).	stimulant
		EDU06	Life-long learning of persons aged 25-64 (%).	stimulant
		EDU07	Graduates (ISCED 5-6) in mathematics, science and technology (per 1 000 inhabitants aged 20-29).	stimulant
	INN	INN08	Persons employed in science and technology (% of total population)	stimulant
		INN09	Researchers in business enterprise sector (per 10 000 employees).	stimulant
		INN10	Total intramural R&D expenditure (% of GDP).	stimulant
	ICT	ICT11	Households with Internet access (%).	stimulant
		ICT12	Persons employed using computers with access to World Wide Web (% of total employment).	stimulant
ED		ED01	Gross domestic product per capita (euro, chain linked volumes - 2010).	stimulant
		ED02	Gross value added per employee (euro, chain linked volumes - 2010).	stimulant
		ED03	Total investment (% of GDP).	stimulant
		ED04	The share of agriculture in gross value added (%).	destimulant
		ED05	The share of industry in gross value added (%).	stimulant

Source: own elaboration

Each of latent variables is defined by a set of indicators based on deductive approach (see Table 1). Data use to specify the model are taken from Eurostat[4] and they refer to period 2000-2014. Indicators of KBE pillars were selected based on the KAM methodology but a key element was data availability. The following items were measured statistically: the variability of indicators (the coefficient of variation above 5%), a correlation level[5]. Missing data were complemented by extrapolation of time series (19 observation from 255 – 7%).

ESTIMATION RESULTS

Model presented on Figure 1 was estimated using the PLS software[6]. Table 2 contains estimates of weights and loadings with regard to external sub-model. All parameters are statistically significant (“2s” rule).

Some results are not consistent with expectations. Indicator REG01 is a stimulant of both REG and KBE latent variables but it has negative weight and loading. The values of this indicator decreased in periods 2000-2003, 2007-2009 and 2011-2014. The average annual rate of decline was 1%. Indicators ED03 and ED04 are stimulant of ED variable but they have negative weights and loadings. It is due to a decrease in the value of these indicators in period 2000-2014 (average annual rates of decline were: 0.3% for ED03 and 2% for ED04).

Table 2. Estimates of weights and loadings of the external sub-model

Latent variable	Indicator	Loading	Standard deviation	Weight	Standard deviation
REG	REG01	-0.6555	0.0008	-0.3172	0.0006
	REG02	0.9755	0.0002	0.4274	0.0002
	REG03	0.9286	0.0003	0.4040	0.0003
EDU	EDU04	0.9870	0.0000	0.2713	0.0000
	EDU05	0.9915	0.0000	0.2714	0.0000
	EDU06	0.8840	0.0000	0.2216	0.0000
	EDU07	0.9894	0.0000	0.2701	0.0000
INN	INN08	0.9795	0.0001	0.3546	0.0000
	INN09	0.9865	0.0000	0.3563	0.0000
	INN10	0.9377	0.0001	0.3212	0.0000
ICT	ICT11	0.9990	0.0000	0.4996	0.0000
	ICT12	0.9990	0.0000	0.5014	0.0000

Source: own calculations

[4] http://ec.europa.eu/eurostat/data/database

[5] Depending on the way a latent variable is defined by indicators (an inductive or a deductive approach), indicators should show low or high correlation respectively.

[6] PLS software was created by J. Rogowski. It is available at Faculty of Economics and Management University of Bialystok.

Table 2. (continued) Estimates of weights and loadings of the external sub-model

Latent variable	Indicator	Loading	Standard deviation	Weight	Standard deviation
KBE	REG01	-0.6928	0.0015	-0.0665	0.0014
	REG02	0.9334	0.0013	0.0864	0.0008
	REG03	0.8824	0.0015	0.0818	0.0011
	EDU04	0.9979	0.0004	0.0949	0.0000
	EDU05	0.9981	0.0003	0.0955	0.0001
	EDU06	0.8149	0.0014	0.0806	0.0003
	EDU07	0.9934	0.0001	0.0953	0.0000
	INN08	0.9898	0.0004	0.0946	0.0002
	INN09	0.9945	0.0004	0.0948	0.0001
	INN10	0.8967	0.0007	0.0831	0.0003
	ICT11	0.9869	0.0001	0.0952	0.0001
	ICT12	0.9905	0.0002	0.0952	0.0000
ED	LED01	0.8400	0.0189	0.2276	0.0362
	LED02	0.9399	0.0166	0.2672	0.0614
	LED03	-0.4341	0.0322	-0.1510	0.0760
	LED04	-0.8598	0.0216	-0.2159	0.0703
	LED05	-0.9457	0.0249	-0.2490	0.0334

Source: own calculations

Other results are consistent with expectations. Stimulants have positive weights and loadings and destimulants have negative ones.

Equations (3) and (4) present estimations of internal relations. Standard deviations calculated basing on Tukey cut method are given in brackets.

$$\hat{KBE} = \underset{(0.0084)}{0.1984}REG + \underset{(0.0024)}{0.3623}EDU + \underset{(0.0105)}{0.2661}INN + \underset{(0.0043)}{0.1851}ICT - \underset{(0.1244)}{0.0999} \tag{3}$$

$$\hat{ED} = \underset{(0.0131)}{0.9733}KBE - \underset{(1.9218)}{2.2056} \tag{4}$$

Signs of estimators are consistent with expectations. Moreover, all latent variable are statistically significant ("2s" rule). Coefficient of determination (R^2) has value 1.0 for the equation (3) and value 0.95 for the equation (4). General Stone-Geisser test is equal to 0.73[7]. The model can be verified positively.

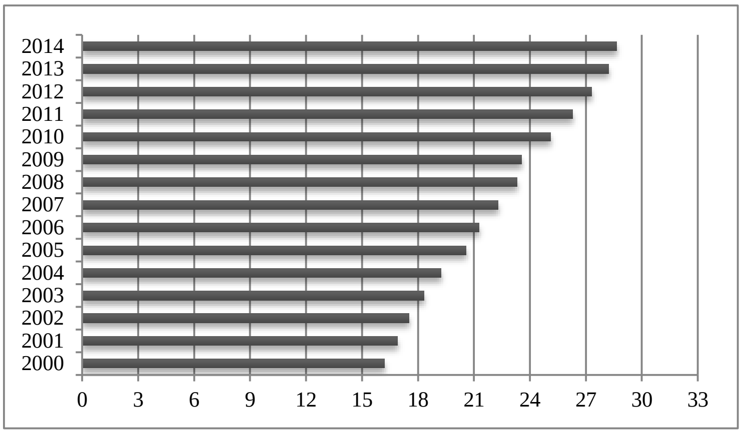
All four pillars influence positively the level of KBE. The strongest impact has pillar "education and training" (0.3623), the lowest – "information infrastructure" (0.1815). Furthermore, equation (4) shows that correlation between the level of KBE development and the level of economic development is positive and strong.

[7] Stone-Geisser test measures prognostic property of soft model. Its values are in the range from $-\infty$ to 1. Positive (negative) value of this test indicates high (poor) quality of model.

CONCLUSIONS

Partial Least Squares method gives estimates of values of KBE latent variable (Figure 2). They can be used to analyze changes in the level of KBE development.

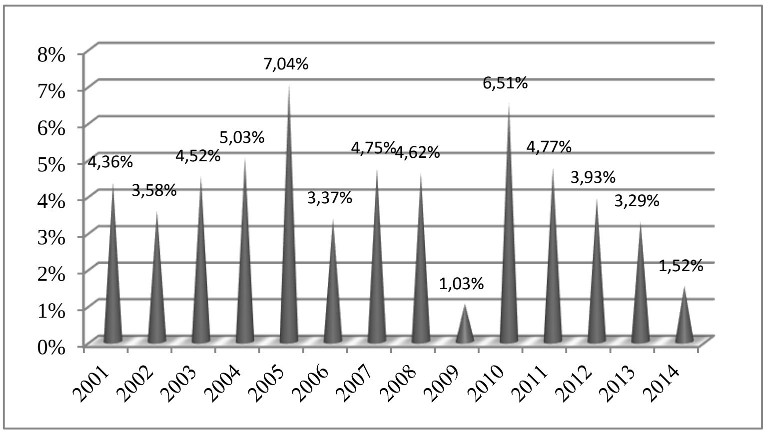
Figure 2. The values of synthetic measure of KBE latent variable



Source: own calculations

The annual rate of change is presented on Figure 3. In the period 2000-2014 KBE grew at an average rate 4% per year. The highest growth took place in 2005 (7.04%), the lowest – in 2009 (1.03%).

Figure 3. The annual rate of change in the level of KBE development in European Union, 2000-2014



Source: own calculations

REFERENCES

APEC Economic Committee (2000) Towards Knowledge-based Economies in APEC.

Becla A. (2010) Wady i zalety metody KAM (Knowledge Assessment Methodology) służącej do identyfikacji poziomu zaawansowania gospodarki opartej na wiedzy, Prace Naukowe UE we Wrocławiu, nr 139, Wrocław.

Chen D. H. C., Dahlman C. J. (2005) The Knowledge Economy, the KAM Methodology and World Bank Operations, World Bank Institute, D.C. 20433, Washington, http://siteresources.worldbank.org/INTUNIKAM/Resources/2012.pdf.

Dworak E. (2014) Gospodarka oparta na wiedzy, innowacyjność i rynek pracy, Wydawnictwo UŁ, Łódź.

European Commission (2005) Working together for growth and jobs – A new start for the Lisbon Strategy, Communication to the Spring European Council, Brussels.

European Commission (2010) Europe 2020, A European strategy for smart, sustainable and inclusive growth, Brussels.

OECD (1996) The Knowledge-based Economy, Paris.

OECD, World Bank (2001) Korea and Knowledge-based Economy, Making the Transition, Paris.

Perło D. (2004) Źródła finansowania rozwoju regionalnego, Wydawnictwo Wyższej Szkoły Ekonomicznej w Białymstoku, Białystok.

Piech K. (2009) Wiedza i innowacje w rozwoju gospodarczym: w kierunku pomiaru i współczesnej roli państwa, Instytut Wiedzy i Innowacji, Warszawa.

Rogowski J. (1990) Modele miękkie. Teoria i zastosowanie w badaniach ekonomicznych, Wydawnictwo Filii UW w Białymstoku, Białystok.

Skrodzka I. (2015) Kapitał ludzki polskich województw – koncepcja pomiaru, Wydawnictwo Uniwersytetu w Białymstoku.

Wold H. (1980) Soft Modelling: Intermediate between Traditional Model Building and Data Analysis, Banach Centre Publication 6, Mathematical Statistics.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 123 – 132

DATA VINTAGE IN TESTING PROPERTIES OF EXPECTATIONS

Emilia Tomczyk
Institute of Econometrics, SGH Warsaw School of Economics
e-mail: Emilia.Tomczyk@sgh.waw.pl

Abstract: Results of quantification procedures and properties of expectations series obtained for two data vintages are described. Volume index of production sold in manufacturing is defined for end-of-sample and real time data, and evaluated against expectations expressed in business tendency surveys. Empirical analysis confirms that while there are only minor differences in quantification results with respect to data vintage, properties of expectations time series obtained on their basis do diverge.

Keywords: end-of-sample (EoS) data, real time (RTV) data, data revisions, quantification procedures, expectations, unbiasedness, orthogonality

INTRODUCTION

Testing properties of economic expectations series constitutes a challenge for many reasons, among them those related to observing and measuring expectations, reliability of survey data, and selection of appropriate statistic and econometric methods for the purposes of empirical analysis. In this paper, I propose to address one of the issues related to quality of data employed to describe and evaluate expectations processes, that is, the subject of data revisions and data vintages.

Data revision is defined as an adjustment introduced after initial announcement had been published. End-of-sample (EoS) data is usually described, following Koenig et al. [2003], as data provided in the most recent announcement. Real time values (RTV) are initial numbers, available to economic agents in real time and (frequently) subject to revisions. The date when a particular dataset was made available is termed "vintage" of that data series. For details on definitions and classifications concerning data revisions, see Tomczyk [2013].

As far as I am aware, the extent of data revisions in Poland and their impact on predictive properties of time series have been addressed in a single paper only [see Syczewska 2013]. General literature pertaining to data revisions and their influence on quality of forecasts or properties of expectations time series is also limited. There is a continuing (if somewhat slow-moving) debate on whether tests of expectations should be based on initial or revised data [see Zarnowitz 1985; Keane and Runkle 1990; Croushore and Stark 2001, Mehra 2002]. Recent econometric analyses on impact of data revisions on forecast quality include Croushore [2011, 2012] and Arnold [2013]. There remain many open questions concerning appropriate data vintage for scaling qualitative survey data, measuring accuracy of expectations with respect to observed values, or testing properties of expectations time series.

In my previous papers [Tomczyk 2013, 2014], review of literature and databases related to economic data revisions, reasons for introducing adjustments to already published economic data, taxonomy of revisions, and comparison of quantification results for initial and revised data on production volume index in Poland are presented. In this paper, I continue this line of research by updating results on quantification procedures and testing properties of expectations obtained for two distinctive data vintages: end-of-sample (EoS) and real time (RTV).

DESCRIPTION OF DATA[1]

Analyses of industrial production are typically based on volume index of production sold in manufacturing provided by the Central Statistical Office (CSO). In Poland, systematic data revisions in the past two decades were due to changes in the base period for the index in 2004, 2009 and 2013. In January 2013, value of reference has been set as the average monthly industrial production of 2010. To extend the sample, observations dating back to January 2005 were recalculated to be consistent with the 2010 base.

To evaluate properties of expectations collected through qualitative business tendency surveys, quantification of survey data is necessary. In this paper, longer data series is used than in an earlier paper [Tomczyk 2014], and an additional issue is addressed: that not only dependent variables in quantification models (that is, CSO data on volume index of industrial production) are subject to revisions, but so are explanatory variables (that is, qualitative data on expectations and assessments of changes in economic variables).

Expectations and subjective assessments of changes in production are collected by the Research Institute for Economic Development (RIED, Warsaw School of Economics) through monthly business tendency surveys. The survey comprises eight questions designed to evaluate both current situation (as compared to last month) and expectations for the next 3 – 4 months by assigning them to one

[1] I would like to thank Mr Konrad Walczyk, PhD (Research Institute for Economic Development, Warsaw School of Economics) for his assistance with compiling the dataset.

of three categories: increase / improvement, no change, or decrease / decline. Previous studies based on RIED survey data show that expectations series defined for three- and four-month horizons exhibit only minor differences, with a slight superiority of the three-month forecast horizon.

Let us define the following:

A_t^1 – percentage of respondents who observed increase between t – 1 and t,
A_t^2 – percentage of respondents who observed no change between t – 1 and t,
A_t^3 – percentage of respondents who observed decrease between t – 1 and t,
P_t^1 – percentage of respondents who expect increase between t and t + 3,
P_t^2 – percentage of respondents who expect no change between t and t + 3,
P_t^3 – percentage of respondents who expect decrease between t and t + 3.

Balance statistics calculated for observed changes:

$$BA_t = A_t^1 - A_t^3$$

and for expectations:

$$BP_t = P_t^1 - P_t^3$$

remain the simplest method of quantification – that is, of converting qualitative business survey data into quantitative time series. More sophisticated procedures can be grouped into probabilistic and regressive quantification methods (for a concise review of basic quantification methods and their modifications, see Pesaran [1989]). In section 3, two versions of regression method are used to compare real time and end-of-sample data vintages.

RIED business survey data is also subject to revisions. Prior to 2012 revisions were sporadic: just a single one in 2010 (in April) and another in 2011 (in October). From 2012 on, adjustments become frequent. In 35 months between January 2012 and November 2014, balance statistics for assessments of changes in production has been revised a total of 19 times. In twelve cases, corrections were positive (that is, final number was larger than initial estimate by, on average, 0.64 of a percentage point). In seven cases, final number was smaller than initial estimate by, on average, 0.51 of a percentage point.

Let us employ the following notation: end-of-sample values will be marked with superscript EoS (for example, $A_t^{1\text{-EoS}}$), and real time values – with superscript RTV (for example, $A_t^{1\text{-RTV}}$). In the next section, both real time and end-of-sample data is used in regression quantification models.

QUANTIFICATION MODELS

Quantification procedures involve scaling qualitative survey data in a manner consistent with observed quantitative values, usually provided by government agencies – that is, widely available and officially endorsed data. In my earlier paper [Tomczyk 2013] I suggested that for quantification purposes, survey data should be compared with final (EoS) data rather than values available in real time because respondents are probably aiming to describe their final assessments and predictions rather than initial estimates subject to revisions. Initial attempt to test this proposition [Tomczyk 2014] has shown that end-of-sample data does indeed appear better suited to quantification of RIED business tendency survey data on volume index of industrial production. However, this conclusion was of limited reliability as none of the quantification models exhibited statistically satisfactory estimation results.

In this paper, I employ two versions of the regression method, introduced by O. Anderson [1952] and D. G. Thomas [1995], respectively. In Anderson's model, the following equation is estimated:

$$_{t-1}y_t = \alpha \cdot A_t^1 + \beta \cdot A_t^3 + \nu_t\,, \tag{1}$$

where $_{t\text{-}1}y_t$ describes relative change in value of variable y published by a statistical agency between t – 1 and t, and ν_t is a white noise error term. Parameters α and β are then estimated by OLS, and on the assumption that the same relationship holds for expectations reported in surveys, quantitative measure of expectations is constructed on the basis of the following equation:

$$_{t-1}\hat{y}_t = \hat{\alpha} \cdot P_t^1 + \hat{\beta} \cdot P_t^3\,, \tag{2}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are OLS estimates of (1) and reflect average change in dependent variable $_{t\text{-}1}y_t$ for respondents expecting, respectively, increase and decrease of dependent variable.

In 1995, D. G. Thomas offered a modification of the basic Anderson model to account for the special case in which normal or typical situation that respondents compare their current circumstances to is subject to a growth rate, making observing (or predicting) decreases in dependent variable more essential than increases:

$$_{t-1}y_t = \gamma + \delta \cdot A_t^3 + \xi_t\,, \tag{3}$$

where δ < 0, constant γ is interpreted as typical growth rate, and ξ_t is a white noise error term. Thomas' quantitative measure of expectations is given by the formula

$$_{t-1}\hat{y}_t = \hat{\gamma} + \hat{\delta} \cdot P_t^3\,, \tag{4}$$

where $\hat{\gamma}$ and $\hat{\delta}$ are OLS estimates obtained on the basis of (3).

For the purpose of comparing data vintages, dependent and explanatory variables in quantification models (1) and (3) may be based on either RTV or EoS data.

In case of real time data, dependent variable in regression quantification models (that is, changes in volume of industrial production) is typically defined on the basis of volume index of industrial production sold available in real time, IP_t^{RTV}. It seems likely that respondents evaluate current changes in production against recent averages, and one quarter appears a plausible observation horizon.

Let us define

$$P_t^{RTV-AV} = \frac{IP_t^{RTV}}{\frac{1}{3}\sum_{s=1}^{3} IP_{t-s}^{RTV}} - 1 \tag{5}$$

for real time data and

$$P_t^{EoS-AV} = \frac{IP_t^{EoS}}{\frac{1}{3}\sum_{s=1}^{3} IP_{t-s}^{EoS}} - 1 \tag{6}$$

for end-of-sample data. Formulas (5) and (6) reflect changes in volume of industrial production sold as compared to the average calculated on the basis of last three months, for real time and end-of-sample data.

All quantification models are estimated by OLS with HAC standard errors – that is, Newey-West heteroskedasticity and serial correlation consistent estimators – to account for possible serial correlation and unstable variance of the error term (due to inertia in processes describing behaviour of macroeconomic variables and probable learning patterns imbedded in expectations formation processes). All models are estimated on sample from April 2005 till November 2014 (n = 116). Estimated equations take the following form:

Anderson's model for real time data: $\widehat{P}_t^{RTV-AV} = 0.2883 \cdot A_t^1 - 0.2473 \cdot A_t^3$

Anderson's model for end-of-sample data: $\widehat{P}_t^{EoS-AV} = 0.2866 \cdot A_t^1 - 0.2458 \cdot A_t^3$

Thomas' model for real time data: $\widehat{P}_t^{RTV-AV} = 0.1241 - 0.4332 \cdot A_t^3$

Thomas' model for end-of-sample data: $\widehat{P}_t^{EoS-AV} = 0.1233 - 0.4304 \cdot A_t^3$

For both data vintages and both quantification models, all estimated parameters exhibit correct signs and are different from zero at 0.01 significance level. RESET test allows to accept functional form of all quantification models as adequate, and coefficients of determination of the models are acceptable. To find basis for selecting either Anderson's or Thomas' models for further analysis, let us note that correlation coefficients between explanatory variables in Anderson's equations, both based on RTV and EoS data, are equal to approximately − 0.87. High degree of multicollinearity in Anderson's models allow to select Thomas' equations as more reliable.

Estimation results do not confirm the preliminary hypothesis that final (EoS) datasets are better suited to modeling assessments of survey respondents. Models estimated for two data vintages are very similar, both from statistical point of view and taking into account their economic interpretation.

To summarize, comparison of regression quantification models across data vintages does not provide immediate recommendations as to whether RTV or EoS data should be used in quantification procedures. In section 4, analysis is continued with expectations series constructed on the basis of the two data vintages.

TESTS OF PROPERTIES OF EXPECTATIONS

In this section, unbiasedness and weak-form orthogonality of expectations are tested. These properties are typically verified within the framework of Rational Expectations Hypothesis, and have been previously analyzed for Polish business survey respondents [see Tomczyk 2011 for review of literature]. Nonetheless, tests of rationality of expectations in Poland have failed to provide conclusive results. Whether expectations on production, prices, employment and general business conditions can be considered rational or not depends on various factors, including sample size, frequency of available data, empirical methods employed, and type of variables included in the analysis. No consistent results on rationality (or, more precisely, its fundamental components: unbiasedness and orthogonality of expectations errors to widely available information) emerge from the literature. Nardo [2003] gives one likely reason for this impasse: “The presence of measurement error in the quantified data is certainly reflected in the general disappointing performance of the standard tests of rationality in the applied literature.” (p. 658) In this section, another possible reason related to data quality in addressed, that is, the issue of selecting appropriate data vintage for empirical analysis of expectations time series.

On the basis of Thomas’ quantification model, expectation series for both data vintages have been constructed. It is assumed that one-month observed changes and three-month expected changes in production are described by the same regression parameters. This simplification constitutes a substantial weakness of regression method, shared by all commonly used quantification methods. It cannot be tested, however, on the basis of dataset available from the RIED business tendency survey as detailed data on individual survey respondents would be required for this purpose.

Two expectations time series have been constructed, that is:

$$E_t^{RTV} = 0.124058 - 0.433176 \cdot P_t^{3-RTV} \tag{7}$$

for real time data and

$$E_t^{EoS} = 0.123343 - 0.430448 \cdot P_t^{3-EoS} \tag{8}$$

for end-of-sample data. To test for unbiasedness, I employ procedure based on unit root tests of expectations and corresponding observed time series [see Liu, Maddala 1992, Maddala, Kim 1998, Da Silva Lopes 1998] which has been extensively used in empirical tests of rationality of expectations. Results of the Augmented Dickey-Fuller test of nonstationarity of expectations series (E_t^{RTV}, E_t^{EoS}) and observed changes in industrial production ($P_t^{RTV\text{-}AV}$, $P_t^{EoS\text{-}AV}$) are presented in

Table 1. All test equations have been estimated with a constant and maximum lag set to 12 on the basis of the modified AIC criterion.

Table 1. Results (*p*-values) of ADF test for expectations and observed production series

		Levels	First differences	Degree of integration
Expectations series	E_t^{RTV}	0.5581	0.0000	I(1)
Observed variable	P_t^{RTV-AV}	0.4237	0.0000	I(1)
Expectations series	E_t^{EoS}	0.3494	0.0000	I(1)
Observed variable	P_t^{EoS-AV}	0.4298	0.0000	I(1)

Source: own calculations

It is clear from Table 1 that all series are integrated of order one. Preliminary condition for expectations series being unbiased predictors of observed series is therefore met, and subsequent conditions may be tested: whether expectations and realized changes in production are cointegrated, and whether the cointegrating parameter is equal to 1 [see Da Silva Lopes 1998]. The following equations are therefore estimated:

$$P_t^{RTV-AV} = \lambda_1 + \mu_1 \cdot E_{t-3}^{RTV} + \varepsilon_{1t} \tag{9}$$

and

$$P_t^{EoS-AV} = \lambda_2 + \mu_2 \cdot E_{t-3}^{EoS} + \varepsilon_{2t}, \tag{10}$$

in which explanatory variables have been lagged three months to account for the 3-month forecast horizon used in RIED business tendency surveys. Models have been estimated by OLS with HAC standard errors. Results of the ADF test for residuals in models for both data vintages, and of the test of linear restriction reflecting the postulated cointegrating vector, are presented in Table 2.

Table 2. Cointegrating regressions

	p-value for ADF test of residuals	*p*-value for restriction
Real time data	$p = 0.3561$	H_0: $\mu_1 = 1$ in (9) $p = 0.0000$
End-of-sample data	$p = 0.0000$	H_0: $\mu_2 = 1$ in (10) $p = 0.0000$

Source: own calculations

In case of real time data, null hypothesis of nonstationarity of the residuals in equation (9) cannot be rejected, that is, expectations and corresponding observed changes in production are not cointegrated. For end-of-sample data, however, null hypothesis is rejected at every typical significance level. It follows that expectations and observed changes in production are in fact cointegrated for series based on the end-of-sample data. Yet, the null hypothesis of cointegrating parameter being equal to one is rejected, and consequently neither of the data vintages lead to unbiased expectations of changes in production. To summarize: there is a notable difference between RTV and EoS data vintages: a cointegrating relation exists only for EoS data. In this case, there is a stable linear combination (that is, expectations and observed series do not diverge in the long run) but it does not support the hypothesis of unbiasedness of expectations.

Unbiasedness tests are considered to be very sensitive to measurement errors and are often supplemented with tests of orthogonality (sometimes also called informational efficiency) of expectations errors with respect to freely available information [see Pesaran 1989, Da Silva Lopes 1998]. Tests of orthogonality are classified as weak, when information set includes only lagged values of variable being forecasted, or strong, when the information set contains additional exogenous variables. I propose to test weak-form orthogonality of expectations errors with respect to production volume data lagged up to three months. I believe that this sets the upper limit on information set of business tendency survey respondents who are not professional forecasters.

The orthogonality hypothesis for RTV data may be therefore written as follows:

H_0: $\kappa_1 = \kappa_2 = \kappa_3 = 0$,

where

$$P_t^{RTV-AV} - E_{t-3}^{RTV} = \kappa_0 + \kappa_1 \cdot P_{t-1}^{RTV-AV} + \kappa_2 \cdot P_{t-2}^{RTV-AV} + \kappa_3 \cdot P_{t-3}^{RTV-AV} + \theta_{1t}, \qquad (11)$$

and for end-of-sample data as

H_0: $\omega_1 = \omega_2 = \omega_3 = 0$,

where

$$P_t^{EoS-AV} - E_{t-3}^{EoS} = \varpi_0 + \varpi_1 \cdot P_{t-1}^{EoS-AV} + \varpi_2 \cdot P_{t-2}^{EoS-AV} + \varpi_3 \cdot P_{t-3}^{EoS-AV} + \theta_{2t}. \qquad (12)$$

Equations (11) and (12) have been estimated by OLS with HAC standard errors. Since three lagged variables are used and therefore multicollinearity of explanatory variables may pose a problem, Variance Inflation Factors are also verified, and found to be equal to 1.18 – 1.22 and to indicate absence of serious multicollinearity. Results of orthogonality tests are presented in Table 3.

Table 3. Results of orthogonality tests

	p-value for restriction
Real time data	H_0: $\kappa_1 = \kappa_2 = \kappa_3 = 0$ in (11) $p = 0.0000$
End-of-sample data	H_0: $\omega_1 = \omega_2 = \omega_3 = 0$ in (12) $p = 0.0000$

Source: own calculations

From Table 3 it is clear that the null hypothesis of insignificance of explanatory variables is rejected. Expectation errors are therefore not orthogonal to easily available information on changes in production index. It follows that RIED business tendency survey respondents do not efficiently make use of available data; specifically, second and third lags of explanatory variables $P_t^{RTV\text{-}AV}$ and $P_t^{EoS\text{-}AV}$ are statistically significant. It seems that when forming their expectations pertaining to volume of industrial production, business tendency survey respondents do not take data older that one month into account.

CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

In this paper, results of quantification procedures and properties of expectations series obtained for two data vintages are described. Empirical analysis confirms that while there are only minor differences in quantification results with respect to data vintage, properties of expectations time series obtained on their basis do diverge. Specifically, there exists a cointegrating regression for one of the vintages only, that is, end-of-sample data. In this case, expectations and observed changes in industrial production exhibit similar long-run properties. Neither of the expectations series, however, constitutes prediction of changes in production that is unbiased or employs available information efficiently.

The research project on impact of data vintage on properties of expectations is continued with the following points considered for further analysis:

- use of other business tendency survey series to scale Central Statistical Office data,
- extending the test of orthogonality to include additional variables in the information set of survey respondents,
- describing and evaluating extent of data revisions in Research Institute for Economic Development business tendency survey data.

Empirical studies of impact of data revisions on expectations promise to assist economists in drawing more general conclusions on behavior and properties of expectations series, including predictive quality, unbiasedness and efficient use of available information. Based on analysis presented in this paper, data vintage does matter in determining basic properties of expectations time series.

REFERENCES

Anderson O. (1952) The business test of the IFO-Institute for Economic Research, Munich, and its theoretical model, Revue de l’Institut International de Statistique 20, 1–17.

Arnold E. A. (2013) The role of revisions and disagreement in professional forecasts, National Bank of Poland Working Paper No. 153, Warsaw 2013.

Croushore D. (2011) Frontiers of real-time data analysis, Journal of Economic Literature 49, pp. 72-100.

Croushore D. (2012) Forecast bias in two dimensions, Federal Reserve Bank of Philadelphia Working Paper No. 12-9.

Croushore D., Stark T. (2001) A real-time data set for macroeconomists, Journal of Econometrics 105, pp. 111-130.

Da Silva Lopes A. C. B. (1998) On the 'restricted cointegration test' as a test of the rational expectations hypothesis, Applied Economics 30, pp. 269-278.

Keane M. P., Runkle D. E. (1990) Testing the rationality of price forecasts: new evidence from panel data, American Economic Review 80, pp. 714-735.

Koenig E. F., Dolmas S., Piger J. (2003) The use and abuse of real-time data in economic forecasting, The Review of Economic and Statistics 85, pp. 618-628.

Liu P. C., Maddala G. S. (1992) Rationality of survey data and tests for market efficiency in the foreign exchange markets, Journal of International Money and Finance 11, pp. 366-381.

Maddala G. S., Kim I.-M. (1998) Unit Roots, Cointegration, and Structural Change, Cambridge University Press, Cambridge.

Mehra Y. P. (2002) Survey Measures of Expected Inflation: Revisiting the Issues of Predictive Content and Rationality, Federal Reserve Bank of Richmond Economic Quarterly, Vol. 88/3.

Nardo M. (2003) The quantification of qualitative survey data: A critical assessment, Journal of Economic Surveys 17 pp. 645-668.

Orphanides A. (2001) Monetary policy rules based on real-time data, American Economic Review 91, pp. 964-985.

Pesaran M. H. (1989) The Limits to Rational Expectations, Basil Blackwell, Oxford.

Syczewska E. M. (2013) Wpływ aktualizacji danych makroekonomicznych bazy AMECO na dokładność prognoz [Influence of updates in AMECO macroeconomic database on forecast accuracy], Badania Statutowe nr KAE/S/07/13, KAE SGH, Warszawa.

Thomas D. G. (1995) Output expectations within manufacturing industry, Applied Economics 27, pp. 403–408.

Tomczyk E. (2011) Oczekiwania w ekonomii. Idea, pomiar, analiza [Expectations in Economics. Definitions, Measurement, Analysis], Oficyna Wydawnicza SGH, Warszawa.

Tomczyk E. (2013) End-of-sample vs. real time data: perspectives for analysis of expectations, [in:] Walczyk K. (ed.) Expectations and Forecasting, Prace i Materiały Instytutu Rozwoju Gospodarczego SGH No. 93, Warsaw School of Economics, pp. 45-69.

Tomczyk E. (2014) Influence of data vintage on quantification of expectations, Applied Econometrics Papers No. 14-04.

Zarnowitz V. (1985) Recent Work on Business Cycles in Historical Perspective: Review of Theories and Evidence, NBER Working Paper No. 1503.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 133 – 141

MODERNIZATION OF MEANS FOR ANALYSES AND SOLUTION OF NONLINEAR PROGRAMMING PROBLEMS

Alexandr Trunov
Petro Mohyla Black Sea State University, Ukraine
e-mail:trunovalexandr@gmail.com

Abstract: The problems of optimization for nonlinear programming (NLP) with constraints inequalities are considered. Definition of condition-indicator as quantitative criterion of the properties of Lagrange function is justified. Application of indicator to increasing degree of completeness for system in NLP for finance and business problems with constraints inequalities are obtained. The new Lagrange function with square of each component of vector Lagrange multipliers for nonlinear objective function simultaneously with criterion-indicator as a source of additional equations is investigated. The conditions, in which the dimensionality of the vector of strategies and the number of constraints doesn't effects on the uniqueness of the optimization problem solution is received and discussed.

Keywords: optimization strategies, NLP, quantitative criterion-indicator, new Lagrange function, source of additional equations.

INTRODUCTION

Applicability of modern methods of optimal design [Akoff and Sasieny 1971, Ventcel. 1980, Germaier 1971, Zaychenko 2000] to the analysis of socio-economic projects, including systems of management [Kondratenko et al. 2013, Kondratenko et al. 2014, Kondratenko 2015] is defined by capabilities of nonlinear modeling techniques [Trunov 2009]. Modernization of its for finding a solution of the maximization problem with constraints inequality for the objective function, structure of which is dynamically changed in according with proposed optimization strategies and decisions, are considered [Trunov 2011, Kondratenko 2015] as actual tasks. In particular, actuality acquires problems of optimal design, selection of solution, when a dynamic systems such as the production company with a stocks and logistic divisions and trading network, or stationary energy generation company, or system of distribution companies, or underwater vehicles, or mobile robots of production systems are designed and defined as systems of objects economic and technical parameters of which are described by nonlinear models [Trunov 2011]. Under these conditions, the task of formulating performance criteria and expressions for their quantitative evaluation and optimal design and decision-making are particularly important from a general methodological and practical point of view. In modern state of development of the NLP and optimal design and management theory, this scientific analysis carried out by expert's and decision making support systems [Zaychenko 2000, Kondratenko et al. 2013, Kondratenko 2015]. The most common approaches, realized in them, lead to the general problem of nonlinear simulation, which usually solves by the method of Lagrange multipliers. The application of these two approaches: profit maximization or losses minimization problems with constraints inequalities, in depending from the number of variables, leads to the problem of finding solutions of independent nonlinear algebraic equations or systems of nonlinear algebraic equations [Trunov 2011].

According to R. Bellman, Kuhn-Tucker, L. V. Kantorovich [Zaychenko 2000] the problem of constructing a standard search algorithm solution of such problems devoted a lot of work, but despite this the applicability of the known methods of successive approximations such as Newton - Kantorovich, Bellman's quasilinearization is limited by the cases, when the first derivative of the pattern, formed by the action of the operator is not zero [Trunov 1999]. The last condition for a long time limited the applicability of quasilinearization - one from the most effective methods to solution of nonlinear problems, in cases when the operator is strictly monotone and its effect on determination plural forms convex set, with one simple root. However, recurrent approximation method, proposed by the author in [Trunov 1999], doesn't impose constraints on the value of the first derivative of operator actions.

However, despite these significant advantages of the method at the present time its applicability to the problems of modeling and designing is constrained by the number of absent equations in the NLP system [Trunov 2011] and non-monotonicity of objective function or its partials derivatives [Trunov 2013]. Thus, despite the success in solving such equations the main problems are don't solved at the present time:

- Building effective generalized algorithm for solution of NLP problems with constraints inequality;
- Determining the conditions, which allow formulate additional independent equations, which complement of system for the Lagrange multipliers method with constraints inequality.

The main purpose of article is to:

- To explore the possibility of using the fundamental properties of the objective function to finding new conditions for the formation of additional equations to complement of NLP system and provide existence and uniqueness of solution of the problem of NLP.

TASK POINTING

Suppose, that $\overline{X}$ is a vector of system states, with *n* – components, that formed and it is defined in the space of real numbers - R^n. Also suppose, that an objective function - $F(\overline{X})$ and limitations provided in the form constraints of inequalities: $g_j(\overline{X}) \le 0$; $j = \overline{1,m}$ are given. Then Lagrange function has the form [Zaychenko 2000]:

$$L(\overline{X}, \overline{\Lambda}) = F(\overline{X}) + \sum_{j=1}^{m} \lambda_j g_j(\overline{X}) . \tag{1}$$

Search for minimum of Lagrange functions, according to Kuhn-Tucker theorem, leads to the problem of the saddle point. The last, under conditions of linearity for restrictions, in general is reduced to a system of equations in which are separated linear and nonlinear operators:

$$L_1(\overline{X}) + L_2(\overline{X}, \overline{\Lambda}) = 0 , \tag{2}$$

where

$$L_1(\overline{X}) = \nabla_x \left[F(\overline{X}) \right]; L_2(\overline{X}, \overline{\Lambda}) = \sum_{j=1}^{m} \lambda_j \nabla_x \left[g_j(\overline{X}) \right]$$

– patterns are created by the action of a linear operator (gradient) in the space of vectors $\overline{X}$ and $\overline{\Lambda}$ determine the formation of the system of equations by the Lagrange multipliers method. In the case of one-dimensionality of the vector $\overline{X}$ and the presence of only one of the constraints in the form of equation (2) are reduced to algebraic equation, which usually nonlinear. In other cases, (2) is reduced to a system of nonlinear equations. Thus, NLP is reduced to the problem of finding a root of nonlinear algebraic equation or roots of systems of nonlinear algebraic equations [Zaychenko 2000]. Generalization of this problem can be done by introducing of the nonlinear objective function $F(\overline{X})$, which is scalar-function of vector and is defined as N times differentiable in the space R^n of n measurable vector strategies $\bar{X}$ under constraints inequalities $g_j\left(\bar{X}\right) \ge 0$, $j = \overline{1, m}$, which will be written compactly due to implementation of vector-functions, i.e. $\bar{G}\left(\bar{X}\right) \le 0$. We pose the problem of maximizing the total benefits from investments under the conditions of existing constraints inequalities:

$$\begin{cases} \max\limits_{\mathrm{X}} \quad F(\overline{X}), \\ \bar{G}\left(\bar{X}\right) \ge 0. \end{cases} \tag{3}$$

and in other cases problem of minimizing, for example, a total loses:

$$\begin{cases} \min\limits_{\bar{X}} \quad F_L(\overline{X}), \\ \bar{G}\left(\bar{X}\right) \le 0. \end{cases} \tag{4}$$

Let's notice, subscript L is indicated changes in general system for problems of minimizing. Assume that each component of the vector-function constraints also are defined as differentiated N times in half-space of real numbers - R^n and allowable region S is formed by intersection of half-space, each of which is the inequality formed by the component of the vector-function of constraints and is the convex polyhedron. Applying to the problem of the Kuhn-Tucker theorem, draw up a Lagrange function [Zaychenko 2000]:

$$L(\bar{X},\bar{\Lambda}) = F(\overline{X}) + \bar{\Lambda}^T\bar{G}\left(\bar{X}\right) \tag{5}$$

and served system of equations maximization problem with constraints inequalities:

$$\begin{cases} \nabla_x L(\bar{X},\bar{\Lambda}) + \bar{V} = 0, \bar{X}^T\bar{V} \ = 0, \\ \nabla_\lambda L(\bar{X},\bar{\Lambda}) - \bar{W} = 0, \bar{\Lambda}^T\bar{W} \ = 0, \end{cases} \tag{6}$$

or minimization problem with constraints inequalities

$$\begin{cases} \nabla_x L(\bar{X},\bar{\Lambda}) - \bar{V} = 0, \bar{X}^T\bar{V} \ = 0, \\ \nabla_\lambda L(\bar{X},\bar{\Lambda}) + \bar{W} = 0, \bar{\Lambda}^T\bar{W} \ = 0, \end{cases}$$

where are indicated:

$\bar{V} = \left[V_1,..V_i,...V_n\right]^T$: $\bar{W} = \left[W_1,..W_j,...W_m\right]^T$; $\bar{\Lambda} = \left[\lambda_1,..\lambda_j,...\lambda_m\right]^T$ - additional relevant n and m - dimensional vectors strategies and vector m - dimensional of Lagrange multipliers. System (6) contains $n+m+2$ equations with $2(n+m)$ unknowns. To increase degree of completeness system, as it is required in work [Zaychenko 2000], there is proposed approximation by MRA [Trunov 2011]. This approach increased number of equations till $2n+m+2$. Certainly that searching for independent and additional m equations, which will be to complete algebraic system (6), is an actual task [Trunov 2013].

JUSTIFICATION OF DEFINITION FORMULAS OF INDICATOR AND SOLUTION OF THE OPTIMIZATION PROBLEM WITH NONLINEAR CONSTRAINTS INEQUALITIES

Realization of this goal is challenging solution of problem, because to create the overalls principles of formation of additional equations it means to find out an conditions or criterions, based on the properties of the objective function and vector function of constraints as a fundamental properties. However, this problem currently not solved, what is stipulated of the difficulties of its nature [Mokhtar et al. 2006].

The search of fundamental properties of function and using them as indicators for solution of finance and business problems by means of NLP with constraints inequalities is discussed in work [Trunov 2011]. Application of comparing operator based on properties of function and derivatives of first and second orders is demonstrated as effective indicators of function states at the point of current approach of vector strategies [Trunov 2013]. Lets, consider another type of indicator for our purposes.

Pointing the goal: theoretically to justify expression of indicator. Define $f(\overline{X})$ in the space of real numbers R^n, as a function of *n* component of the state vector in a neighborhood of approach $\overline{X_n}$, assume that $f(\overline{X})$ is continuous and differentiated three times, while assuming that all approaches have one initial value, following to the method of recurrent approximation [Trunov 1999, Trunov 2011], decompose its in series:

$$f(\overline{X_{n+1}}) = f(\overline{X_n}) + \sum_{i=1}^{n} \frac{\partial f(\overline{X_n})}{\partial x_i} \Delta x_i + \frac{1}{2} \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f(\overline{X_n})}{\partial x_j \partial x_i} \Delta x_j \Delta x_i + R(\overline{X_n}). \quad (7)$$

Finding partial derivative of (7), with respect to one of the variable x_i involved and the other variable the components of the vector $\overline{X}$ being treated, for example, as constants we can be put in the form:

$$\frac{\partial f(\overline{X_{n+1}})}{\partial x_i} = \sum_{i=1}^{n} \frac{\partial f(\overline{X_n})}{\partial x_i} + \frac{1}{2} \frac{\partial}{\partial x_i} \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f(\overline{X_n})}{\partial x_j \partial x_i} \Delta x_j \Delta x_i + \frac{\partial R(\overline{X_n})}{\partial x_i}. \quad (8)$$

Assuming, that in the direction of x_i - components of vector of the objective function becomes local extremum in according of it necessary condition is written:

$$\frac{1}{2} \frac{\partial}{\partial x_i} \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f(\overline{X_n})}{\partial x_j \partial x_i} \Delta x_j \Delta x_i + \frac{\partial R(\overline{X_n})}{\partial x_i} = 0 \quad (9)$$

or

$$\frac{\partial^2 f(\overline{X_n})}{\partial x_i^2} \Delta x_i + \frac{\partial^2 f(\overline{X_n})}{\partial x_j \partial x_i} \Delta x_j + \frac{\partial R(\overline{X_n})}{\partial x_i} = 0. \quad (10)$$

Differentiating (8) in the direction of the vector components x_j obtain:

$$\frac{\partial^2 f(\overline{X_n})}{\partial x_j^2} \Delta x_j + \frac{\partial^2 f(\overline{X_n})}{\partial x_j \partial x_i} \Delta x_i + \frac{\partial R(\overline{X_n})}{\partial x_j} = 0. \quad (11)$$

Solving the latter with respect to Δx_i and substituting it to (10) after algebraic transforms establish that the simultaneous execution of necessary conditions of local extremum for Δx_i and Δx_j directions and provided with:

$$\frac{\partial^2 f(\overline{X_n})}{\partial x_i^2}\left[\frac{\partial^2 f(\overline{X_n})}{\partial x_j^2}+\frac{\partial R(\overline{X_n})}{\Delta x_j \partial x_j}\right]-\left[\frac{\partial^2 f(\overline{X_n})}{\partial x_j \partial x_i}\right]^2-\frac{\partial R(\overline{X_n})}{\Delta x_j \partial x_i}\frac{\partial^2 f(\overline{X_n})}{\partial x_j \partial x_i}=0\,. \tag{12}$$

The latter form can be reduced only in two cases:

- $f(\overline{X})$ is the quadratic form;
- the remainder as negligibly small in comparing with another terms, value of which is determined by the maximum of module third order partial derivatives. For these conditions relation (12) can be rewritten:

$$\frac{\partial^2 f(\overline{X_n})}{\partial x_i^2}\frac{\partial^2 f(\overline{X_n})}{\partial x_j^2}-\left[\frac{\partial^2 f(\overline{X_n})}{\partial x_j \partial x_i}\right]^2=0\,, \tag{13}$$

what coinciding with the expression conditions of saddle point. It should be noted that this expression is obtained directly, based only on the properties of continuity and differentiability of objective function.

DISCUSSION OF RESULTS

Let's consider an example of application equation (12) as necessary conditions of simultaneous execution of local extremum in Δx_i and Δx_j directions to Lagrange function. For this case can be written:

$$\left[\frac{\partial^2 F(\overline{X_n})}{\partial x_i^2}+\sum_{i=1}^{m}\lambda_j\frac{\partial^2 g_j(X_n)}{\partial x_i^2}\right]\left[\frac{\partial^2 F(\overline{X_n})}{\partial x_j^2}+\sum_{i=1}^{m}\lambda_j\frac{\partial^2 g_j(X_n)}{\partial x_j^2}+\frac{\partial R(\overline{X_n})}{\Delta x_j \partial x_j}\right]-\left[\frac{\partial^2 F(\overline{X_n})}{\partial x_j \partial x_i}+\sum_{i=1}^{m}\lambda_j\frac{\partial^2 g_j(X_n)}{\partial x_j \partial x_i}\right]^2-$$

$$-\frac{\partial R(\overline{X_n})}{\Delta x_j \partial x_i}\left[\frac{\partial^2 F(\overline{X_n})}{\partial x_j \partial x_i}+\sum_{i=1}^{m}\lambda_j\frac{\partial^2 g_j(X_n)}{\partial x_j \partial x_i}\right]=0 \tag{14}$$

Let's consider other example of application equation (12) at saddle point in $\overline{X}$ and $\overline{\Lambda}$ directions. This case of problem can be applied to solution of finance or economic problems of optimization for NLP with linear constraints inequalities. For this type of problems according with Lagrange function as function of vector strategies and vector of Lagrange multipliers are reduced due to of it properties to

$$\left[\frac{\partial^2 F(\overline{X_n})}{\partial x_i^2}\right]\left[\frac{\partial R(\overline{X_n})}{\Delta x_j \partial \lambda_j}\right]-\left[\frac{\partial g_j(X_n)}{\partial x_i}\right]^2-\frac{\partial R(\overline{X_n})}{\Delta x_j \partial \lambda_j}\left[\frac{\partial\, g_j(X_n)}{\partial \lambda_j\, \partial x_i}\right]=0 \tag{15}$$

or

$$\begin{cases} \dfrac{\partial^2 L(\overline{X_n},\overline{\Lambda})}{\partial \lambda_j^2} = 0, \\ \dfrac{\partial g_j(\overline{X_n})}{\partial x_i} = 0, \qquad i=\overline{1,n} \qquad j=\overline{1,m} \end{cases} . \tag{16}$$

Latest results (16) can be interpreted as condition of searching of solution along of the axis, which is parallel for x_i axis, other result (15) can be interpreted as searching in along of two intersected directions, but for linear function $g_j(\overline{X_n})$ they demonstrate violation (16). That indicates: components of vector Lagrange multipliers had to be excluded from consideration in another words: equation (14), for linear constraints, can generate additional equation for system (6) only for objective function and number of them is limited.

For resolution of this contradiction will be introduced as analog of Lagrange function, with non-linear function $g_j(\overline{X_n})$:

$$L(\overline{X}) = F(\overline{X}) + \sum_{i=1}^{m} \lambda_j^2 g_j(\overline{X}) .$$

For this case under condition $g_j(\overline{X_n}) \geq 0$, and $\dfrac{\partial^2 F(\overline{X})}{\partial x_i^2} \leq 0$, and $\dfrac{\partial^2 g_j(\overline{X})}{\partial x_i^2} \leq 0$ equation (13) transform to operator:

$$D_{ij}(\overline{X}) = \left(\frac{\partial^2 F(\overline{X})}{\partial x_i^2} + \sum_{i=1}^{m} \lambda_j^2 \frac{\partial^2 g_j(\overline{X})}{\partial x_i^2} \right) \left(2 g_j(\overline{X}) \right) - \left(2\lambda_j \frac{\partial\, g_j(\overline{X})}{\partial x_i} \right)^2 \leq 0, \qquad i=\overline{1,n}, j=\overline{1,m} \tag{17}$$

and in the case of existence its simultaneously solution system (6) and equation (17) at point (λ_j^*) will be saddle point for any of i -th component of vector strategies and other than j - th component of vector Lagrange multipliers. Thus, the system (6) can be added by $n \times m$ additional equations, thus now the system will be consists from $(n+m+n\times m)+2$ of equations with $2\ (n+m)$ unknowns, i.e. the dimension of the vector of strategies and vector of the restrictions doesn't effect onto existence and uniqueness of optimization problem solution and the number of constraints, that provides a unique solution doesn't limited by two. This result is demonstrate, that application of new form of Lagrange function simultaneously with indicator - operator (17) is a source of additional equations. It should be noted, that the equations (13) are satisfy the conditions of continuity and uniqueness and consequently to them are preliminary quadratic recurrent approximation for new form of Lagrange function, completed by objective function of the vector strategies and sum of products of square of each components of vector Lagrange multipliers on components of vector constraints, will be satisfy of condition existence and uniqueness, since created by linear operator. Also it should be noted, that for linear constraints and new form of Lagrange function in condition (14) generates additional equations for system (6) of NLP resolved in simple form for value λ_j :

$$\lambda_j \geq \left[\frac{\partial\ g_j(\overline{X})}{\partial x_i}\right]^{-1} \sqrt{\frac{\partial^2 F(\overline{X})}{2\partial x_i^2} g_j(\overline{X})}, \qquad i=\overline{1,n}, j=\overline{1,m} \ .$$

This approach, open new modern instruments for solution of optimization problem and control on the bases of this solutions for social-economy project.

CONCLUSION

1. The application of recurrent approximations of Lagrange function simultaneously with condition overlaid on the value of partial derivatives of first, second and third order can increase the degree of completeness of the NLP system, in cases of monotonicity, differentiability of Lagrange function, but for linear constraints can generate only limited number of additional equations.
2. In the case of linear or non-linear constraints due to obtaining new Lagrange functions with square of each component of vector Lagrange multipliers for non-linear objective function simultaneous with indicator-operator (17) provide the conditions in which the dimensionality of the vector of strategies doesn’t effects on the uniqueness of the solution for optimization problem and the number of constraints, that provides a unique of solution of NLP system, doesn’t limited by two.
3. Condition of saddle point (13), which is usually used is approximate expression, accuracy of which determined by maximum of third order derivative of objective function and only for quadratic form of objective function it take exact form and it’s a source of additional equations.

REFERENCES

Akoff R., Sasieny M. (1971) Bases of Operational Research: Translation from Engl. – M. Wold.

Germaier J. B. (1971) Introduction in the Theory of Operational Research, M.:Science.

Kondratenko Y., Klymenko L., Kondratenko V., Kondratenko G., Shvets E. (2013) Slip Displacement Sensors For Intelligent Robots: Solutions and Models, in Proc. of the 2013 IEEE 7th Int. Conf. on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), Berlin, Vol. 2, Sept. 12-14, pp. 861-866.

Kondratenko Y. P., Sidenko Ie. V. (2014) Decision-Making Based on Fuzzy Estimation of Quality Level for Cargo Delivery, [In book:] Recent Developments and New Directions in Soft Computing, Studies in Fuzziness and Soft Computing 317, Zadeh L. A. et al. (eds), Springer International Publishing Switzerland, pp. 331-344.

Kondratenko Y. P. (2015) Robotics, Automation and Information Systems: Future Perspectives and Correlation with Culture, Sport and Life Science [in:] Decision Making and Knowledge Decision Support Systems, Lecture Notes in Economics and Mathematical Systems, Vol. 675, A. M. Gil-Lafuente, C. Zopounidis (eds.), Springer International Publishing Switzerland, pp. 43–56.

Mokhtar S. Bazaraa, Hanif D. Sherali, C. M. Shetty (2006) Nonlinear Programming: Theory and Algorithms, John Wiley & Sons Inc.

Trunov A. N. (1999) Application of methods recurrence approximation for solution of non-linear problems, Scientific Transaction of MB NU KMA, Mykolayiv, pp.135-142.

Trunov A. N. (2009) Investigation of convergence solution for problems of non-linear programming and optimal design, ISSN 1609-7742, Scientific Transaction of BSSU - Computer Science, Mykolayiv, BSSU, Vol. 117, No. 104, pp. 68-79.

Trunov A. N. (2011) Recurrence approximation in problems of modeling and design // Monografy – Mykolayiv.

Trunov A. N. (2013) Intellectualization of the models' transformation process to the recurrent sequence, European Applied Sciences, Studgard,. Ort Publishing, # 9, Vol. 1, pp. 123-130.

Ventcel E. S. (1980) Operational research: problems, principal, methodology, M.:Science.

Zaychenko J. P. (2000) Operational Research, Text books.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 142 – 151

MACROECONOMIC INDICATORS FORECASTS ACCURACY AND REACTION OF INVESTORS ON THE WSE

Tomasz Wójtowicz[1]
Department of Applications of Mathematics in Economics
AGH University of Science and Technology in Krakow
e-mail: twojtow@agh.edu.pl

Abstract: Every day analysts and news agencies publish forecasts of important macroeconomic indicators. When the announced value of an indicator differs from its forecast, investors must revise their strategies. The strength of investors' reaction depend on the difference between expectations and the true value of the indicator. In this paper we analyze the reaction of investors on the WSE to U.S. macroeconomic news announcements. We compare the strength of the reaction when forecasts are based on information from different financial services.

Keywords: macroeconomic news announcements, WSE, event study

INTRODUCTION

It is well known that publications of various macroeconomic data impact stock markets. In particular, data describing the U.S. economy imply strong reaction of investors all around the world. It is clearly visible in the case of European stock markets, because U.S. macroeconomic news is released during trading hours of stock markets in Europe. The impact of U.S. data is even stronger than impact of news from European economies. It is because news from European economies is mainly released before opening of stock markets or after their closure. This fact is confirmed by Nikkinen and Sahlström [Nikkinen and Sahlström 2004] who study the impact of monthly announcements of CPI, PPI and Unemployment Rate on German and Finnish stock markets on the basis of data from January 1996 to December 1999. They show that the strongest reaction

[1] Financial support for this paper from the National Science Centre of Poland (Research Grant DEC-2012/05/B/HS4/00810) is gratefully acknowledged.

of both markets is implied by announcements about unemployment and PPI in U.S. Moreover, both markets react only on announcements of U.S. data and they are unaffected by domestic news.

These results are extended by Nikkinen et al. [Nikkinen et al. 2006] who analyze reaction of developed and emerging markets in various parts of the world. They show that announcements of some U.S. macroeconomic indicators significantly impact developed European markets, while emerging markets in Central and Eastern Europe (including Poland) remain unaffected. On the other hand, [Cakan et al. 2015] show the strong impact of U.S. news on volatility on emerging markets (including Poland, Russia and Turkey). The effect of U.S. macroeconomic data announcements solely on the Warsaw Stock Exchange is examined by Gurgul et al. [Gurgul et al. 2012]. They show that CPI and Industrial Production announcements significantly impacts daily returns of WIG20, but there is no significant reaction to announcements about unemployment.

The above papers are based on daily data, however, U.S. macroeconomic news announcements mostly impact intraday returns. It is clearly showed by Harju and Hussain [Harju and Hussain 2011] who examine intraday pattern in volatility and returns of CAC40, DAX30, FTSE100 and SMI. They find that U.S. macroeconomic news announcements induce an immediate and significant reaction of European developed markets. 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.

Quite different results concerning European emerging economies are provided by Hanousek et al. [Hanousek et al. 2009] who study reaction of stock markets in the Czech Republic, Hungary and Poland. They show that the strongest reaction of 5-minute returns is observed on the stock market in Prague, while the Warsaw Stock Exchange seems to be unaffected by U.S. macroeconomic news.

Quite different conclusions follow from study of Gurgul et al. [Gurgul et al. 2013]. On the basis of intraday data, they show very strong and immediate reaction of WIG20 to unexpected news from the U.S. economy. Significant changes in the main index of WSE are observed in the first five minutes after announcements about industrial production, durable goods orders, retail sales and nonfarm payrolls. The later implies the strongest reaction. These results are strengthened by Gurgul and Wójtowicz [Gurgul and Wójtowicz 2014] who prove that indices of WSE react significantly to U.S. macroeconomic data even in the first minute after news announcements. Once again, the strongest reaction is implied by Nonfarm Payrolls.

This is in line with results of Suliga and Wójtowicz [Suliga and Wójtowicz 2013] who study reaction of WIG20 to announcements of different indices included in the Employment Report describing the U.S. labor market. The strongest reaction is connected with announced values of Nonfarm Payrolls. It is even much more important to investors than Unemployment Rate. The importance of Nonfarm Payrolls is also underlined by Andersen et al. [Andersen et al. 2007].

In the above papers there are similar definitions of good and bad news. Usually, news is good when an announced value of a macroeconomic indicator is greater than its forecast. News is bad when the value of the indicator is less than expected[2]. However, when the difference between real and expected value of the indicator is small, investors may treat information as in line with expectations. Hence, results of an analysis of the impact of macroeconomic news announcements may depend on definition of good and bad news. In similar way, the choice of a source of macroeconomic forecasts may impacts results of such analysis.

In this paper we analyze the impact of unexpected news implied by publication of the Employment Report by the U.S. Bureau of Labor Statistics on the Warsaw Stock Exchange. We study how investors reaction depends on the difference between announced values of Nonfarm Payrolls and their forecasts published by various news agencies and internet services. There are two main aims of this analysis. First, we examine how large should be the discrepancy between value of the announced indicator and its forecast to describe investors' reaction properly. The second aim, is the comparison of practical value of macroeconomic forecasts published on different websites. To do this we compare reaction of WIG20 returns to Nonfarm Payrolls announcements on the basis of forecasts provided by different financial services.

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

The Employment Report published monthly by the U.S. Bureau of Labor Statistics describes the U.S. labor market in the month prior to release date. It is usually released on the first Friday of the month at 8:30 EST (Eastern Standard Time) i.e. at 14:30 CET (Central European Time)[3]. The Report is one of the most important publication containing macroeconomic data. Its importance comes from the fact that it is usually first official publication in the month that describes U.S. economy. Because it precedes other macroeconomic indicators announcements, values of these indicators can be partially forecasted on the basis of information contained in the Report.

The Employment Report contains four important indicators: Unemployment Rate, Average Hourly Earnings, Average Workweek and Nonfarm Payrolls (NFP). Previous studies [Suliga and Wójtowicz 2013] show that reaction of stock markets depends mostly on unexpected news contained in NPF. Hence, an analysis of the impact of NFP announcements gives the most visible results.

[2] Opposite definitions of good and bad news are applied in the case of unemployment.

[3] 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.

In this paper we study the impact of NFP announcements on 1-minute log-returns of WIG20 on the basis of data from July 2008 to April 2015. In this period, there were 77 announcements of the Employment Report that took place during trading session on the Warsaw Stock Exchange.

In order to compare investors' reaction to NFP announcements when their expectations are based on different sources, we take into account macroeconomic forecasts published by several financial services. Forecasts come from the following websites:

- bloomberg.com,
- briefing.com,
- yahoo.com,
- deltastock.com,
- forexfactory.com,
- investing.com,
- wbponline.com,
- macronext.com.

It should be noted here that macroeconomic forecasts published by Bloomberg are provided by Econoday service. Forecast made by Briefing are published also by Yahoo! Finance whereas data provided by macroNEXT are published by several important Polish financial services, for example by biznes.pl or parkiet.com.

To investigate the impact of NFP announcements on intraday WIG20 returns we apply event study methodology. For each of the announcements we use 1-minute WIG20 returns from a window that starts 185 minutes before the announcement and ends 60 minutes after the announcement. These returns are divided into two groups: a pre-event window that contains first 180 returns and an event window that starts five minutes before the announcement and ends one hour after it. In the whole window we define abnormal returns (AR_t) as differences between observed returns and the average of returns from the pre-event window. The total effect of announced news is better described by cumulative abnormal returns (CAR_t) defined as a sum of abnormal returns from the announcement time to a given time t.

News about NFP has impact on WIG20 if mean of abnormal returns or cumulative abnormal returns are significantly different from zero after announcements. To test the significance of these means in the event window we apply the nonparametric generalized rank test of Kolari and Pynnönen [Kolari and Pynnönen 2011] with a correction for event-implied volatility[4]. This test does not need any assumption about the normality of abnormal returns and it has relatively high power.

[4] See also Gurgul and Wójtowicz [2015] for detailed information about test procedure.

EMPIRICAL RESULTS

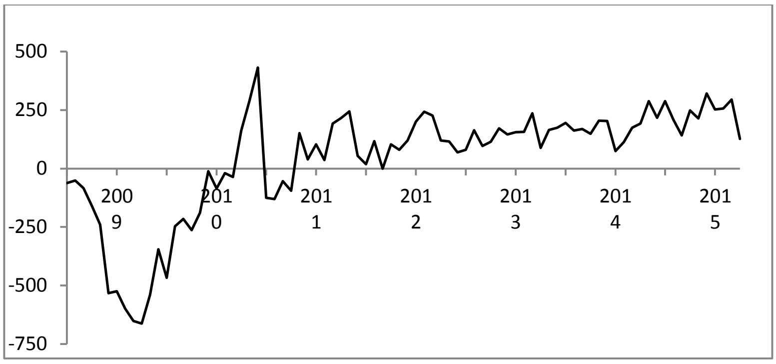
Different definition of good and bad news

In the first part of the empirical analysis we study how the impact of NFP announcements (measured by AR_t and CAR_t) depends on definition of good and bad news. The definition is based on consensus values published by Bloomberg a few days before each announcement.

In the literature, unexpected news is defined usually as the difference between announced value of an indicator and its expected value published few days before the announcement. Good news is when released value of NFP is greater than forecast while NFP smaller than expected is bad news. Good news is followed by positive returns, while bad news is followed by negative returns. However, when difference between forecast and announced value is small, the announcement may be ignored by investors irrespective of the sign of the difference. Determination of a suitable threshold that will separate important and unimportant news may be important to properly describe investors reaction to NFP announcements.

Values of NFP released in the period under study presented in Figure 1 range from -663 000 to 431 000 with median equal to 113 000. However, threshold value depends rather on forecasts' accuracy. Absolute values of differences between NFP values and Bloomberg forecasts in this period range from 0 to 233 000 with median 42 000.

Figure 1. Nonfarm Payrolls changes (in thousands) in the period July 2008 - April 2015



Source: Author's computation

Taking it into account we consider six threshold values, namely (in thousands): 0, 10, 20, ..., 50 and we define good news when the difference between announced NFP and Bloomberg consensus is greater than the given threshold. Bad news is defined analogously, when the difference is smaller than "-threshold". For different values of threshold we report in Table 1 averages of abnormal returns in first minute after news announcements together with the number of announcements in each cluster. Panel A contains results for "good news" clusters, while Panel B contains results for "bad news" clusters. Kolari-Pynnönen tests confirm that all computed means are significant at least at the 1% level. It means that regardless of the definition of good and bad news, reaction of WIG20 is significant in the first minute after news announcements. However, the strength of the reaction depends on this definition. The strongest reaction to good news is observed for threshold 40K, while the strongest reaction to bad news is observed for threshold 30K. To determine whether means of abnormal returns immediately after news announcements depend on assumed threshold we apply bootstrap methods to test the significance of differences between means of abnormal returns reported in Table 1. Means of abnormal returns after good news do not differ significantly. On the other hand, after bad news only significant difference is between mean for thresholds 0 and 30K (approximate p-value is 0.026). It follows that in the case of good news, the choice of threshold and definition of good news has no visible impact on WIG20. However, the choice is important for the definition bad news. Taking into account only announcements that are smaller than forecast more than 30K significantly improves mean of abnormal returns.

Table 1. Averages of abnormal returns in first minute after NFP announcements when good and bad news are defined for different values of threshold

	Threshold					
	0	10K	20K	30K	40K	50K
Panel A: good news						
AR_1 (in %)	0.268	0.299	0.292	0.296	0.299	0.268
number of events	42	38	28	23	22	21
Panel B: bad news						
AR_1 (in %)	–0.262	–0.302	–0.323	–0.398	–0.370	–0.381
number of events	33	29	27	22	18	16

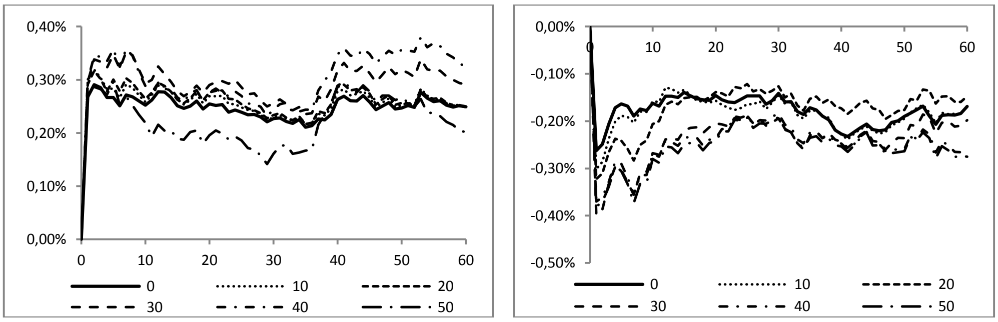
Source: Author's computation

Publication of the Employment Report impacts WIG20 not only in first minutes after announcements. Kolari-Pynnönen tests confirm significance of cumulative abnormal returns in almost the whole event window[5]. Thus, it is justified to examine also differences between cumulative abnormal returns computed for different values of threshold. Figure 2 presents averages of CAR_ts in the first hour after good (left panel) and bad (right panel) news announcements. Differences between means of CAR_ts after good news do not differ significantly for any time t in the event window. On the other hand, means of cumulative abnormal returns

[5] For example, when threshold is equal to 0 CAR_ts are significant at the 5% level in the whole event window after bad news and up to 56 minutes after good news announcements.

computed for threshold 0 and 30K are significantly different up to eight minutes after bad news announcements. The other means can be seen as equal.

Figure 2. Averages of cumulative abnormal returns (CAR_ts) in first 60 minutes after NFP announcements computed for different values of threshold after good (left panel) and bad (right) panel news announcements



Source: Author's computation

Different source of information

In the second part of the empirical analysis we study differences in WIG20 abnormal returns when good and bad news are defined on the basis of NFP forecasts provided by different financial services mentioned in the previous section. As before, we first consider the case when threshold is equal to 0 i.e. news is good if announced value of NFP is simply greater than its forecast. Table 2 presents averages of abnormal returns in the first minute after the announcements[6]. As previously, all means of abnormal returns are significant at least at 1% level. When good news is announced three differences are significant: Bloomberg-Briefing, Bloomberg-macroNEXT and WBP Online-Briefing. After bad news announcements only information from Bloomberg and macroNEXT leads to significantly different means of abnormal returns[7]. The other means of abnormal returns do not differ significantly.

As a comparison, Table 3 presents averages of AR_1s for threshold equal to 30K. As above, Kolari-Pynnönen tests indicate that all the means are significantly different from zero. When good news is announced mean based on Bloomberg forecasts does not differ significantly from the other means. The only significant differences are between results based on Briefing and Yahoo! or between Briefing and WBP Online. On the other hand, after bad news mean abnormal returns

[6] Information from three services (DeltaStock, Forex Factory and Investing) give the same "good news" and "bad news" clusters. Hence, we present their averages in one column.

[7] It should be noted here that in the bootstrap procedure not only difference between means is important, but also number of events in clusters and number of common events play important role.

implied by macroNEXT forecasts differ significantly from means based on information from Bloomberg and Forex Factory.

Table 2. Averages of abnormal returns in first minute after NFP announcements when good and bad news are defined on the basis of forecasts provided by different financial services

	Financial service					
	Bloomberg	Briefing.com	Yahoo!	DeltaStock Forex Factory Investing	WBP Online	makroNEXT
Panel A: good news						
AR_1	0.268	0.172	0.245	0.226	0.250	0.224
number of events	33	37	35	35	36	35
Panel B: bad news						
AR_1	–0.244	–0.206	–0.249	–0.233	–0.265	–0.231
number of events	44	40	42	42	41	42

Source: Author’s computation

Table 3. Averages of abnormal returns in first minute after NFP announcements when good and bad news are defined for threshold 30K on the basis of forecasts provided by different financial services

	Financial service							
	Bloomberg	Briefing.com	Yahoo!	DeltaStock	Forex Factory	Investing	WBP Online	makroNEXT
Panel A: good news								
AR_1 (in %)	0.296	0.178	0.307	0.284	0.266	0.266	0.312	0.266
number of events	22	23	23	22	21	21	23	21
Panel B: bad news								
AR_1 (in %)	–0.398	–0.343	–0.390	–0.360	–0.397	–0.353	–0.390	–0.352
number of events	23	26	25	25	24	25	25	24

Source: Author’s computation

When we compare total impact of unexpected news based on information provided by the financial services under study, two main conclusions arise. First, for threshold 0 the differences between means of CAR_ts after bad news are almost indistinguishable with the only exception CAR_ts implied by forecasts from

Briefing. Means based on Briefing forecasts diverge from the other means towards zero. Similar pattern is observed after bad news announcements.

When we consider threshold equal to 30, means based of Briefing forecasts are significantly different after good news. Bad news are followed by almost identical means irrespective on the source of information.

CONCLUSIONS

In this paper we study how results of the analysis of stock market reaction to U.S. macroeconomic news announcements depend on definition of good and bad news. We also study the robustness of such analysis to the choice of a source of macroeconomic forecasts. The empirical analysis in the paper is performed on example of Nonfarm Payrolls announcements between July 2008 and April 2015. On the basis of 1-minute log-returns of WIG20 we show that reaction of investors on WSE is significant in first minutes after news announcements irrespective of the definition of good and bad news. However, the strength of the reaction depends on what we define as unimportant news. The best results are obtained when threshold is between first quartile and the median of absolute values of differences between announced and forecasted values of a macroeconomic indicator. In general, the conclusion about the significance of reaction to news announcements also do not depend on source of forecasts. However, application of data from some financial services can lead to significantly different means of cumulative abnormal results.

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.

Cakan E., Doytch N., Upadhyaya K. P. (2015) Does U.S. macroeconomic news make emerging financial markets riskier?, Borsa Istanbul Rewiev, Vol. 15(1), pp. 37-43.

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, Vol. 22, pp. 795–817.

Gurgul H., Wójtowicz T. (2015) The Response of Intraday ATX Returns to U.S. Macroeconomic News, Finance a úvěr – Czech Journal of Economics and Finance, Vol. 65(3), pp. 230-253.

Hanousek J., Kočenda E., Kutan A. M. (2009) The reaction of asset prices to macroeconomic announcements in new EU markets: evidence from intraday data, Journal of Financial Stability, Vol. 5(2), pp. 199-219.

Harju K., Hussain S. (2011) Intraday seasonalities and macroeconomic news announcements, European Financial Management, Vol. 17, pp. 367–390.

Kolari J., Pynnönen S. (2011) Nonparametric rank tests for event studies, Journal of Empirical Finance, Vol. 18, pp. 953-971.

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. XVI, No. 2, 2015, pp. 152 – 158

A CONFIDENCE INTERVAL FOR PROPORTION IN FINITE POPULATION DIVIDED INTO TWO STRATA: A NUMERICAL STUDY

Wojciech Zieliński
Department of Econometrics and Statistics
Warsaw University of Life Sciences – SGGW
e-mail: wojciech_zielinski@sggw.pl

Abstract: Consider a finite population of *N* units. Let θ∈(0,1) denotes the fraction of units with a given property. The problem is in interval estimation of θ on the basis of a sample drawn due to the simple random sampling without replacement. Suppose, that the population is divided into two (disjoint) strata. In the paper the confidence interval for θ is proposed based on samples from two strata.

Keywords: confidence interval, sample size, fraction, finite population

The problem of the interval estimation of the fraction (proportion) θ is very old. The first solution was given by Clopper and Pearson [1934] and since then many authors deals with the problem. An exhaustive presentation of the problem along with the very rich literature may be found in the textbook by Koronacki and Mielniczuk [2009]. Presented solutions are valid for infinite populations. In many applications (economic, social, etc.) we deal with the finite population, so we are interested in interval estimation of θ in such finite populations. Remarks on differences in statistical inference in infinite and finite populations may be found in Bracha [1996].

Consider a population $U = \{u_1, \dots, u_N\}$ containing the finite number *N* units. Let *M* denotes an unknown number of objects in population which has an interesting property. We are interested in an interval estimation of *M*, or equivalently, the fraction $\theta = \frac{M}{N}$. The sample of size *n* is drawn due to the simple random sampling without replacement. Let ξ be a random variable describing a number of objects with the property in the sample. On the basis of ξ we want to construct a confidence interval for θ at the confidence level δ.

The random variable ξ has the hypergeometric distribution [Johnson and Kotz 1969, Zieliński 2010]

$$P_{\theta,N,n}\{\xi = x\} = \frac{\binom{\theta N}{x}\binom{(1-\theta)N}{n-x}}{\binom{N}{n}}$$

for integer x from the interval $\langle max\{0, n-(1-\theta)N\}, min\{n, \theta N\}\rangle$. Let $f_{\theta,N,n}(\cdot)$ be the probability distribution function, i.e.

$$f_{\theta,N,n}(x) = \begin{cases} P_{\theta,N,n}\{\xi = x\}, & for\ integer\ x \in \langle \max\{0, n-(1-\theta)N\}, \min\{n, \theta N\}\rangle \\ 0, & elsewhere \end{cases}$$

and let

$$F_{\theta,N,n}(x) = \sum_{t \le x} f_{\theta,N,n}(t)$$

be the cumulative distribution function of ξ. The CDF of θ may be written as

$$1 - \frac{\binom{n}{x+1}\binom{N-n}{\theta N-x-1}}{\binom{N}{\theta N}} \cdot \ {}_3F_2[\{1, x+1-\theta N, x+1-n\}, \{x+2, (1-\theta)N+x+2-n\}; 1],$$

where

$${}_3F_2[\{a_1, a_2, a_3\}, \{b_1, b_2\}; t] = \sum_{k=0}^{\infty} \left(\frac{(a_1)_k (a_2)_k (a_3)_k}{(b_1)_k (b_2)_k}\right)\left(\frac{t^k}{k!}\right)$$

and $(a)_k = a(a+1)\cdots(a+k-1)$.

A construction of the confidence interval at a confidence level δ for θ is based on the cumulative distribution function of ξ. If $\xi = x$ is observed then the ends

$$\theta_L = \theta_L(x-1, N, n, \delta_1) \text{ and } \theta_U = \theta_U(x, N, n, \delta_2)$$

of the confidence interval are the solutions of the two following equations

$$F_{\theta_L,N,n}(x-1) = \delta_1, F_{\theta_U,N,n}(x) = \delta_2.$$

The numbers δ_1 and δ_2 are such that $\delta_1 - \delta_2 = \delta$. In what follows we take $\delta_1 = (1+\delta)/2$ and $\delta_2 = (1-\delta)/2$. For $\xi = 0$ the left end is taken to be 0, and for $\xi = n$ the right end is taken to be 1. Analytic solution is unavailable. However, for given x, n and N, the confidence interval may be found numerically.

The hypergeometric distribution is analytically and numerically untractable. Hence different approximations are applied. There are at least two approximations commonly used in applications: Binomial and Normal. But using those approximations may lead to wrong conclusions [c.f Zieliński 2011]. So in what follows the exact distribution of ξ will be used.

Suppose that the population is divided into two strata: $U_1 = \{u_{11}, \dots, u_{1N_1}\}$ and $U_2 = \{u_{21}, \dots, u_{2N_2}\}$. Of course, $N_1 + N_2 = N$ and $U_1 \cap U_2 = \emptyset$. Let θ_1 and θ_2 be fractions of marked out units in the first and the second strata, respectively. The fraction of marked out units in the whole population equals

$$\theta = w_1\theta_1 + w_2\theta_2,$$

where $w_1 = N_1/N$ and $w_2 = N_2/N$.

It is known that for stratified populations it is better (sometimes) to estimate the proportion in the whole population using the information of stratification. Let n_1 and n_2 be the sizes of the samples drawn from the first and second strata due to the simple random sampling without replacement scheme, and let ξ_1 and ξ_2 be the random variables describing a number of "successes" in the first and second sample, respectively. The whole sample size is $n = n_1 + n_2$. The unbiased with the minimal variance estimator of θ is of the form

$$\hat{\theta}_s = w_1\hat{\theta}_1 + w_2\hat{\theta}_2,$$

where

$$\hat{\theta}_1 = \frac{\xi_1}{n_1} \quad \text{and} \quad \hat{\theta}_2 = \frac{\xi_2}{n_2}.$$

The variance of that estimator equals

$$var(\hat{\theta}_s) = w_1^2 \frac{\theta_1(1-\theta_1)}{n_1}\frac{N_1 - n_1}{N_1 - 1} + w_2^2 \frac{\theta_2(1-\theta_2)}{n_2}\frac{N_2 - n_2}{N_2 - 1},$$

while the variance of the estimator

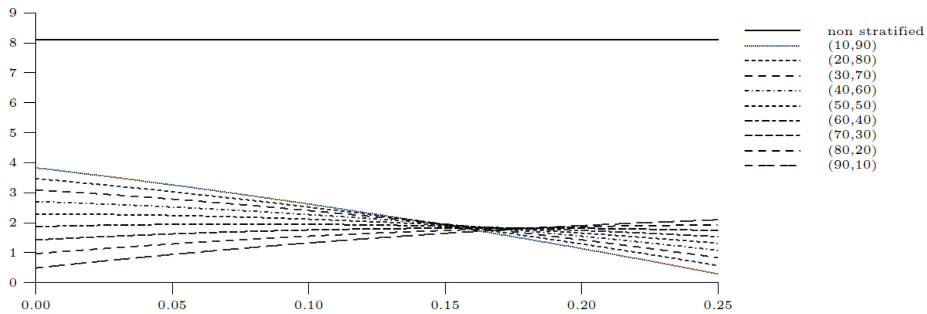
$$\hat{\theta} = \frac{\xi}{n}$$

of θ in non stratified population is

$$var(\hat{\theta}) = \frac{\theta\ (1-\theta\ )}{n}\frac{N\ -n}{N\ -1}.$$

The comparison of those variances (see Figure 1) shows that for given θ and for all values of θ_1 (x-axis) and different allocations (n_1, n_2) the stratified estimator is better than non stratified ($N = 1000$, $N_1 = 400$, $N_2 = 600$, $n = 100$, $\theta = 0.1$). Of course, $\theta_2 = (\theta - w_1\theta_1)/w_2$.

Figure 1. Variances of $\hat{\theta}$ and $\hat{\theta}_s$ for $\theta = 0.1$



Source: own preparation

Now the question is, how to construct a confidence interval for θ on the basis of observed values x_1 and x_2 of r.v's ξ_1 and ξ_2 respectively. It may be expected that such confidence interval may be "better" than the confidence interval in the whole (non stratified) population. Let

$$\theta_L^1 = \theta_L^1(x_1 - 1, N_1, n_1, \gamma_{11}) \ and \ \theta_U^1 = \theta_U^1(x_1, N_1, n_1, \gamma_{21})$$

and

$$\theta_L^2 = \theta_L^2(x_2 - 1, N_2, n_2, \gamma_{12}) \ and \ \theta_U^2 = \theta_U^2(x_2, N_2, n_2, \gamma_{22})$$

be confidence intervals for θ_1 and θ_2 respectively. The confidence levels of those intervals are γ_1 and γ_2, i.e.

$$\gamma_{11} - \gamma_{21} = \gamma_1 \ \ and \ \ \gamma_{12} - \gamma_{22} = \gamma_2.$$

Consider the interval with the ends

$$\theta_L^s = w_1\theta_L^1 + w_2\theta_L^2 \ \ and \ \ \theta_U^s = w_1\theta_U^1 + w_2\theta_U^2.$$

The interval above may be considered as a confidence interval for θ constructed on the basis of two samples drawn from two strata.
The confidence level of the above interval equals

$$P_\theta\{\theta \in (\theta_L^s, \theta_U^s)\} = \frac{1}{H}\sum_{\theta_1=L}^{U}\sum_{x_1,x_2} P_{\theta_1}\{\xi_1 = x_1\}P_{\frac{\theta-w_1\theta_1}{w_2}}\{\xi_2 = x_2\}\mathbf{1}_{(w_1\theta_L^1(x_1)+w_2\theta_L^2(x_2), w_1\theta_U^1(x_1)+w_2\theta_U^2(x_2))}(\theta)\ ,$$

where

$$L = max\left\{0, \frac{\theta - w_2}{w_1}\right\}, \qquad U = min\left\{1, \frac{\theta}{w_1}\right\}, \qquad H = min\{1, \frac{\theta}{w_1}\} - max\{0, \theta - \frac{w_{2)}}{w_1}\} + \frac{1}{N_1}$$

and

$$\mathbf{1}_A(\theta) = \begin{cases} 1, & if\ \theta \in A \\ 0, & if\ \theta \notin A. \end{cases}$$

The expected length of the confidence level equals

$$d(\theta) = \sum_{x_1,x_2}(\theta_U^s - \theta_L^s)P_\theta\{\xi_1 = x_1, \xi_2 = x_2\}\mathbf{1}_{(\theta_L^s,\theta_U^s)}(\theta).$$

The main problem is to find γ_1 and γ_2 such that the confidence level of the confidence interval for θ is at least δ. The problem seems to be unsolvable analytically, so an appropriate numerical study was performed.

Numerical study

In the numerical study the following values were employed:

$$N = 1000, \quad N_1 = 400, \qquad N_2 = 600.$$

The overall sample size were taken $n = 100$. As the confidence level of the confidence interval for θ the value $\delta = 0.95$ was taken.

The numerical study had two aims. Firstly, we want to determine values of γ_1 and γ_2 such that the confidence level of the confidence interval for θ is δ . We assume that $\gamma_1 = \gamma_2 = \gamma$.

The second aim was to compare the lengths of the confidence intervals obtained for different allocations of the sample with the length of the confidence interval obtained for the non stratified population.

In Table 1 there are given confidence levels for different values of γ and different sample allocations. It may be seen that none of the proposed $\gamma's$ gives the prescribed confidence level δ.

Table 1. Confidence levels

(n_1,n_2)	γ=0.85	γ=0.86	(n_1,n_2)	γ=0.80	γ=0.81	(n_1,n_2)	γ=0.80	γ=0.81
(10,90)	0.94255	0.96743	(20,80)	0.94493	0.95373	(30,70)	0.94862	0.95121
(n_1,n_2)	γ=0.83	γ=0.84	(n_1,n_2)	γ=0.81	γ=0.82	(n_1,n_2)	γ=0.81	γ=0.82
(40,60)	0.94359	0.96172	(50,50)	0.94983	0.95568	(60,40)	0.94539	0.95059
(n_1,n_2)	γ=0.83	γ=0.84	(n_1,n_2)	γ=0.83	γ=0.84	(n_1,n_2)	γ=0.83	γ=0.84
(70,30)	0.94998	0.95687	(80,20)	0.94532	0.9519	(90,10)	0.93834	0.95489

Source: the Author's calculations

Because of the discreetness of the r.v's ξ, no more accurate results are available. For example, for allocation (10,90) there exists $0.85 < \gamma^* < 0.86$ such that for $\gamma \le \gamma^*$ the confidence level equals 0.94255 and equals 0.96743 otherwise. For length comparison we took the probability γ such that the confidence level is as near 0.95 as possible. Chosen values of γ for different allocations are given in Table 2.

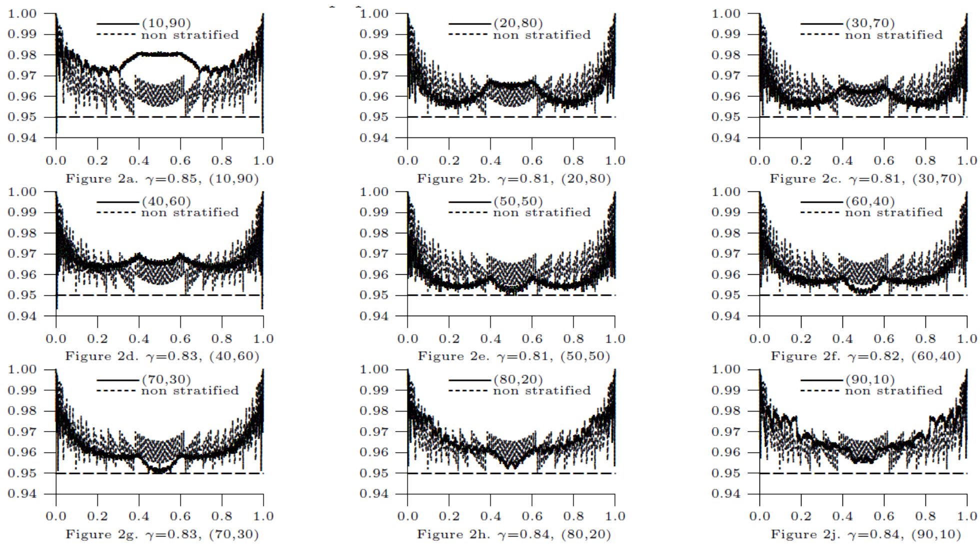
Table 2. Confidence levels

(n_1, n_2)	(10,90)	(20,80)	(30,70)	(40,60)	(50,50)	(60,40)	(70,30)	(80,20)	(90,10)
γ	0.85	0.81	0.81	0.83	0.81	0.82	0.83	0.84	0.84

Source: the Author's calculations

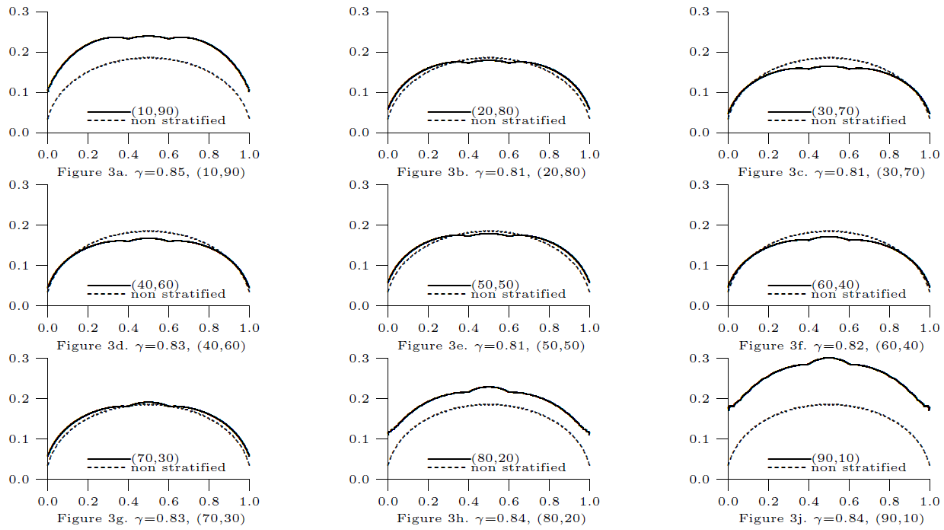
In the following figures there are presented confidence levels of the confidence intervals for stratified population as well as for the non stratified population.

Figure 2. Confidence levels of confidence intervals



Source: the Author's preparation

Figure 3. Comparison of confidence intervals



Source: the Author's preparation

In Figure 3 lengths of the proposed confidence interval are compared with the length of the non stratified confidence interval. It is seen that the use of the information of stratification gives worse results for $n_1 = 10, 80, 90$; comparable lengths for $n_1 = 20, 50, 70$ and shorter confidence interval otherwise.

Final remarks

Due to the Author knowledge confidence intervals for θ in stratified population were never considered, so the presented confidence interval is the first such proposition. There arises some very important questions with respect to the proposed confidence interval. The first one concerns of choosing γ_1 and γ_2: what values they should take on to obtain the prescribed confidence level of the confidence interval for θ in the whole population. The second question is of the optimal sample size and its allocation between two strata. The last but not least problem is the generalization of presented confidence interval for θ to the case of more than two strata.

REFERENCES

Bracha Cz. (1996) Teoretyczne podstawy metody reprezentacyjnej, PWN, Warszawa.

Johnson N. L., Kotz S. (1969) Discrete distributions: distributions in statistics, Houghton Mifflin Company, Boston.

Koronacki J., Mielniczuk J. (2009) Statystyka dla studentów kierunków technicznych i przyrodniczych, WNT, Warszawa.

Neyman J. (1934) On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection, Journal of the Royal Statistical Society, Vol. 97, No. 4, pp. 558-625.

Zieliński W. (2010) Estymacja wskaźnika struktury, Wydawnictwo SGGW, Warszawa.

Zieliński W. (2011) Comparison of confidence intervals for fraction in finite populations, Quantitative Methods in Economics, Vol. XII, pp. 177-182.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XVI, No. 2, 2015, pp. 159 – 169

THE INFLUENCE OF PUBLIC AND PRIVATE HIGHER EDUCATION IN POLAND ON THE ECONOMIC GROWTH OF THE COUNTRY

Joanna Żyra
Institute of Economy, Sociology and Philosophy
Cracow University of Technology
e-mail: jzyra@pk.edu.pl

Abstract: This paper seeks to investigate empirically the relationship between the enrolment at public and private higher education institutions (HEIs) and regional growth in Poland. Based on the panel data of 16 voivodeships for the period of 2000-2011, it is established that an increase in the number of students at both types of HEIs has a positive effect on the regional growth. Regardless of the specifications of the regression model, our results demonstrate, supporting the Nelson–Phelps hypothesis, that accumulation of human capital is one of the factors behind economic growth, with no differences detected between public and private HEIs.

Keywords: higher education, the Nelson–Phelps hypothesis, regional growth, public and private higher education institutions

INTRODUCTION

As a source of human capital, higher education used to be viewed as an important source of economic growth [An and Iyigun 2004, Barro 2002, Daren 2007, Lee 2010, Miller 2007, Sianesi and Reenen 2003]. There are two main approaches to explanation of the macroeconomic relationship between education and output: (i) interpretation of economic growth as the result of an increase in the stock of human capital as a factor of production (it means that the rate of economic growth depends on the *changes* in the education variable) or (ii) looking at the stock of human capital as the most important source of innovation and implementation of new technologies (it means that the rate of economic growth depends on the *level* of the education variable). The former approach is based on the proposals by Lucas [1988] and thus treats education in the similar way as physical capital. As a consequence, the rate of output growth is accelerated by an increase in the rate of student enrolment, which is viewed in the similar way as an investment in human capital. On the other hand, the Nelson–Phelps approach [1966] is based on the assumption that the R&D activities are stimulated by accumulation of immaterial factors of inventions and knowledge resulting from a larger number of educated workers, which in turn contributes to economic growth.

Despite numerous arguments in favour of the benefits of education, including external effects [Bredt and Sycz 2007, Karamalla-Gaiball 2006, Moretti 2004], an implied positive relationship between education and the long-term level of output or its rate of growth is lacking strong empirical support [Benhabib and Spiegel 1994, Islam 1995, Levine and Renelt 1992, Pritchett 2001]. Voices expressing concern regarding a remarkable increase in the number of students in Poland are not rare [Chałasińska-Macukow 2009, Kuciński 2009, Papuzińska 2009, Tomusk 2001], especially in the context of financing public and private HEIs or their location outside large and long-established academic centres. As mentioned by Minkiewicz [2007], quantitative and structural changes in higher education are stemming mainly from higher educational aspirations of the youth and higher returns on education, but low costs of higher education at public HEIs are very important as well.

As found by Żyra [2013] based on annual data on Poland's economy over the period of 1988 – 2010, an increase in the number of students in either dynamic (in first differences) or static form (in levels) leads to a slower output growth across all fields of study. However, a growth stimulating effect of the number of students has been confirmed for the estimates in levels for a shorter sample of regional data over the period of 2000 – 2010, yet a negative effect is still observed for estimates in first differences. A positive relationship between the number of students and regional growth in Poland has been demonstrated by Bronisz and Heijman [2010]. For Poland, various aspects of higher education and research sector growth effects are studied by Florczak [2006] and Welfe [2008]. Considering the ongoing discussion on higher education in Poland, it is of particular interest to estimate education effects across public and private HEIs.

THE ROLE OF EDUCATION IN ECONOMIC GROWTH

The influence of education on growth has been studied from many academic angles. Endogenous growth models imply increasing returns on human capital. Lucas [1988] demonstrated how investments in education could contribute to acceleration of economic growth. In a similar way, Romer [1990] identified human capital with research activities. As suggested by Psacharopoulos and Patrinos [2004 a], both Lukas and Romer growth models imply that (i) output is not constrained by the constant return to scale assumption, contrary to the neoclassical

Solow model, and (ii) ”knowledge” is considered a public good, with favourable educational externalities. Mueller [2007] suggests that such externalities should be viewed as an argument in favour of expansion of higher education, regardless of considerable expenses on HEIs.

There are numerous arguments that positive educational externalities are created in a situation of sharing knowledge and skills during formal and informal contacts or skill-based endogenous technological progress, which brings about non-monetary benefits as well [Moretti 2004]. Educational externalities could also be found in other forms, such as better health conditions, stronger social relationships, rational election choices etc., creativity or new approaches to professional activities [Bredt and Sycz 2007]. Education contributes to a wide range of social activities, which makes it easier to cooperate on the basis of positive interpersonal relationships [Karamalla-Gaiball 2006]. Based on the empirical findings for Poland, Bronisz and Heijman [2010] argue that both economic growth and regional competitiveness are positively correlated with the level of social capital.

Marattin [2007] established that economic growth is stimulated by accumulation of human capital, no matter how it is financed – with private or public funds:

$$y = \frac{\gamma}{(1+\beta+\gamma)(1+n)^2}\big(1+(1-\tau_k)r\big)(1-\tau_w)w_t E^{\eta} h_t^{1-\eta}, \tag{1}$$

where y is the rate of economic growth, γ is the share of expenditure on education by parents, β is a discount factor, n is the rate of population growth, τ_k and τ_w are tax rates on labour and capital, respectively, r is the interest rate, w_t is wage, E_t is the expenditure on education, h_t is the stock of human capital.

The model implies that private and public expenditure on education is complementary. Thus it is expedient to tax labour in order to increase the total expenditure on education. Taxation of capital promotes growth only if the share of households in the total expenditure on education, η, is high enough and the share of capital in the production function is between 0.3 and 0.4, which is relevant to the actual figures for industrial countries.

According to the Nelson–Phelps approach [1966], an increase in the level of education is almost immediately translated into a decrease in the distance between the levels of (1) technology in practice, which measures the best-practice level of technology or the average technology level “embodied” in the representative assortment of capital goods currently being purchased, and (2) theoretical technology, which is the best-practice level of technology that would prevail if technological diffusion were completely instantaneous (it is a measure of the stock of knowledge that is available to innovators). The rate of innovation growth is increased by the level of education and the gap between technology in practice and theoretical technology. As the equilibrium gap is a decreasing function of educational attainment, a higher level of education increases the path of the technology in practice in the long run.

The Nelson–Phelps approach envisages that the rates of productivity growth and technological progress are higher in line with an increase in the number of educated persons, especially university graduates, as it increases the number of potential researchers/inventors in the economy. The payoff to increased educational attainment is greater, the more technologically advanced is the economy. This kind of regularity is supported by empirical findings, which indicate an inverse relationship between the demand for education and amortisation of the stock of physical capital. For technologically backward countries, education allows an import of technologies from advanced countries in attempts to achieve an appropriate rate of productivity growth.

A class of economic growth models relates the impact of education to a certain *threshold level*, which decides on whether a higher level of education leads to a higher output growth or to a stagnation trap [Aghion and Hewitt 1998]. As it is not expedient to invest in education while in the stagnation trap, the economy is stuck indefinitely in the equilibrium with a low output growth. Assuming complementarity between investments in physical and human capital, Bassetti [2009] provides evidence of the same education-motivated stagnation trap within the framework of the Solow model.

Education subsidies are among solutions for avoiding the stagnation trap, providing a possibility of attaining a higher path of economic growth. A positive impact of public educational expenditure on economic growth is found in empirical studies by Barro and Sala-i-Martin [1995] and Baldacci et al. [2008], though Blankenau and Simpson [2004], by calibration of the endogenous growth model, obtain a less optimistic result. A positive effect on growth of public expenditure on education is seriously weakened or even reversed if it is controlled by such factors as the level of government expenditure, composition of taxes and production technologies. Lin [1998] argues that an increase of public expenditure on education, for example in the form of educational subsidies, is useful for accumulation of human capital, though on condition that it does not lead to an increase in the long-term interest rate.

All said, educational expenditure is a necessary but not sufficient factor for achieving a higher output growth rate. The efficiency of education as a growth factor is dependent on such factors as: (i) openness of the economy, (ii) quality of institutions, or (iii) functioning of the labour market.

MODELS OF INTERACTIONS BETWEEN PUBLIC AND PRIVATE UNIVERSITIES

Geiger [1987] identifies three models of interactions between public and private HEIs:

1. *Monopolisation of educational services by several public universities* (usually with high academic credentials). In such a situation, public HEIs attract students on the basis of low tuition costs, being set up in response to a high demand for higher education. It is common that private HEIs are characterized by modest financial capabilities and employment of faculty on the part-time basis.
2. *Parallel coexistence of public and private universities*. Among necessary conditions for that kind of model are such factors as relevant group interests, a unified national educational standard for diplomas issued by HEIs and significant co-financing of private HEIs by the state.
3. *Peripheral status of private HEIs*. In this case private HEIs are marginalized, lacking public financial support and thus being not able to match academic standards of public universities.

As for now, a dynamic development of private HEIs is observed in Australia [Edwards and Ali Radloff 2013] and China [Zha 2006]. Expansion of the Chinese private HEIs is based upon the pillars of strong economic growth and decentralisation policies. However, there is considerable concern related to the equal access to higher education for all social groups.

The main arguments in favour of public education could be summarized as follows: (i) redistributional effects; (ii) using tax revenues for strengthening the foundations of the economic growth (in particular, it is important for developing new technologies, which is not possible outside strong HEIs and requires substantial public expenditure); (iii) complementarity with expenditure borne by households.

Using the *overlapping generations* model, Ayed Zambaa and Ben Hassen [2013] show that public HEIs can be more effective in attaining a higher rate of economic growth. Angelopoulos et al. [2007] present a model that implies favourable externalities in the case of higher expenditure on public education. For the USA, it has been demonstrated that such expenditure stimulates economic growth and the sense of well-being in the society. Although public expenditure is characterized by partial crowding out of private consumption, nevertheless it is expedient to modify the structure of public expenditure in favour of education.

In a slightly different setting, Arcalean and Schiopu [2007] suggest that an increase of public expenditure on education results in “crowding out” of private expenditure on higher education, while there is an increase of expenditure on education in schools. On the other hand, an increase of public expenditure on schooling induces higher expenditure on higher education by households.

Northern Cyprus, suffering from lack of international recognition since 1974, is a good example of using HEIs as a source of economic growth. In the absence of other alternatives, it had been decided to invest in the development of tourism and higher education [Katircioĝlu et al. 2010]. Since the beginning of the 1990s, local HEIs are oriented towards students from Turkey and African countries. Universities organize academic conferences, while being engaged in cultural events and sport competitions as well.

Two arguments in favour of private HEIs are as follows:

- General benefits from private ownership,
- An increase in accumulation of human capital.

As stated by Yamada [2005], in societies with a strong quest for social status privately financed education could improve the allocation of human resources, thus increasing the growth rate in the private finance regime compared with the public finance regime. The explanation is that private cost dissuades *wrong* agents from participating in growth enhancing activities. As a consequence, the allocation of human resources is improved. It is worth noting that under strong preferences for social status the growth rate declines in the public finance regime even though tax revenues are used for productivity augmenting expenditure.

DATA AND STATISTICAL MODEL

The annual data for 16 *voivodships* of Poland over the period of 2000–2011 are used. The panel dataset is balanced, and the number of observations depend on the lags for dependent and independent (explanatory) variables. Educational variable is the number of students at public and private HEIs, according to the information provided by the GUS (Central Statistical Office of Poland).

Our statistical model is as follows:

$$\Delta \log Y_{it} = a_1 \Delta \log S_{it} + a_2 \log S_{it} + a_3 \log I_{it} + a_4 \log W_{it} + \eta_i + \tau_t + \varepsilon_{it}, \tag{1}$$

where Y_{it} is the regional output per capita (in PLN), S_{it} is the number of students per 1000 of population, I_{it} – is the level of investments per 1000 of population (in *PLN*), W_{it} is the nominal wage (in PLN), η_i and τ_t are variables for controlling regional and time effects, ε_{it} is the stochastic factor.

Microeconomic Mincer-style models of wages and some endogenous growth models, for example by Lucas [1988], imply positive returns on *changes* in educational attainment, not in the *level* of education. Such a relationship is not always supported by empirical studies. For example, Benhabib and Spiegel [1994] find that there are no positive effects of changes in educational attainment on economic growth, or that there is even an inverse relationship between education and economic growth, although there exists an expected positive link between the level of education and the rate of GDP growth. Krueger and Lindahl [2001] provide arguments that using *first differences* of educational variable is responsible for weakening of the signal and strengthening of the white noise, which could yield biased estimates of the educational effects upon economic growth. In order to improve empirical assessment of educational effects, it is suggested to include into regression modes the educational variable in both first differences and levels.

Along the lines of endogenous growth models, it is possible to assume that an increase in the number of students, regardless of whether it is in the first differences or in the level of educational variable(the number of students), should contribute to a higher regional growth per capita $(a_1, a_2 > 0)$. However, numerous empirical studies suggest a possibility of negative coefficients, especially for a_1. For example, such results are obtained in the study of educational effects on economic growth in Poland by Żyra [2013]. Changes in educational attainment potentially reflect general equilibrium effects on the national level and eliminate all kinds of effects by permanent technological shocks [Kruger and Lindahl 2001], while the level of educational variable controls the effects of an increase in the stock of human capital.

Variable I_t captures the effects of an increase in the stock of physical capital, as one of the production function components. It is standard for neoclassical growth models (the Solow model) or endogenous growth models (the Lucas and Romer models) to claim that investments contribute to economic growth $(a_3 > 0)$. Kruger and Lindahl [2001] acknowledge that the probability of finding an inverse relationship between education and economic growth is higher in specifications, with investments in physical capital included. Controlling the growth effects by investments in physical capital, it is possible to obtain estimates of the impact of education upon economic growth that are not biased by the possible link between investments in education and physical capital.

The nominal wages W_{it} are for the labour market effects. The impact of wages on economic growth is dependent to a large extent on the interplay between supply and demand for labour. If the wage level is below equilibrium, it brings about an increase in the demand for labour, with a positive link between wages and economic growth to be implied $(a_4 > 0)$. Otherwise it is likely to expect the inverse relationship between wages and economic growth, as the lack of demand would inhibit employment $(a_4 < 0)$. Of course, all considerations of the sign on the coefficient on W_{it} are just the opposite for the supply-driven developments on the labour market.

ESTIMATION RESULTS

In order to estimate the magnitude of educational effects, fixed effects (FE) is used. Our results are presented in Table 1. The baseline regression model of specification I includes only investments in the physical capital, while the extended regression model of specification II corresponds to the full set of explanatory variables from equation (1). According to the R^2 statistics, the FE estimates explain from 29 to 35 percent of changes in regional growth in specifications with the number of students at public HEIs, and from 19 to 33 percent in specification with the number of students and private HEIs. It is clear that the extended specification has a better explanatory power.

If control for investment in physical capital, there is a negative effect of the changes in the number of students (in *first differences*) at public HEIs on regional growth (coefficients are statistically significant at the 1 percent level). For private HEIs, there is a positive and statistically significant link between $\Delta \ln S_{it}$ and regional growth. Regardless of the type of HEIs — public or private, an increase in the number of students (in *levels*) has a positive effect on regional growth, which is compliant with the estimates for nationwide data [Żyra 2013]. The magnitude of the coefficient for the number of students is higher for public HEIs.

Table 1. Determinants of regional economic growth (baseline regression model)

Variables	Public HEIs		Private HEIs	
	I	II	I	II
$\Delta \ln S_{it}$	–0.373 (–6.45***)	–0.253 (–3.76*)	0.070 (2.20*)	0.070 (2.42**)
$\ln S_{it}$	0.108 (2.74***)	0.079 (2.04**)	0.072 (5.21***)	0.041 (2.92*)
$\ln I_{it}$	–0.021 (–2.01**)	–0.088 (–3.77*)	–0.016 (–2.76***)	–0.011 (–5.08*)
$\ln W_{it}$		0.068 (3.18*)		0.097 (5.01*)
Year dummies	Yes	Yes	Yes	Yes
R^2	0.29	0.35	0.19	0.33

Notes: t-statistics in parenthesis: ***, **, * significant at the 1, 5 and 10 percent levels respectively.

Source: Author's calculations

It is confirmed that there is a stronger positive impact on regional growth of the number of students at public HEIs (in levels), which could be indicative of a better position in respect to accumulation of human capital. In the context of the Nelson–Phelps approach, our results could be explained by the fact that study programmes in the fields of technology and natural sciences are predominantly the domain of public HEIs, while being hardly noticed at private HEIs due to substantially higher costs of studying. Estimates for the educational effect in the extended specification are similar to those of the baseline model (Table 1), as the coefficients on $\Delta \ln S_{it}$ favour private HEIs.

Our results can reflect the selection problem. To put it in the simplest way, students at private HEIs are those who otherwise would have been unemployed or employed as low-skilled labour being unable to enrol at public HEIs. A decrease in the number of such persons as they enrol at private HEIs may contribute to an increase in the stock of human capital, thus leading to a higher rate of economic growth. Another line of reasoning implies higher motivation of students at private

HEIs, as it is necessary to pay tuition fees. One more argument refers to a lower probability of educational mismatches in the case of choosing fields of study at private HEIs. It is quite often that the choice is made by persons who are already employed in a specific profession.

Among other results, investments in physical capital do not seem to stimulate regional growth, running counter to standard predictions of neoclassical and endogenous growth models. Our results provide evidence that investments in physical capital in Poland may not be effective or may be confined only to the long-term growth effects. An increase in the level of nominal wage is a pro-growth factor, as the coefficient on $\ln W_{it}$ is positive and statistically significant at the 1 percent level. Such a result means that the level of nominal wage is below equilibrium, so an increase in the wage level is associated with higher employment.

CONCLUSIONS

This article presents an analysis of the relationship between the number of students at public and private HEIs and regional economic growth in Poland. Although a higher enrolment at the public HEIs is associated with lower private costs of education and a higher stock of high-skilled labour, which is supposed to accelerate economic growth, it should be noted that public expenditure on education can crowd out investments in the stock of physical capital and weaken incentives for *learning for doing* in the workplace, which could be detrimental to economic growth. Among significant benefits of private education for economic growth, the intertemporal transfer of human capital is frequently mentioned.

Based on the annual panel data of 16 voivodeships for the period of 2000-2011, it has been found – with the use of FE estimator – that an increase in the number of students (in *levels*) at public HEIs is of stronger positive effect on regional economic growth if compared with the student enrolment at private HEIs. For private HEIs, this result is confirmed for the estimates of educational variable in *first differences*. Our findings are robust to changes in specification of the regression model that controls for the level of nominal wage. Regardless of the specification of regression model, our results are in favour of the Nelson–Phelps approach, which implies that accumulation of the human capital stock is one of the factors behind economic growth. Positive growth effects are associated with an increase in the number of students at either public or private HEIs.

REFERENCES

Aghion P., Hewitt P. (1998) Endogenous Growth Theory, MIT Press, Cambridge, MA.

An G., Iyigun M. (2004) The export skill content, learning by exporting and economic growth, Economics Letters, Vol. 84, Issue 1, pp. 29–34.

Angelopoulos K., Malley J., Philippopoulos A. (2007) Public Education Expenditure, Growth and Welfare, CESifo Economic Studies Conference on Productivity and Growth (22-23 June, 2007), CESifo, Munich.

Ayed Zambaa A., Ben Hassen L. (2013) Education financing and growth using an overlapping generations model: A theoretical perspective, Economic Annals, Vol. LVIII, No. 197.

Baldacci E., Clements B., Gupta S., Cui Q. (2008) Social Spending, Human Capital, and Growth in Developing Countries, World Development, Vol. 36, No. 8, pp. 1317–1341.

Barro R. (2002) Quantity and Quality of Economic Growth [in:] Economic Growth: Sources, Trends, and Cycles, N. Loayza, R. Soto, K. Schmidt-Hebbel (eds.), Central Bank of Chile, Santiago, Edition 1, Vol. 6, Chapter 5, pp. 135–162.

Barro R., Sala-i-Martin X. (1995) Economic Growth, McGraw-Hill, N.Y.

Bassetti T. (2009) Education and Poverty in a Solow Growth Model [in:] Long-Run Growth, Social Institutions and Living Standards, N. Salvadori, A. Opocher (eds.), Edward Elgar, Cheltenhem, UK, pp. 133–157.

Benhabib J., Spiegel M. (1994) The Role of Human Capital in Economic Development: Evidence From Aggregate Cross-Country Data, Journal of Monetary Economics, Vol. 34, Issue 2, pp. 143–174.

Blankenau W., Simpson N. (2004) Public education expenditures and growth, Journal of Development Economics, Vol. 73, No. 2, pp. 583–605.

Bredt J., Sycz C. (2007) Education and Economic Growth: A Literature Review, Working Paper No. 50, Department of Education, Training and the Arts of Queensland Government, Brisbane.

Bronisz U., Heijman W. (2010) The Relationship Between Social Capital and Regional Competitiveness, Applied Studies in Agrobusiness and Commerce, No. 1, pp. 43–48.

Chałasińska-Macukow K. (2009) Uniwersytety są traktowane jak huty, Rzeczpospolita, 22 grudnia.

Daren C. (2007) Education and Economic Growth: Is There a Link?, MPRA Paper No. 18176, Munich Personal RePEc Archive, Munich.

Edwards D., Radloff A. (2013) Higher Education enrolment growth, change and the role of private HEPs, Australian Council of Educational Research, Hawthorn, VA.

Florczak W. (2006) Modelowanie sektora nauki i edukacji – makroekonomiczne ujecie systemowe [w:] Innowacyjność w budowaniu gospodarki wiedzy w Polsce, E. Okoń-Horodyńska, S. Pangsy-Kania (eds), Instytut Wiedzy i Innowacji, Warszawa, str. 16–32.

Geiger R. (1987) Privatization of Higher Education: International Trends & Issues, International Council for Educational development, Princeton, NJ.

Islam N. (1995) Growth Empirics: a Panel Data Approach, Quarterly Journal of Economics, Vol. 110, No. 4, pp. 1127–1170.

Karamalla-Gaiballa E. (2006) Oddziaływanie edukacji na przedsiębiorstwo – kierunki i uwarunkowania, Nauka i Szkolnictwo Wyższe, nr 1/27, str. 56–68.

Katircioĝlu S., Fethi S., Kilin C. (2010) A Long Run Equilibrium Relationship Between International Tourism, Higher Education and Economic Growth in Northern Cyprus, Ekonomska Istraživanja, Vol. 23, No. 1, pp. 86–96.

Krueger A., Lindahl M. (2001) Education for Growth: Why and For Whom?, Journal of Economic Literature, Vol. XXXIX, No. 4, pp. 1101–1136.

Kuciński K. (2009) Knowledge as a Factor of Growth and Production, Economics & Business Administration Journal, No. 1, pp. 34–41.

Lee C. (2010) Education and Economic Growth: Further Empirical Evidence, European Journal of Economics, Finance and Administrative Sciences, Vol. 23, No. 8, pp. 161–169.

Levine R., Renelt D. (1992) A sensitivity analysis of cross-country growth regressions, American Economic Review, Vol. 82, No. 4, pp. 942–963.

Lin S. (1998) Government education spending and human capital formation, Economics Letters, Vol. 61, Issue 3, pp. 391–393.

Lucas R. (1988) On the mechanics of economic development, Journal of Monetary Economics, Vol. 22, No. 1, pp. 3–42.

Marattin L. (2007) Optimal Fiscal Policy with Private and Public Investment in Education, University of Bologna, Bologna.

Miller R. (2007) Education and Economic Growth: From the 19th to the 21st Century, CISCO Systems, San Jose.

Minkiewicz B. (2007) Szkolnictwo wyższe [w:] Edukacja dla pracy. Raport o rozwoju społecznym Polska 2007, UNDP, Warszawa, str. 42–46.

Moretti E. (2004) Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data, Journal of Econometrics, Vol. 121, No. 1–2, pp. 175–212.

Mueller N. (2007) (Mis-) Understanding Education Externalities**,** MPRA Paper No. 5331, Munich Personal RePEc Archive, Munich.

Nelson R., Phelps E. (1966) Investment in Humans, Technological Diffusion and Economic Growth, American Economic Review, Vol. 61, No. 1/2, pp. 69–75.

Papuzińska M. (2009) Szkoły z kasą, Polityka, Rok 52, Nr. 40, str. 12–14.

Pritchett L. (2001) Where Has All the Education Gone?, World Bank Economic Review, Vol. 15, No. 3, pp. 367–391.

Psacharopoulos G., Patrinos H. (2004) Human capital and rates of return [in:] International Handbook on the Economics of Education, G. Johnes, J. Johnes (eds.), Edgar Elgar, Cheltenham, UK, pp. 1–57.

Romer P. (1990) Endogenous technological change, Journal of Political Economy, Vol. 98, No. 10, pp. S71–S102.

Sianesi B., van Reenen J. (2003) The Returns to Education: Macroeconomics, Journal of Economic Surveys, Vol. 17, Issue 2, pp. 157–200.

Tomusk V. (2001) Enlightenment and minority cultures: Central and East European higher education reform ten years later, Higher Education Policy, Vol. 14, Issue 1, pp. 61–73.

Welfe W. (2008) A Knowledge-Based Economy: New Directions of Macromodelling, International Advanced Economic Research, Vol. 14, No. 2, pp. 167–180.

Yamada K. (2005) Public versus Private Education in an Endogenous Growth Model with Social Status, Economics Bulletin, Vol.15, No. 11, pp. 1–9.

Zha Q. (2006) The Resurgence and Growth of Private Higher Education in China, University of Toronto, Toronto, pp. 54–68.

Żyra J. (2013) Efekty ekonomiczne niedopasowań strukturalnych i kompetencyjnych w obszarze edukacji, Politechnika Krakowska im. Tadeusza Kościuszki, Kraków.

REVIEWERS COOPERATING WITH THE JOURNAL IN 2015

Jarosław Becker
Michał Bernardelli
Zbigniew Binderman
Dariusz Błaszczuk
Jarosław Bojarski
Agata Boratyńska
Bolesław Borkowski
Jadwiga Bożek
Ryszard Budziński
Mariola Chrzanowska
Hanna Dudek
Konrad Furmańczyk
Stanislaw Gędek
Urszula Grzybowska
Mariusz Hamulczuk
Beata Jackowska-Zduniak
Piotr Jałowiecki
Stanisław Jaworski
Dorota Juszczak
Dariusz Kacprzak
Andrzej Karpio
Marek Karwański

Joanna Kisielińska
Marek Kociński
Krzysztof Kompa
Grzegorz Koszela
Jolanta Kotlarska
Monika Krawiec
Justyna Kujawska
Karol Kukuła
Tomasz Kuszewski
Joanna Landmesser
Jarosław Lira
Andrzej Łodziński
Piotr Łukasiewicz
Wanda Marcinkowska-Lewandowska
Iwona Markowicz
Jerzy Marzec
Aleksandra Matuszewska-Janica
Iwona Müller-Frączek
Joanna Muszyńska
Rafik Nafkha
Zdzisław Odrzygódź
Maria Parlińska

Marian Podstawka
Artur Prędki
Ewa M. Syczewska
Alexander Prokopenya
Marcin Rudź
Elżbieta Saganowska
Victor Shevchuk
Stanisław Stańko
Jacek Strojny
Wiesław Szczesny
Emilia Tomczyk
Dorota Witkowska
Jolanta Wojnar
Jacek Wolak
Mirosław Woźniakowski
Adam Zaremba
Jan Zawadzki
Tomasz Ząbkowski
Monika Zielińska-Sitkiewicz
Wojciech Zieliński
Dorota Żebrowska-Suchodolska

Vasily Dikusar (Dorodnitsyn Computing Center RAS, Moscow, Russia)
Paolo Gajo (University of Florence, Italy)
Vasile Glavan (Moldova State University, Chisinau, Moldova)
Vyacheslav S. Kharchenko (National Aerospace University "KhAI", Kharkiv, Ukraine)
Galyna V. Kondratenko (Black Sea State University, Mykolaiv, Ukraine)
Yuriy Kondratenko (Black Sea State University, Mykolaiv, Ukraine)
Vassilis Kostoglou (Alexander Technological Educational Institute of Thessaloniki, Greece)
Yochanan Shachmurove (The City College of The City University of New York, USA)
Ievgen V. Sidenko (Petro Mohyla Black Sea State University, Mykolaiv, Ukraine)
Mirbulat B. Sikhov (al-Farabi Kazakh National University, Almaty, Kazakhstan)
Nikolas Olenev (Dorodnitsyn Computing Center RAS, Moscow, Russia)
Alexandr Trunov (Black Sea State University, Mykolaiv, Ukraine)